

UNIVERSITÀ DEGLI STUDI “ROMA TRE”
DIPARTIMENTO DI ECONOMIA AZIENDALE



XXXVII CICLO DI DOTTORATO DI RICERCA

Mercati, impresa e consumatori
Curriculum “Diritto dell’economia e dei consumatori”

Tesi di dottorato

IL CREDITO BANCARIO NELL’ERA DELL’AI

Tutor:

Prof.ssa Maddalena Rabitti

Dottorando:

Federico M. Gabbricci

Sommario

INTRODUZIONE	2
1. PROFILI ECONOMICI DEL CREDITO, PROBLEMI INFORMATIVI E INTELLIGENZA ARTIFICIALE	7
1. L'ATTIVITÀ BANCARIA: PROFILI ECONOMICI	8
2. ASIMMETRIA INFORMATIVA, SELEZIONE AVVERSA E MORAL HAZARD NELLA CONCESSIONE DEL CREDITO.	10
3. TASSO DI INTERESSE, SELEZIONE AVVERSA E RAZIONAMENTO DEL CREDITO	13
4. STRUMENTI DI ATTENUAZIONE DEL RAZIONAMENTO DEL CREDITO: COLLATERALI E COVENANTS.	19
5. DATI E MACHINE LEARNING: LA NASCITA DI UN NUOVO COLLATERALE	22
6. OPPORTUNITÀ E RISCHI CYBER DEI MODELLI ML APPLICATI AL CREDITO	29
2. IL CREDIT SCORING ALGORITMICO	36
1. LA VALUTAZIONE DEL MERITO CREDITIZIO TRAMITE ALGORITMI DI ML	37
2. CREDIT SCORING ALGORITMICO: QUALI VANTAGGI E RISCHI PER GLI UTENTI?	44
3. L'APPLICAZIONE DEL GDPR AL CREDIT SCORING ALGORITMICO: LE SENTENZE "SCHUFA" E "D&B"	53
3.1. CORTE DI GIUSTIZIA DELL'UNIONE EUROPEA IN CAUSA C-634/21 "SCHUFA".	56
3.2. I PROBLEMI RIMASTI APERTI	63
3.3. CORTE DI GIUSTIZIA DEL 27 FEBBRAIO 2025 IN CAUSA C- 203/22 (D&B)	66
4. CREDIT SCORING E TRASPARENZA BANCARIA: L'IMPATTO DELLA CCD2	70
5. IL CREDIT SCORING COME AI AD ALTO RISCHIO: I REQUISITI PREVISTI DALL'AI ACT	79
6. LA GOVERNANCE DELL'INTELLIGENZA ARTIFICIALE IN UNIONE EUROPEA ED IN ITALIA: IL COORDINAMENTO CON LA VIGILANZA BANCARIA.....	99
3. RISCHIO DI CREDITO E MACHINE LEARNING.....	106
1. IL RISCHIO DI CREDITO NELL'AMBITO DELLA DISCIPLINA PRUDENZIALE DELLE BANCHE.	107
2. MODELLI IRB BASATI SUL MACHINE LEARNING.....	121
3. IL QUADRO REGOLATORIO DEI MODELLI IRB-ML	126
4. LA TECNOLOGIA COME STRUMENTO DI TRASPARENZA: EXAI.....	134
5. EXAI, REGOLAMENTAZIONE E NEUTRALITÀ TECNOLOGICA.	142
CONCLUSIONI	146
BIBLIOGRAFIA.....	155

INTRODUZIONE

Il mercato del credito, come ampiamente sottolineato dalla letteratura economica¹, si caratterizza per la presenza di asimmetrie informative che possono portare a situazioni di fallimento del mercato, inefficiente allocazione delle risorse, razionamento del credito nonché erronee “prezzature del rischio”².

La prassi bancaria, per mitigare l’asimmetria ed i connessi problemi di *adverse selection* e *moral hazard*, ha fatto ricorso sia a procedure di valutazione del merito creditizio (basate su modelli matematici lineari e dati storici provenienti dalle centrali rischi e dai bilanci delle società³) che a soluzioni contrattuali quali *covenants* e collaterali.

Queste soluzioni, seppur efficaci nel ridurre il *credit risk* per l’intermediario, possono escludere dalla concessione del credito soggetti che, seppur in grado di onorare gli impegni finanziari con la banca, si vedono denegata la loro domanda di credito per la mancanza di una storia creditizia⁴ o l’assenza di beni da poter dare in garanzia⁵.

In tale contesto si è inserita in maniera disruptiva l’intelligenza artificiale (AI) termine con cui ci si riferisce a «*sistemi che mostrano un comportamento*

¹ La teoria economica in materia, originata dalle ricerche di G. A. Akerlof, *The Market for Lemons: Quality Uncertainty and the Market Mechanism*, in *The Quarterly Journal of Economics*, 84(3), 488–500 e di J. E. Stiglitz, Andrew Weiss, *Credit Rationing in Markets with Imperfect Information*, in *The American Economic Review*, Vol. 71, No. 3 (Jun, 1981) 393–410, annovera vari studi tra cui si citano D. De Meza, D. C. Webb, *Too Much Investment: A Problem of Asymmetric Information*, in *The Quarterly Journal of Economics*, Volume 102, Issue 2, May 1987, Pages 281–292; W.G. Gale, “Federal Lending and the Market for Credit,” in *Journal of Public Economics*, 1990, 42 (2), 177–193; N. G. Mankiw, “The Allocation of Credit and Financial Collapse,” in *Quarterly Journal of Economics*, 1986, 101 (3), 455–470.

² G. S. Crawford, N. Pavaniz, F. Schivardi, *Asymmetric Information and Imperfect Competition in Lending Markets*, in *American Economic Review*, vol. 108, no. 7, July 2018 (pp. 1659–1701)

³ S. Zedda, *Credit scoring: Does XGboost outperform logistic regression? A test on Italian SMEs*, in *Research in International Business and Finance*, Vol 70, Part B, (2024) p.2

⁴ Consumer Financial Protection Bureau, *Data Point: Credit Invisibles*, 2015, https://files.consumerfinance.gov/f/201505_cfpb_data-point-credit-invisibles.pdf

⁵ W. G. Gale, *Collateral, Rationing, and Government Intervention in Credit Markets*, in *Asymmetric Information, Corporate Finance, and Investment* a cura di R. G. Hubbard, University of Chicago Press (1990) pp 43–62

intelligente analizzando il proprio ambiente e compiendo azioni, con un certo grado di autonomia, per raggiungere obiettivi specifici⁶».

I sistemi AI hanno offerto alle banche nuovi strumenti per colmare le lacune informative e migliorare l'inclusione finanziaria: l'utilizzo di algoritmi di *machine learning*, in grado di analizzare fonti di dati alternative (c.d. *soft data*), permette una migliore profilazione del rischio, estendendo l'accesso al credito anche a soggetti privi di una storia creditizia sufficientemente lunga o strutturata per essere valutata⁷. Al contempo, l'intelligenza artificiale permette di ridurre l'asimmetria *ex post* ed i connessi problemi di *moral hazard* grazie ad un monitoraggio più approfondito dei mutuatari (ad esempio, dati testuali, immagini e video relativi alle operazioni aziendali di un'impresa potrebbero essere raccolti e utilizzati per valutare il rispetto dei vincoli del prestito⁸).

Tali tecnologie innovative possono dunque consentire alle banche di aumentare i tassi di approvazione dei prestiti senza incrementare proporzionalmente il rischio e di dare accesso al credito ad un'ampia fascia di soggetti che storicamente hanno sofferto di razionamento del credito per la mancanza di una storia creditizia e di attivi tangibili da offrire in garanzia⁹.

Ma non solo, le banche potrebbero integrare tecniche di intelligenza artificiale nei loro modelli IRB per calcolare il rischio di credito, rispetto al quale ponderare i fondi propri conformemente a quanto previsto dal primo pilastro di

⁶ Commissione Europea, *Comunicazione della Commissione al Parlamento Europeo, al Consiglio Europeo, al Consiglio, al Comitato Economico e Sociale Europeo e al Comitato delle Regioni: Piano Coordinato sull'Intelligenza Artificiale* (COM 2018/795), 2018 par. 1

⁷ J. Jagtiani, C. Lemieux, *The roles of alternative data and machine learning in fintech lending: Evidence from the LendingClub consumer platform*, <https://www.philadelphiafed.org/-/media/frbp/assets/working-papers/2018/wp18-15r.pdf>

⁸ Majid Bazarbash, *FinTech in Financial Inclusion: Machine Learning Applications in Assessing Credit Risk*, in IMF Working Paper (WP/19/109), 2019, <https://www.imf.org/-/media/files/publications/wp/2019/wp19109.pdf> p.27

⁹ OECD, *AI in Business and Finance 2021: OECD Business and Finance Outlook, 2021* https://www.oecd.org/content/dam/oecd/en/publications/reports/2021/09/oecd-business-and-finance-outlook-2021_377c2c18/ba682899-en.pdf pp 38 ss.

Basilea¹⁰; la maggiore accuratezza¹¹ di tali tecniche può infatti ottimizzare i valori della ponderazione media delle attività rispetto ai modelli Logit¹² liberando riserve di capitale che potranno essere reimmesse nel sistema economico sotto forma di nuovi prestiti¹³ incrementando la liquidità complessiva del sistema economico.

Pur presentando molti benefici, l'utilizzo dell'intelligenza artificiale non è immune da rischi quali discriminazione algoritmica, bias nei dati di addestramento dell'algoritmo nonché opacità strutturale della decisione (c.d. *black box*) che possono riflettersi sia sui richiedenti che sulla stabilità complessiva del sistema finanziario.

Tali problematiche sono divenute centrali nel dibattito istituzionale europeo: la giurisprudenza della Corte di Giustizia con le sentenze Schufa e D&B ha definito il diritto alla spiegazione di una decisione algoritmica alla luce del GDPR; parallelamente, il legislatore comunitario è intervenuto prima con la Direttiva (UE) 2023/2225 (*Consumer Credit Directive II*), che all'art. 18 subordina il credit scoring algoritmico al diritto del consumatore di ottenere un intervento umano ed una spiegazione della decisione chiara e comprensibile e, successivamente con il Regolamento (UE) 2024/1689 (*AI Act*) che annovera il credit scoring fra le applicazioni ad alto rischio facendo soggiacere il suo utilizzo a stringenti obblighi di governance e trasparenza.

Il tema è in pieno divenire anche alla luce del fatto che le normative sono appena entrate o devono ancora entrare compiutamente in vigore: la CCD2 è

¹⁰ Comitato di Basilea per la vigilanza bancaria, *Basilea 3 – Schema di regolamentazione internazionale per il rafforzamento delle banche e dei sistemi bancari*, https://www.bis.org/publ/bcbs189_it.pdf

¹¹ Intesa come la proporzione di previsioni corrette (sia positivi che negativi) sul totale delle osservazioni considerate in un modello di classificazione.

¹² H. Fraisse, M. Laporte, *Return on Investment on AI: The Case of Capital Requirement*, in Banque de France Working Paper n. 809

¹³ J. Bridges, D. Gregory, M. Nielsen, S. Pezzini, A. Radia, M. Spaltro, *The impact of capital requirements on bank lending*, Working Paper No. 486, Bank of England, 2014 <https://www.bankofengland.co.uk/-/media/boe/files/working-paper/2014/the-impact-of-capital-requirements-on-bank-lending.pdf>

stata recepita dal Decreto Legislativo n. 212 del 31/12/2025 mentre la maggior parte delle norme dell'AI Act inizierà ad applicarsi il 2 agosto 2026¹⁴.

Obiettivo dell'indagine è pertanto quello di studiare l'impatto dell'intelligenza artificiale nell'ambito della concessione del credito approfondendo i benefici ed i rischi che la stessa possa apportare alle banche, alla clientela ed al sistema finanziario nel suo complesso individuando al contempo le appropriate misure normative e tecniche per mitigare le criticità che saranno rilevate.

Per comprendere appieno un fenomeno di tale portata, è impensabile limitarsi alla sola prospettiva giuridica ma è necessario adottare un approccio ibrido, capace di condensare diritto, tecnica ed economia. L'indagine che seguirà si muove proprio lungo queste tre direttrici.

Nel primo capitolo si esamineranno i profili economici della concessione del credito ed il funzionamento concreto delle tecniche di machine learning, evidenziando come i dati analizzati mediante tali tecniche possano divenire un "nuovo collaterale" in grado di mitigare il razionamento del credito ed apportare benefici al sistema economico.

Nel secondo capitolo l'indagine si concentrerà innanzitutto sul funzionamento pratico di tale collaterale analizzando i benefici ed i rischi tecnici, sociali ed economici del *credit scoring algoritmico* per poi andare ad approfondire come l'ordinamento giuridico si sia rapportato nei confronti di tali tecniche, partendo dalla giurisprudenza della Corte di Giustizia dell'Unione Europea e soffermandosi sulla *Consumer Credit Directive II* e sull'*AI Act* nonché sul recepimento in Italia di quest'ultimo e delle modalità di coordinamento tra le autorità di vigilanza designate.

¹⁴ Il legislatore con la Legge n. 132/2025 ha designato quali Autorità nazionali per l'Intelligenza Artificiale l'Agenzia per l'Italia digitale (AgID) e l'Agenzia per la cybersicurezza nazionale (ACN); i decreti attuativi della legge sono al momento in fase di elaborazione.

Il terzo ed ultimo capitolo sposterà l'attenzione dall'individuo alla banca, esaminando i benefici e i rischi che l'integrazione dell'AI nei modelli IRB può apportare al sistema e come vada coordinata la normativa prudenziale rispetto a tali modelli innovativi ed alla disciplina dell'AI Act. L'analisi verrà chiusa focalizzandosi nuovamente sulla tecnologia con l'analisi delle tecniche di eXplainable AI (XAI) al fine di comprendere se le stesse possano aiutare a rendere intellegibili i modelli risolvendo i problemi di opacità dei modelli così da garantire la compliance normativa.

A guidare l'intero lavoro vi è la consapevolezza di trovarsi di fronte a una trasformazione epocale che, come affermato dal premio Nobel Demis Hassabis, *«sarà dieci volte più grande della Rivoluzione Industriale e, forse, dieci volte più veloce¹⁵»* e ridefinirà in modo permanente le nostre vite e le dinamiche del mercato bancario.

¹⁵ <https://www.theguardian.com/technology/2025/aug/04/demis-hassabis-ai-future-10-times-bigger-than-industrial-revolution-and-10-times-faster>

1. PROFILI ECONOMICI DEL CREDITO, PROBLEMI INFORMATIVI E INTELLIGENZA ARTIFICIALE

1. L'attività bancaria: profili economici

Ai sensi dell'articolo 10, comma 1 del D.lgs. 1° settembre 1993, n. 385 (c.d. TUB) l'attività bancaria si esplica nella raccolta del risparmio tra il pubblico e nell'esercizio del credito. La banca esercita tali attività in maniera "coordinata" acquisendo fondi rimborsabili, tipicamente sotto forma di depositi, dalla clientela e utilizzandoli per erogare il credito.

Nello svolgere tali operazioni la banca esercita una funzione di intermediazione, fra chi ha eccedenze di fondi e chi necessita di finanziamenti facilitando il trasferimento di risorse¹⁶: senza il mercato bancario i singoli risparmiatori dovrebbero individualmente valutare l'affidabilità e i progetti dei debitori, con costi e difficoltà molto elevati; di converso la banca, accentrando tale attività, riduce i costi di informazione, valutazione e monitoraggio dei debitori per conto dei risparmiatori.

Osservata dal punto di vista microeconomico la funzione essenziale della banca è proprio la gestione dell'informazione e il monitoraggio dei debitori¹⁷: le banche esistono perché sono in grado di mitigare i costi di transazione e in particolare quelli derivanti dalle asimmetrie informative che rendono i mercati finanziari incompleti e imperfetti.

Poiché i finanziatori (*rectius* i depositanti) sono (tipicamente) piccoli, numerosi e dispersi essi non hanno né le risorse per monitorare efficacemente i debitori né sono incentivati a farlo: se ogni prestatore, infatti, decidesse di monitorare individualmente il proprio investimento, il costo aggregato totale della attività di sorveglianza sarebbe la somma dei costi individuali con una conseguente moltiplicazione dello sforzo palesemente inefficiente dal punto di vista aggregato poiché l'informazione sullo stato di salute finanziaria del debitore verrebbe

¹⁶ A. Ferrari, E. Gualandri, A. Landi, *Il sistema finanziario: funzioni, mercati e intermediari*, Giappichelli, 2024, Ottava Edizione, p. 20

¹⁷ X. Freixas, J.C. Rochet, *Microeconomics of Banking*, Mit Pr, I edizione, 1/10/1997 pp 1-18

prodotta X volte quando sarebbe sufficiente produrla esclusivamente una sola volta¹⁸.

Per esemplificare se il costo di un audit completo fosse di 200 euro e l'impresa fosse finanziata da 10.000 piccoli obbligazionisti, il costo totale del monitoraggio duplicato raggiungerebbe i 2 milioni di euro¹⁹ con conseguente impraticabilità del monitoraggio individuale su larga scala.

Il finanziamento diretto porta, dunque, ad un'inefficienza di costo lorda che rende il sistema insostenibile e apre la strada al problema del *free-riding*: ogni prestatore razionale nella situazione sopra descritta adotterà infatti come strategia quella di astenersi dal monitorare, sperando che siano gli altri a sostenere il costo; ma poiché ogni prestatore ragionerà allo stesso modo, il monitoraggio non avrà luogo. Un siffatto risultato, sebbene razionale dal punto di vista individuale, è sub-ottimale per il gruppo nel suo complesso poiché l'assenza di monitoraggio aumenta il rischio di insolvenza e, di conseguenza, il costo del capitale per il debitore. Se infatti il monitoraggio serve a verificare che il debitore utilizzi i fondi in conformità con quanto pattuito e che gestisca l'impresa con la dovuta diligenza la sua assenza comporta un maggiore rischio per i creditori che per essere disposti a concedere il prestito in tale contesto richiederanno una remunerazione maggiore come compensazione.

È qui che la banca interviene a risolvere il problema del *free riding* agendo come un "monitore delegato", un agente specializzato che raccoglie i fondi da molti depositanti per prestarli a numerosi debitori; in questo modo, la banca sfrutta le economie di scala nel monitoraggio e riduce i costi informativi complessivi del sistema.

¹⁸ D. W. Diamond, "Financial Intermediation and Delegated Monitoring in" *The Review of Economic Studies* 51, no. 3 (1984): 393–414. <https://doi.org/10.2307/2297430>.

¹⁹ D. W. Diamond, *Financial Intermediation as Delegated Monitoring: A Simple Example*, in *Economic Quarterly*, Summer 1996 https://www.richmondfed.org/-/media/richmondfedorg/publications/research/economic_quarterly/1996/summer/pdf/diamond.pdf pp. 59-60

Tuttavia, le banche, pur realizzando, grazie alla loro struttura, una mitigazione dei problemi informativi e una riduzione dei costi di transazione nelle operazioni di finanziamento²⁰, non riescono mai ad ottenere un'informazione completa sui soggetti da finanziare; tale informazione imperfetta porta a problemi di selezione avversa e di azzardo morale che possono influenzare il prezzo, le quantità e la composizione del credito erogato.

2. Asimmetria informativa, Selezione Avversa e Moral Hazard nella concessione del credito.

Il rischio è una componente fisiologica ed ineliminabile nei mercati all'interno dei quali i soggetti economici operano prendendo decisioni che possono caratterizzarsi per una elevata incertezza a causa di una informazione incompleta ed asimmetrica²¹.

L'asimmetria informativa, definibile come la condizione nella quale le parti di una transazione dispongono di livelli di informazione differenti per assumere la decisione economica²², rappresenta infatti una delle principali cause di fallimento del mercato poiché impedisce il raggiungimento di un equilibrio efficiente e fa sì che l'agente meno informato possa assumere decisioni subottimali, da cui si avvantaggerà il soggetto maggiormente informato.

L'asimmetria informativa può verificarsi sia *ex ante* che *ex post* alla decisione economica: nel primo caso il divario informativo precede l'operazione finanziaria (ad esempio la scelta di effettuare o meno un dato investimento); nel secondo la differenza di conoscenza attiene all'esito di un evento associato alla decisione economica assunta.

²⁰ C. Brescia Morra, *Il diritto delle banche*, Il Mulino, 2 edizione (2016), pp 15 ss.

²¹ M. Biasin, S. Cosma, M. Oriani, *La Banca. Istituzione e impresa*, Isedi (2013) pp. 42 ss.

²² G. Nicodano, *Asimmetria Informativa*, in *Dizionario di Economia e Finanza* (2012), [https://www.treccani.it/enciclopedia/asimmetria-informativa_\(Dizionario-di-Economia-e-Finanza\)](https://www.treccani.it/enciclopedia/asimmetria-informativa_(Dizionario-di-Economia-e-Finanza))

La concessione del credito è un tipico caso in cui si presenta l'asimmetria informativa dal momento che la transazione economica differisce dai normali scambi commerciali (dove a fronte del pagamento di un prezzo si ottiene la merce corrispondente) per la presenza di un elevato grado di incertezza circa la bontà del progetto e il suo effettivo esito: si pensi a un imprenditore (soggetto in deficit) che richiede un finanziamento per realizzare un progetto industriale e a un potenziale finanziatore (soggetto in surplus) che deve valutare l'opportunità di concederglielo; il trasferimento delle risorse finanziarie si verifica nel tempo T_1 a fronte della promessa di restituzione futura del capitale e dei relativi interessi in un tempo T_2 , condizionata dal risultato del progetto; la conoscenza del progetto da finanziare sarà diversa tra il soggetto richiedente il prestito (più informato e magari interessato a omettere talune informazioni) rispetto al soggetto erogante (meno informato). È questo un caso di asimmetria *ex ante* laddove il richiedente avrà una migliore conoscenza del progetto, avendolo ideato e sviluppato, rispetto al concedente che dovrà cercare di diminuire lo svantaggio conoscitivo acquisendo informazioni al fine di valutare l'affidabilità del debitore e verificare, per quanto possibile, le informazioni fornitegli in sede precontrattuale²³.

Alla asimmetria *ex ante* si correla strettamente il problema della "selezione avversa" (*adverse selection*). Il fenomeno, analizzato da Akerlof nel suo studio sul mercato delle auto usate²⁴, si verifica laddove la parte meno informata (il compratore) non è in grado di distinguere la qualità intrinseca dei beni o delle controparti presenti sul mercato; conseguentemente egli sarà disposto a pagare un prezzo basato sulla qualità media attesa di tutte le auto disponibili il quale risulterà troppo basso per i proprietari di auto di alta qualità che saranno incentivati a ritirarsi dal mercato (non vedendo riconosciuto il reale valore del

²³ F. S. Mishkin, S. G. Eakins, E. Beccalli, *Istituzioni e mercati finanziari*, Pearson, 9 edizione pp. 24-25, (2019)

²⁴ G. A. Akerlof, *The Market for "Lemons": Quality Uncertainty and the Market Mechanism*, in *The Quarterly Journal of Economics*, 84(3), 488-500 <https://doi.org/10.2307/1879431> (1970)

loro bene). Il processo innescherà una spirale negativa: il ritiro dei venditori migliori abbasserà la qualità media del parco auto rimanente, spingendo i compratori a ridurre ulteriormente il prezzo offerto, inducendo l'uscita di altri venditori e così via fino a poter arrivare ad una situazione nella quale solo i prodotti di qualità peggiore ("i limoni") vengono scambiati. La tesi di Akerlof è trasponibile nel mercato del credito in cui per le banche è difficile separare a priori i debitori solvibili da quelli che non lo sono proprio in ragione del manifestarsi della asimmetria informativa.

Nel mercato del credito, inoltre, il divario informativo può riguardare non solo caratteristiche nascoste della controparte ma anche le azioni che questa compie successivamente alla stipula del contratto.

I problemi di asimmetria *ex post* possono creare una situazione c.d. di azzardo morale (*moral hazard*) per la quale l'*agent*/debitore, avendo incentivi diversi dal *principal*/creditore (il quale subisce le conseguenze dell'azione dell'*agent*) potrebbe alterare il proprio comportamento a svantaggio di quest'ultimo: un imprenditore finanziato potrebbe ad esempio utilizzare i fondi per investimenti più rischiosi di quelli dichiarati nel business plan (*asset substitution*), massimizzando i propri potenziali profitti in caso di successo e scaricando le perdite sul creditore in caso di fallimento. Tali problematiche per essere gestite dalla banca, richiedono, da un lato che questa colmi la mancanza di informazioni, impiegando mezzi e risorse per raccogliere dati e notizie sul progetto in esame e per interpretarle adeguatamente, dall'altro lato che essa svolga un attento monitoraggio nella fase dell'esecuzione del finanziamento²⁵.

²⁵ M. Biasin, S. Cosma, M. Oriani, *La Banca. Istituzione e impresa*, Isedi (2013) pp. 404 ss.

3. Tasso di interesse, Selezione Avversa e Razionamento del Credito

Nel precedente paragrafo abbiamo evidenziato come nel mercato del credito la selezione avversa si manifesti laddove la banca non sia in grado di distinguere *ex ante* la qualità dei richiedenti (merito di credito, rischio di progetti, probabilità di insolvenza).

A questo punto è necessario ampliare l'analisi, è fondamentale introducendo un ulteriore elemento peculiare dell'attività bancaria: il tasso d'interesse.

L'*interest rate*, ossia il costo pagato per poter prendere in affitto i fondi²⁶ non agisce semplicemente come prezzo del capitale, ma anche come un potente, seppur imperfetto, meccanismo di selezione. La radice del problema risiede nell'incapacità dell'intermediario di osservare direttamente la rischio di progetti di investimento proposti dai richiedenti (selezione avversa); la banca, a fronte dell'incertezza, è costretta a formare delle aspettative sulla qualità media dei potenziali debitori e, di conseguenza, tende ad applicare un tasso "medio" omogeneo a un insieme eterogeneo di prenditori²⁷. Questa strategia di "*pooling*", tuttavia, genera una significativa distorsione allocativa: il tasso medio risulta infatti relativamente elevato per i debitori di qualità elevata (che meriterebbero condizioni migliori) e troppo favorevole per i debitori più rischiosi (per i quali il tasso "corretto" dovrebbe essere più alto), con l'effetto di espellere dal mercato i prenditori virtuosi e di trattenere/attrarre quelli peggiori. L'effetto "perverso" di questo meccanismo è che, all'aumentare del tasso d'interesse, si innesca un processo di autoselezione negativa: i richiedenti meno

²⁶ F. S. Mishkin, *The Economics of Money, Banking, and Financial Markets*, Pearson Education Limited 2022, XIII Edizione Globale, https://api.pageplace.de/preview/DT0400.9781292409566_A42815947/preview-9781292409566_A42815947.pdf P.51,

²⁷ J. E. Stiglitz, Andrew Weiss, *Credit Rationing in Markets with Imperfect Information*, in *The American Economic Review*, Vol. 71, No. 3 (Jun 1981), pp. 393- 410

rischiosi rinunciano al credito, dato che le condizioni praticate dalla banca pregiudicano la redditività attesa dell'investimento da finanziare con debito, oppure si spostano su progetti di investimento maggiormente rischiosi senza informarne la banca; di contro i prenditori peggiori hanno un chiaro incentivo a richiedere credito sia perché offerto a condizioni vantaggiose sia poiché in caso di fallimento le perdite saranno "socializzate", ricadendo principalmente sulla banca. I «cattivi» prenditori scacciano dunque quelli «buoni» che non riescono a ottenere credito alle condizioni (migliori di quelle dei prenditori «cattivi») che verrebbero loro applicate in condizioni di simmetria informativa ovvero di corretta conoscenza del loro profilo di rischio²⁸.

In termini più generali, se la banca manovra al rialzo il tasso d'interesse (applicandolo a tutti i prenditori), la quota di «cattivi» prenditori sul totale del credito erogato tenderà a crescere, perché veicola i progetti più rischiosi; a essi è tuttavia associata una probabilità di insolvenza che peggiora la qualità media del credito erogato e che può più che compensare gli effetti positivi sui profitti della banca dati dall'aumento del tasso di interesse, con conseguente riduzione della redditività attesa della banca (poiché l'aumento di prezzo non riesce più a compensare il maggior rischio sostenuto).

Un esempio può meglio aiutare a comprendere la situazione: si ipotizzi che nel mercato esistano due tipologie di prenditori di fondi, i prenditori "buoni" (B) e i prenditori "cattivi" (C) che vengono distinti in funzione del profilo di rischio dei progetti da finanziare e che la banca non sia in grado (senza intraprendere adeguate azioni di contrasto, ma costose) di discernere tra queste due categorie. I progetti di entrambe le categorie richiedono un investimento iniziale di 100, da finanziare interamente con debito; qualora il debitore non fosse in grado di far

²⁸ M. Biasin, S. Cosma, M. Oriani, *La Banca. Istituzione e impresa*, Isedi (2013) pp. 402 ss.

fronte al rimborso del finanziamento e al pagamento degli interessi, la banca può rivalersi sui ricavi del progetto.

L'iniziativa dei prenditori di tipo B può avere un esito di 120 con una probabilità (p) del 50% e di 110 con p sempre del 50%; il progetto dei prenditori di tipo C un esito di 150 o di 80 entrambi con p del 50%.

Il rischio relativo è tuttavia diverso, evidentemente maggiore per C e minore per B: in particolare, in caso di output negativo del progetto B (80), il debitore non è in grado di rimborsare il prestito con conseguente perdita in capo alla banca di (-20).

Supponiamo che la banca abbia la possibilità di applicare i seguenti tassi di interesse:

	r1	r2	r3	r4	r5	r6
Tasso r	10%	15%	20%	>20%	25%	30%

Qualora la banca fissasse il tasso di interesse in un intervallo compreso tra $0\% < r < 20\%$ tutti i prenditori hanno convenienza a domandare credito (dove di per sé la banca ottiene profitti superiori con i debitori di tipo B rispetto al tipo C, dato che questi ripagano sempre il debito e parte degli interessi).

Al raggiungimento del tasso soglia del 20%, tuttavia, i prenditori migliori (B) raggiungono la soglia di indifferenza, rinunciano all'investimento (e quindi al finanziamento) per ogni livello di tasso superiore, dato che vedono annullato il rendimento netto atteso (il tasso di interesse uguaglia infatti il rendimento atteso del progetto).

Per ogni livello di tasso $>20\%$ la banca calamiterà pertanto solamente debitori rischiosi (tipo C), i quali richiederanno comunque credito, dato che in caso di esito positivo dell'investimento, il rendimento netto risulta comunque positivo. L'uscita dal mercato dei prenditori migliori (B) comporta però una drastica riduzione dei profitti della banca (a 7,5 per soglie immediatamente superiori a 20%):

	r1	r2	r3	r4	r5	r6
Tasso r	10%	15%	20%	$>20\%$	25%	30%
Interessi Banca (B)	10	12.5	15	--	--	--
	$=0.5*10$ $+0.5*10$	$=0.5*15+$ $0.5*10$	$=0.5*20+$ $0.5*10$	--	--	--
Interessi Banca (C)	2.5	5	7.5	7.5	10	12.5
	$=0.5*10$ $+0.5*(-5)$	$=0.5*15+$ $0.5*(-5)$	$=0.5*20+$ $0.5*(-5)$	$=0.5*20+0.5$ $*(-5)$	$=0.5*25+$ $0.5*(-5)$	$=0.5*30+0.5*(-5)$
Totale r	12.5	17.5	22.5	7.5	10	12.5

In sintesi, un aumento del tasso di interesse da parte della banca comporta per la stessa due effetti contrapposti sui ricavi attesi:

- (1) un effetto positivo dato dall'aumento del prezzo del credito;
- (2) un effetto negativo dato dall'uscita dal mercato dei debitori a basso rischio.

Il rendimento atteso di un prestito è infatti una funzione decrescente della sua rischiosità²⁹ per cui ogni qualvolta i prestatori migliori rinunciano al credito in quanto il tasso di interesse è divenuto per loro eccessivamente oneroso, peggiora (cioè aumenta) la rischiosità media del portafoglio, assieme ai ricavi potenziali della banca; l'effetto netto generato sul rendimento dipenderà dalla grandezza relativa.

Ora la banca, in quanto agente economico razionale, è consapevole che la relazione tra il tasso applicato e il proprio profitto atteso non è monotona: esiste un tasso che massimizza il rendimento atteso r^* , oltre il quale l'aumento del tasso riduce tale rendimento³⁰ poiché oltre una certa soglia, ogni ulteriore aumento del tasso (l'effetto prezzo positivo) è più che compensato dal deterioramento della qualità del portafoglio (l'effetto rischio negativo).

Il profitto atteso della banca seguirà una curva a U invertita (Figura 1); conseguentemente la strategia ottimale non è alzare il tasso all'infinito, ma di fissarlo nel punto di massimo di tale curva e, a quel punto, allocare i fondi disponibili tramite meccanismi non di prezzo.

La banca non ha pertanto più interesse a concedere credito oltre tale soglia, anticipando il comportamento atteso dei debitori e ciò si traduce nel negare il credito a una parte dei richiedenti, anche a coloro che sarebbero apparentemente identici a chi lo riceve e disposti a pagare un tasso superiore.

²⁹ J. E. Stiglitz, A. Weiss, *Credit Rationing in Markets with Imperfect Information*, in *The American Economic Review*, Jun. 1981, Vol. 71, No. 3, p. 397

³⁰ Ivi, pp. 393- 410

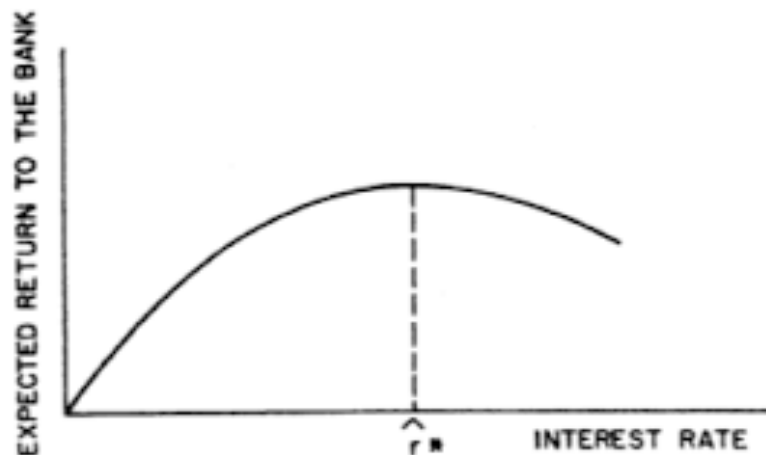


Figura 1 (Fonte: J.E. Stiglitz, A. Weiss, *Credit Rationing in Markets with Imperfect Information*)

Le banche, pertanto, al fine di massimizzare il profitto atteso, ottimizzano la selezione dei prenditori anziché alzare indiscriminatamente i tassi, lasciando insoddisfatta parte della domanda di prestiti anche quando vi sarebbe offerta potenziale.

Questa dinamica conduce al "razionamento del credito"³¹ il quale rappresenta una delle più significative deviazioni dal modello di mercato concorrenziale perfetto; se infatti in una situazione di "mercato standard", un eccesso di domanda verrebbe riassorbito da un aumento del prezzo fino a raggiungere un punto di equilibrio ciò non si verifica nel mercato del credito dove il tasso d'interesse non è in grado di operare come meccanismo di "market clearing". Nel

³¹ Il tema del razionamento del credito è stato oggetto di grande interesse nella dottrina economica degli anni Sessanta e Settanta, oltre che da Stiglitz e Weiss. Per un approfondimento sul tema si rimanda a:

- D.M Jaffee, F. Modigliani, "A Theory and Test of Credit Rationing" in *The American Economic Review*, Vol. 59 (1969).

- D. M. Jaffee, T. Russell, *Imperfect Information, Uncertainty, and Credit Rationing* in *The Quarterly Journal of Economics*, Volume 90, Issue 4, November 1976.

- W. R. Keeton, *Equilibrium Credit Rationing*, Garland Press, New York (1979).

mercato permarrà quindi una domanda di credito di soggetti disposti a pagare un tasso di interesse superiore a quello massimo praticato dalla banca, ma che non sarà soddisfatta dall'intermediario che massimizza il suo profitto in corrispondenza del tasso soglia che è tuttavia inferiore rispetto a quello di equilibrio teorico³² (ossia il tasso che eguaglia la domanda e l'offerta di prestiti). Il razionamento del credito emerge quindi non come un'anomalia, ma come una risposta di equilibrio alle imperfezioni informative; tale equilibrio, tuttavia, può generare conseguenze macroeconomiche profonde: esso può portare a un sotto-investimento sistematico nell'economia, penalizzando in particolare le piccole e medie imprese e i progetti innovativi, con un impatto negativo sulla crescita economica complessiva.

4. Strumenti di attenuazione del razionamento del credito: collaterali e covenants.

Nel precedente paragrafo è stato evidenziato come in presenza di informazione imperfetta, il mercato del credito non si equilibri attraverso il solo meccanismo del prezzo (il tasso di interesse): invece di un equilibrio dato dall'incontro di domanda ed offerta nel mercato del credito, emerge un equilibrio di razionamento che fa sì che non tutti i soggetti che richiedono il credito lo ottengono anche qualora questi fossero disposti a pagare un tasso di interesse superiore; tale equilibrio è l'effetto di un comportamento razionale della banca, la quale cerca di massimizzare i propri profitti in un contesto di incertezza.

Di fronte all'inefficacia del tasso di interesse quale strumento di screening la teoria economica e la prassi bancaria hanno cercato meccanismi alternativi per superare il problema dell'asimmetria informativa; due sono in particolare gli strumenti su cui è opportuno soffermarsi: i loan covenant ed i collateral.

³² M. Biasin, S. Cosma, M. Oriani, *La Banca. Istituzione e impresa*, Isedi (2013) pp. 403-404.

I covenants possono essere definiti come un insieme di clausole, inserite nei contratti di debito, finalizzato a limitare l'autonomia imprenditoriale del debitore imponendogli determinati parametri finanziari o limiti operativi che permettono di ridurre i problemi di asimmetria informativa ex post ed il moral hazard³³.

Guardando alla funzione della clausola, i covenants possono essere ricompresi in quattro categorie:

- i. covenants che impongono all'impresa l'adozione di predeterminati ed efficienti comportamenti ai fini della conservazione del patrimonio netto;
- ii. covenants di bilancio che obbligano il debitore al rispetto di dati e rapporti di bilancio;
- iii. covenants che pongono a condizione del default il comportamento di terzi;
- iv. covenants informativi che obbligano il debitore a comunicare al finanziatore l'intenzione di compiere determinate operazioni³⁴.

L'inadempimento delle clausole costituisce un "technical default"³⁵, anche se il mutuatario è in regola con i pagamenti; l'evento non porta necessariamente alla risoluzione del contratto o alla decadenza dal beneficio del termine con obbligo di rimborsare immediatamente il prestito, ma può trasferire poteri di controllo, dandogli il diritto di rinegoziare i termini del contratto (ad esempio, aumentando il tasso di interesse), richiedere garanzie aggiuntive o, nei casi più gravi, esigere il rimborso anticipato.

Il covenant è uno strumento i cui risultati possono essere ottimizzati solo se questo viene concepito ed applicato «su misura» del singolo finanziato (tenendo così conto del suo *business plan* e del connesso piano finanziario, da sottoporre ad

³³ G. Barbara, Covenants e "fattispecie aggregative" Contributo alla teoria del controllo e del collegamento "Esterni", G. Giappichelli Editore, 2022 pp. 3 ss

³⁴ U. Patroni Griffi, I covenants finanziari. Note introduttive, in Rivista di Diritto Societario, Fascicolo 3 – 2009, Giappichelli pp. 602 ss.

³⁵ Situazione in cui viene a trovarsi il finanziato a seguito del mancato rispetto di uno degli impegni contrattuali (ad es. i covenant) inseriti all'interno del contratto di finanziamento.

analisi di sensibilità)³⁶ e se viene applicato non in maniera rigida o come strumento per imporre oneri maggiori al debitore.

Una simile attività per i costi che presenta è peraltro difficile possa essere condotta su mutui dal valore limitato; inoltre, la pressione per evitare la violazione dei covenant può portare i manager delle aziende finanziate a prendere decisioni operative sub-ottimali che sacrificano il valore a lungo termine al fine di evitare l'attivazione dei campanelli di allarme previsti nel covenant (c.d. *trip wires*) generando inefficienze economiche³⁷.

Il secondo strumento adottato nella prassi per superare i problemi di asimmetria informativa è la richiesta di garanzie reali a fronte del prestito c.d. collaterale, il quale è in grado di svolgere una duplice funzione: di screening contro la selezione avversa; di incentivo contro l'azzardo morale (la minaccia di perdere il bene pignorato, in caso di insolvenza, è un forte incentivo affinché il mutuatario si comporti in modo prudente e profonda il massimo sforzo per onorare il debito)³⁸.

Dal punto di vista della funzione di screening il collaterale riesce a creare un meccanismo di autoselezione dei prestiti inducendo i mutuatari a rivelare il proprio tipo di rischio: i richiedenti ad alto rischio, avendo una maggiore probabilità di insolvenza e quindi di perdita della garanzia, saranno infatti meno disposti a impegnare beni di valore rispetto ai mutuatari a basso rischio; sfruttando questa diversa propensione, le banche possono progettare "contratti separatori" che attraggano in modo differenziato i diversi tipi di mutuatari, (ad esempio, offrendo una combinazione di tassi di interesse più bassi in cambio di

³⁶ D. U. Santosuosso, Covenants finanziari: rischio d'impresa e responsabilità gestionali, in Rivista di Diritto Societario, Fascicolo 3 – 2009, Giappichelli p. 638

³⁷ I. D. Dichev, D. J. Skinner, Large-Sample Evidence on the Debt Covenant Hypothesis, in Journal of Accounting Research, Sep. 2002, Vol. 40, No. 4 (Sep. 2002), pp.1114 ss.

³⁸ C. Yuk-Shee, G. Kanatas "Asymmetric Valuations and the Role of Collateral in Loan Agreements, in Journal of Money, Credit and Banking, vol. 17, no. 1, 1985, pp. 84-95

un collaterale più elevato) consentendo in tal modo di riuscire a distinguere tra i soggetti richiedenti³⁹.

Tuttavia, l'efficacia del collaterale non è priva di ambiguità e presenta limitazioni sia teoriche che pratiche: basti notare come esso filtri i mutuatari non tanto in base alla qualità intrinseca dei loro progetti di investimento quanto piuttosto in base alla loro ricchezza iniziale: si pensi ipoteticamente ad un imprenditore con un progetto altamente redditizio e a basso rischio ma che è privo di beni da offrire in garanzia; sulla base della mera applicazione del collaterale esso rischierebbe di vedersi denegato il finanziamento e di essere escluso dal mercato del credito. Di converso un individuo facoltoso con un progetto più rischioso potrebbe ottenere il finanziamento proprio in virtù della sua capacità di fornire un collaterale adeguato. A ben vedere dunque il collaterale pur risolvendo parzialmente il problema della selezione avversa per il prestatore rischia di generare una allocazione del capitale subottimale che favorirà i soggetti già "ricchi" a discapito di altri soggetti i cui progetti, seppur economicamente validi non saranno finanziati poiché essi sono privi di collateral⁴⁰.

5. Dati e Machine Learning: la nascita di un nuovo collaterale

L'ascesa delle tecnologie digitali ha innescato una trasformazione radicale e nel mercato del credito; i big data e l'intelligenza artificiale stanno infatti ridefinendo i processi di valutazione della concessione del credito (c.d. credit scoring) innovando i modelli tradizionali di valutazione della possibilità di default del richiedente.

³⁹ D. Besanko, A. V Thakor, *Collateral and Rationing: Sorting Equilibria in Monopolistic and Competitive Credit Markets* in *International economic review* vol 28 n.3 (1987) pp. 671–689

⁴⁰ W. G. Gale, *Collateral, Rationing, and Government Intervention in Credit Markets*, in *Asymmetric Information, Corporate Finance, and Investment* a cura di R. G. Hubbard, University of Chicago Press (1990) pp 43-62

I modelli tradizionali di credit scoring, come avremo modo di approfondire nel successivo capitolo, si basano su di un insieme ristretto di “*hard data*”, principalmente provenienti dalle centrali rischi e dai bilanci che vengono elaborati attraverso modelli matematici lineari⁴¹.

Gli sviluppi della tecnologia permettono ora di sfruttare due elementi rivoluzionari: l'accesso a una mole immensa e variegata di “dati alternativi” e la potenza computazionale degli algoritmi di machine learning.

Per dati alternativi ci si riferisce a tutta una serie di dati che sono originati dalla c.d. “impronta digitale” del richiedente quali, ad esempio, le attività sui social media, le sue abitudini o propensioni di acquisto nonché le preferenze personali manifestate attraverso le ricerche sui motori di ricerca; tali dati, indicati con il termine di “*soft data*”, possono integrarsi con gli “*hard data*” di cui sopra e migliorare le previsioni⁴².

Così facendo i dati divengono in grado di sostituire il collaterale tradizionale e di svolgere un ruolo cruciale nel mitigare i problemi di azzardo morale e selezione avversa derivanti dall'asimmetria informativa. Il “collaterale informativo” non si limita peraltro a mitigare le conseguenze dell'asimmetria informativa ma agisce direttamente per ridurre l'esistenza stessa: se una garanzia reale funge da “assicurazione” contro l'incertezza, i dati operano come uno strumento di riduzione ab origine dell'incertezza a monte rendendo il comportamento del debitore più trasparente e prevedibile⁴³.

Questo nuovo paradigma può funzionare grazie alla combinazione di big data e tecnologie di *data analytics* basate sull'intelligenza artificiale: i big data, ossia un quantitativo di informazioni caratterizzate dalla presenza delle cosiddette cinque

⁴¹ S. Zedda, *Credit scoring: Does XGboost outperform logistic regression? A test on Italian SMEs*, in *Research in International Business and Finance*, Vol 70, Part B, (2024) p.2

⁴² M. Rabitti, *Credit scoring via machine learning e prestito responsabile*, in *Rivista di Diritto Bancario*, Fascicolo I, Anno 2023, Sezione seconda, p. 182

⁴³ L. Gambacorta, Y. Huang, Z. Li, H. Qiu, S. Chen, *Data versus Collateral*, in *Review of Finance*, Volume 27, Issue 2, March 2023, pp. 369–398

V⁴⁴ (volume, velocità, varietà, veridicità e valore), richiedono, per essere compresi ed utilizzati nei processi decisionali, di innovative tecniche di elaborazione che sono ora rese possibili grazie all'utilizzo di modelli di *machine learning* (ML), termine con cui ci si riferisce ad un sottoinsieme delle tecnologie di intelligenza artificiale, caratterizzato dalla creazione di algoritmi in grado di apprendere automaticamente da dati ed esperienze pregresse così da migliorare le proprie prestazioni⁴⁵.

Il machine learning viene suddiviso, dai ricercatori sulla base dei modelli di apprendimento in tre tipologie:

- i. ***Supervised Learning***: il modello viene addestrato su un set di dati "etichettato", ovvero un insieme di input per i quali il risultato corretto (la variabile di output o "target") è già noto. L'obiettivo è che il modello impari a mappare gli input agli output, per poi effettuare previsioni autonomamente su nuovi dati non etichettati⁴⁶.

Questo è il paradigma più comune nelle applicazioni di credit scoring, dove l'algoritmo impara a prevedere la probabilità di default (la variabile target) basandosi su dati storici di clienti in default e non⁴⁷.

- ii. ***Unsupervised Learning***: l'addestramento della macchina avviene tramite un set di dati non "etichettati"; la macchina opererà in modo esplorativo, cercando di scoprire strutture, pattern o raggruppamenti intrinseci nei dati⁴⁸. L'attività di raggruppamento ossia la suddivisione in gruppi in base a somiglianze o differenze dei dati (c.d. clustering) viene applicata in vari

⁴⁴ B. Indovina, *Informatica, diritto, intelligenza artificiale*, Egea S.p.a., 2024 pp 24 ss.

⁴⁵ Hicham Sadok, Fadi Sakka, Mohammed El Hadi El Maknouzi, *Artificial intelligence and bank credit analysis: A review in Cogent Economics & Finance*, 2022 p.3

⁴⁶ C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, (2006) p.3

⁴⁷ Y. Li, W. Chen, *A Comparative Performance Assessment of Ensemble Learning for Credit Scoring*, in *Mathematics*, 8(10), 2020 pp. 1756.

⁴⁸ B. Baesens, *Analytics in a big data world: the essential guide to data science and its Applications*, John Wiley & Sons, 2014 pp. 87 ss

ambiti, quali la segmentazione della clientela e l'identificazione di gruppi di comportamento anomalo per la prevenzione delle frodi⁴⁹.

- iii. ***Reinforcement Learning***: la macchina impara interagendo con un ambiente dinamico, senza che vengano forniti dati di addestramento, tramite una funzione di ricompensa che guida il suo processo decisionale.

Più specificatamente la macchina osserva lo stato corrente dell'ambiente e seleziona un'azione; in risposta a tale azione, l'ambiente transita in un nuovo stato e fornisce all'agente un segnale di feedback scalare, denominato ricompensa. Attraverso tentativi ed errori, l'algoritmo impara a compiere azioni per massimizzare la ricompensa totale⁵⁰.

Sui modelli citati si basano, tra gli altri, due categorie di algoritmi significativi per la loro diffusione o le loro intrinseche caratteristiche:

- **algoritmi di *ensemble learning***: combinano le previsioni di più modelli, spesso deboli, per ottenere un esito (generalmente ottenuto come la previsione media o maggioritaria dei modelli) che ha una accuratezza superiore rispetto a un singolo modello⁵¹.

Un caso tipico è quello in cui i singoli modelli da combinare sono i *decision trees* (Figura 2) ossia modelli di apprendimento supervisionato che utilizzano una struttura gerarchica ad albero per prendere decisioni di classificazione o regressione; l'algoritmo funziona in modo intuitivo, ponendo una serie di domande sequenziali sulle variabili dei dati (ad esempio, "il reddito del cliente è superiore a X?") per dividere il set di dati in sottoinsiemi sempre più omogenei. Data la loro struttura a diagramma di flusso essi sono facilmente comprensibili e interpretabili.

⁴⁹ Google Cloud, Che cos'è l'apprendimento non supervisionato?, <https://cloud.google.com/discover/what-is-unsupervised-learning>

⁵⁰ R. S. Sutton, A. G. Barto, Reinforcement Learning: An Introduction, second edition, The MIT Press Cambridge, Massachusetts (2018), pp 1 ss

⁵¹ M. A. Jabbar, L. Gaur, A. Haqiq, *Ensemble Machine Learning: Advances in Research and Applications*, Nova Science Publishers, 2024 pp. 6 ss

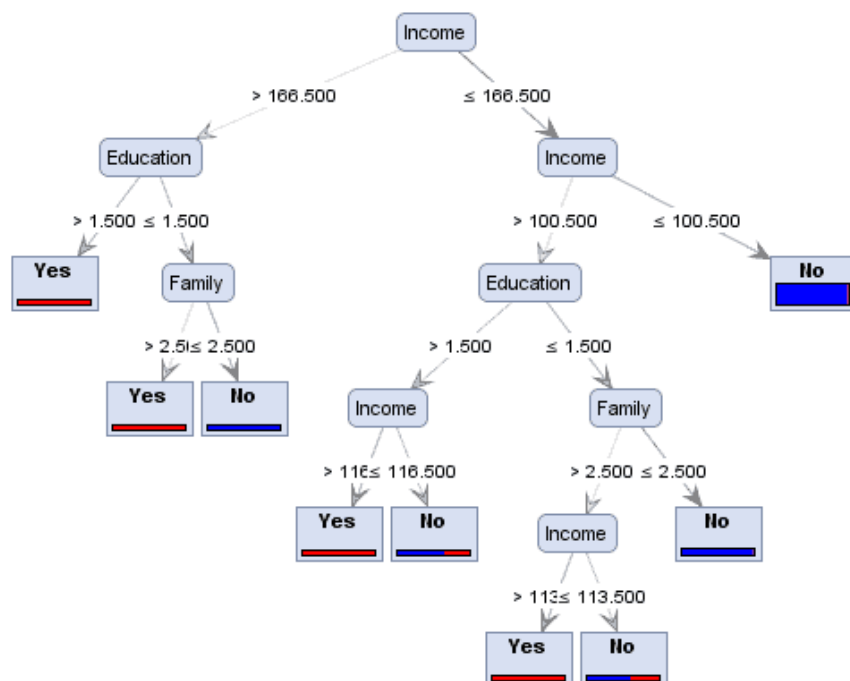


Figura 2 (Fonte <https://www.simafore.com/how-to-use-decision-trees-for-credit-scoring-part-1-of-4/>)

Le tecniche di *ensemble learning* possono essere a loro volta suddivise in due sottocategorie a seconda che esse siano basate sul *bagging* (apprendimento simultaneo di modelli indipendenti tra loro, ciascuno caratterizzato dallo sfruttamento di porzioni dell'informazione complessiva) o sul *boosting* (sequenze di modelli che vanno progressivamente a raffinare il processo di apprendimento)⁵².

Due applicazioni dei citati approcci, riconducibili al contesto dell'apprendimento supervisionato, sono il *random forest* e il *gradient boosting*:

⁵² M. A. Jabbar, L. Gaur, A. Haqiq, *Ensemble Machine Learning: Advances in Research and Applications*, Nova Science Publishers, 2024 pp. 8 ss

- i. il *random forest* è un algoritmo di ensemble learning che combina le decisioni di più *decision trees* per raggiungere un singolo risultato che sarà dato dalla media dei risultati di ciascun albero (Figura 3);

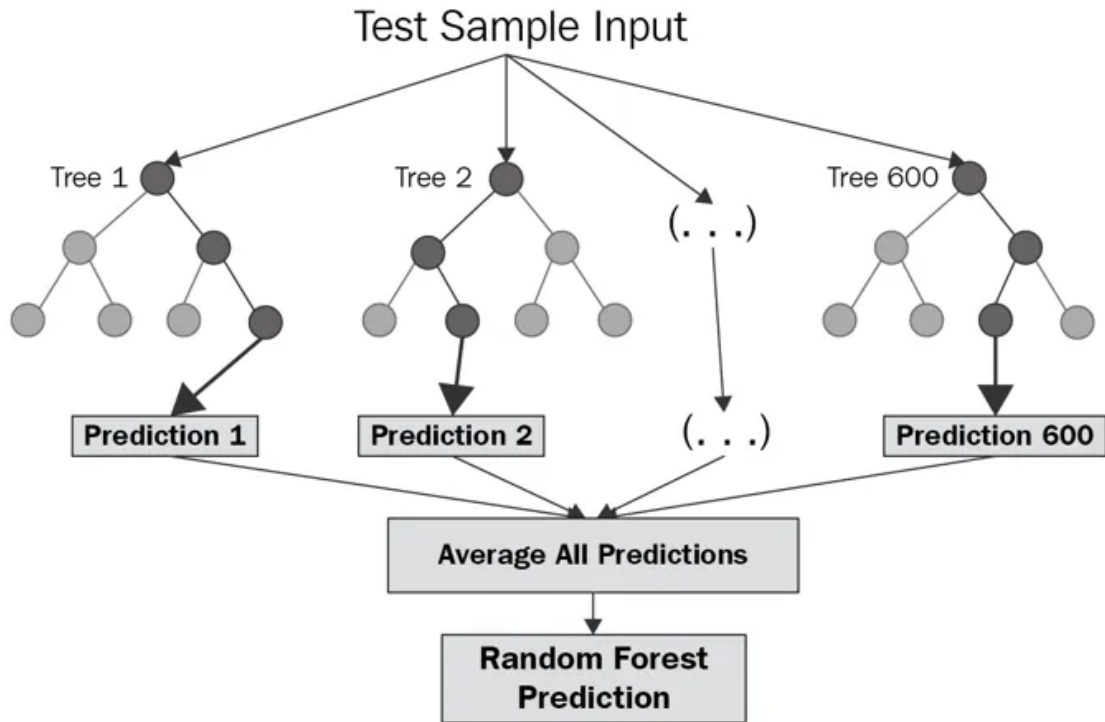


Figura 3 (Fonte <https://corporatefinanceinstitute.com/resources/data-science/random-forest/>)

- ii. il *gradient boosting* è un algoritmo di ensemble learning che combina più alberi decisionali in un unico modello sequenziale, allenando ogni nuovo albero per correggere gli errori del modello già costruito (Figura 4). Questo processo iterativo ha l'obiettivo di minimizzare la "funzione di perdita"⁵³ e di affinare la precisione predittiva.

⁵³ Metrica usata per quantificare il grado di errore nell'output di un modello ML.

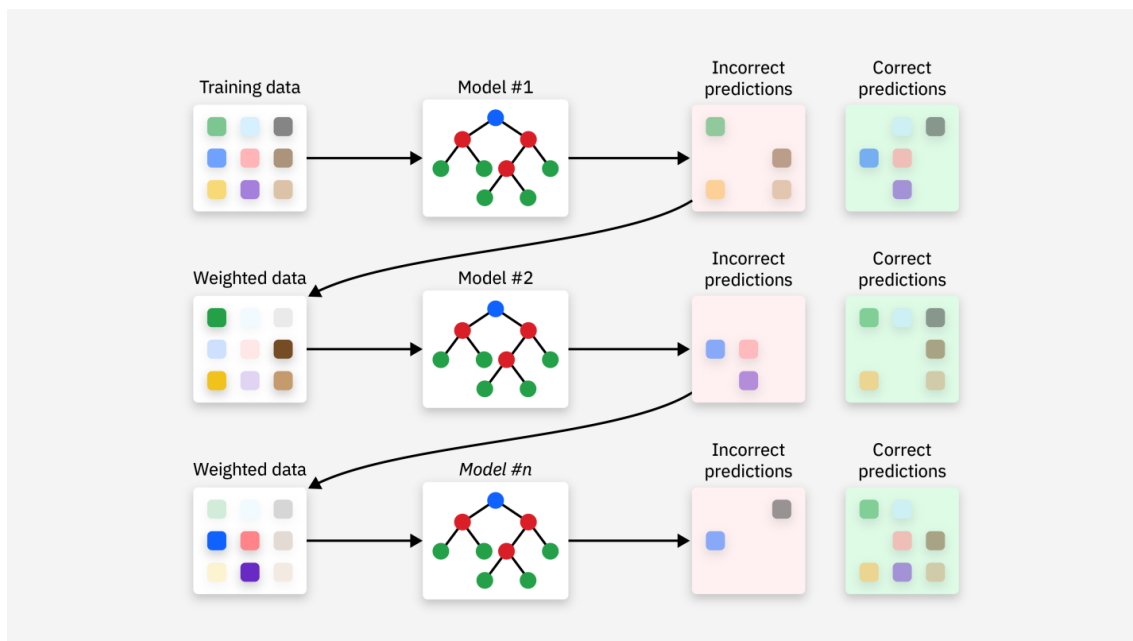


Figura 4 (Fonte B. Clark, F. Lee, *What is Gradient Boosting*, IBM)

Algoritmi di *deep learning*: si tratta di una famiglia di algoritmi di ML i cui processi elaborativi, ispirandosi al comportamento dei neuroni del cervello umano, si basano su reti che interconnettono nodi organizzati in livelli successivi (cosiddette reti neurali); in tale tipologia di algoritmo ad ogni livello della rete (Figura 5) corrisponderà una fase dell'apprendimento di concetti via via sempre più complessi⁵⁴. Tale tipologia di algoritmo è quella maggiormente efficace, fra le tecniche di *machine learning*, quando si trova ad elaborare grandi quantità di

⁵⁴ E. Bonaccorsi di Patti, F. Calabresi, B. De Varti, F. Federico, M. Affinito, M. Antolini, F. Lorizzo, S. Marchetti, I. Masiani, M. Moscatelli, F. Privitera e G. Rinna, *Intelligenza artificiale nel credit scoring. Analisi di alcune esperienze nel sistema finanziario italiano*, https://www.bancaditalia.it/publicazioni/qef/2022-0721/QEF_721_IT.pdf pp. 9 ss.

dati⁵⁵.

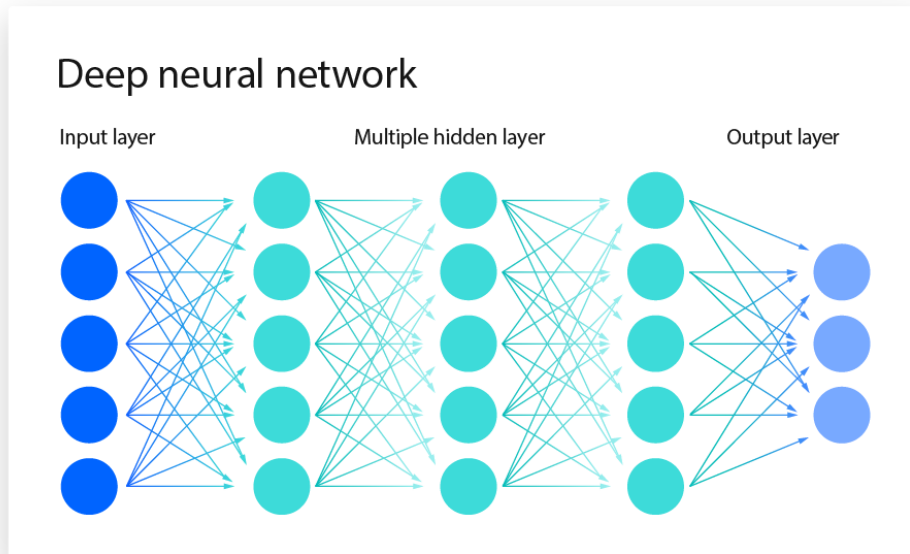


Figura 5 (Fonte <https://www.ibm.com/it-it/think/topics/neural-networks>)

6. Opportunità e rischi cyber dei modelli ML applicati al credito

Al di là delle categorizzazioni l'elemento che accomuna i modelli di *machine learning* è la loro capacità di individuare "corrispondenze invisibili" e tendenze complesse che l'occhio umano o i modelli pre-programmati non sarebbero in grado di cogliere. L'abilità del ML di analizzare set di dati vasti e diversificati per rilevare "segnali deboli", siano essi interazioni o non-linearità tra le variabili esplicative, permette di ottenere miglioramenti nelle previsioni rispetto alle misure convenzionali di solvibilità; questo focus sulle relazioni non lineari e sui pattern nascosti, piuttosto che sulla mera convalida di ipotesi predefinite, fa sì

⁵⁵ A. Md Zahangir, T. M. Taha, C. Yakopic, S. Westberg, P. Sidike, M. S. Nasrin, M. Hasan, B. C. Van Essen, A. A. S. Awwal, V. K. Asari, "A State-of-the-Art Survey on Deep Learning Theory and Architectures" in Electronics 8, no. 3, p. 292 2019

che tale tecnologia sia potenzialmente in grado di offrire un livello di analisi precedentemente inaccessibile⁵⁶.

Il portato di tale evoluzione tecnologica consente alle banche di ridurre i costi operativi, aumentare i tassi di approvazione senza incrementare proporzionalmente il rischio e conseguentemente dare l'accesso al credito ad un'ampia fascia di soggetti che storicamente hanno sofferto di razionamento del credito per la mancanza di attivi tangibili da offrire in garanzia⁵⁷.

L'inclusione finanziaria, tuttavia, non è il solo vantaggio che tali tecnologie permettono di perseguire; esse, infatti, possono anche apportare una maggiore stabilità del ciclo del credito: come noto, il credito bancario è prociclico motivo per cui si espande durante le fasi di boom economico e si contrae durante le recessioni, con conseguente amplificazione delle fluttuazioni macroeconomiche⁵⁸; inoltre il collaterale tradizionale è anch'esso legato al ciclo economico laddove il valore delle garanzie aumenta durante le espansioni (facilitando il credito) e diminuisce durante le crisi (implementando il credit crunch). Un recente studio della Banca d'Italia⁵⁹ ha riscontrato come l'utilizzo dell'intelligenza artificiale nella attività di concessione del credito possa mitigare proprio gli effetti delle fluttuazioni macroeconomiche: analizzando i dati delle banche italiane prima e durante la crisi COVID-19, è stato infatti rilevato come le banche che investono maggiormente in strumenti di intelligenza artificiale ("AI banks") presentino un'offerta di credito meno sensibile agli shock

⁵⁶ Financial Stability Board, *Artificial intelligence and machine learning in financial services: Market developments and financial stability implications*, 2017, <https://www.fsb.org/uploads/P011117.pdf>, pp. 4 ss.

⁵⁷ OECD, *AI in Business and Finance 2021: OECD Business and Finance Outlook*, 2021 https://www.oecd.org/content/dam/oecd/en/publications/reports/2021/09/oecd-business-and-finance-outlook-2021_377c2c18/ba682899-en.pdf pp 38 ss.

⁵⁸ F. Panetta, P. Angelini (coordinators), U. Albertazzi, F. Columba, W. Cornacchia, A. Di Cesare, A. Pilati, C. Salleo, G. Santini, *La prociclicità del sistema finanziario: lezioni dalla crisi*, https://www.bancaditalia.it/pubblicazioni/qef/2009-0044/OEF_44.pdf pp. 38 ss.

⁵⁹ L. Gambacorta, F. Sabatini, S. Schiaffi, *Artificial intelligence and relationship lending*, https://www.bancaditalia.it/pubblicazioni/temi-discussione/2025/2025-1476/en_tema_1476.pdf?language_id=1

macroeconomici e maggiormente tarata sulle condizioni specifiche della singola impresa.

Ciò comporta due conseguenze: da un lato, in tempi normali, le "AI banks" tendono a offrire più credito a tassi di interesse più bassi rispetto alle banche tradizionali a parità di durata della relazione, mitigando così l'estrazione di rendita che può caratterizzare il *relationship lending*⁶⁰; dall'altro lato durante una crisi, pur non offrendo una protezione aggiuntiva rispetto al *relationship lending* (che tende a proteggere i clienti di lunga data), il loro comportamento è meno volatile, contribuendo a smorzare gli effetti contro ciclici e a stabilizzare l'offerta di credito, con impatti positivi su investimenti e occupazione delle imprese clienti⁶¹.

L'adozione di sistemi basati su intelligenza artificiale nella concessione del credito introduce tuttavia nuovi rischi tecnici e sociali che richiedono una attenta disamina delle minacce che spaziano da problemi di discriminazione algoritmica e opacità della decisione a rischi di natura informatica che possono compromettere l'integrità, la riservatezza e la disponibilità dei dati e dei sistemi informatici.

Per quanto concerne gli aspetti di discriminazione questi saranno approfonditi in maniera dettagliata nel successivo capitolo che avrà modo di concentrarsi sul *credit scoring* algoritmico. In questo paragrafo appare invece opportuno, esaminare i rischi di cybersicurezza che possono presentarsi.

Se è innegabile che l'intelligenza artificiale possa contribuire a migliorare la gestione del rischio di credito dall'altro non va sottovalutato come la stessa possa

⁶⁰ Come ben sintetizzato da R. Corigliano, *Banca e impresa in Italia: caratteri evolutivi del relationship lending e sostegno dello sviluppo in Banca Impresa Società*, Fascicolo 1, 2007, Il Mulino, il *relationship lending* è un modello di concessione del credito, basato su una relazione stabile tra banca e cliente, ed alimentato da contratti pregressi di prestito e deposito; la banca acquisisce informazioni dirette e riservate sul cliente che vanno oltre i dati pubblici, arricchendole continuamente attraverso l'estensione della relazione ad altri servizi. Grazie a tale patrimonio informativo, di esclusiva conoscenza della banca, il merito creditizio viene valutato più approfonditamente ed il rapporto commerciale tra banca e cliente assume una ampia personalizzazione.

⁶¹ Ivi, p. 20

ampliare i potenziali scenari di vulnerabilità, facilitando fenomeni come l'intensificarsi degli attacchi informatici, lo sviluppo di malware sempre più sofisticati e l'automazione di tecniche di ingegneria sociale e disinformazione (ad esempio, phishing e impersonificazione); focalizzandosi proprio sulle minacce informatiche atte a colpire i modelli basati sull'AI il *Cyber Experts Group* del G7 ha avuto modo di indicarne tre in particolare⁶²:

- i. *Data Poisoning*: la manipolazione dei dati di addestramento da parte di un soggetto malevolo finalizzato a corrompere il processo di apprendimento dell'algoritmo e alterarne il comportamento futuro al fine di degradare le prestazioni complessive del modello (attacco non mirato) o di creare "backdoor" specifiche che possono essere sfruttate in seguito (attacco mirato)⁶³.
- ii. *Data Leaks*: sono possibili fughe di dati sensibili sia involontarie (a causa una erronea anonimizzazione dei dati di addestramento o del verificarsi di un overfitting⁶⁴) che volontarie (ad opera di utente malevoli che attraverso query attentamente formulate riescono ad estrarre informazioni sensibili dal modello).
- iii. *Prompt Injection*: Gli aggressori possono sfruttare la sequenza di istruzioni di ingresso che l'utente fornisce come input ad un algoritmo di apprendimento automatico (c.d. prompt⁶⁵) per manipolare gli output o recuperare informazioni sensibili.

Allo stesso tempo emergono forme di dipendenza sistemica dall'infrastruttura tecnologica e dalle piattaforme che raccolgono, controllano ed elaborano i flussi

⁶² G7 Cyber Expert Group, *Statement on AI and Cybersecurity*, <https://home.treasury.gov/system/files/136/G7-Cyber-Expert-Group-Statement-AI-and-Cybersecurity-2025.pdf>

⁶³ <https://www.ibm.com/it-it/think/topics/data-poisoning>

⁶⁴ Problematica del machine learning a causa della quale il modello memorizza il set di addestramento in modo troppo dettagliato e perde così la capacità di generalizzare e fare previsioni accurate su nuovi dati sconosciuti

⁶⁵ <https://www.treccani.it/vocabolario/prompt/>

di dati: la stabilità del sistema creditizio diventa dipendente dalla sicurezza informatica, dalle politiche di condivisione dei dati e dalla stabilità operativa dei *third party providers*, cui le banche si affidano per poter rimanere al passo con l'evoluzione dell'innovazione tecnologica, con conseguente possibilità di rischi di contagio tra il settore finanziario e quello tecnologico⁶⁶.

Una singola vulnerabilità in questo ecosistema può infatti avere effetti a cascata: un attacco di *data poisoning* che compromettesse un fornitore di dati di mercato non rappresenterebbe un rischio isolato per una singola banca. Se infatti più istituti utilizzano la stessa fonte di dati compromessa o lo stesso modello vulnerabile, essi diventano tutti simultaneamente suscettibili allo stesso attacco. Un singolo evento informatico può così diventare un potenziale rischio sistemico, che compromette contemporaneamente le capacità di valutazione del credito di numerosi attori chiave del mercato, con possibili impatti sulla stabilità finanziaria complessiva.

A riprova di tali considerazioni la Banca d'Italia ha avuto modo di rilevare l'aumento degli incidenti informatici nel sistema finanziario e la interconnessione degli stessi con i servizi forniti dai *third party providers*, posto che due terzi degli incidenti informatici del 2024 hanno coinvolto proprio i fornitori di servizi ICT.⁶⁷ Affrontare la complessa matrice di rischi associati all'IA nel credito richiede un approccio proattivo che combini soluzioni tecnologiche avanzate con solidi quadri di governance organizzativa.

Un'efficace difesa contro la minaccia cyber richiede l'implementazione di numerosi presidi e misure di sicurezza che i singoli operatori devono adottare, sia all'interno della propria organizzazione che in relazione ai servizi esternalizzati.

⁶⁶ Financial Stability Board, *The Financial Stability Implications of Artificial Intelligence*, <https://www.fsb.org/uploads/P14112024.pdf> pp. 15 ss

⁶⁷G. Siani, *Cyber Risk e Sistema finanziario*, Intervento al convegno "Sadibaquarantotto", 1/04/2025 Milano, <https://www.bancaditalia.it/pubblicazioni/interventi-vari/int-var-2025/SIANI-Bologna-28.03.2025.pdf> p.2

Per garantire che gli operatori raggiungano un livello adeguato di preparazione, sono fondamentali regole e linee guida riguardanti: politiche di governance e strategie aziendali; definizione di ruoli, responsabilità e sistemi di controllo; implementazione di misure efficaci di rilevazione, prevenzione e contrasto dei rischi e delle minacce; predisposizione di adeguate misure di continuità operativa; assegnazione di risorse adeguate e aggiornate; azioni volte a innalzare le conoscenze e la consapevolezza del personale.

Le regole assumono una loro importanza in un'ottica di risposta ultranazionale ai cyberattacchi dal momento che possono realizzarsi casi di minacce che hanno un impatto sistemico su più paesi.

Di fronte alla rapida evoluzione dell'IA e ai rischi associati, l'Unione Europea ha adottato un approccio normativo proattivo e completo, volto a creare un equilibrio tra la promozione dell'innovazione e la protezione dei diritti fondamentali, della sicurezza e della stabilità finanziaria.

Questo quadro si articola principalmente su due pilastri legislativi: Il Regolamento (UE) 2024/1689 del Parlamento Europeo e del Consiglio (c.d. AI Act), che stabilisce regole armonizzate per l'intelligenza artificiale in base ad un approccio risk based ed il Regolamento (UE) 2022/2554 del Parlamento europeo e del Consiglio del 14 dicembre 2022 Digital Operational Resilience Act (c.d. DORA) che rafforza la resilienza digitale del settore finanziario.

DORA in particolare stabilisce requisiti armonizzati per la gestione del rischio ICT per le varie tipologie di operatori del settore finanziario e introduce un regime di sorveglianza sui fornitori critici di servizi ICT. A tale normativa specificatamente destinata al settore finanziario si affiancano la direttiva Network and Information System Security (c.d. NIS2) e la direttiva Critical Entities Resilience (CER), che si applicano ai maggiori settori economici europei, tra i quali figurano il settore del credito e quello delle infrastrutture di mercato.

La risposta dell'Unione Europea va oltre una mera regolamentazione tecnica, ma si configura come una visione strategica per il futuro del mercato del credito. Diventa cruciale, a questo punto, decodificare il complesso intreccio di obblighi e opportunità che questo nuovo quadro normativo presenta per il credito bancario.

L'analisi che seguirà si propone di mappare in dettaglio questo percorso, esaminando nello specifico il credit scoring algoritmico, l'applicazione del *machine learning* per il calcolo del rischio di credito nel capitale nelle banche, le disposizioni normative nonché le sfide della vigilanza rispetto a tali nuove tecnologie.

2. IL CREDIT SCORING ALGORITMICO

1. La valutazione del merito creditizio tramite algoritmi di ML

L'accesso al credito costituisce uno dei pilastri delle moderne economie di mercato agendo come catalizzatore per lo sviluppo individuale e collettivo.

Una sua corretta allocazione è essenziale per sostenere i consumi, finanziare gli investimenti e promuovere l'innovazione; è in ragione di ciò che la concessione del credito viene preceduta da una complessa attività finalizzata a discernere quali siano i soggetti meritevoli della fiducia creditizia in quanto idonei a rendere effettiva la prospettiva restitutoria.

Se infatti l'attività di erogazione di moneta bancaria, attuata nell'esercizio della funzione creditizia, si sostanzia nell'accensione di relazioni giuridiche attive con i soggetti richiedenti, l'ordinamento bancario impone alle banche il compito di provvedere alla accurata cernita di quei soggetti, attraverso un procedimento, diretto ad accertarne il c.d. «merito creditizio» così da ridurre le probabilità di insolvenza⁶⁸.

La valutazione del merito creditizio, anche per effetto delle previsioni contenute nella disciplina consumeristica e in quella relativa al sovraindebitamento, si sostanzia in un obbligo teso non solo ad assicurare il rispetto, da parte dell'intermediario, dei principi di stabilità e sana e prudente gestione, di cui all'articolo 5 del Testo Unico Bancario, previsti per tutelare la stabilità del sistema finanziario nel suo complesso, ma anche a proteggere l'integrità della situazione patrimoniale del debitore assicurando la sostenibilità del prestito in modo da scongiurare l'esposizione di quest'ultimo ad una situazione di sovraindebitamento⁶⁹. Quest'ultima finalità, incentrata sulla tutela del contraente debole, ha acquisito una centralità crescente nel panorama normativo europeo come si evince dall'articolo 18 della Direttiva (UE) 2023/2225 (c.d. Consumer

⁶⁸ L. Di Brina, L'esercizio del credito: profili generali, in L'attività delle Banche, a cura di U. Urbani, Wolters Kluwer Italia, Seconda Edizione (2019) pp. 147 ss.

⁶⁹ E. Capobianco, Intelligenza artificiale e rapporti bancari, in Rivista di diritto bancario, (2024) pp. 1008 ss.

Credit Directive II) ai sensi del quale la valutazione del merito viene “*effettuata nell’interesse del consumatore*”.

Lo svolgimento delle attività di verifica del merito creditizio trova esplicitazione, nella prassi, nella attribuzione automatica, ai richiedenti il finanziamento, di un “punteggio di affidabilità” calcolato utilizzando sistemi di *credit scoring*, ossia metodi statistici finalizzati a prevedere la probabilità che un aspirante prestatore possa divenire insolvente e conseguentemente non ripagare il prestito ottenuto⁷⁰. Ogni aspirante prestatore appartiene ad una classe (quella dei solvibili/*good risk* o dei non solvibili/*bad risks*) sconosciuta al momento della concessione del finanziamento che pertanto necessita di essere prevista. La costruzione di un modello predittivo efficace, si basa sull’utilizzo di un campione di addestramento composto da dati storici, ossia una porzione di soggetti per i quali l’informazione relativa alla classe di appartenenza è nota poiché la loro storia creditizia è stata osservata fino alla chiusura del finanziamento o fino all’intervallo temporale entro il quale vi è interesse ad effettuare la classificazione. Il *credit scoring* utilizza le informazioni apprese dal campione per costruire una regola che consenta di assegnare correttamente il dato inserito (la richiesta di prestito) alla sua classe appartenenza⁷¹.

Tramite il *credit scoring*, il concedente ottiene un’indicazione dell’affidabilità creditizia normalmente espressa numericamente: l’ottenimento da parte di un soggetto di un elevato punteggio indicherà un’alta probabilità di rimborso completo delle obbligazioni di debito viceversa, un rating basso suggerirà che la controparte ha difficoltà nell’onorare le sue obbligazioni, motivo per cui il

⁷⁰ G. Albareto, R. Felici, E. Sette, *Does credit scoring improve the selection of borrowers and credit quality?*, in Temi di Discussione della Banca d’Italia, n. 1090, 2016, https://www.bancaditalia.it/pubblicazioni/temi-discussione/2016/2016-1090/en_tema_1090.pdf

⁷¹ E. Stanghellini, *Introduzione ai Metodi statistici per il credit scoring*, Springer (2009) pp. 5 ss.

prestito o gli verrà negato o concesso dietro pagamento di un tasso di interesse più elevato rispetto ad un soggetto con un credit scoring maggiore⁷².

Più nel dettaglio il calcolo del punteggio è composto da due elementi: i dati e le modalità con cui vengono analizzati.

I sistemi di valutazione del merito creditizio “tradizionali” utilizzano dati e informazioni relative al profilo economico-finanziario del cliente (c.d. hard data). Tali dati si distinguono per la loro oggettività e verificabilità in quanto riconducibili a valori numerici univoci⁷³ e possono essere classificati in tre categorie principali:

- i) dati anagrafici e sociodemografici: quali, ad esempio, l’età lo stato civile, l’anzianità lavorativa;
- ii) dati reddituali e patrimoniali: quali, ad esempio, il reddito dichiarato, le proprietà immobiliari, la sussistenza di altri impegni finanziari, lo scopo del prestito;
- iii) dati comportamentali creditizi: quali, ad esempio lo storico dei pagamenti passati, e dell’utilizzo delle linee di credito, la passata abilità di ripagare i debiti, le movimentazioni dei conti correnti⁷⁴.

I dati possono essere acquisiti sia direttamente dal richiedente che da fonti interne o esterne all’intermediario come le banche dati creditizie. I sistemi tradizionali analizzano i dati tramite modelli statistico-matematici che attribuiscono, a ciascun singolo richiedente, un punteggio che, sulla base di una correlazione statistica, esprime in termini predittivi il profilo di rischio, l’affidabilità e la puntualità nei pagamenti.

⁷² World Bank, Credit scoring approaches guidelines, <https://thedocs.worldbank.org/en/doc/935891585869698451-0130022020/original/CREDITSCORINGAPPROACHESGUIDELINESFINALWEB.pdf> p.3

⁷³ A. Davola, Algoritmi decisionali e trasparenza Bancaria, Utet Giuridica 2020 pp. 129 ss.

⁷⁴ European Banking Authority, *Guidelines on Loan Origination and Monitoring*, <https://www.eba.europa.eu/activities/single-rulebook/regulatory-activities/credit-risk/guidelines-loan-origination-and-monitoring#activity-versions,29/05/2020>, pp. 71 ss

Dal punto di vista matematico-statistico tali modelli rientrano nel novero della regressione lineare: essi creano una relazione tra una variabile target di cui si voglia prevedere l'esito (normalmente l'insolvenza di un debitore) e altre variabili, i c.d. "predittori", ritenute come statisticamente esplicative⁷⁵.

Storicamente, approcci come l'analisi discriminante lineare e, in modo predominante, la regressione logistica hanno rappresentato lo standard di settore per la valutazione del rischio di credito grazie alla loro affidabilità e trasparenza (*rectius* interpretabilità)⁷⁶: dal momento che il punteggio è il risultato di un modello lineare basato sulla somma ponderata di variabili note (ad esempio, reddito, livello di indebitamento, storico dei pagamenti) è relativamente facile spiegare il motivo del rigetto di un finanziamento. La linearità del modello, tuttavia, è anche la sua più grande limitazione, dal momento che non permette di identificare e modellare interazioni complesse e dinamiche non lineari⁷⁷, sempre più presenti nei moderni dataset finanziari con il conseguente rischio di ottenere una minore accuratezza predittiva e di denegare credito a clienti meritevoli o di approvare clienti che potrebbero diventare insolventi.

L'evoluzione tecnologica ha reso possibile il superamento dei modelli tradizionali grazie all'avvento dei big data, all'incremento esponenziale della potenza computazionale (conformemente a quanto affermato dalla c.d. legge di Moore⁷⁸) e all'avvento di framework di calcolo distribuiti che hanno permesso l'addestramento di modelli complessi di *machine learning* capaci di offrire un potere predittivo superiore ai modelli tradizionali di credit scoring in ragione

⁷⁵ P. Gaggero, C. A. Valenza, Le moderne tecniche di credit scoring tra GDPR, disciplina di settore e AI Act, in Rivista di Diritto Bancario (2024) pp. 833 ss.

⁷⁶ L. Thomas, J. Crook, J. N., D. Edelman, *Credit Scoring and its Applications* (Second Edition), SIAM (2017) pp. 26 ss.

⁷⁷ J. Gareth, D. Witten, T. Hastie, R. Tibshirani, *An Introduction to Statistical Learning: with Applications in R*, Second edition, 2021, Springer pp. 24 ss.

⁷⁸ Legge empirica formulata da Gordon Moore, cofondatore di Intel, secondo la quale il numero di transistor all'interno di un microprocessore raddoppia circa ogni diciotto mesi, con un corrispondente aumento della potenza di calcolo e una riduzione dei costi per transistor.

della loro abilità di selezionare in maniera automatica le variabili rilevanti e catturare le interrelazioni fra le stesse⁷⁹.

La modifica delle modalità di analisi ha portato ad un ampliamento delle fonti di dati utilizzate: non più solo dati finanziari strutturati ma anche dati finanziari non strutturati (analisi delle informazioni transazionali e di quelle derivate da open banking), dati non finanziari strutturati (dati di tipo sociodemografico ottenuti anche da fonti terze) e dati non finanziari non strutturati (dati di navigazione, *digital footprint*, informazioni conferite sui social network). Il ricorso a modelli di ML, che sfruttano le citate fonti di dati alternative, permette di considerare variabili che non hanno una chiara interpretazione/relazione economica e che pertanto non verrebbero considerate in un modello tradizionale ma che l'algoritmo identifica come statisticamente significative grazie al suo peculiare processo di addestramento e calibrazione di questi modelli⁸⁰.

L'integrazione di informazioni alternative, complementari a quelle impiegate nei modelli tradizionali, porta ad un miglioramento dell'accuratezza predittiva migliorando la stabilità del sistema finanziario nel suo complesso, favorendo una gestione, sia preventiva che reattiva dell'indebitamento dei consumatori e, contestualmente, amplia la capacità di valutazione del merito creditizio nei confronti di soggetti che altrimenti sarebbero esclusi o penalizzati nell'accesso al credito⁸¹. A riprova di tali affermazioni si può citare il caso della "*digital footprint*" termine con cui si definisce l'insieme delle informazioni che le persone "lasciano online" in maniera passiva semplicemente visitando o registrandosi su di un sito internet. Tale impronta è caratterizzata dal fatto che può venire collezionata a

⁷⁹ M. Moscatelli, S. Narizzano, F. Parlapiano, G. Viggiano, *Corporate default forecasting with machine learning*, in *Temi di discussione della Banca d'Italia*, n. 1256, 2019, https://www.bancaditalia.it/pubblicazioni/temi-discussione/2019/2019-1256/en_Tema_1256.pdf?language_id=1

⁸⁰ E. Bonaccorsi di Patti, F. Calabresi, B. De Varti, F. Federico, M. Affinito, M. Antolini, F. Lorzio, S. Marchetti, I. Masiani, M. Moscatelli, F. Privitera e G. Rinna, *Intelligenza artificiale nel credit scoring. Analisi di alcune esperienze nel sistema finanziario italiano*, https://www.bancaditalia.it/pubblicazioni/qef/2022-0721/QEF_721_IT.pdf pp. 14 ss.

⁸¹ A. Davola, *Algoritmi decisionali e trasparenza Bancaria*, Utet Giuridica 2020 p. 140

costo quasi nullo da qualunque operatore digitale: basta che un utente interagisca con un sito e una serie di informazioni standard (device, browser, email, ora di accesso, ecc.) verranno generate automaticamente e potranno essere immediatamente acquisite dall'intermediario senza che debbano essere compilati moduli e formulari o vada sottoposta documentazione finanziaria o autorizzati controlli invasivi da parte del cliente; in altre parole l'impronta digitale è disponibile *by default* nel mondo online.

Le variabili che compongono la *footprint* seppur eterogenee possono essere così categorizzate:

- i. **tracce socioeconomiche:** quali il sistema operativo del dispositivo utilizzato (l'ecosistema *iOS*, ad esempio, è idoneo ad indicare soggetti con un reddito superiore rispetto agli utilizzatori di *Android*) o il *provider* di posta elettronica (certi domini e-mail possono segmentare la clientela per età o status: si pensi, ad es., a quelli a pagamento);
- ii. **tracce "comportamentali":** quali il canale di acquisizione del cliente (ad esempio un utente proveniente da un comparatore di prezzi presenta un rischio di insolvenza dimezzato rispetto a chi risponde a pubblicità a pagamento), la cronometria delle transazioni (gli acquisti notturni raddoppiano il rischio di default) e l'analisi dell'indirizzo e-mail (la presenza del nome e cognome dell'acquirente nell'indirizzo o errori di battitura sono predittori del comportamento di pagamento futuro).

Le considerazioni esposte sono il risultato di uno studio⁸² condotto su un dataset, di oltre 250.000 transazioni, appartenente ad una società tedesca di *e-commerce* il cui modello aziendale si basa sulla vendita di beni con pagamento posticipato dopo la spedizione del prodotto, cosicché ogni acquisto dei clienti si sostanzia, nei fatti, in un piccolo prestito commerciale che l'azienda concede fino al saldo.

⁸² T. Berg, V. Burg, A. Gombović, M. Puri, *On the Rise of FinTechs: Credit Scoring Using Digital Footprints in The Review of Financial Studies*, Volume 33, Issue 7, July 2020, Pages 2845–2897

La società, in ragione di tale *business model*, ha da sempre avuto la necessità di valutare l'affidabilità della clientela per evitare insolvenze e la ha soddisfatta, storicamente, utilizzando un sistema di *credit score*, fornito da una agenzia privata di informazioni creditizie, per stimare il rischio di ogni transazione. Ad un certo punto, tuttavia, l'azienda ha deciso di affiancare a tale metodo anche un insieme di variabili di digital footprint del cliente per decidere se accettare l'ordine oppure rifiutarlo; in seguito alla introduzione delle variabili digitali si è riscontrato un calo significativo dei tassi di insolvenza senza che ciò abbia comportato la riduzione dell'accesso complessivo al credito: in un sottocampione di transazioni per cui precedentemente si usava il credit score e successivamente si è aggiunto il footprint il default rate è sceso dal 3,62% al 2,33% dopo l'adozione del digital footprint (una riduzione di circa 42%); al tempo stesso, il tasso di accettazione degli ordini è leggermente aumentato passando dal 39,0% al 40,1%. Ulteriore risultato significativo emerso dallo studio è come la *digital footprint* permetta di ottenere risultati affidabili anche nei confronti dei richiedenti *unscorable* (privi di una storia creditizia formale) cosicché il suo utilizzo permette di estendere la valutazione del merito di credito anche a soggetti che altrimenti verrebbero esclusi o penalizzati nell'accesso al credito, come i c.d. "*thin-file*" (con storico creditizio limitato) o "*no-file*", tra cui rientrano ad esempio i giovani ed i migranti⁸³.

Lo *scoring* di questi individui è reso possibile dalla velocità con cui vengono generate queste nuove tipologie di dati: si assiste infatti al passaggio da una valutazione prevalentemente retrospettiva della storia finanziaria a una sorveglianza prospettica e in tempo reale del comportamento permessa dalla capacità dei modelli di machine learning di sfruttare i dati comportamentali in tempo reale per fare previsioni sul comportamento finanziario futuro.

⁸³ L. Ammannati, G. L. Greco, *Il credit scoring "intelligente": esperienze, rischi e nuove regole*, in *Rivista di diritto bancario* p. 465 ss.

Il merito creditizio, di conseguenza, subisce una metamorfosi concettuale passando dall'essere un attributo statico, basato sul comportamento passato del debitore e su dati patrimoniali, ad un processo dinamico dipendente da chi il debitore appare essere sulla base delle tracce di dati generate nella vita quotidiana.

Emergono dunque vantaggi dall'impiego dell'intelligenza artificiale nell'ambito del credit scoring e dalla dataficazione del comportamento individuale quali la riduzione dei costi operativi e di transazione nonché delle asimmetrie informative ma dall'altro lato si stagliano nuovi rischi quali la difficile spiegabilità dei modelli, la possibile mancanza di equità, la discriminazione algoritmica; argomenti questi che saranno meglio affrontati nelle prossime pagine.

2. Credit Scoring Algoritmico: quali vantaggi e rischi per gli utenti?

Come si è potuto iniziare a verificare nel precedente paragrafo i modelli di credit scoring basati sull'intelligenza artificiale presentano il vantaggio di rendere più facili e spedite le operazioni di concessione del credito fornendo una immagine più completa del richiedente.

Vari studi hanno confrontato la capacità delle tecniche di ML di prevedere il default delle imprese e dei privati con l'approccio statistico econometrico tradizionale (basato sulla stima della probabilità di default)⁸⁴. Dalle loro risultanze è emerso come i modelli di ML siano maggiormente accurati nel prevedere le insolvenze grazie all'estensione della gamma delle forme funzionali

⁸⁴ Per un'accurata sintesi degli studi condotti in letteratura dal 2018 al 2024 si rinvia a H. Ayari, Ramzi Guetari, N. Kraïem, *Machine learning powered financial credit scoring: a systematic literature review*, in *Artificial Intelligence Review* Vol. 59, 2026

e alla capacità di individuare interrelazioni tra le diverse variabili considerate dal modello; ma non solo, la capacità del ML di gestire ed elaborare grandi quantità di dati in volume (numero di osservazioni) e ricchezza (numero di variabili, tipi di dato), sfruttando l'aumentata potenza di calcolo resasi disponibile negli anni più recenti ha permesso di sfruttare fonti di dati alternative e considerare variabili che non hanno una chiara interpretazione/relazione economica permettendo di aumentare l'accuratezza previsiva e al contempo estendere la possibilità di valutare il merito creditizio⁸⁵. Tali sistemi sono idonei a migliorare l'inclusione finanziaria permettendo a fasce di popolazione che non hanno facile accesso ai servizi finanziari, a causa della mancanza delle informazioni tradizionalmente utilizzate nella valutazione del merito di credito, di divenire "scrutinabili" a costi ridotti dagli intermediari ai fini della concessione di un credito⁸⁶.

Ai modelli tradizionali di valutazione del merito creditizio (come il punteggio Fair Isaac Corporation c.d. FICO⁸⁷ ampiamente utilizzato negli Stati Uniti) basati su di un insieme ristretto di dati finanziari che hanno portato all'esclusione dall'accesso al credito una porzione significativa della popolazione, in quanto priva di una storia creditizia sufficientemente lunga o strutturata per essere valutata, si contrappongono i nascenti sistemi di credit scoring algoritmico che possono, svolgere un ruolo a tratti "sociale". La letteratura empirica

⁸⁵ E. Bonaccorsi di Patti, F. Calabresi, B. De Varti, F. Federico, M. Affinito, M. Antolini, F. Lorzio, S. Marchetti, I. Masiani, M. Moscatelli, F. Privitera e G. Rinna, *Intelligenza artificiale nel credit scoring. Analisi di alcune esperienze nel sistema finanziario italiano*, https://www.bancaditalia.it/publicazioni/qef/2022-0721/QEF_721_IT.pdf p.15

⁸⁶ M. Bianco, *Intelligenza artificiale nel credit scoring: analisi di alcune esperienze nel sistema finanziario italiano*, *Presentazione del QEF in pubblicazione*

⁸⁷ Come spiegato da A. Davola, op. cit, pp. 130 ss. il FICO si basa sulla ripartizione delle informazioni relative al merito di credito di un individuo in cinque categorie (storia creditizia del cliente; reddito; lunghezza della storia creditizia; tipologia di prestiti già richiesti; nuovo tipo di credito per il quale si presenta richiesta); per ciascuna categoria viene assegnato al richiedente un punteggio individuale e successivamente tramite una somma algebrica, ponderata per il tipo di categoria, si ottiene un punteggio complessivo fra i 350 e gli 800 punti rispetto al quale il richiedente verrà collocato all'interno di una fra cinque "fasce di meritevolezza" in base alle quali si stabilisce il rischio di insolvenza del soggetto e le connesse condizioni del prestito.

internazionale, analizzando l'operato di piattaforme FinTech, quali *LendingClub*, ha confermato che l'uso di dati alternativi e algoritmi di ML permette di estendere l'accesso al credito a debitori che i sistemi tradizionali classificherebbero come subprime⁸⁸. Molti di questi soggetti si rivelano, infatti, nella realtà "buoni debitori" il cui rischio, grazie alla tecnologia, viene "prezzato" in modo più accurato, offrendo loro condizioni di finanziamento più vantaggiose con conseguente aumento del tasso di inclusione. Le considerazioni enunciate non si limitano ad applicarsi ai soggetti "*thin-file*" o "*no-file*" ma si estendono anche a categorie di soggetti vulnerabili, quali minoranze etniche come gli afroamericani e gli ispanici rispetto ai quali è stata verificata sul mercato statunitense⁸⁹ la riduzione delle disparità di accesso al credito grazie all'AI credit scoring.

Nel contesto italiano, l'adozione di sistemi di credit scoring basati sull'intelligenza artificiale, sebbene sia ancora in una fase non pienamente matura, è guidata dalla duplice promessa di una maggiore accuratezza predittiva e di una significativa spinta verso l'inclusione finanziaria. Un'approfondita indagine della Banca d'Italia⁹⁰ ha delineato un mercato del credito nel quale il ricorso a strumenti di IA nella valutazione del credit risk è in rapida espansione. Tale transizione, pur procedendo con cautela, è animata dalla consapevolezza che tali tecnologie possono rappresentare uno strumento cruciale per superare i limiti dei modelli tradizionali: l'indagine ha rilevato come gli intermediari comincino «a sfruttare il potenziale in termini di leva inclusiva che i sistemi di AI-ML

⁸⁸ J. Jagtiani, C. Lemieux, *The roles of alternative data and machine learning in fintech lending: Evidence from the LendingClub consumer platform*, <https://www.philadelphiafed.org/-/media/frbp/assets/working-papers/2018/wp18-15r.pdf>

⁸⁹ R. Bartlett, A. Morse, R. H Stanton, N. Wallace, *Consumer-Lending Discrimination in the Fintech Era*, in NBER Working Paper No. w25943, (June 2019) p.2.

⁹⁰ E. Bonaccorsi di Patti, F. Calabresi, B. De Varti, F. Federico, M. Affinito, M. Antolini, F. Loriglio, S. Marchetti, I. Masiani, M. Moscatelli, F. Privitera e G. Rinna, *Intelligenza artificiale nel credit scoring. Analisi di alcune esperienze nel sistema finanziario italiano*, https://www.bancaditalia.it/pubblicazioni/qef/2022-0721/QEF_721_IT.pdf

possono esprimere accedendo a fonti dati alternative e sviluppando modelli in grado di trattarli più efficacemente⁹¹» e ha riscontrato le prime esperienze di allargamento dell'offerta di credito a fasce di clienti, tradizionalmente esclusi per la loro scarsa storia creditizia, dimostrando così che i benefici del credit scoring non sono meramente teorici ma trovano concretizzazione nella prassi operativa.

Non può tuttavia non evidenziarsi l'esistenza di un *trade-off* fra i benefici ed i rischi, seppur non ancora compiutamente percepiti dagli intermediari, che richiedono di bilanciare l'incremento dell'accuratezza delle previsioni con il costo in termini di minore spiegabilità del modello ed il rischio che possano verificarsi casi di discriminazione e distorsione algoritmica⁹².

Il rischio discriminatorio si presenta quando bias cognitivi, derivanti da malfunzionamenti degli algoritmi, portano a una rappresentazione distorta, se non addirittura ingiusta, della capacità finanziaria delle persone. In ambito informatico il termine "discriminare" indica la pratica di distinguere, separare o differenziare i dati di input attraverso l'applicazione di un modello discriminatorio (*pattern discrimination*); la discriminazione cui ci riferiamo ha dunque una valenza meramente tecnica e riguarda il processo di filtraggio di informazioni specifiche ricavate da dati, con la finalità di produrre output riferibili a un'analisi mirata o a un trattamento differenziato sulla base di determinati criteri⁹³.

Gli algoritmi alla base delle tecnologie di intelligenza artificiale, seppur in prima battuta possano far sorgere l'idea di una loro neutralità e obiettività a ben vedere, altro non sono che artefatti umani e, pertanto sono influenzati dal sistema di significati e idee appresi attraverso il linguaggio; l'intelligenza artificiale apprende dai dati che le vengono forniti e tali dati in virtù della loro origine

⁹¹ Ivi, p.7

⁹² Ivi, pp. 15-39

⁹³ V. Barone, *La discriminazione ai tempi dell'intelligenza artificiale: la questione algoritmica*, in *Diritto e Tecnologie Informatiche*, a cura di T. Casadei e S. Pietropaoli, Wolters Kluwer, 2024 p. 285

"umana" possono essere lacunosi, parziali e, ovviamente discriminatori: se i dati contengono disuguaglianze, gli algoritmi possono replicarle e persino amplificarle generando risultati discriminatori⁹⁴.

Gli algoritmi non sono dunque neutri e oggettivi, ma intrinsecamente plasmati dalla prospettiva di chi li progetta, addestra e impiega.

Più nel dettaglio i *bias*⁹⁵ atti a generare un comportamento discriminatorio della macchina possono verificarsi in:

- i. **fase di design:** la fase in cui si individuano delle variabili obiettivo (*target variables*⁹⁶) e si stabilisce cosa la macchina dovrà prevedere o classificare. La definizione del "problema" che l'algoritmo è chiamato a risolvere e la conseguente selezione delle variabili non sono atti neutrali, ma incarnano i valori, gli scopi e i pregiudizi dell'organizzazione che sviluppa il sistema. È in questa fase che un *bias* può essere codificato in modo strutturale rendendone particolarmente insidiosa la successiva eradicazione⁹⁷.
- ii. **fase di addestramento:** la fase in cui si utilizzano dati storici per insegnare alla macchina a fare previsioni. L'addestramento è basato sul fondamentale presupposto che i dati storici siano un riflesso sufficientemente accurato della realtà cosicché questi possano essere usati per prevedere correttamente il futuro⁹⁸; per tale motivo è essenziale che

⁹⁴ K. Crawford, *Atlas of AI: power, politics, and the planetary costs of artificial intelligence*, Yale University Press 2021 pp. 211 ss

⁹⁵ Con il termine *bias* ci si riferisce ad una distorsione sistematica del giudizio o del processo decisionale, che porta a conclusioni, stime o interpretazioni errate o deviate dalla realtà oggettiva.

⁹⁶ La variabile obiettivo può essere definita come la variabile dipendente che il modello cerca di spiegare o prevedere utilizzando le informazioni in ingresso (le variabili indipendenti che nel caso di specie sono i soft e gli hard data)

⁹⁷ F. Z. Borgesius, *Discrimination, Artificial Intelligence and Algorithmic*, Directorate General of Democracy, Council of Europe (2018), <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73> pp. 15 ss

⁹⁸ Peraltro, come osservato correttamente da H. Suresh e J.V. Guttag, *A Framework for Understanding Unintended Consequences of Machine Learning*, 2021 <https://dl.acm.org/doi/10.1145/3465416.3483305>, anche se i dati sono misurati e campionati perfettamente, la "realtà del nostro mondo" può portare un modello a produrre risultati indesiderati. Ad esempio, anche se avessimo accesso alla caratteristica "criminalità" misurata in modo perfetto, essa potrebbe comunque riflettere fattori storici che hanno portato a una maggiore criminalità nei quartieri più poveri.

siano selezionati dati rappresentativi e non distorti al fine di evitare che il modello apprenda informazioni parziali o sbagliate e che le amplifichi nella fase di output⁹⁹.

Durante l'addestramento non tutti i dati disponibili vengono utilizzati; i progettisti eseguono infatti una selezione (c.d. *feature selection*) e una trasformazione (*feature engineering*) delle variabili che ritengono essere maggiormente rilevanti all'interno del set di dati. Se chi seleziona le *features* ha preconcetti o compie assunzioni errate sulle relazioni tra variabili e risultati ciò può causare l'introduzione di un bias che riflette tali preconcetti piuttosto che la realtà dei dati¹⁰⁰.

L'algoritmo peraltro potrebbe non limitarsi a replicare il *bias* presente nei dati bensì amplificarlo. I modelli di classificazione, per massimizzare la loro accuratezza complessiva tendono infatti a concentrarsi sui pattern più forti e a rafforzare le correlazioni maggioritarie, anche a costo di commettere errori sistematici sui gruppi minoritari. Ad esempio, immaginiamo un dataset di addestramento in cui un gruppo minoritario presenta un tasso di insolvenza leggermente superiore (6%) rispetto al gruppo maggioritario (4%). L'algoritmo, per ottimizzare la propria performance predittiva potrebbe "imparare" da quella minima differenza del 2% una regola decisionale molto più rigida, e negare il credito al 40% dei richiedenti del gruppo minoritario e solo al 15% di quelli del gruppo maggioritario¹⁰¹. La possibilità che l'algoritmo trasformi una piccola discrepanza statistica nei dati di input in una vasta disparità di

⁹⁹ K. Shrishak, *AI-Complex Algorithms and effective Data Protection Supervision - Bias evaluation*, https://www.edpb.europa.eu/system/files/2025-01/d1-ai-bias-evaluation_en.pdf, pp. 5 ss.

¹⁰⁰ F. Z. Borgesius, *Discrimination, Artificial Intelligence and Algorithmic*, Directorate General of Democracy, Council of Europe (2018), <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73> p. 20

¹⁰¹ M.Hall, L. van der Maaten, L. Gustafson, M. Jones, A. Adcock, *A Systematic Study of Bias Amplification*, <https://arxiv.org/abs/2201.11706>

trattamento nell'output è motivata dal fatto che dal punto di vista puramente matematico, applicare uno standard più severo al gruppo più piccolo aiuta il modello a minimizzare l'errore totale seppur a discapito dell'equità.

- iii. **fase di impiego:** la fase in cui l'algoritmo, una volta progettato ed addestrato, viene applicato nella pratica. Questa fase non è una semplice esecuzione passiva di istruzioni bensì un momento dinamico in cui il modello può sia generare nuovi *bias*, a causa dell'interazione con un ambiente sociale in costante mutamento, che accentuarli creando un feedback circolare, in cui i risultati discriminatori influenzano i dati futuri, con conseguente perpetuazione ed amplificazione degli eventuali pregiudizi iniziali¹⁰².

I modelli, inoltre, col passare del tempo possono deteriorarsi poiché le variabili significative che emergevano dai dati con cui sono stati addestrati (ad es. il lavorare in un determinato settore economico) potrebbero non cogliere più l'evoluzione della attuale realtà e continuare a dare un output viziato da un pregiudizio storico. Per esemplificare si pensi ad un modello il quale, in virtù dei dati con cui è stato addestrato, ha appreso che un determinato settore lavorativo è poco produttivo e genera precariato cosicché ai lavoratori in quell'ambito vengono erogati meno prestiti; ad un certo punto si verifica una significativa crescita economica in quel settore con conseguente stabilizzazione degli impieghi. L'algoritmo, tuttavia ignaro del cambiamento della situazione,

¹⁰² Il problema del *feedback loop* è stato ben spiegato in *Bias in Algorithms – Artificial Intelligence and Discrimination* dalla European Union Agency for Fundamental Rights, che, prendendo ad esempio i sistemi che aiutano a determinare quali quartieri dovrebbero essere pattugliati sulla base dell'analisi relativa ai crimini registrati (c.d. sistemi di polizia predittiva) ha osservato come il *feedback loop* possa creare una "profezia autoavverante". Se il sistema "rilevasse" più crimini nel distretto A e decidesse di inviare lì più pattuglie il loro maggior numero porterebbe a registrare più crimini nel distretto A e tali dati reinseriti nel sistema, rafforzerebbero la "convinzione" del sistema che nel distretto A ci sia più criminalità.

continuerà ad applicare le vecchie regole, assegnando bassi punteggi di affidabilità a soggetti che si trovano ora in condizione di ripagare il credito.

Il bias può peraltro essere voluto dal programmatore stesso che potrebbe introdurre una selezione parziale dei dati atta a danneggiare specifiche categorie di persone (c.d. *masking*)¹⁰³. Ad esempio, si può immaginare un modello di credit scoring nel quale un programmatore, introducendo un bias nell'algoritmo, decida intenzionalmente di penalizzare richiedenti che abitino in una determinata regione perché li ritiene meno affidabili.

Soffermandoci sul rischio di opacità della decisione va osservato come nell'AI credit scoring, avendo l'algoritmo la possibilità di selezionare ed utilizzare una pluralità di informazioni solo indirettamente collegate alla sfera finanziaria dell'individuo, il processo di valutazione del merito creditizio del soggetto appare foriero di scarsa trasparenza. Tale problema, seppur in parte già presente nei modelli di scoring tradizionali viene qui esasperato: se, infatti, pur in presenza di margini di incertezza, l'utilizzo di dati direttamente collegati al profilo patrimoniale (specie se, come nel sistema FICO, sono conosciuti i principali criteri di ponderazione di questi) permetteva ai richiedenti una comprensione, seppur parziale, delle dinamiche di base attraverso le quali il loro score era stato calcolato, l'introduzione di un numero estremamente più ampio di variabili da considerare e l'impiego di modelli algoritmici sempre più complessi tali da essere delle vere e proprie *black box* rende difficile se non addirittura impossibile ricostruire e spiegare in modo intelligibile, all'esterno, le ragioni specifiche di una determinata decisione come il diniego del credito. Ciò da un lato complica enormemente gli sforzi per individuare e correggere i *bias* e la conseguente discriminazione che potrebbero generare e dall'altro è idoneo a

¹⁰³ V. Barone, *La discriminazione ai tempi dell'intelligenza artificiale: la questione algoritmica*, in *Diritto e Tecnologie Informatiche*, a cura di T. Casadei e S. Pietropaoli, Wolters Kluwer, 2024 p. 285

precludere *ab origine* agli interessati la possibilità di contestare rifiuti che appaiano ingiusti o ingiustificati¹⁰⁴.

La difficoltà nello spiegare puntualmente i motivi della decisione algoritmica è da rinvenirsi nel processo che i sistemi basati sul machine learning (soprattutto quelli di *deep learning*) seguono per consegnare l'*output* più efficace per gli obiettivi impartiti: essi ricercano all'interno dei dati forniti una regolarità statistica, una correlazione, dei pattern, dei modelli di ricorrenza tramite i quali estrarre informazioni predittive funzionali all'effetto richiesto. Si tratta di un processo analitico iterativo che permette al sistema di progredire e generare continuamente nuovi modelli sulla base dei dati di *feedback* così da adeguare, modificare, perfezionare autonomamente le proprie azioni, in termini che seppur funzionali non sempre sono prevedibili¹⁰⁵.

Aspetto non scevro da rischi è poi il ricorso all'*outsourcing* da parte delle banche per avvalersi di tali strumenti. Gli istituti di credito, non possedendo adeguate competenze specialistiche in materia di AI sono spinti ad esternalizzare la funzione verso fornitori non vigilati, complicando ulteriormente il governo e il controllo dei rischi e delle logiche decisionali sottostanti in quanto le analisi si basano su dati e metodologie di proprietà dei fornitori che da sempre hanno rifiutato di condividerle con i destinatari degli output dei sistemi AI¹⁰⁶.

Risulta, inoltre, dall'indagine della Banca d'Italia¹⁰⁷ come i meccanismi di governance e controllo degli intermediari non siano ancora aggiornati rispetto all'utilizzo delle tecniche AI: è emerso in particolare un "*mancato ricorso a tecniche*

¹⁰⁴ A. Davola, *Algoritmi decisionali e trasparenza Bancaria*, Utet Giuridica 2020 pp. 142-143

¹⁰⁵ M. Rabitti, *Credit scoring via machine learning e prestito responsabile*, in *Rivista di Diritto Bancario* Fascicolo I, Anno 2023, Sezione seconda, p. 190

¹⁰⁶ European Banking Authority, *Report On Big Data And Advanced Analytics*, https://www.eba.europa.eu/sites/default/files/document_library/Final%20Report%20on%20Big%20Data%20and%20Advanced%20Analytics.pdf p. 45

¹⁰⁷ E. Bonaccorsi di Patti, F. Calabresi, B. De Varti, F. Federico, M. Affinito, M. Antolini, F. Lorizzo, S. Marchetti, I. Masiani, M. Moscatelli, F. Privitera e G. Rinna, *Intelligenza artificiale nel credit scoring. Analisi di alcune esperienze nel sistema finanziario italiano*, https://www.bancaditalia.it/pubblicazioni/qef/2022-0721/QEF_721_IT.pdf

esplicitamente mirate alla riduzione della distorsione algoritmica”, come il bilanciamento del dataset o il controllo del bias storico.

Queste lacune e criticità hanno portato alla necessità di dare una adeguata regolazione ad un fenomeno *disruptive* quale l’intelligenza artificiale.

La risposta a tale esigenza è arrivata dapprima tramite un’interpretazione estensiva, da parte della giurisprudenza europea della legislazione già in vigore, in materia di trattamento dei dati personali e, successivamente, dall’aggiornamento della normativa bancaria, ad opera della Direttiva sul Credito ai Consumatori, e alla adozione di una normativa europea specifica in materia di intelligenza artificiale: l’AI Act.

3. L’applicazione del GDPR al *credit scoring algoritmico*: le sentenze “Schufa” e “D&B”.

La valutazione del merito creditizio tramite algoritmi non si limita a digitalizzare procedure esistenti, ma configura un vero e proprio “processo decisionale automatizzato” basato sulla profilazione¹⁰⁸. In tale processo, la discrezionalità umana e la valutazione qualitativa dell’intermediario vengono sostituite, integralmente o parzialmente, da un’inferenza statistica: l’algoritmo elaborando una vasta mole di dati assegna al soggetto lo *score* che rappresenta una stima probabilistica del suo rischio di insolvenza.

Quando tale punteggio determina l’esito della richiesta (concessione o diniego) senza che vi sia un intervento umano significativo, ci troviamo di fronte a una fattispecie giuridicamente rilevante per il Regolamento Generale sulla Protezione dei Dati Personali, Regolamento n. 2016/679 (in seguito GDPR).

¹⁰⁸ Gruppo di Lavoro Articolo 29 per la Protezione dei Dati, *Linee guida sul processo decisionale automatizzato relativo alle persone fisiche e sulla profilazione ai fini del regolamento 2016/679*, 2018, pp. 7 ss

Il GDPR prevede infatti una tutela ai sensi del combinato disposto degli articoli 13 c. 2 lett f), 15 e 22 rispetto a tale tipo di trattamento dei dati; più nello specifico, l'art. 22 stabilisce un divieto generale di ricorrere a processi decisionali automatizzati da parte dei titolari del trattamento salvo che:

- i. la decisione sia necessaria per la conclusione o l'esecuzione di un rapporto negoziale tra l'interessato e il titolare del trattamento;
- ii. la decisione sia autorizzata dal diritto nazionale dello Stato membro o dal diritto dell'Unione;
- iii. l'interessato abbia esplicitamente prestato il consenso alla decisione

Inoltre, il citato articolo vieta di basare le decisioni automatizzate sui "dati sensibili" ex art. 9 del GDPR salvo *«che non sia d'applicazione l'articolo 9, paragrafo 2, lettere a) o g), e non siano in vigore misure adeguate a tutela dei diritti, delle libertà e dei legittimi interessi dell'interessato»*.

Gli articoli 13 e 14 del regolamento europeo disciplinano le informazioni che il titolare del trattamento deve fornire all'interessato tra le quali rientrano l'obbligo di comunicare *«l'esistenza di un processo decisionale automatizzato, compresa la profilazione»* e quello di fornire *«informazioni significative sulla logica utilizzata, nonché l'importanza e le conseguenze previste di tale trattamento per l'interessato»*. Infine, l'articolo 15, disciplinante il diritto di accesso, riproducendo il testo delle disposizioni appena citate, riconosce all'interessato il diritto di accedere al processo decisionale automatizzato al fine di conoscere la logica sottostante alle decisioni che lo riguardano nonché l'importanza e le conseguenze delle stesse su di esso.

La ratio di tali norme si rinviene nella necessità di fornire tutele ai soggetti rispetto all'assunzione di decisioni automatizzate basate sul trattamento dei loro dati personali, tutela che si esplica da un lato, limitando i casi in cui il titolare del trattamento possa legittimamente adottare un processo decisionale

automatizzato nei loro confronti e, dall'altro, garantendo un adeguato livello di trasparenza del trattamento stesso¹⁰⁹.

Il GDPR, nel disciplinare i casi in cui si può ricorrere ad un processo decisionale automatizzato, omette tuttavia di fornirne una definizione esplicita limitandosi a definire il prodotto del processo decisionale automatizzato¹¹⁰ come «una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che [...] riguardano [l'interessato] o che incida in modo analogo significativamente sulla sua persona¹¹¹».

A ben vedere, tale mancata definizione si presta ad una lettura restrittiva e letterale dell'art. 22 che facendo leva sull'avverbio “unicamente” può portare ad affermare che tale disposizione si applicasse esclusivamente alle decisioni completamente automatizzate ossia prive di qualsivoglia intervento umano¹¹². Tale interpretazione restrittiva comporterebbe tuttavia il rischio di escludere fattispecie più complesse o articolate rispetto a quanto astrattamente previsto

¹⁰⁹ In tal senso si esprime il considerando 71 del GDPR il quale afferma come «L'interessato dovrebbe avere il diritto di non essere sottoposto a una decisione, che possa includere una misura, che valuti aspetti personali che lo riguardano, che sia basata unicamente su un trattamento automatizzato [...] Tuttavia, è opportuno che sia consentito adottare decisioni sulla base di tale trattamento, compresa la profilazione, se ciò è espressamente previsto dal diritto dell'Unione o degli Stati membri cui è soggetto il titolare del trattamento, [...] o se è necessario per la conclusione o l'esecuzione di un contratto tra l'interessato e un titolare del trattamento, o se l'interessato ha espresso il proprio consenso esplicito. In ogni caso, tale trattamento dovrebbe essere subordinato a garanzie adeguate, che dovrebbero comprendere la specifica informazione all'interessato e il diritto di ottenere l'intervento umano, di esprimere la propria opinione, di ottenere una spiegazione della decisione conseguita dopo tale valutazione e di contestare la decisione. [...] Al fine di garantire un trattamento corretto e trasparente nel rispetto dell'interessato, tenendo in considerazione le circostanze e il contesto specifici in cui i dati personali sono trattati, è opportuno che il titolare del trattamento utilizzi procedure matematiche o statistiche appropriate per la profilazione, metta in atto misure tecniche e organizzative adeguate al fine di garantire, in particolare, che siano rettificati i fattori che comportano inesattezze dei dati e sia minimizzato il rischio di errori e al fine di garantire la sicurezza dei dati personali secondo una modalità che tenga conto dei potenziali rischi esistenti per gli interessi e i diritti dell'interessato e che impedisca tra l'altro effetti discriminatori nei confronti di persone fisiche sulla base della razza o dell'origine etnica, delle opinioni politiche, della religione o delle convinzioni personali, dell'appartenenza sindacale, dello status genetico, dello stato di salute o dell'orientamento sessuale, ovvero che comportano misure aventi tali effetti. Il processo decisionale automatizzato e la profilazione basati su categorie particolari di dati personali dovrebbero essere consentiti solo a determinate condizioni».

¹¹⁰ Valeria Caforio, Federica Paolucci, *Processi decisionali automatizzati, tutela dei dati personali e del segreto commerciale nell'Unione europea: spunti dal caso “SCHUFA” (causa C-634/21)*, in *Il Diritto industriale* 3/2025 pp. 265 ss.

¹¹¹ Reg. Ue n. 2016/679, Art. 22

¹¹² F. Palmiotto, *When Is a Decision Automated? A Taxonomy for a Fundamental Rights Analysis*, in *German Law Journal*, 2024 pp. 6 ss.

dalla norma; in particolare se una decisione è basata prevalentemente su un trattamento automatizzato, ma vi è un intervento umano parziale, questa ricade o meno nelle previsioni di cui all'art. 22? È chiaro che stabilire se una determinata decisione sia o meno automatizzata è questione dirimente ai fini dell'applicazione della tutela prevista dall'art. 22¹¹³.

3.1. Corte di Giustizia dell'Unione Europea in causa C-634/21 “Schufa”.

La questione interpretativa sulla definizione di decisione automatizzata ex art. 22 del GDPR è stata risolta dalla Corte di Giustizia dell'Unione Europea, nella causa C-634/21, proprio in materia *di credit scoring*: in Germania opera la *Schufa Holding AG* (di seguito Schufa), una società che gestisce un sistema automatizzato di valutazione dell'affidabilità creditizia e fornisce servizi di credit scoring alle imprese che le trasmettono le richieste di finanziamento dei loro clienti. Nel 2018, una cittadina tedesca si rivolgeva a un istituto di credito per ottenere la concessione di un prestito, ma vedeva respinta la sua richiesta dall'intermediario che basava la sua decisione sulle risultanze ricavate dal credit scoring elaborato consultando il sistema della Schufa.

La signora si rivolgeva alla Schufa chiedendo, da un lato, l'accesso alle informazioni utilizzate per fondare la decisione e, dall'altro, la cancellazione dei dati inesatti; la società le comunicava il punteggio di scoring, calcolato in relazione alla sua posizione e, le forniva, una descrizione generale del suo sistema di calcolo, senza tuttavia indicarle il peso attribuito alle singole variabili. La Schufa riteneva, infatti, di non essere tenuta a divulgare i propri metodi di calcolo, in quanto coperti, a suo dire, da segreto professionale, industriale e

¹¹³ Valeria Caforio, Federica Paolucci, *Processi decisionali automatizzati, tutela dei dati personali e del segreto commerciale nell'Unione europea: spunti dal caso “SCHUFA” (causa C-634/21)*, in *Il Diritto industriale* 3/2025 p. 266.

commerciale; inoltre, la società evidenziava il suo ruolo di mero fornitore delle informazioni rilevanti ai soggetti richiedenti il servizio, questi ultimi gli assuntori in concreto della decisione finale sulla “meritevolezza” del credito e della decisione sulla erogazione o meno del finanziamento. A seguito di tale diniego la richiedente presentava reclamo al garante per la protezione dei dati personali tedesco (il Bundesbeauftragte, in seguito BfDI) chiedendo che questi ordinasse alla società di fornire le informazioni richieste e di cancellare i dati inesatti. Il BfDI rigettava la richiesta ritenendo che non vi fosse motivo di agire ulteriormente nei confronti della società, poiché essa aveva rispettato i requisiti imposti dalla legge federale tedesca sulla protezione dei dati.

La decisione del garante veniva impugnata dinanzi al *Verwaltungsgericht Wiesbaden* (Tribunale amministrativo di Wiesbaden, Germania) il quale sollevava un rinvio pregiudiziale alla Corte di giustizia dell’Unione europea sottoponendole le seguenti questioni:

«1) Se l’articolo 22, paragrafo 1, del [GDPR] debba essere interpretato nel senso che il calcolo automatizzato di un tasso di probabilità relativo alla capacità di un interessato di saldare in futuro un debito costituisce già una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produce effetti giuridici che riguardano l’interessato o che incide in modo analogo significativamente sulla sua persona, qualora tale tasso, calcolato sulla base di dati personali relativi all’interessato, sia trasmesso dal titolare del trattamento a un terzo titolare del trattamento e quest’ultimo basi prevalentemente su tale tasso la sua decisione sulla stipulazione, sull’attuazione o sulla cessazione di un contratto con l’interessato.

2) In caso di risposta negativa alla prima questione pregiudiziale: se l’articolo 6, paragrafo 1, e l’articolo 22 del [GDPR] debbano essere interpretati nel senso che ostano a una normativa nazionale ai sensi della quale il ricorso a un tasso di probabilità – nella fattispecie relativo alla solvibilità e alla disponibilità a pagare di una persona fisica, che

includa informazioni sui crediti – di un certo comportamento futuro di una persona fisica, allo scopo di decidere sulla stipulazione, sull’attuazione o sulla cessazione di un contratto con tale persona (scoring), è consentito solo se sono soddisfatte determinate ulteriori condizioni, meglio specificate nella motivazione della domanda di pronuncia pregiudiziale¹¹⁴».

In altre parole, il giudice del rinvio chiedeva alla Corte di Giustizia di accertare se il credit scoring costituisse *“una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione”* di cui all’art. 22 del GDPR anche qualora il punteggio, come nel caso di specie, fosse stato calcolato da un soggetto titolare del trattamento (la Schufa) e successivamente trasmesso all’intermediario che poi lo ha utilizzato come presupposto essenziale per basare la sua decisione sulla stipulazione, sull’attuazione o sulla cessazione di un contratto con l’interessato.

In caso di risposta affermativa a questa domanda, la Schufa sarebbe tenuta a dare accesso ai dati supplementari cui l’interessato ha diritto ai sensi dell’art. 15, par. 1, lett. h), del GDPR, ossia le informazioni circa l’esistenza di un processo decisionale automatizzato, sulla logica utilizzata, nonché sull’importanza e le conseguenze del già menzionato trattamento per l’interessato. Di converso, tale obbligo non sorgerebbe laddove si dovesse ritenere che, in situazioni come quella in esame, soltanto il provvedimento (di diniego del credito) assunto dal terzo finanziatore possa qualificarsi, qualora ricorrano le condizioni di applicazione dell’art. 22 GDPR, come una decisione adottata in esito ad un trattamento interamente automatizzato relativo a persone fisiche¹¹⁵.

La Corte di Giustizia con sentenza in data 7 dicembre 2023 in causa C-634/21 ha accolto una interpretazione estensiva dell’articolo 22 del GDPR e si è pronunciata statuendo che: *«L’articolo 22, paragrafo 1, del regolamento (UE) 2016/679 del*

¹¹⁴ Causa C-634/21, Conclusioni dell’Avvocato Generale Priit Pikamäe, <https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:62021CC0634> p. 7

¹¹⁵ F. Ciraolo, *Le valutazioni automatizzate del merito creditizio nel quadro regolatorio europeo. Quale futuro per il credit scoring algoritmico?*, in *Rivista di Diritto Bancario*, Fascicolo I - Anno 2025 - Sezione Prima, pp 110 ss.

Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (regolamento generale sulla protezione dei dati), deve essere interpretato nel senso che: il calcolo automatizzato, da parte di una società che fornisce informazioni commerciali, di un tasso di probabilità basato su dati personali relativi a una persona e riguardanti la capacità di quest'ultima di onorare in futuro gli impegni di pagamento costituisce un "processo decisionale automatizzato relativo alle persone fisiche", ai sensi di tale disposizione, qualora da tale tasso di probabilità dipenda in modo decisivo la stipula, l'esecuzione o la cessazione di un rapporto contrattuale con tale persona da parte di un terzo, al quale è comunicato tale tasso di probabilità¹¹⁶».

I giudici del Lussemburgo hanno motivato la loro decisione perimetrando l'ambito di applicazione dell'art. 22 del GDPR alla sussistenza di tre condizioni cumulative¹¹⁷ che sono state ritenute sussistere nel caso di specie:

- i) **l'esistenza di una decisione:** la Corte ha scelto, richiamandosi alle conclusioni dell'Avvocato Generale, di perimetrare il concetto di decisione in un senso "ampio" tale da «ricomprendere una molteplicità di atti che possono incidere sull'interessato in diversi modi»¹¹⁸ così che deve ritenersi compreso «il risultato del calcolo della solvibilità di una persona sotto forma di tasso di probabilità relativo alla capacità di tale persona di onorare impegni di pagamento in futuro»¹¹⁹.
- ii) **Il fatto che tale decisione sia basata unicamente sul trattamento automatizzato, compresa la profilazione:** la Corte, anche qui aderendo

¹¹⁶ Corte di Giustizia dell'Unione Europea, Sentenza - 7/12/2023 - SCHUFA Holding (Scoring) Causa C-634/21 <https://curia.europa.eu/juris/documents.jsf?num=C-634/21>, Paragrafo 75

¹¹⁷ Corte di Giustizia dell'Unione Europea, Sentenza - 7/12/2023 - SCHUFA Holding (Scoring) Causa C-634/21 <https://curia.europa.eu/juris/documents.jsf?num=C-634/21>, Paragrafo 43

¹¹⁸ Conclusioni dell'Avvocato Generale Priit Pikamäe in Causa C-634/21, <https://eur-lex.europa.eu/legal-content/it/TXT/?uri=CELEX:62021CC0634>, Paragrafo 38

¹¹⁹ Corte di Giustizia dell'Unione Europea, Sentenza - 7/12/2023 - SCHUFA Holding (Scoring) Causa C-634/21 <https://curia.europa.eu/juris/documents.jsf?num=C-634/21> Paragrafo 46

alle conclusioni dell'Avvocato Generale, ha ritenuto che l'attività di *credit scoring* della Schufa rispondesse alla definizione di profilazione ex art. 4, punto 4 del GDPR poiché, come emerge dal fascicolo della causa la società utilizza una procedura che fornisce un risultato da cui possono trarsi conclusioni in merito alla solvibilità dell'interessato e che «*detta procedura utilizza dati personali per valutare determinati aspetti relativi alle persone fisiche per analizzare o prevedere aspetti riguardanti la situazione economica, l'affidabilità e il probabile comportamento di dette persone*».¹²⁰

iii) **Il fatto che tale decisione produca effetti giuridici riguardanti l'interessato o che incidano in modo analogo significativamente sulla sua persona:** la Corte ha ritenuto sussistere nel caso di specie anche questa condizione poiché, pur essendo vero che la decisione sulla concessione del finanziamento viene effettuata dall'intermediario, questa, come emerso dai fatti di causa, è guidata in modo decisivo dal credit scoring che gli viene trasmesso cosicché una valutazione di scarsa affidabilità da parte della Schufa comporterà in quasi tutti i casi il rifiuto da parte della banca di concedere il prestito richiesto.

Il credit scoring effettuato dalla Schufa non è dunque un mero atto preparatorio, ma un elemento decisivo che influisce direttamente sull'esito della richiesta di credito; il punteggio che viene assegnato è infatti parte integrante del processo decisionale automatizzato, e come tale deve essere soggetto alle regole di trasparenza e spiegazione previste dal GDPR, nonché a tutti i diritti degli interessati, tra cui il diritto d'accesso o, ancora, il diritto alla cancellazione¹²¹.

¹²⁰ Causa C-634/21, Conclusioni dell'Avvocato Generale Priit Pikamäe, <https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:62021CC0634>, Paragrafo 33

¹²¹ Valeria Caforio, Federica Paolucci, *Processi decisionali automatizzati, tutela dei dati personali e del segreto commerciale nell'Unione europea: spunti dal caso "SCHUFA" (causa C-634/21)*, in *Il Diritto industriale* 3/2025, p. 269

La decisione della Corte è pienamente coerente con il dettato dell'art. 22 e con le finalità del Regolamento sulla protezione dei dati personali: la tutela dell'interessato viene perseguita mediante la previsione di un divieto di principio, che può però essere superato al verificarsi di determinate eccezioni alle quali si correla l'applicazione di particolari garanzie (previste dall'art. 22, par. 3 e 4 nonché dagli artt. 13-15 GDPR). È nei requisiti maggiormente stringenti stabiliti per le decisioni automatizzate che si riflette la *ratio* della disposizione, tesa a proteggere i soggetti dai rischi che determinati trattamenti di dati possono comportare per i loro diritti e per le loro libertà. Di converso, una interpretazione restrittiva della norma, tale per cui il calcolo del tasso di probabilità dovrebbe essere considerato un mero atto preparatorio e solo l'atto adottato dal terzo (*rectius* la banca) potrebbe, se del caso, essere qualificato come «*decisione*», ai sensi dell'articolo 22, paragrafo 1, porterebbe al rischio di una elusione dell'articolo 22 del GDPR e, di conseguenza, ad una lacuna nella protezione giuridica¹²²: se da un lato la società di valutazione del credito, non fosse tenuta a fornire le informazioni all'interessato, in forza dell'articolo 15, paragrafo 1, lettera h), del GDPR, poiché essa non effettua in senso stretto un «*processo decisionale automatizzato*», dall'altro lato l'istituto finanziario che adottasse la sua decisione in maniera automatizzata sulla base del punteggio di scoring calcolato (e che pertanto dovrebbe fornire le informazioni richieste ai sensi dell'articolo 15, paragrafo 1, lettera h del GDPR) non potrebbe ottemperare a tale obbligo in quanto tali informazioni non sono in suo possesso ma della società¹²³. Conseguentemente, in caso di contestazione della decisione, l'istituto finanziario non sarebbe in grado di controllare la valutazione della solvibilità del soggetto

¹²² F. D'Orazio, *Il credit scoring e l'art. 22 del GDPR al vaglio della Corte di giustizia* in La nuova giurisprudenza civile commentata 2024, n° 2, p. 410-421.

¹²³ V. Pietrella, S. Racioppi, *Il credit scoring e la protezione dei dati personali: commento alle sentenze della Corte di giustizia dell'Unione europea del 7 dicembre 2023*, in Rivista Italiana di Informatica e Diritto, Fascicolo 1, 2024, p. 186

che richiede il credito, come impone l'articolo 22, paragrafo 3, del GDPR, né di garantire un trattamento corretto, trasparente e non discriminatorio mediante procedure matematiche o statistiche appropriate e misure tecniche e organizzative adeguate, conformemente a quanto affermato dal considerando 71 del GDPR¹²⁴.

Ma non solo! A ben vedere l'unico soggetto in grado di dar seguito ad ulteriori richieste dell'interessato, fondate su diritti garantiti dal GDPR (il diritto di rettifica, ex art. 16, e il diritto alla cancellazione ex art. 17) altri non è che la società che effettua il credit scoring dal momento che quando tali attività sono esternalizzate¹²⁵ la banca non partecipa né alla raccolta di detti dati, né alla profilazione sicché va escluso che essa possa garantire in maniera efficace il rispetto dei diritti in materia di dati personali. Per non far subire, all'interessato dal trattamento, le conseguenze sfavorevoli della esternalizzazione delle attività non si può quindi che *«considerare l'agenzia di valutazione del credito responsabile in ragione del calcolo del punteggio di scoring – e non in ragione del suo successivo utilizzo – è il modo più efficace per garantire il diritto alla protezione dei dati di carattere personale, di cui all'articolo 8 della Carta dei diritti fondamentali dell'Unione europea ma anche del diritto al rispetto della sua vita privata, di cui all'articolo 7 della Carta, poiché detta attività rappresenta, in definitiva, l'origine di ogni eventuale pregiudizio¹²⁶»*.

¹²⁴ (79)[...] Al fine di garantire un trattamento corretto e trasparente nel rispetto dell'interessato, tenendo in considerazione le circostanze e il contesto specifici in cui i dati personali sono trattati, è opportuno che il titolare del trattamento utilizzi procedure matematiche o statistiche appropriate per la profilazione, metta in atto misure tecniche e organizzative adeguate al fine di garantire, in particolare, che siano rettificati i fattori che comportano inesattezze dei dati e sia minimizzato il rischio di errori e al fine di garantire la sicurezza dei dati personali secondo una modalità che tenga conto dei potenziali rischi esistenti per gli interessi e i diritti dell'interessato e che impedisca tra l'altro effetti discriminatori nei confronti di persone fisiche sulla base della razza o dell'origine etnica, delle opinioni politiche, della religione o delle convinzioni personali, dell'appartenenza sindacale, dello status genetico, dello stato di salute o dell'orientamento sessuale, ovvero che comportano misure aventi tali effetti. Il processo decisionale automatizzato e la profilazione basati su categorie particolari di dati personali dovrebbero essere consentiti solo a determinate condizioni.

¹²⁵ E. Falletti, C. Gallese, *Credit scoring and transparency between the AI Act and the Court of Justice of the European Union*, AIMMES, 2024 p.3

¹²⁶ Causa C-634/21, Conclusioni dell'Avvocato Generale Priit Pikamäe, <https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:62021CC0634>, par 48 ss.

Con la sua decisione la Corte fa venire meno il presupposto dietro cui società, come la Schufa, che vendono a terze valutazioni circa l'affidabilità creditizia di potenziali utenti di servizi, si sono sempre rifugiate per non applicare l'art. 22 del GDPR: esse non sono dei meri prestatori di servizi, nell'ambito di una filiera più ampia che non le vede coinvolte nella decisione finale ma soggetti capaci di incidere in maniera dirimente sull'erogazione del credito. La reale lesione del diritto/interesse del richiedente il finanziamento si realizza infatti ben prima dell'inevitabile e successivo "formale" diniego del prestito da parte dell'intermediario finanziario, ossia nel momento in cui la società di rating aveva reso il suo giudizio¹²⁷.

3.2. I problemi rimasti aperti

La sentenza Schufa ha lasciato aperte alcune questioni cruciali riguardanti la reale portata dei diritti del soggetto interessato dal trattamento dei dati personali che possono venire riassunte in due domande:

- i) Quale tipo di spiegazione può pretendere l'interessato a titolo di informazioni significative?
- ii) Come risolvere il conflitto qualora le informazioni da fornire contengano dati personali di terzi o segreti commerciali¹²⁸?

La Corte non ha esplicitato se l'art. del 15 GDPR conferisca un vero e proprio diritto dell'interessato ad ottenere una spiegazione della decisione algoritmica né, in caso di risposta affermativa, quali siano le modalità idonee a darvi concreta attuazione, anche alla luce del fenomeno della *black box* che caratterizza gli algoritmi di machine learning.

¹²⁷ F. Mattassoglio, *La Corte di giustizia europea, algoritmi e credit scoring. L'apertura del vaso di Pandora delle società che si "limitano" a elaborare gli scoring*, in Dialoghi di diritto dell'Economia, gennaio 2025, <https://www.dirittobancario.it/wp-content/uploads/2025/01/2025-Mattassoglio-Credit-scoring.pdf>

¹²⁸ B. Parenzo, *"Informazioni significative" nei processi decisionali automatizzati: parla la Corte di Giustizia*, in *La nuova giurisprudenza civile commentata*, 4/2025, Wolters Kluwer, pp 972 ss.

Il diritto alla spiegazione ha suscitato, in passato, un acceso dibattito in dottrina tra le opposte tesi di chi (minoritariamente) riteneva che il GDPR riconoscesse all'interessato un mero diritto ad essere informato sulla logica generale del sistema automatizzato, ma non ad ottenere una spiegazione per la decisione specifica, e quanti sostenevano che andava garantita sia la comprensione del meccanismo decisionale automatizzato sia la spiegazione della specifica decisione¹²⁹.

I fautori della tesi minoritaria basano le loro argomentazioni sulla natura, vincolante e non, tra gli articoli del regolamento e i suoi considerando: se la tesi di un "diritto alla spiegazione" deriva principalmente dalla lettura del considerando 71 del GDPR, il quale afferma che le garanzie adeguate per l'interessato dovrebbero includere il diritto di «*ottenere una spiegazione della decisione conseguita*» va osservato come, da un punto di vista normativo, i considerando non siano giuridicamente vincolanti ma possano essere utilizzati al solo scopo di interpretare il testo degli articoli senza tuttavia creare un nuovo obbligo giuridico. La norma cogente è l'articolo 22 GDPR che al comma 3 elenca le garanzie minime per l'interessato, che prevedono «*almeno il diritto di ottenere l'intervento umano [...] di esprimere la propria opinione e di contestare la decisione*»; ciò che il GDPR garantirebbe in modo vincolante è “un limitato diritto ad essere informati” derivante dagli obblighi di notifica cui agli artt. 13 e 14 e dal diritto di accesso di cui all'art. 15. Tali disposizioni impongono al titolare del trattamento di fornire all'interessato informazioni sull'esistenza di un trattamento automatizzato nonché sulla logica utilizzata, l'importanza e le conseguenze previste da tale trattamento; sarebbe tuttavia erroneo equiparare tali obblighi ad una vera e propria spiegazione della decisione specifica¹³⁰ poiché da un lato si

¹²⁹ F. Ciraolo, *Le valutazioni automatizzate del merito creditizio nel quadro regolatorio europeo. Quale futuro per il credit scoring algoritmico?*, in *Rivista di Diritto Bancario*, Fascicolo I - Anno 2025 - Sezione Prima p. 118

¹³⁰ G. Finocchiaro, *Intelligenza Artificiale e protezione dei dati personali*, in *Giurisprudenza Italiana*, Luglio 2019, p. 1675, afferma come l'art. 15 non rifletta un diritto alla spiegazione bensì un diritto ad “avere

tratterebbe di una spiegazione preventiva e generale (posto che gli artt. 13 e 14 si riferiscono a notifiche fornite al momento della raccolta dei dati, prima che una decisione venga presa) e dall'altro lato l'art. 15 utilizza un linguaggio orientato al futuro che suggerirebbe come l'obbligo riguardi la logica generale, gli scopi e gli impatti previsti del sistema, piuttosto che la motivazione *ex post* di un singolo risultato¹³¹.

A tale tesi si oppongono le argomentazioni di coloro¹³² i quali sostengono che le finalità protettive dell'art. 15 del GDPR possono realizzarsi unicamente nel caso in cui l'interessato venga posto nella condizione di comprendere, mediante la comunicazione di informazioni espresse in un linguaggio chiaro, preciso e facilmente accessibile, l'intero processo decisionale automatizzato e le motivazioni sottostanti alla singola decisione adottata. Argomentando diversamente l'interessato vedrebbe precluso l'effettivo esercizio dei diritti riconosciutigli dall'art. 22 GDPR e, segnatamente, il diritto di esprimere la propria opinione e di contestare la decisione, attività questa che necessita la piena comprensione delle ragioni alla stessa sottese. In linea con le finalità degli articoli 13, 14 e 15 del GDPR, deve essere garantita sia la comprensione *ex ante* del meccanismo decisionale automatizzato (comprese le sue finalità, le caratteristiche tecnico-operative, i criteri utilizzati per le decisioni, la loro ponderazione e i tipi di dati analizzabili) sia la spiegazione *ex post* della decisione algoritmica (specificando i dati, i criteri e le ragioni alla base di un provvedimento

informazioni"; tale diritto concerne l'accedere ad informazioni sulla esistenza di un processo decisionale automatizzato, compresa la profilazione di cui all'art. 22, par. 1 e 4, e ad ottenere informazioni significative sulla logica utilizzata, nonché l'importanza e le conseguenze previste di tale trattamento per l'interessato.

¹³¹ S. Wachter, B. Mittelstadt, L. Floridi, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation* in *International Data Privacy Law*, Volume 7, Issue 2, Maggio 2017, pp. 76–99, <https://doi.org/10.1093/idpl/ix005>

¹³² G. Malgieri, G. Comandé, *Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation* in *International Data Privacy Law*, vol. 7, Issue 3, (2017) rinviengono dall'interpretazione del combinato disposto degli articoli del GDPR un "diritto alla leggibilità dell'algoritmo che può essere raggiunto attraverso la divulgazione di informazioni significative sulla logica coinvolta, sul significato e sulle conseguenze previste delle decisioni automatizzate.

specifico) sì da consentirne la verifica della correttezza e della legittimità nonché l'esercizio del diritto di contestazione, ove applicabile¹³³.

Con riguardo invece alla seconda questione lasciata in sospeso, essa concerne l'individuazione delle modalità atte a risolvere il conflitto fra l'esercizio dei diritti del GDPR e la tutela degli interessi dei terzi affinché siano bilanciati il diritto di accesso alle informazioni per l'interessato (ex art. 15 del GDPR) con il diritto alla tutela dei dati personali di terzi (tutelati in quanto tali dallo stesso GDPR) e dei segreti commerciali (tutelati dalla direttiva n. 2016/943 UE c.d. direttiva *know-how*) alla luce del diritto di privacy sul codice sorgente¹³⁴, il quale, è alla base del sistema decisionale automatizzato e precluderebbe, in linea di principio, la rivelazione della relativa formula matematica alla base dell'algoritmo¹³⁵.

3.3. Corte di Giustizia del 27 febbraio 2025 in causa C-203/22 (D&B)

Alle questioni sopra rappresentate prova a dare risposta la recentissima pronuncia della Corte di Giustizia del 27 febbraio 2025 in causa C-203/22 (D&B) che ha fornito importanti chiarimenti sulla portata del diritto di accesso alle informazioni, spettanti ai sensi dell'art. 15 del GDPR, per un soggetto sottoposto a credit scoring automatizzato, e sul rapporto tra tali informazioni e il diritto alla tutela dei segreti commerciali.

La controversia trae origine dal diniego opposto alla signora CK, da parte di un operatore di telefonia mobile austriaco, alla stipula di un contratto telefonico. Tale rifiuto, concernente un'obbligazione contrattuale di modesta entità (pari a

¹³³ F. Ciraolo, *Le valutazioni automatizzate del merito creditizio nel quadro regolatorio europeo. Quale futuro per il credit scoring algoritmico?*, in *Rivista di Diritto Bancario*, Fascicolo I - Anno 2025 - Sezione Prima p. 119

¹³⁴ Il codice sorgente è l'insieme di istruzioni scritte in un linguaggio di programmazione leggibile dall'uomo.

¹³⁵ B. Parenzo, *"Informazioni significative" nei processi decisionali automatizzati: parla la Corte di Giustizia*, in *La nuova giurisprudenza civile commentata*, 4/2025, Wolters Kluwer, pp 976 ss.

dieci euro mensili) veniva motivato sulla scorta di una valutazione automatizzata dello scoring creditizio dell'interessata, effettuata dalla società D&B, che ne attestava una insufficiente solvibilità finanziaria.

La CK adiva la *Datenschutzbehörde* (l'Autorità Austriaca per la Protezione dei Dati Personali) che ingiungeva alla D&B di fornire alla ricorrente le "informazioni significative sulla logica utilizzata", nel processo decisionale automatizzato basato sui dati personali relativi a CK, ai sensi dell'articolo 15, paragrafo 1, lettera h), del GDPR.

La D&B proponeva gravame avverso tale decisione dinanzi al *Bundesverwaltungsgericht* (Corte amministrativa federale austriaca) eccependo la sussistenza di un segreto commerciale idoneo a paralizzare l'istanza di accesso dell'interessata nonché asserendo di avere già fornito tutte le informazioni dovute.

Con decisione¹³⁶ passata in giudicato del 23 ottobre 2019, il *Bundesverwaltungsgericht* statuiva che la D&B fosse inadempiente rispetto all'obbligo di cui all'art. 15, paragrafo 1, lettera h) del GDPR. Rilevava, in particolare, il giudice austriaco l'insufficienza delle spiegazioni fornite alla CK inidonee a permettere all'interessata di comprendere come fosse stato calcolato il suo credit score.

Tuttavia, la successiva fase di esecuzione forzata della decisione, incardinata presso l'autorità esecutiva competente (l'amministrazione comunale della città di Vienna) veniva respinta, con la motivazione che la D&B avesse sufficientemente adempiuto ai propri obblighi informativi; a questo punto l'interessata si rivolgeva al *Verwaltungsgericht Wien* (Tribunale amministrativo di Vienna).

¹³⁶ Bundesverwaltungsgericht, W256 2217011-1 del 23/10/2019, reperibile al seguente link: <https://www.ris.bka.gv.at/Ergebnis.wxe?Abfrage=Bvbwg&Entscheidungsart=Undefined&SucheNachRechtsatz=True&SucheNachText=True&GZ=W256+2217011-1&VonDatum=23.10.2019&BisDatum=21.11.2025&Norm=&ImRisSeitVonDatum=&ImRisSeitBisDatum=&ImRisSeit=Undefined&ResultPageSize=100&Suchworte=&Position=1&SkipToDocumentPage=true>

Il *Verwaltungsgericht Wien* sospendeva il giudizio¹³⁷ e sottoponeva alla Corte di Giustizia una serie di questioni pregiudiziali, attinenti alla corretta interpretazione dell'art. 15 GDPR chiedendo di chiarire quale sia la portata delle informazioni che devono essere fornite all'interessato affinché il suo diritto di accesso previsto dal Regolamento Generale sulla protezione dei Dati Personali possa essere pienamente garantito nonché di precisare se e in quale misura l'eccezione relativa all'esistenza di un segreto commerciale possa limitare l'esercizio di tale diritto. Per tale motivo il giudice del rinvio chiedeva se l'interessato possa «pretendere dal titolare del trattamento, “a titolo di informazioni significative sulla logica utilizzata”, una spiegazione esaustiva della procedura e dei principi concretamente applicati per usare, con mezzi automatizzati, i dati personali relativi a tale persona al fine di ottenerne un determinato risultato, quale un profilo di solvibilità¹³⁸».

Adottando un approccio prevalentemente funzionalistico, la Corte ha precisato¹³⁹ che il diritto di accesso, sancito dall'articolo 15 GDPR, debba essere interpretato, nella sua applicazione concreta, in linea con l'obiettivo primario perseguito dalla norma, ossia garantire all'interessato la possibilità di esercitare i diritti a lui riconosciuti dalla normativa vigente, quali il diritto alla rettifica, alla cancellazione e alla limitazione del trattamento (artt. 16, 17 e 18), il diritto all'opposizione al trattamento stesso (articolo 21), nonché, nel contesto di un processo decisionale automatizzato, il diritto di esprimere il proprio parere e di contestare la decisione (art. 22)¹⁴⁰. Per consentire all'interessato di esercitare in modo efficace tali diritti la spiegazione della decisione algoritmica deve essere

¹³⁷ Verwaltungsgericht Wien, VGW-101/042/791/2020-44 dell'11/02/2022 https://verwaltungsgericht.wien.gv.at/Content.Node/rechtsprechung/101-042-791-2020_EuGH-Antrag.pdf

¹³⁸ Corte di Giustizia dell'Unione Europea, Sentenza - 27/02/2025 - Dun & Bradstreet Austria Causa C-203/22, par. 38 <https://curia.europa.eu/juris/liste.jsf?num=C-203/22>

¹³⁹ Corte di Giustizia dell'Unione Europea, Sentenza - 27/02/2025 - Dun & Bradstreet Austria Causa C-203/22, par. 54-55 <https://curia.europa.eu/juris/liste.jsf?num=C-203/22>

¹⁴⁰ F. Ciraolo, *Le valutazioni automatizzate del merito creditizio nel quadro regolatorio europeo. Quale futuro per il credit scoring algoritmico?*, in *Rivista di Diritto Bancario*, Fascicolo I - Anno 2025 - Sezione Prima p. 119

fornita mediante informazioni pertinenti, redatte in forma concisa, trasparente, intelligibile e facilmente accessibile; al fine di soddisfare tali requisiti non può ritenersi sufficiente né la semplice comunicazione di una formula matematica complessa, come un algoritmo, né la descrizione dettagliata di tutte le fasi di un processo decisionale automatizzato, in quanto nessuno di questi metodi sarebbe idoneo a costituire una spiegazione sufficientemente concisa e comprensibile (tanto più se fornita in linguaggio informatico); il titolare del trattamento deve descrivere la procedura e i principi concretamente applicati nella valutazione del merito creditizio del cliente così da consentire all'interessato di comprendere quali suoi dati personali siano stati utilizzati, e in che modo, nel processo decisionale automatizzato: ad esempio, informando l'interessato se, e come, una variazione dei dati personali presi in considerazione avrebbe condotto a un risultato diverso¹⁴¹.

Concludeva pertanto la Corte di Giustizia affermando che *«l'interessato può pretendere dal titolare del trattamento, a titolo di "informazioni significative sulla logica utilizzata", che quest'ultimo gli spieghi, mediante informazioni pertinenti e in forma concisa, trasparente, comprensibile e facilmente accessibile, la procedura e i principi concretamente applicati per utilizzare, con mezzi automatizzati, i dati personali relativi a tale interessato al fine di ottenerne un risultato determinato, come un profilo di solvibilità¹⁴²»*.

La prima parte della sentenza affronta e risolve la questione della conoscibilità, senza soffermarsi sul bilanciamento tra il diritto d'accesso dell'interessato e il diritto d'autore sul software esaminando la questione in termini di effettività rispetto alle esigenze di trasparenza e tutela dei diritti individuali sancite dal GDPR: un dato non deve essere comunicato solo qualora esso sia ininfluenza, per

¹⁴¹ M. Lembo, L'intelligenza artificiale e i contratti bancari: il caso del credit scoring, in *I contratti: rivista di dottrina e giurisprudenza*, n.4/2025, Wolters Kluwer p. 406

¹⁴² Corte di Giustizia dell'Unione Europea, Sentenza - 27/02/2025 - Dun & Bradstreet Austria Causa C-203/22, par. 66 <https://curia.europa.eu/juris/liste.jsf?num=C-203/22>

complessità e inintelligibilità, ai fini della spiegazione della decisione algoritmica. Qualora, di converso, un dato riguardante informazioni di terzi protette dal GDPR o dal segreto commerciale andasse osteso per poter efficacemente ottemperare a quanto previsto dall'art. 15 GDPR la Corte muovendo dal presupposto che il diritto alla protezione dei dati personali non costituisca una prerogativa assoluta, ma vada temperato con altri diritti, in ossequio al principio di proporzionalità¹⁴³, ha affermato che dovrà essere il giudice competente o l'autorità di controllo a valutare quali dati personali delle parti o di terzi debbano essergli comunicati, per consentirgli di ponderare, con cognizione di causa e nel rispetto del citato principio di proporzionalità, gli interessi in gioco determinando la portata del diritto di accesso dell'interessato previsto dall'art. 15 del GDPR. Sarà dunque l'autorità adita a stabilire, secondo le circostanze del caso, se vada autorizzata la divulgazione (integrale o parziale) di dati riservati, ivi incluso il codice sorgente dell'algoritmo, senza che ciò comporti un eccessivo e immotivato sacrificio dei diritti e delle libertà altrui (nella specie, la proprietà intellettuale e il segreto commerciale)¹⁴⁴.

4. Credit Scoring e Trasparenza Bancaria: l'impatto della CCD2

Le risultanze così enunciate vanno lette in combinato disposto con la disciplina settoriale prevista per gli intermediari finanziari.

La valutazione del merito creditizio, in apparenza paradossalmente disciplinato nell'art. 35-duo-decies T.U.F. con riferimento alle scelte di investimento da parte

¹⁴³ La necessità di temperare il diritto alla protezione dei dati personali con altri diritti in ossequio al principio di proporzionalità viene affermato nel considerando 4 del GDPR ed è stato ribadito in varie sentenze della Corte di Giustizia dell'Unione Europea, tra cui si possono citare le pronunce Google/CNIL (C507/17) par. 60; Facebook/Schrems (C311/18) par. 172; RW/Österreichische (C154/21) par 47.

¹⁴⁴ F. Ciraolo, *Le valutazioni automatizzate del merito creditizio nel quadro regolatorio europeo. Quale futuro per il credit scoring algoritmico?*, in *Rivista di Diritto Bancario*, Fascicolo I - Anno 2025 - Sezione Prima P. 121

dei gestori di OICR¹⁴⁵, per lungo tempo, è stato un tema poco attenzionato sia dal Legislatore che dal Regolatore (Banca d'Italia). Se i rule maker, nell'enucleare il principio di *"sana e prudente gestione"* sancito dall'art. 5 T.U.B si sono storicamente focalizzati più sul fine (la sana e prudente gestione) che sul mezzo (i sistemi di credit scoring) l'impatto disruptive di big data, intelligenza artificiale e, più in generale, lo sviluppo generalizzato e progressivo del fintech hanno però reso sempre più attuale l'esigenza di un impianto normativo ad hoc¹⁴⁶.

Riservandoci di esaminare la normativa prudenziale nel prossimo capitolo, appare opportuno soffermarsi sulla normativa prevista dalle norme che disciplinano la c.d. trasparenza bancaria.

La direttiva sul credito ai consumatori (Direttiva 2008/48/CE) e la direttiva sul credito immobiliare ai consumatori (Direttiva 2014/17/CE) hanno posto un obbligo di valutare approfonditamente il merito creditizio sulla base di un "set informativo" adeguato ed il T.U.B. ha recepito tale vincolo agli artt. 120-undecies (per il credito immobiliare ai consumatori) e 124-bis (per il credito al consumo):

- i. in materia di credito immobiliare ai consumatori, l'art. 120-undecies del TUB, relativo alla verifica del merito creditizio, afferma, al comma 1, che *«[...] La valutazione del merito creditizio è effettuata sulla base delle informazioni sulla situazione economica e finanziaria del consumatore necessarie, sufficienti e proporzionate e opportunamente verificate»* cosicché vengono posti dunque dei requisiti alle caratteristiche dei dati utilizzati dall'intermediario ai fini del credit scoring. Il comma 5 aggiunge che *«quando la domanda di credito è*

¹⁴⁵ Art. 35- duodecies: 1. I gestori, per la valutazione del merito di credito dei beni in cui investono gli Oicr, adottano sistemi e procedure che non prevedono l'affidamento esclusivo o meccanico alle valutazioni emesse da agenzie di rating del credito.

2. Tenendo conto della natura, della portata e della complessità delle attività degli Oicr, la Banca d'Italia e la Consob, nell'ambito delle rispettive competenze, verificano l'adeguatezza dei sistemi e delle procedure adottati dai gestori ai sensi del comma 1 e valutano che l'utilizzo, nell'ambito delle politiche di investimento degli Oicr, dei riferimenti ai rating del credito emessi da agenzie di rating del credito, sia effettuato in modo da ridurre l'affidamento esclusivo o meccanico agli stessi.

¹⁴⁶ M. Lembo, L'intelligenza artificiale e i contratti bancari: il caso del credit scoring, in *I contratti: rivista di dottrina e giurisprudenza*, n.4/2025 p. 401 ss., Wolters Kluwer

*respinta, il finanziatore informa il consumatore senza indugio del rifiuto e, se del caso, del fatto che la decisione è basata sul trattamento automatico di dati», evidenziando un diritto del consumatore ad essere informato qualora il rifiuto della sua domanda di credito immobiliare sia stato basato sull'utilizzo di forme di *algorithmic credit scoring*¹⁴⁷.*

- ii. Con riguardo al credito ai consumatori l'art. 124-bis¹⁴⁸ del TUB prevedeva in capo ai finanziatori un semplice onere di valutare il merito creditizio del consumatore acquisendo "*informazioni adeguate*" e ricorrendo ove necessario anche a "*banche dati pertinenti*".

Se la disciplina di cui all'art. 120-undecies, in ragione del fatto che la sua redazione è stata influenzata dalla crisi dei mutui *subprime* è molto più dettagliata, le previsioni normative di cui all'art. 124-bis TUB sono per certi versi scarse e si limitavano a sancire l'obbligatorietà della valutazione del merito di credito, senza fornire indicazioni ulteriori sulle modalità con cui essa dev'essere svolta e sulle informazioni su cui deve basarsi¹⁴⁹. Nel complesso la CCD1 ha mostrato dei limiti che sono da ricondurre sia a fattori endogeni alla stessa direttiva che esogeni: con riguardo ai limiti interni va osservato come la direttiva abbia subito, proprio a causa della formulazione ambigua e scarna di alcuni articoli, un'interpretazione e un'applicazione disomogenea a livello degli Stati membri con conseguenti distorsioni della concorrenza tra gli intermediari dei singoli Stati e la limitazione

¹⁴⁷ F. M. Gabbricci, *La vigilanza sull'intelligenza artificiale in ambito bancario, in L'applicazione dell'AI Act in Italia e la tutela del consumatore. Il ruolo delle autorità indipendenti*, a cura di M. Rabitti, F. Bassan <https://romatrepress.uniroma3.it/wp-content/uploads/2025/07/DECA-n.-10-2025-ebook.pdf> p. 107 e ss.

¹⁴⁸ Art. 124 bis. (Verifica del merito creditizio) 1. *Prima della conclusione del contratto di credito, il finanziatore valuta il merito creditizio del consumatore sulla base di informazioni adeguate, se del caso fornite dal consumatore stesso e, ove necessario, ottenute consultando una banca dati pertinente.*

2. *Se le parti convengono di modificare l'importo totale del credito dopo la conclusione del contratto di credito, il finanziatore aggiorna le informazioni finanziarie di cui dispone riguardo al consumatore e valuta il merito creditizio del medesimo prima di procedere ad un aumento significativo dell'importo totale del credito.*

3. *La Banca d'Italia, in conformità alle deliberazioni del CICR, detta disposizioni attuative del presente articolo.*

¹⁴⁹ C. A. Valenza, *Obbligo di Valutazione del Merito Creditizio del Consumatore e Regole in Materia di Banche Dati*, in *Rivista Trimestrale di Diritto dell'Economia*, Supplemento al n. 4 /2024, p. 248

delle opportunità per i consumatori di usufruire della crescente offerta di credito transfrontaliero. A ciò si è aggiunto il fattore esterno dato dall'impatto dirompente della digitalizzazione nel mercato del credito al consumo, che ha determinato significative trasformazioni sia sul fronte dell'offerta che su quello della domanda e ha portato alla creazione di nuovi prodotti e all'evoluzione del comportamento e delle preferenze dei consumatori. Tali circostanze hanno pertanto fatto sorgere la necessità di dare una nuova armonizzazione al quadro giuridico di riferimento e aggiornarlo ai nuovi prodotti o servizi di credito, al fine di garantire la massima tutela del cliente¹⁵⁰. Ciò è avvenuto ad opera dell'emanazione della Direttiva (UE) 2023/2225 c.d. *Consumer Credit Directive II* (in seguito anche CCD2): finalizzata ad ampliare le tutele per i consumatori, introducendo offerte più trasparenti e obblighi più rigorosi di valutazione del merito creditizio, con una maggiore responsabilità per i finanziatori in caso di concessione imprudente del credito¹⁵¹, la CCD2 introduce norme di armonizzazione e abroga la direttiva 2008/48/CE (in seguito CCD1) che si è dimostrata solo parzialmente efficace nel garantire un elevato livello di tutela dei consumatori e nel favorire lo sviluppo di un mercato unico del credito.

La nuova direttiva dedica, alla valutazione del merito creditizio, l'art. 18 nel quale viene esplicitato l'obbligo in capo ai creditori di condurre una valutazione e una verifica meticolosa del merito creditizio di un consumatore, esaminando informazioni che non devono essere semplicemente "adeguate" (come richiedeva genericamente la CCD1) ma altresì "pertinenti e accurate" in relazione al reddito, alle spese del consumatore e ad altre circostanze finanziarie ed

¹⁵⁰ Ufficio Legislativo Economia, *Analisi Tecnico Normativa allo schema di decreto legislativo recante il recepimento della direttiva (UE) 2023/2225 del Parlamento europeo e del Consiglio, del 18 ottobre 2023, relativa ai contratti di credito ai consumatori e che abroga la direttiva 2008/48/CE*, https://documenti.camera.it/apps/nuovoSito/attiGoverno/schedaLavori/getTesto.ashx?leg=XIX&file=0321_F001.pdf p.1

¹⁵¹ D. Siclari, R. Lo Conte, *Informazioni da Inserire nel Contratto di Credito*, in *Rivista Trimestrale di Diritto dell'Economia*, Supplemento al n. 4 /2024, pp. 319-320

economiche. Tali informazioni, conformemente a quanto previsto dal principio di minimizzazione, sancito dall'art. 5 del GDPR, dovranno essere necessarie e proporzionate rispetto alla natura, alla durata, al valore e ai rischi del credito per il consumatore¹⁵².

I dati dovranno essere acquisiti da fonti pertinenti, sia interne che esterne (viene precisato che tra di esse non possono esservi ricompresi i social network), inclusi i consumatori stessi, e, ove necessario, mediante consultazione di banche dati.

Per quanto riguarda la valutazione del merito creditizio tramite strumenti automatizzati (come il credit scoring algoritmico) il comma 8 dell'art. 18 ne legittima l'utilizzo a condizione che i consumatori abbiano il diritto di chiedere e ottenere l'intervento umano del creditore che consiste nel diritto di chiedere ed avere, da parte del creditore, di una spiegazione chiara e comprensibile di come sia stata compiuta la valutazione del merito creditizio nella quale vengano indicati anche «la logica e i rischi derivanti dal trattamento automatizzato dei dati personali nonché la rilevanza e gli effetti sulla decisione» nonché nel diritto di esprimere il proprio punto di vista e di contestare la valutazione stessa chiedendone un riesame¹⁵³. Tale previsione è stata pensata proprio per porre rimedio alla scarsa trasparenza decisionale che si riscontra in tutti quei casi in cui la verifica del merito creditizio del consumatore sia svolta impiegando tecniche di *credit scoring algoritmico*; tuttavia tale soluzione è limitata ai soli casi in cui vi sia un trattamento automatizzato di dati personali e dietro espressa richiesta dello stesso non costituendo un obbligo generale in capo al finanziatore di informare il consumatore delle risultanze della verifica e di spiegargliele¹⁵⁴.

¹⁵² C. A. Valenza, *Obbligo di Valutazione del Merito Creditizio del Consumatore e Regole in Materia di Banche Dati*, in *Rivista Trimestrale di Diritto dell'Economia*, Supplemento al n. 4 /2024, p.251

¹⁵³ Servizio Studi della Camera dei Deputati, *Dossier relativo al recepimento della direttiva (UE) 2023/2225, relativa ai contratti di credito ai consumatori e che abroga la direttiva 2008/48/CE*, <https://temi.camera.it/leg19/dossier/OCD18-22272/recepimento-della-direttiva-ue-2023-2225-relativa-ai-contratti-credito-ai-consumatori-e-che-abroga-direttiva-2008-48-ce.html>

¹⁵⁴ M. De Poli, *Spiegazioni Adeguate*, in *Rivista Trimestrale di Diritto dell'Economia*, Supplemento al n. 4/2024, p. 212

L'articolo 19 disciplina le banche dati utilizzabili per l'acquisizione delle informazioni sulla solvibilità del richiedente prevedendo, da un lato, l'obbligo di garantirne la libera consultabilità agli operatori comunitari al fine di non penalizzare la concessione di credito transfrontaliero e, dall'altro lato, imponendo obblighi sul tipo di dati che queste possono contenere e sulle modalità di trattamento degli stessi: nel ribadire l'obbligo di rispettare integralmente il GDPR viene fatto divieto di trattare le "categorie particolari di dati personali" di cui all'art. 9 senza che siano ammessi i casi di esenzione del comma 2 del suddetto articolo. Le banche dati non potranno pertanto contenere dati che rivelino *«l'origine razziale o etnica, le opinioni politiche, le convinzioni religiose o filosofiche, o l'appartenenza sindacale, nonché trattare dati genetici, dati biometrici intesi a identificare in modo univoco una persona fisica, dati relativi alla salute o alla vita sessuale o all'orientamento sessuale della persona (art. 9)»* nonché i dati personali provenienti dai social network. La previsione dimostra una diffidenza, da parte del legislatore europeo, verso l'utilizzo dei *soft data* e la predilizione affinché, la valutazione del merito creditizio si basi su informazioni correlate alla situazione economico-finanziaria del soggetto¹⁵⁵.

Le informazioni raccolte nelle banche dati devono essere aggiornate e corrette ed al consumatore spetta il diritto di essere informato, entro 30 giorni, in merito registrazione dei dati relativi al ritardato rimborso, nonché di ottenere gli estremi identificativi della banca dati consultata e delle informazioni prese in considerazione, in caso di diniego del finanziamento.

L'articolo 4 della legge n. 91 del 13/06/2025 (c.d. Legge di delegazione europea 2024) ha attribuito la delega al Governo per il recepimento della CCD2 fissandone i principi ed i criteri direttivi per l'esercizio; tale delega è stata esercitata con il Decreto Legislativo n. 212 del 31/12/2025, che ha modificato il Testo Unico

¹⁵⁵ C. A. Valenza, *Obbligo di Valutazione del Merito Creditizio del Consumatore e Regole in Materia di Banche Dati*, in *Rivista Trimestrale di Diritto dell'Economia*, Supplemento al n. 4 /2024, p.285

Bancario¹⁵⁶, al fine di recepire gli articoli 18 e 19 della Direttiva, rivedendo gli articoli 124-bis e 125:

- i. il riformato articolo 124-bis TUB prevede che la valutazione del merito creditizio debba essere svolta dal finanziatore in maniera approfondita, nell'interesse del consumatore, tenendo conto di informazioni pertinenti alla verifica delle prospettive di adempimento degli obblighi imposti a questo dal contratto. Al fine di garantire un coordinamento tra la disciplina della valutazione del merito di credito prevista dalla CCD2 e la normativa prudenziale, ove applicabile ai finanziatori, la disposizione rimanda altresì al rispetto delle finalità di sana e prudente gestione¹⁵⁷. Sono stati altresì aggiunti i commi da 1-bis a 1-*quinquies* che riproducono quanto previsto nell'art. 18 della direttiva: la necessità che la valutazione del merito creditizio sia effettuata sulla base di informazioni necessarie e proporzionate rispetto alla natura, alla durata, al valore e ai rischi senza che possano essere usati i dati di cui all'articolo 9 del GDPR; l'obbligo di ottenere le informazioni sulla situazione economica e finanziaria del consumatore che siano necessarie e proporzionate rispetto alla natura del credito utilizzando fonti interne od esterne con esplicito divieto di usare i social network quali fonti esterne. Si impone altresì ai finanziatori di creare, documentare e rivedere periodicamente le proprie procedure per valutare il merito creditizio nonché di mantenere aggiornate le

¹⁵⁶ Il decreto rimanda altresì alla disciplina regolamentare della Banca d'Italia, (*"Disposizioni di trasparenza delle operazioni e dei servizi bancari e finanziari; correttezza delle relazioni tra intermediari e clienti"*) per la specificazione di elementi di dettaglio; essendo tuttavia la direttiva di massima armonizzazione e particolarmente dettagliata, le Disposizioni di Trasparenza della Banca d'Italia recepiranno gli aspetti strettamente tecnici, operativi e di coordinamento con le altre disposizioni

¹⁵⁷ Camera dei Deputati, *Relazione Illustrativa allo Schema di decreto legislativo recante recepimento della direttiva (UE) 2023/2225 del Parlamento europeo e del Consiglio, del 18 ottobre 2023, relativa ai contratti di credito ai consumatori e che abroga la direttiva 2008/48/CE* https://documenti.camera.it/apps/nuovoSito/attiGoverno/schedaLavori/getTesto.ashx?leg=XIX&file=0321_F001.pdf p. 10

informazioni utilizzate per lo scoring. I commi *1-quater* e *1-quinquies* stabiliscono che l'erogazione del credito è subordinata a una valutazione positiva del merito creditizio e che un calcolo errato di quest'ultimo non può essere usato come pretesto per apportare modifiche unilaterali svantaggiose per il consumatore o per risolvere il contratto di credito, salvo che siano state fornite informazioni false o vi siano state omissioni.

Particolare attenzione meritano poi i nuovi commi da *2-bis* a *2-sexies*, atti a regolare la valutazione del merito creditizio basata, anche parzialmente, sul trattamento automatizzato di dati personali del consumatore: le disposizioni riconoscono al consumatore il diritto di richiedere ed ottenere un intervento umano volto a fornire una spiegazione chiara e comprensibile dei criteri adottati nella valutazione del merito creditizio, ivi incluse le informazioni relative alle modalità e ai rischi connessi al trattamento automatizzato; si attribuisce inoltre al consumatore la facoltà di esprimere le proprie opinioni al finanziatore e di richiedere un intervento umano, al finanziatore, affinché venga riesaminata la sua valutazione (comma *2-bis*).

Qualora la valutazione si basi sul trattamento automatico dei dati da parte di un soggetto terzo, di cui si avvale il finanziatore, quest'ultimo sarà obbligato ad acquisire dal terzo, tutte le informazioni necessarie a fornire la spiegazione in merito al risultato della valutazione del merito creditizio (comma *2-ter*). Inoltre, il finanziatore è tenuto ad informare il consumatore dei diritti che gli spettano ai sensi del comma *2-bis* e, se del caso, del fatto che il trattamento automatizzato dei suoi dati verrà svolto da un terzo (comma *2-quater*); con un esplicito rimando esterno

il comma 2-*quinquies* fa esplicitamente salvi i diritti esercitabili ai sensi del GDPR sui dati personali. Il comma 2-*sexies* dispone in merito alle informazioni da rendere al consumatore in caso di diniego del finanziamento e sui conseguenti diritti del consumatore, ivi compreso il diritto ad essere informato sul fatto che tale diniego si sia basato su di un trattamento automatizzato dei dati e sulla modalità per chiedere un riesame della decisione¹⁵⁸.

- ii. L'articolo 125 del Tub, in materia di banche dati, viene emendato con l'aggiunta dei commi da 1-bis ad 1-*quinquies* e 4-bis con i quali si riproduce il divieto per i finanziatori e gli intermediari del credito di trattare ai fini della valutazione del merito creditizio le categorie particolari di dati di cui all'articolo 9, paragrafo 1, del GDPR e i dati personali ottenuti dai social network; nonché l'obbligo per i gestori delle banche dati di adottare procedure atte a verificare che le informazioni inserite nelle stesse siano corrette ed aggiornate.

Viene altresì inserita al Capo III del Titolo VI TUB una nuova disposizione, di portata generale, l'articolo 127-ter¹⁵⁹, che estende a tutti i contratti di credito le garanzie relative all'intervento umano, qualora la valutazione del merito creditizio si fondi, anche parzialmente, sul trattamento automatizzato dei dati personali del consumatore. Si prevede, a tal fine, che l'applicazione di quanto stabilito dal novellato articolo 124-bis, commi 2, 2-bis, 2-ter, 2-quater, 2-*quinquies*, 2-*sexies* e 3, anche ai fini dell'articolo 22, paragrafo 2, lettera b), del

¹⁵⁸ Camera dei Deputati, *Relazione Illustrativa allo Schema di decreto legislativo recante recepimento della direttiva (UE) 2023/2225 del Parlamento europeo e del Consiglio, del 18 ottobre 2023, relativa ai contratti di credito ai consumatori e che abroga la direttiva 2008/48/CE* https://documenti.camera.it/apps/nuovoSito/attiGoverno/schedaLavori/getTesto.ashx?leg=XIX&file=0321_F001.pdf pp. 10-11

¹⁵⁹ Art. 127-ter (Valutazione del merito creditizio fondata sul trattamento automatizzato di dati personali del cliente): 1. *Quando la valutazione del merito creditizio relativa a un contratto di finanziamento si fonda, anche solo in parte, sul trattamento automatizzato di dati personali del cliente, si applica quanto stabilito dall'articolo 124-bis, commi 2, 2-bis, 2-ter, 2-quater, 2-quinquies, 2-sexies e 3, anche ai fini dell'articolo 22, paragrafo 2, lettera b), del regolamento (UE) 2016/679, del Parlamento europeo e del Consiglio, del 27 aprile 2016.*

regolamento (UE) 2016/679, GDPR. Come affermato nella Relazione Illustrativa allo schema di decreto legislativo¹⁶⁰, l'estensione delle garanzie previste dalla CCD2 per il cliente in tutti i contratti di finanziamento nei quali le decisioni si fondino su trattamento automatizzato, è giustificata nell'ottica di applicare la normativa in maniera armonizzata con le previsioni del GDPR (art. 22, par. 2, lett. a) applicate conformemente alle statuizioni della giurisprudenza della Corte di Giustizia di cui si ha avuto modo di discutere nelle pagine precedenti. Dal punto di vista cronologico, infatti, l'emanazione della CCD2 è successiva alla Sentenza Schufa (C-634/21) motivo per cui il legislatore europeo ha optato per esplicitare compiutamente le risultanze della giurisprudenza nella normativa consumeristica bancaria¹⁶¹.

Il legislatore italiano dal canto suo ha scelto di ampliare ulteriormente l'ambito a tutti i contratti così da garantire una maggiore coerenza ed armonizzazione altresì con la disciplina prevista dal Regolamento Europeo sull'Intelligenza Artificiale che dedica una specifica attenzione alla disciplina in materia di credit scoring.

5. Il credit scoring come AI ad alto rischio: i requisiti previsti dall'AI Act

Il combinato disposto dell'articolo 22 del GDPR, interpretato alla luce delle sentenze della Corte di Giustizia, e degli articoli 18 e 19 della CCD2 può così

¹⁶⁰ Camera dei Deputati, *Relazione Illustrativa allo Schema di decreto legislativo recante recepimento della direttiva (UE) 2023/2225 del Parlamento europeo e del Consiglio, del 18 ottobre 2023, relativa ai contratti di credito ai consumatori e che abroga la direttiva 2008/48/CE* https://documenti.camera.it/apps/nuovoSito/attiGoverno/schedaLavori/getTesto.ashx?leg=XIX&file=0321_F001.pdf p. 15

¹⁶¹ In tal senso si è espresso anche il Dipartimento del Tesoro nel Documento per la consultazione pubblica sullo schema di decreto legislativo per il recepimento della Direttiva, consultabile al seguente link https://www.dt.mef.gov.it/export/sites/sitodt/modules/documenti_it/regolamentazione_bancaria_finanziaria/consultazioni_pubbliche/Documento-per-la-consultazione-pubblica-2107.pdf

declinarsi: a valle di un processo di valutazione della affidabilità creditizia adottato in maniera automatizzata, l'interessato ha diritto di ricevere una spiegazione che enunci le ragioni a fondamento della decisione presa nei suoi confronti e che sia sufficientemente trasparente e intellegibile da consentirgli di esercitare il suo diritto a contestarla; tale decisione nell'ambito degli intermediari finanziari, soggetti alle previsioni del TUB, qualora si basi sul trattamento di dati personali non potrà basarsi né su i dati "personalissimi" di cui all'art. 9 del GDPR né su dati presi dai social network del richiedente.

Per poter garantire all'interessato di ottenere una spiegazione comprensibile è tuttavia necessario che il funzionamento (*rectius* la "procedura") del sistema di IA impiegato nel processo decisionale sia spiegabile. Il tema giuridico della spiegazione è strettamente legato a quello tecnologico poiché un algoritmo di machine learning, programmato e addestrato dall'essere umano, con il passare del tempo si auto-evolve, imparando ed adattandosi, fino a divenire sempre più sofisticato e ad applicare una logica inferenziale che non riesce più ad essere integralmente compresa e conseguentemente spiegata ai terzi¹⁶².

Del problema tecnologico ha tenuto conto il legislatore eurounitario nel "Regolamento (UE) 2024/1689 del Parlamento Europeo e del Consiglio del 13 giugno 2024" c.d. *Regolamento sull'intelligenza artificiale*¹⁶³ (in seguito anche AI Act o il Regolamento) il cui apparato normativo si muove sul piano della regolazione della architettura delle tecnologie basate sull'intelligenza artificiale con l'esplicito intento di garantire una piena comprensibilità dei sistemi da parte dei soggetti che ne fanno uso¹⁶⁴.

¹⁶² Z. C. Lipton, *The Mythos of Model Interpretability*, in *Communications of the ACM*, Vol. 61, Issue 10, 2018

¹⁶³ L'AI Act rappresenta il culmine di un lungo processo iniziato nel 2020 con la presentazione da parte della Commissione Europea del "Libro bianco sull'intelligenza artificiale" nel quale venivano definiti gli elementi chiave su cui basare una regolamentazione in materia di intelligenza artificiale che riguardasse in particolare quei sistemi che implicino rischi significativi per la sicurezza ed i diritti fondamentali dei cittadini.

¹⁶⁴ B. Parenzo, "Informazioni significative" nei processi decisionali automatizzati: parla la Corte di Giustizia, in *La nuova giurisprudenza civile commentata*, 4/2025, Wolters Kluwer, pp 974 ss.

Dal punto di vista del rapporto con il GDPR, l'AI Act è complementare rispetto a questo in quanto non pregiudica la disciplina in materia di protezione dei dati, né i diritti e gli strumenti di ricorso da essa previsti: da un lato infatti si prevede che il regolamento lasci impregiudicati gli obblighi dei fornitori e dei deployer dei sistemi di IA nel loro ruolo di titolari del trattamento o responsabili del trattamento di dati personali impiegati nella progettazione, nello sviluppo o nell'uso di tali sistemi; dall'altro lato viene fatta salva l'applicazione dei diritti e delle garanzie conferite dal GDPR rispetto ai processi decisionali automatizzati basati sul trattamento dei dati personali ed affermata l'inidoneità dell'AI Act a costituire, di per sé, una base giuridica che autorizza il ricorso a tali decisioni automatizzate¹⁶⁵ (considerando 140).

Entrando più nel dettaglio dell'AI Act va osservato come esso si strutturi attorno agli approcci *human centric* e *risk based*: l'approccio antropocentrico vuole garantire che i sistemi di intelligenza artificiale¹⁶⁶ siano sviluppati e impiegati come strumenti al servizio delle persone per migliorarne il loro benessere, rispettando al contempo i diritti fondamentali protetti dalla Convenzione europea sui diritti dell'uomo (CEDU) e dalla Carta dei diritti fondamentali dell'Unione europea. La consapevolezza che tali tecnologie possano ledere i diritti fondamentali, tramite meccanismi decisionali opachi e discriminazioni ha giustificato l'emanazione di un insieme di regole vincolanti da applicare nel rispetto di un approccio basato sul rischio (conformemente a quanto prevede già il GDPR) guidato dalla consapevolezza che l'utilizzo dell'AI comporta pericoli

¹⁶⁵ D. Palazzo, *Il diritto alla spiegazione delle decisioni automatizzate. Trasparenza e tutela del segreto commerciale alla luce del GDPR e dell'AI Act*. Nota a Corte di Giustizia dell'Unione europea, sentenza 27 febbraio 2025, C-203/22, CK c. Magistrat der Stadt Wien, <https://www.rivistadirittoesocieta.it/il-diritto-alla-spiegazione-delle-decisioni-automatizzate-trasparenza-e-tutela-del-segreto-commerciale-alla-luce-del-gdpr-e-dellai-act-nota-a-corte-di-giustizia-dellunione-europea/> p.12

¹⁶⁶ L'articolo 3 par. 1, n. 1, del Regolamento, fornisce una nozione omnicomprensiva di sistema di i.a. definendolo come: «un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali»

differenti a seconda delle finalità e dei settori cui essa viene impiegata; ciò giustifica dunque una differenziazione delle misure di mitigazione a seconda della gravità del rischio¹⁶⁷.

In base alla minaccia che pongono alla sicurezza, alla salute, ai diritti fondamentali dei cittadini ed al mercato unico i sistemi sono divisi tra quattro categorie di rischio cui corrisponde un determinato regime giuridico:

- i. **IA a rischio inaccettabile:** si tratta di una serie di pratiche elencate all'articolo 5 del Regolamento quali la manipolazione cognitivo-comportamentale, la polizia predittiva, il riconoscimento delle emozioni sul luogo di lavoro e negli istituti di istruzione, il social scoring¹⁶⁸ che vengono espressamente vietate poiché ritenute, in ogni caso, non rispettose dei diritti fondamentali degli individui¹⁶⁹.
- ii. **IA ad alto rischio:** si tratta di applicazioni dell'intelligenza artificiale che possono comportare rischi elevati per la salute, la sicurezza o i diritti fondamentali delle persone. Prima di immettere una IA ad alto rischio nel mercato comunitario o di utilizzarla, i fornitori dovranno dimostrare che il loro sistema rispetti una serie di requisiti obbligatori, atti a mitigare i rischi e a garantirne l'affidabilità, in relazione a qualità dei dati, documentazione e tracciabilità, trasparenza, sorveglianza umana, accuratezza, cybersicurezza e robustezza¹⁷⁰.
- iii. **IA con specifici obblighi di trasparenza:** si tratta di applicazioni dell'intelligenza artificiale, che, se non fossero sufficientemente

¹⁶⁷ P. Deriu, S. Racioppi, *Riflessioni in tema di intelligenza artificiale e attività di vigilanza*, in Quaderni Fintech Consob, <https://www.consob.it/documents/d/area-pubblica/fintech15> pp. 22-23

¹⁶⁸ Meccanismo di valutazione dei cittadini, in base alle attività da questi compiute, al fine di assegnare loro un punteggio dal quale deriveranno o meno limitazioni dei loro diritti.

¹⁶⁹ La Commissione Europea ha emanato degli *Orientamenti sulle pratiche di intelligenza artificiale vietate ai sensi del regolamento (UE) 2024/1689 (C 2025/5052)* al fine di migliorare la chiarezza giuridica e fornire indicazioni sull'interpretazione dei divieti di cui all'articolo 5 del regolamento sull'IA così da garantirne l'applicazione coerente, efficace e uniforme.

¹⁷⁰ Dipartimento per il programma di Governo, *Analisi del quadro normativo in materia di Intelligenza artificiale*, <https://www.programmagoverno.gov.it/media/je0lo4i0/focus-ia.pdf> p.4

trasparenti, nei confronti degli utenti, potrebbero ingenerare, in essi, confusione; è il caso dei chatbot e dei deepfake rispetto ai quali l'utente, se non informato specificatamente, potrebbe non comprendere di avere a davanti a sé un risultato generato da una macchina. Al fine di tutelare l'autodeterminazione informativa dei soggetti e di prevenire pratiche ingannevoli i fornitori di sistemi di IA, che generano contenuti audio, video o di testo dovranno garantire che gli output siano contrassegnati e rilevabili come creati artificialmente¹⁷¹.

- iv. **IA con rischi minimi o assenti:** si tratta di applicazioni dell'intelligenza artificiale che presentano rischi limitati o addirittura assenti nei confronti dei diritti dei cittadini (ad es. gli algoritmi contenuti nei filtri antispam). Il legislatore europeo prevede che tali utilizzi siano leciti e non vadano assoggettati ad obblighi specifici. Il considerando 165 e l'art. 95 del Regolamento, tuttavia attribuisce il compito all'ufficio europeo per l'IA ed agli Stati membri di incoraggiare e agevolare l'elaborazione di codici di auto condotta per gli sviluppatori e gli utilizzatori che ricalchino le disposizioni previste in materia di IA ad alto rischio, scelta questa basata sulla convinzione che la conformità dei sistemi di IA diversi da quelli ad alto rischio alle regole più restrittive previste dal Regolamento possa aiutare a diffondere un'IA etica e affidabile e ad appianare il divario esistente tra la regolamentazione dei sistemi ad alto rischio e gli altri sistemi di AI¹⁷².

Al di fuori della quadripartizione piramidale basata sul rischio il Regolamento, agli articoli 51 e seguenti, prevede norme specifiche per i *“modelli I.A. per finalità generali”* (c.d. GPAL) ossia quei modelli che sono *«in grado di svolgere con*

¹⁷¹ Per un approfondimento sugli obblighi di cui all'art. 50 dell'AI Act si rinvia a J. Barata, *Transparency Obligations for all AI Systems: Article 50 of the AI Act*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5902402

¹⁷² A. Mantelero, G. Resta, G. M. Ricico, *Intelligenza Artificiale Commentario*, I Edizione, Wolters Kluwer, settembre 2025 p. 819

*competenza un'ampia gamma di compiti distinti*¹⁷³» quali, ad esempio, la generazione di testi, il riconoscimento delle immagini, la scrittura di codice informatico...

Rispetto a tali sistemi, il regolamento impone una serie di obblighi: alcuni applicabili a tutti i modelli di GPAI, altri solo a quelli che presentano rischi sistemici; ai sensi dell'articolo 53 tutti i fornitori di modelli GPAI devono rispettare specifici obblighi di trasparenza e documentazione tecnica e sono tenuti a adottare una *policy* che garantisca il rispetto della normativa in materia di diritto d'autore. Per quanto riguarda invece i GPA che presentano rischi sistemici¹⁷⁴ i fornitori soggiacciono a misure di gestione del rischio più invasive quali l'esecuzione di valutazioni del modello; la valutazione e l'attenuazione dei possibili rischi sistemici; la documentazione e comunicazione all'Ufficio per l'IA di qualsiasi "incidente grave" che si verifichi; l'adozione di adeguati sistemi di protezione informatica sia per il modello che per l'infrastruttura fisica di supporto¹⁷⁵.

Soffermando la nostra analisi sui sistemi di intelligenza artificiale ad alto rischio va innanzitutto evidenziato come in tale ambito il legislatore europeo vi abbia ricompreso due "categorie":

¹⁷³ Reg. (UE) 2024/1689, art. 3, c. 1, n. 63

¹⁷⁴ Reg. (UE) 2024/1689, art. 51: « 1. Un modello di IA per finalità generali è classificato come modello di IA per finalità generali con rischio sistemico se soddisfa una delle condizioni seguenti:

a) presenta capacità di impatto elevato valutate sulla base di strumenti tecnici e metodologie adeguati, compresi indicatori e parametri di riferimento;

b) sulla base di una decisione della Commissione, ex officio o a seguito di una segnalazione qualificata del gruppo di esperti scientifici, presenta capacità o un impatto equivalenti a quelli di cui alla lettera a), tenendo conto dei criteri di cui all'allegato XIII.

2. Si presume che un modello di IA per finalità generali abbia capacità di impatto elevato a norma del paragrafo 1, lettera a), quando la quantità cumulativa di calcolo utilizzata per il suo addestramento misurata in operazioni in virgola mobile è superiore a 10^{25} »

¹⁷⁵ European Commission, *General-Purpose AI Models in the AI Act – Questions & Answers*, <https://digital-strategy.ec.europa.eu/en/faqs/general-purpose-ai-models-ai-act-questions-answers>

a) sistemi che costituiscono componenti di un prodotto, che già sono soggetti alle normative di armonizzazione settoriali indicate all'Allegato I del regolamento (ad esempio autoveicoli, giocattoli, ascensori, dispositivi medici...)¹⁷⁶;

b) sistemi "*stand alone*" che presentano un rischio significativo di danno per la salute, la sicurezza o i diritti fondamentali delle persone fisiche e che sono ricompresi nell'Allegato III¹⁷⁷. L'allegato, attualmente,¹⁷⁸ individua otto aree ad alto rischio: biometria, infrastrutture critiche, istruzione e formazione professionale, occupazione, gestione dei lavoratori e accesso al lavoro autonomo, accesso a servizi privati essenziali e a prestazioni e servizi pubblici essenziali e fruizione degli stessi, attività di contrasto, migrazione, asilo e gestione del controllo delle frontiere, amministrazione della giustizia e processi democratici. I sistemi *high risk* sono soggetti sia ad obblighi di tipo "*sostanziale*" (il cui rispetto è necessario per la commercializzazione) che di tipo "*organizzativo*", relativi alle modalità "*dinamiche*" di gestione del rischio che operano successivamente alla messa in commercio dello stesso e a cui soggiacciono sia le imprese che le autorità di vigilanza¹⁷⁹.

I vincoli sostanziali sono esplicitati negli articoli da 10 a 15 del Regolamento e possono essere distinti in tre principali ambiti: (i) Trasparenza (ii) Sorveglianza umana (iii) Governance dei dati, robustezza, accuratezza e cybersicurezza.

(i) La trasparenza dei sistemi ad alto rischio si realizza nel principio della "*explainability by design*": per il Regolamento un sistema AI ad alto rischio può essere impiegato, e la sua decisione utilizzata, esclusivamente laddove l'*output* generato possa essere compiutamente compreso e spiegato; di converso qualora

¹⁷⁶ P. Deriu, S. Racioppi, *Riflessioni in tema di intelligenza artificiale e attività di vigilanza*, in Quaderni Fintech Consob, <https://www.consob.it/documents/d/area-pubblica/fintech15> p. 24

¹⁷⁷ Ai sensi dell'art. 6 c.4 del Regolamento AI è onere del fornitore provare alle Autorità competenti prima della messa in commercio che il sistema AI ricompreso nell'Allegato III non sia ad alto rischio.

¹⁷⁸ L'art. 7 dell'AI Act attribuisce alla Commissione il potere di adottare atti delegati al fine di modificare l'allegato III aggiungendo o modificando i casi d'uso dei sistemi di IA.

¹⁷⁹ B. Fragasso, *Il regolamento europeo sull'I.A. (AI ACT) e i primi tentativi di regolazione in Italia e all'estero*, in *Intelligenza artificiale e responsabilità penale*, Giappichelli Editore, 2025 p. 165

ciò non si verifichi il suo utilizzo deve intendersi vietato per contrarietà alla normativa in oggetto¹⁸⁰. Per ottemperare a tale obbligo è necessario, per la messa in commercio, che sia intrinseco nella progettazione dei sistemi un "principio di spiegabilità"¹⁸¹ che si declina nella:

- a) **Sorvegliabilità:** permettere da un lato alle autorità di vigilanza di verificare il rispetto dei requisiti per la messa in commercio dei sistemi di I.A., attraverso l'imposizione di obblighi di documentazione tecnica (art. 11) e dall'altro agli sviluppatori di poter supervisionare i sistemi durante il loro utilizzo (art. 13)
- b) **Tracciabilità:** avere lo storico degli eventi (log) che derivino dall'attivazione e dall'utilizzo del sistema (art. 12)
- c) **Intellegibilità:** garantire che il funzionamento del sistema sia sufficientemente trasparente da consentire ai deployer¹⁸² di «interpretare l'output del sistema e utilizzarlo adeguatamente (art. 13)»

Più nel dettaglio l'art. 11 prevede un obbligo di redigere la documentazione tecnica, di un sistema IA ad alto rischio prima dell'immissione sul mercato o della messa in servizio di tale sistema nonché di aggiornarla durante tutto il ciclo di vita del sistema; tale documentazione andrà sottoposta a un organismo notificato¹⁸³ che ne verificherà la conformità rispetto ai requisiti del Regolamento e, in caso di esito positivo rilascerà un certificato di conformità necessario ai fini

¹⁸⁰ B. Parenzo, "Informazioni significative" nei processi decisionali automatizzati: parla la Corte di Giustizia, in *La nuova giurisprudenza civile commentata*, 4/2025, Wolters Kluwer, pp 974 ss.

¹⁸¹ S. Sapienza, M. Palmirani, Algorithmic Knowability: A Unified Approach to Explanations in the AI Act, in R. Guidotti, U. Schmid, L. Longo, *Explainable Artificial Intelligence, Communications, Computer and Information Science*, vol 2576. Springer p. 193

¹⁸² Ai sensi dell'art. 3, c.4 dell'AI Act per deployer si intende «una persona fisica o giuridica, un'autorità pubblica, un'agenzia o un altro organismo che utilizza un sistema di IA sotto la propria autorità, tranne nel caso in cui il sistema di IA sia utilizzato nel corso di un'attività personale non professionale».

¹⁸³ Si tratta di organismi indipendenti che effettueranno la valutazione della conformità dei sistemi AI prima dell'immissione sul mercato; la designazione ed il controllo è in capo alla autorità nazionale di notifica.

della messa in commercio¹⁸⁴. Lo scopo di tale documentazione è duplice: da un lato assume una funzione probatoria permettendo ai fornitori di dimostrare che i loro sistemi AI sono conformi ai requisiti previsti dal regolamento; dall'altro lato fornisce ausilio alla attività di vigilanza fornendo alle autorità nazionali competenti e agli organismi notificati le informazioni necessarie per valutare il rispetto dei requisiti normativi.

I contenuti della documentazione tecnica sono delineati nell'allegato IV del Regolamento e comprendono aspetti rilevanti sulla progettazione e la gestione del sistema di IA ad alto rischio: si tratta di informazioni finalizzate a fornire una panoramica completa della macchina così da permettere ai fornitori di dimostrarne la conformità e alle autorità di valutarla.

Nella documentazione viene coperto l'intero ciclo di vita del sistema di IA, incluso il processo di sviluppo, la struttura e le prestazioni, le misure di mitigazione del rischio adottate e le attività previste per il monitoraggio del funzionamento del sistema. Ad ogni buon conto, va precisato come l'elenco di cui all'allegato IV non sia esaustivo poiché delinea solo le informazioni minime necessarie per comprendere lo specifico sistema di IA: il Regolamento stabilisce, infatti, che la documentazione tecnica debba includere "*almeno*" le informazioni specificamente indicate, ove pertinenti al caso concreto, lasciando però ai fornitori l'obbligo di integrarla con tutte le ulteriori informazioni necessarie a garantire una completa conoscenza¹⁸⁵. I contenuti minimi indicati nel regolamento dovranno inoltre essere integrati da standard tecnici e linee guida al fine di individuare in maniera omogenea le soluzioni tecniche e gli approcci da adottare per soddisfare tali esigenze documentali.

¹⁸⁴ Osservatorio Accredia, Norme tecniche e valutazione della conformità accreditata per lo sviluppo dei sistemi di Intelligenza Artificiale, https://www.accredia.it/wp-content/uploads/2024/11/01_2024_Osservatorio_ACCREDIA_Intelligenza_Artificiale_WEB.pdf pp 28 ss.

¹⁸⁵ A. Mantelero, G. Resta, G. M. Ricico, *Intelligenza Artificiale Commentario*, I Edizione, Wolters Kluwer, settembre 2025 p.240 ss.

L'art. 12 impone un obbligo di registrazione automatica degli eventi («log») per l'intera durata del ciclo di vita del sistema. Il «logging» è uno strumento indispensabile per garantire la tracciabilità *ex post* del funzionamento del sistema¹⁸⁶; fornendo uno strumento fondamentale per (i) verificare la conformità del sistema al regolamento, consentendo alle autorità di vigilanza e agli organismi notificati di appurare, *ex post*, che il sistema abbia operato rispettando l'AI Act e le istruzioni d'uso; (ii) fornire elementi indispensabili per indagare su incidenti gravi o malfunzionamenti, permettendo di risalire alla causa specifica che ha generato un output anomalo; (iii) procurare una base probatoria oggettiva per ricostruire la catena causale che ha portato a una determinata decisione algoritmica.

L'art. 13 enuclea un requisito funzionale, volto a consentire al deployer non solo di «interpretare l'output del sistema», ma soprattutto di «utilizzarlo adeguatamente». Il deployer deve infatti essere in grado di rispettare i propri obblighi di conformità, primo fra tutti quello relativo alla sorveglianza umana.

Affinché ciò si realizzi concretamente sul *provider* grava un duplice obbligo: da un lato, un obbligo di progettazione (art. 13 c. 1), in virtù del quale il sistema deve essere ideato e sviluppato *by design* in modo tale da permettere una facile interpretazione dei suoi output ed un utilizzo appropriato degli stessi; dall'altro lato, un obbligo informativo (art. 13 c. 2) che si sostanzia nell'accompagnare il sistema con delle istruzioni per l'uso chiare, concise e complete. Tali istruzioni rispondono ad una logica funzionale: se il deployer deve utilizzare il sistema di IA e monitorare il funzionamento, anche in un'ottica di gestione dei rischi è necessario che egli abbia contezza delle potenzialità e dei limiti del sistema. Tale contezza viene garantita proprio dal documento di istruzioni, fornito dal provider al deployer, nel quale vengono descritte le caratteristiche, le capacità e

¹⁸⁶ P. Marano, *L'Intelligenza Artificiale nella Realizzazione e Distribuzione dei Prodotti Assicurativi: Verso una Compliance Integrata tra AI Act, IDD e POG*, in [Rivista](#) di Diritto Bancario, Fascicolo II, 2025, p.291

i limiti di prestazione del sistema (come i livelli di accuratezza attesi e i rischi prevedibili), il suo scopo e le metriche di robustezza e cybersicurezza¹⁸⁷.

(ii) La sorveglianza umana è uno dei perni del Regolamento e la Commissione Europea la ha ritenuta essere una misura di salvaguardia fondamentale per garantire un'IA "antropocentrica" che sia al servizio delle persone, e garantisca la sicurezza e il rispetto dei diritti fondamentali¹⁸⁸. La supervisione umana ricalca nell'AI Act, l'approccio già seguito nella regolamentazione dei prodotti che vede nel ruolo umano di sorveglianza il mezzo per controllare i potenziali rischi della macchina¹⁸⁹.

La *human oversight* è disciplinata dall'articolo 14 dell'AI Act ai sensi del quale vige l'obbligo di progettare e sviluppare i sistemi «anche con strumenti di interfaccia uomo-macchina» in modo tale da consentire un'efficace supervisione da parte delle persone fisiche¹⁹⁰. Il comma 4 dell'articolo dettaglia le funzionalità tecniche che i fornitori, «ove opportuno e proporzionato» devono implementare al fine di permettere ad operatori umani dotati di competenze, formazione e autorità necessaria all'espletamento del compito di sorveglianza¹⁹¹ di:

- a) comprendere appieno le capacità e i limiti del sistema e monitorarne il funzionamento;
- b) rimanere consapevoli della potenziale tendenza ad affidarsi eccessivamente sull'output prodotto (c.d. "distorsione dell'automazione");
- c) interpretare correttamente l'output del sistema;

¹⁸⁷ A. Mantelero, G. Resta, G. M. Ricico, *Intelligenza Artificiale Commentario*, I Edizione, Wolters Kluwer, settembre 2025 p.251

¹⁸⁸ Commissione Europea, *Explanatory Memorandum*, par. 1.1. <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021PC0206>

¹⁸⁹ A. Mantelero, G. Resta, G. M. Ricico, *Intelligenza Artificiale Commentario*, I Edizione, Wolters Kluwer, settembre 2025 p.257

¹⁹⁰ L. Enqvist, "Human oversight' in the EU artificial intelligence act: what, when and by whom? in *Law, Innovation and Technology*, 15(2), pp. 508–535, 2023

¹⁹¹ AI Act, Art. 26 c.2 «I deployer affidano la sorveglianza umana a persone fisiche che dispongono della competenza, della formazione e dell'autorità necessarie nonché del sostegno necessario».

- d) decidere di non utilizzare il sistema o di ignorare, modificare o ribaltare l'output ottenuto;
- e) intervenire sul funzionamento o interrompere il sistema, mediante una procedura di arresto che consenta di fermare il sistema in condizioni di sicurezza.

L'obbligo di sorveglianza di cui all'art. 14 ha un carattere prettamente "negativo": la norma non impone al sorvegliante di sostituirsi al sistema AI, nello svolgimento dei compiti cui questo era preordinato ma gli attribuisce il compito di arrestare il sistema o di non prendere in considerazione la raccomandazione fornita dalla macchina qualora rilevi malfunzionamenti¹⁹².

Presupposto indefettibile per l'effettività della sorveglianza umana è l'intellegibilità di cui all'art. 13: senza la trasparenza progettuale e le informazioni da fornire ai sensi di tale articolo, il *deployer* (o le persone fisiche che operano per suo conto) non sarebbe materialmente in grado di esercitare adeguatamente il controllo e di implementare le misure di sorveglianza richieste dalla norma. In altre parole, non potrebbe essere compreso appieno l'output del sistema né valutata la sua correttezza nel contesto specifico e, di conseguenza, diverrebbe difficile decidere, in maniera consapevole, di non utilizzare, annullare o intervenire sull'output del sistema ad alto rischio¹⁹³.

(iii) L'AI Act nell'istituire un quadro normativo rigoroso per lo sviluppo dei sistemi ad alto rischio ha posto al centro della strategia europea la qualità dei dati come elemento imprescindibile per garantire innovazione e fiducia. In tale contesto, il legislatore europeo ha richiesto, all'art. 10, che i dati utilizzati per addestrare, convalidare e testare i sistemi ad alto rischio soddisfino stringenti criteri di qualità che devono essere garantiti attraverso il rispetto di pratiche

¹⁹² B. Fragasso, *Il regolamento europeo sull'I.A. (AI ACT) e i primi tentativi di regolazione in Italia e all'estero*, in *Intelligenza artificiale e responsabilità penale*, Giappichelli Editore, 2025 p. 173

¹⁹³ D. Schnurr, *Effective Implementation of Requirements for High-Risk AI Systems Under the AI Act: Transparency and Appropriate Accuracy*, Centro on Regulation in Europe (2025), pp 17-18

di governance e di gestione dei dati. Le pratiche di governance devono essere adeguate alle finalità che l'algoritmo persegue e strutturate per coprire l'intero ciclo di vita del dato e garantirne la qualità: dalle scelte progettuali e la raccolta, fino alla preparazione (che si esplica nella annotazione, etichettatura pulizia, aggiornamento, arricchimento e aggregazione) e alla formulazione di ipotesi su ciò che i dati dovrebbero misurare. Si vuole così prevenire il fenomeno del "garbage in, garbage out" termine con cui ci si riferisce alla generazione di output algoritmici distorti o discriminatori dovuti all'utilizzo di dati di bassa qualità (perché inesatti, incompleti o non rappresentativi)¹⁹⁴.

I dati devono infatti possedere «*le proprietà statistiche appropriate, anche, ove applicabile, per quanto riguarda le persone o i gruppi di persone relativamente ai quali il sistema di IA ad alto rischio è destinato a essere usato*¹⁹⁵». A riprova della centralità della qualità dei dati, il Regolamento introduce una significativa deroga al GDPR: l'art. 10, par. 5, consente infatti ai fornitori, laddove strettamente necessario, di trattare categorie particolari di dati personali al fine di rilevare e correggere i *bias*; tale facoltà viene esercitata nel rispetto di rigorose garanzie quali l'adozione di misure tecniche come la pseudonimizzazione, il rispetto di obblighi di *accountability*, che impongono ad esempio di documentare nei registri le motivazioni per cui l'uso di tali dati è ritenuto indispensabile e non sostituibile con dati sintetici¹⁹⁶.

Strettamente legata, seppure concettualmente distinta, dalla precisione e dalla qualità dei dati utilizzati per addestrare il modello è l'accuratezza intesa come la capacità del sistema di compiere il compito per cui è stato progettato fornendo

¹⁹⁴ B. Hanson, Sh. Stall, J. Cutcher-Gershenfeld, K. Vrouwenvelder, C. Wirz, Y. Rao, G. Peng, *Garbage in, garbage out: mitigating risks and maximizing benefits of AI in research*, in *Nature*, Vol 623, n. 28, p 29

¹⁹⁵ AI Act, art. 10 c. 3

¹⁹⁶ A. Mantelero, G. Resta, G. M. Ricico, *Intelligenza Artificiale Commentario*, I Edizione, Wolters Kluwer, settembre 2025 pp. 226 ss.

risultati corretti ed affidabili¹⁹⁷; l'art 15 la indica fra i requisiti¹⁹⁸ che devono guidare la progettazione dei sistemi AI assieme alla robustezza e la cybersicurezza. Tali requisiti sono finalizzati alla realizzazione di sistemi che siano resilienti *«sia per quanto riguarda errori, guasti o incongruenze che possono verificarsi all'interno del sistema o nell'ambiente in cui esso opera, in particolare a causa della loro interazione con persone fisiche o altri sistemi»* sia rispetto *«ai tentativi di terzi non autorizzati di modificarne l'uso, gli output o le prestazioni sfruttando le vulnerabilità del sistema»*.

L'aumento crescente degli attacchi informatici¹⁹⁹ impone la necessità di adottare soluzioni di mitigazione dei rischi cyber; il legislatore europeo, anche in virtù della consapevolezza della rapidità del progresso tecnologico non elenca, nel concreto, le misure tecniche da adottare per assicurare un adeguato livello di sicurezza dei sistemi di IA, ma indica le caratteristiche che queste devono avere per essere considerate adeguate ad affrontare i rischi nel caso concreto. Inoltre, il riferimento alle “soluzioni tecniche” di cui al comma 5 dell'art. 15 va interpretato in senso estensivo così da ricomprendervi anche misure di tipo organizzativo imprescindibili per garantire la più ampia sicurezza dei sistemi di IA. Allo stato della tecnica esempi di tali misure possono essere: l'utilizzo della crittografia per tutelare i dati sensibili durante la trasmissione e l'archiviazione ed impedire l'accesso non autorizzato alle informazioni; l'aggiunta di rumore artificiale²⁰⁰ ai

¹⁹⁷ O. Fernández, *Accuracy, Robustness and Cybersecurity (Article 15)*, in *The EU regulation on Artificial Intelligence: A commentary*, a cura di A. J. Huergo Lora (ed.), G. M. Díaz González, Wolters Kluwers Italia, p.274

¹⁹⁸ Tali requisiti sono espressione di un approccio, già seguito nel GDPR, di tipo proattivo e preventivo, mirato a identificare e mitigare i rischi prima che possano arrecare pregiudizi.

¹⁹⁹ Agenzia per la Cybersicurezza Nazionale, *Operational Summary 1° semestre 2025*, https://www.acn.gov.it/portale/documents/20119/989383/Operational_Summary_H1_2025_CLEAR.pdf/93dfccd6-695f-5e5c-f166-03f7550b2430?t=1754322719758

²⁰⁰ Tecnica di mascheramento dei dati che consiste nell'inserimento di elementi casuali in dati sensibili, al fine di ostacolare la decifrazione delle informazioni da parte di utenti non autorizzati. Un esempio in tal senso è la c.d. *privacy differenziale* una metodologia, adottata per la salvaguardia delle informazioni sensibili, che consiste nell'introdurre del rumore o delle variazioni nei dati complessivi, rendendo arduo per un soggetto esterno estrapolare informazioni specifiche relative a un singolo individuo a partire dai dati aggregati.

dati di addestramento per rendere il modello maggiormente resiliente; la pulizia dei dati di addestramento per rimuovere dati anomali; la verifica dell'accuratezza delle etichette. Tra le misure mitigative di tipo organizzativo si possono annoverare la garanzia dei controlli degli accessi; l'implementazione di sistemi di monitoraggio continuo per rilevare e rispondere rapidamente a eventuali anomalie o difetti del modello durante il suo utilizzo.

Per agevolare l'individuazione delle soluzioni appropriate per la progettazione e lo sviluppo dei sistemi di IA, sotto il profilo anche dell'accuratezza, della robustezza e della sicurezza, l'art. 40 prevede l'adozione di specifiche tecniche da far elaborare agli organismi europei di normazione competenti. La definizione dei soli requisiti essenziali ed il rimando alle organizzazioni europee di armonizzazione per lo sviluppo dei relativi standard è una tecnica legislativa da tempo adottata nell'UE che permette di creare un quadro normativo più efficace per i prodotti e i servizi all'interno del mercato unico²⁰¹. Gli standard armonizzati, se rispettati volontariamente, oltre a semplificare la procedura di valutazione della conformità, consentiranno di attenuare gli obblighi previsti, per i sistemi di IA ad alto rischio, anche sotto il profilo probatorio: il loro rispetto rappresenterà infatti una presunzione di conformità dei sistemi di IA ai requisiti legali (cfr. Considerando n. 121 e artt. 40 e 41) fermo restando la possibilità per i fornitori dei sistemi di implementare misure specifiche differenti da quelle previste dalle norme armonizzate a patto di dimostrarne la conformità ai requisiti previsti dal Regolamento²⁰².

Agli obblighi sostanziali dei sistemi di intelligenza artificiale si affiancano vincoli organizzativi per fornitori (*providers*) ed utilizzatori (*deployers*) che

²⁰¹ Per un approfondimento sul ruolo degli standard nella legislazione europea si rimanda a M. McFadden, K. Jones, E. Taylor, G. Osbornpp, *Harmonising Artificial Intelligence: The Role of Standards in the EU AI Regulation*, 2021, Oxford Commission on AI and Good Governance, 2021, 8 ss.

²⁰² A. Mantelero, G. Resta, G. M. Ricico, *Intelligenza Artificiale Commentario*, I Edizione, Wolters Kluwer, settembre 2025 pp.275 ss.

delineano una governance del rischio condivisa tra i diversi attori della catena del valore.

Rilievo centrale assume l'obbligo per il fornitore di istituire, attuare, documentare e mantenere un sistema di gestione dei rischi conformemente alle previsioni di cui all'articolo 9 del Regolamento.

Il sistema di gestione dei rischi non si configura come una valutazione statica da compiere *“una tantum”* ma si realizza in *«un processo iterativo continuo pianificato ed eseguito nel corso dell'intero ciclo di vita di un sistema di IA ad alto rischio, che richiede un riesame e un aggiornamento costanti e sistematici»* (art. 9, par. 2). Tale processo, finalizzato a garantire un elevato livello di affidabilità e di tutela dei diritti fondamentali, si articola in varie fasi tra loro interconnesse:

- a) **identificazione e analisi dei rischi:** il fornitore deve mappare i rischi noti e ragionevolmente prevedibili per la salute, la sicurezza e i diritti fondamentali, che possono essere generati durante l'uso del sistema conforme alla sua finalità prevista;
- b) **stima e valutazione dei rischi:** il fornitore deve stimare e valutare i rischi emergenti sia in caso di uso del sistema conformemente alla sua finalità prevista sia in caso ne venga fatto un uso improprio ma ragionevolmente prevedibile (c.d. *“misuse”*); il fornitore deve pertanto anticipare comportamenti umani o interazioni con altri sistemi che, pur non essendo l'intento originale, sono prevedibili date le caratteristiche del sistema stesso;
- c) **valutazione post immissione sul mercato:** il fornitore, sulla base dell'analisi dei dati raccolti dal sistema di monitoraggio successivo all'immissione sul mercato deve valutare la sussistenza di ulteriori rischi e se nel caso integrare il sistema di gestione dei rischi;

- d) **adozione di misure di gestione dei rischi:** una volta individuati i rischi il fornitore deve implementare opportune azioni mirate ad affrontare i rischi individuati²⁰³.

L'architettura del sistema di *risk management* si fonda su di un concetto di "rischio residuo accettabile" (art. 9 c.5) tale per cui le misure per la sua gestione devono garantire che residuino solo rischi considerati ammissibili. Determinare la soglia di accettabilità del rischio è tuttavia un processo complesso dal momento che non è pensabile che si possa compiere una mera valutazione cumulativa degli interessi eterogenei in gioco (ad es. non è pensabile che la sicurezza del sistema possa compensare un rischio elevato per la tutela dei diritti fondamentali o viceversa). Dal momento che nel Regolamento non viene indicata una metodologia da seguire per compiere l'attività di valutazione del rischio, l'unica interpretazione plausibile appare essere quella di esaminare i vari rischi non dal punto di vista aggregativo ma combinatorio, non prevedendo un unico indice di rischio bensì considerando i diversi impatti che possono esserci sui singoli diritti fondamentali e sulla sicurezza nel suo complesso e valutando successivamente che tutti i rischi, nella loro singola dimensione, rientrino all'interno della soglia di accettabilità²⁰⁴.

Gli obblighi in capo al fornitore ex art. 9 si integrano sinergicamente con quelli previsti in capo al deployer dall'art 26. Dal momento che è impensabile che il fornitore riesca a prevedere tutti i possibili contesti d'uso e i relativi rischi specifici sorge in capo al deployer un obbligo di adottare «*idonee misure tecniche e organizzative per garantire di utilizzare tali sistemi conformemente alle istruzioni per l'uso che accompagnano i sistemi*²⁰⁵» ed a monitorare il funzionamento del sistema,

²⁰³ Federation of European Risk Management Associations, *The EU's Artificial Intelligence Act*, https://www.ferma.eu/wp-content/uploads/2024/10/FERMA-Policy-Note-EU-AI-Act_v0.2.pdf?utm_source=chatgpt.com p.3

²⁰⁴ A. Mantelero, G. Resta, G. M. Ricico, *Intelligenza Artificiale Commentario*, I Edizione, Wolters Kluwer, settembre 2025 pp.219

²⁰⁵ Regolamento (Ue) 2024/1689, Art. 26 c.1.

sulla base delle istruzioni per l'uso e, informare i fornitori qualora ritengano che il sistema presenti un rischio per la salute, la sicurezza od i diritti fondamentali delle persone.

Il ruolo del deployer trascende quello di mero utilizzatore passivo, e si configura come parte attiva nella gestione del rischio²⁰⁶ motivo per cui egli deve:

- i. **Garantire la sorveglianza umana:** incaricando soggetti che dispongano non solo della competenza e della formazione necessarie, ma anche di autorità e del sostegno operativo adeguati (26 c.2).
- ii. **Controllare i dati di input:** nei limiti in cui esercita controllo su di essi, il *deployer* deve assicurare che siano «*pertinenti e sufficientemente rappresentativi* (26 c.4)» rispetto alla finalità del sistema, al fine di prevenire *bias* del sistema dovuti da sottorappresentazione o sovrarappresentazione²⁰⁷.
- iii. **Monitorare e segnalare:** il funzionamento del sistema sulla base delle istruzioni d'uso e qualora, durante l'uso conforme, emerga motivo di ritenere che il sistema presenti un rischio per la salute, la sicurezza o i diritti fondamentali il deployer ha l'obbligo di informare senza indebito ritardo il fornitore o il distributore e l'autorità di vigilanza, sospendendo contestualmente l'uso del sistema (26 c.5).

L'allegato III dell'AI Act al punto 5 lett b) ricomprende fra i sistemi di AI ad alto rischio i «*sistemi di IA destinati a essere utilizzati per valutare l'affidabilità creditizia delle persone fisiche o per stabilire il loro merito di credito*». Il loro inserimento

²⁰⁶ Tale ruolo è chiaramente esplicitato al considerando 91 dell'AI Act laddove viene affermato che «*in considerazione della natura dei sistemi di IA e dei possibili rischi per la sicurezza e i diritti fondamentali associati al loro utilizzo, [...] è opportuno stabilire responsabilità specifiche per i deployer... È in particolare opportuno che i deployer adottino misure tecniche e organizzative adeguate per garantire di utilizzare i sistemi di IA ad alto rischio conformemente alle istruzioni per l'uso e che siano previsti alcuni altri obblighi in materia di monitoraggio del funzionamento dei sistemi di IA e conservazione delle registrazioni...*»

²⁰⁷ A. Mantelero, G. Resta, G. M. Ricico, *Intelligenza Artificiale Commentario*, I Edizione, Wolters Kluwer, settembre 2025 p.328

nella categoria *high risk* è giustificato, nel considerando 58 del Regolamento, sulla base del fatto che i sistemi di credit scoring algoritmico possono avere un ruolo decisivo decisioni sull'accesso ai servizi essenziali²⁰⁸ e che un loro utilizzo non regolato potrebbe *«portare alla discriminazione fra persone o gruppi e perpetuare modelli storici di discriminazione, come quella basata sull'origine razziale o etnica, sul genere, sulle disabilità, sull'età o sull'orientamento sessuale, o poss[an]no dar vita a nuove forme di impatti discriminatori»*.

Proprio in ragione dell'impatto che i sistemi di credit scoring algoritmico possono avere sui diritti fondamentali il Regolamento, all'art. 27, prevede per il loro utilizzo la necessità di rispettare di un ulteriore obbligo che si esplicita nella valutazione di impatto sui diritti fondamentali (c.d. Fria): prima di utilizzare l'intelligenza artificiale in settori sensibili, quale quello della concessione del credito, il *deployer* deve dimostrare di aver valutato anticipatamente l'impatto che l'utilizzo da parte sua di tale tecnologia possa avere sui diritti fondamentali delle persone, sanciti dalla Carta di Nizza, quali dignità, libertà, uguaglianza, non discriminazione. Si tratta a ben vedere di un ampliamento a tutti i diritti fondamentali di un istituto già previsto nel GDPR (art. 35) ossia la valutazione di impatto sulla protezione dei dati personali (DPIA)²⁰⁹. La previsione di un tale obbligo in capo all'utilizzatore si spiega alla luce della diversità degli *output* che può produrre un sistema a seconda dell'ambiente con cui interagisce nonché delle modalità con cui viene implementato dal *deployer* circostanze queste che portano a notevoli variazioni dal punto di vista degli interessi tutelati e dei soggetti i cui diritti fondamentali possono essere messi a rischio. Si spiega quindi

²⁰⁸ Ivi, p. 337

²⁰⁹ Come affermato dal Gruppo di Lavoro Articolo 29 per la Protezione dei Dati nelle *“Linee guida in materia di valutazione d'impatto sulla protezione dei dati e determinazione della possibilità che il trattamento “possa presentare un rischio elevato” ai fini del regolamento (UE) 2016/679”*, per DPIA si intende un processo inteso a descrivere il trattamento dei dati, valutarne la necessità e la proporzionalità, nonché a contribuire a gestire i rischi per i diritti e le libertà delle persone fisiche derivanti dal trattamento di dati personali, valutando detti rischi e determinando le misure per affrontarli.

perché è necessario far compiere, accanto alla valutazione effettuata dal fornitore, anche una specifica valutazione da parte dell'utilizzatore. Quest'ultimo, come osservato dal considerando 93, è nella posizione migliore per comprendere come il sistema ad alto rischio «sarà utilizzato concretamente» e può pertanto individuare «potenziali rischi significativi non previsti nella fase di sviluppo, in ragione di una conoscenza più puntuale del contesto di utilizzo e delle persone o dei gruppi di persone che potrebbero essere interessati, compresi i gruppi vulnerabili».

Dal punto di vista degli obblighi che il *deployer* deve ottemperare la FRIA gli richiede di compiere una mappatura del contesto di utilizzo, di identificare le persone che possono essere impattate da tale sistema e con quali conseguenze, nonché di descrivere le modalità di supervisione umana ed i meccanismi rimediali adottati²¹⁰. Tale valutazione accompagna il sistema per l'intera vita del sistema dovendo essere condotta prima della messa in uso, aggiornata al verificarsi di modifiche sostanziali e riesaminata con periodicità. Applicata al *credit scoring*, la FRIA richiederà al *deployer* di esaminare i *dataset* per individuare variabili che possano generare *bias* nel sistema; garantire la trasparenza nonché assicurare un intervento umano nelle decisioni²¹¹.

La Fria, una volta redatta, va trasmessa alla autorità nazionale di vigilanza del mercato, la quale dovrà verificare il rispetto delle previsioni dell'AI Act.

²¹⁰ Sulle modalità pratiche per svolgere tale attività si segnala A. Cosentini, O. Pollicino, G. De Gregorio, A. Ermellino, D. Fontanella, I. Nicole, F. Paolucci, I. G. Penco, D. Regoli, S. Tessaro Trapani, *Assessing the Impact of Artificial Intelligence Systems on Fundamental Rights Available* i quali suggeriscono di adottare un'impostazione "risk-based" articolata in tre fasi tra loro consequenziali: (1) una fase preliminare di classificazione del rischio e di analisi della finalità previste dall'uso del sistema AI e del contesto in cui viene utilizzato, volta a verificare la qualificazione del sistema come high-risk e a delimitarne l'uso concreto, così da costituire la base giustificativa delle scelte successive; (2) una fase di mappatura tecnica e monitoraggio, nella quale le caratteristiche del sistema vengono esaminate e tradotte in un indicatore sintetico di rischio costruito mediante un questionario che attribuisce a ciascuna risposta un livello di rischio in una scala da 1 (rischio minimo) a 10 (rischio massimo); (3) una fase di valutazione dell'impatto sui diritti fondamentali, che integra i risultati del questionario con una matrice di valutazione per stimare, per ciascun diritto considerato, gravità e probabilità degli impatti, così da ottenere uno score per singolo diritto su cui basarsi per adottare misure di mitigazione coerenti e proporzionate.

²¹¹ F. Paoloucci, O. Pollicino, *Diritti fondamentali Valutazione d'impatto per banche, assicurazioni, enti o servizi pubblici*, in *Norme & Tributi*, Il Sole 24 Ore 30 ottobre 2025

Con riguardo all'individuazione delle autorità di vigilanza il Regolamento AI traccia un complesso quadro di governance di cui si cercherà di dare conto nel prossimo paragrafo.

6. La governance dell'Intelligenza Artificiale in Unione Europea ed in Italia: il coordinamento con la vigilanza bancaria

Al fine di applicare uniformemente il regolamento e mantenere la competitività dell'Unione Europea in materia di intelligenza artificiale l'AI Act ha previsto un sistema di governance multilivello nel quale collaborano Commissione europea, Stati membri, autorità nazionali competenti e organismi tecnici²¹².

In ambito europeo sono previsti quattro organismi:

- i. **l'Ai Office** (art. 64): istituito, con Decisione C/2024/1459 presso la Commissione Europea²¹³, ha il compito, ai sensi del regolamento, di assicurare l'applicazione coerente ed armonizzata del regolamento anche mediante la predisposizione di codici di condotta, linee guida, atti delegati e regolamenti attuativi. Con la Decisione summenzionata la Commissione Europea ha altresì attribuito all'Ufficio il compito di rafforzare un approccio unitario ed efficace dell'Unione alle iniziative globali sull'IA, promuovere iniziative e politiche che sfruttino i benefici sociali ed economici dell'IA, sostenere lo sviluppo di soluzioni di IA affidabili e utili, monitorare le tendenze e gli scenari di sviluppo dei mercati e delle tecnologie di IA²¹⁴.

²¹² S. Pellegrini, R. Pupella, *Competenze suddivise tra le Ue e le autorità degli Stati Membri*, in *Norme & Tributi*, Il Sole 24 Ore 30 ottobre 2025

²¹³ L'Ufficio è inserito all'interno della struttura amministrativa della Direzione Generale delle Reti di comunicazione, dei contenuti e delle tecnologie.

²¹⁴ A. Mantelero, G. Resta, G. M. Ricico, *Intelligenza Artificiale Commentario*, I Edizione, Wolters Kluwer, settembre 2025 p. 602

- ii. **Il Consiglio per l'IA europeo per l'intelligenza artificiale** (artt. 65-66): è composto da un rappresentante per Stato membro e sostenuto dall'Ufficio per l'IA che ne svolge la funzione di segretariato. Modellato sui compiti e le responsabilità del Comitato europeo per la protezione dei dati (ex art. 68 GDPR), il Consiglio fornisce consulenza e assistenza alla Commissione e agli Stati membri al fine di agevolare l'applicazione coerente ed efficace del regolamento anche discutendo orientamenti, progetti di atti delegati e di esecuzione. Ma il Consiglio non si limita alla sola attività di regolamentazione, esso ha anche il compito di garantire all'interno dell'UE l'effettiva attuazione della normativa sull'IA, coordinando e garantendo la cooperazione tra gli Stati membri dell'UE, in particolare con le altre istituzioni comunitarie e fornendo assistenza allo sviluppo delle competenze necessarie per l'attuazione del Regolamento²¹⁵;
- iii. il **Forum consultivo** (art. 67): composto da membri scelti dalla Commissione, fra differenti stakeholders, fornisce consulenza e competenze tecniche al consiglio per l'AI e alla Commissione europea, al fine di aiutarli a meglio comprendere e applicare la normativa sull'IA, garantendo al contempo una varietà di prospettive e competenze, cruciali per un approccio inclusivo ed equo all'implementazione del regolamento;
- iv. il **Gruppo di esperti scientifici indipendenti** (art. 68): composto da personalità indipendenti dotate di elevate conoscenze e capacità tecnico/scientifiche nel campo dell'AI ha il compito di fornire consulenza e sostegno all'ufficio per l'AI.

La Commissione e l'Ufficio per l'IA si coordinano con le Autorità Nazionali Competenti (ANC) cui è assegnato il compito di supervisionare l'attuazione e l'applicazione dell'AI Act a livello nazionale: il Regolamento, all'art. 70, prevede

²¹⁵ Commissione Europea, *Consiglio per l'IA*, <https://digital-strategy.ec.europa.eu/it/policies/ai-board>

che, per garantire l'applicazione uniforme della normativa, ciascuno Stato membro debba istituire o designare come ANC almeno un'autorità di notifica e un'autorità di vigilanza del mercato, separando funzionalmente le fasi di controllo.

Più nello specifico le autorità di notifica sono chiamate a designare e controllare gli organismi notificati ossia entità terze indipendenti responsabili di valutare e certificare *ex ante* il rispetto dei requisiti dell'AI Act per i sistemi di intelligenza artificiale ad alto rischio, condizione imprescindibile per la loro marcatura CE e la conseguente immissione sul mercato²¹⁶. Alle autorità di vigilanza del mercato, è invece attribuito il compito di vigilare *ex post* sul corretto rispetto delle norme, da parte di produttori, distributori ed utilizzatori dei sistemi AI. Alle autorità di vigilanza sono attribuiti penetranti poteri di intervento qualora i sistemi pongano rischi per la salute, la sicurezza o i diritti fondamentali ovvero non siano conformi al regolamento; nel loro potere di vigilanza e indagine, esse possono accedere alla documentazione tecnica dei fornitori, ai set di dati di addestramento e validazione utilizzati, nonché, a determinate condizioni²¹⁷ al codice sorgente. L'accesso al codice sorgente costituisce una misura di ultima istanza per verificare la conformità di un sistema di IA ad alto rischio ai requisiti dell'AI Act. L'art. 74, par. 13, ne delimita infatti l'operatività, subordinandola all'esaurimento degli altri strumenti istruttori a disposizione della autorità di vigilanza del mercato siano state esaurite²¹⁸. Dal punto di vista dell'esercizio del potere amministrativo l'uso di una simile prerogativa sarà subordinato all'emanazione

²¹⁶J. A. Carrillo Donaire, *Notifying Authorities and Notified Bodies (Chapter III, Section 4). Standards, Conformity Assessment, Certificates, Registration (Chapter III, Section 5)*, in *The EU regulation on Artificial Intelligence: a commentary* a cura di G. M. Díaz González; A. J. Huergo Lora, Wolters Kluwer Italia, 2025 p. 356-357

²¹⁷ Ai senso dell'art. 74 c. 13 le autorità di vigilanza del mercato possono accedere al codice sorgente al ricorrere congiunto di due condizioni: a) l'accesso al codice sorgente deve essere necessario per valutare la conformità di un sistema di IA ad alto rischio ai requisiti dei sistemi enunciati nel capo III, sezione 2 del Regolamento; le procedure di prova o di audit e le verifiche basate sui dati e sulla documentazione presentati dal fornitore sono state esaurite o si sono dimostrate insufficienti.

²¹⁸ D. R. Cembellin, *Enforcement (Chapter IX, Section 3)*, in *The EU regulation on Artificial Intelligence: a commentary* a cura di G. M. Díaz González; A. J. Huergo Lora, Wolters Kluwer Italia, 2025 pp 484-485.

di un provvedimento amministrativo motivato nel quale l'autorità di sorveglianza del mercato dovrà dare conto della presenza dei presupposti necessari per accedere al codice sorgente.

Al contempo, tuttavia, l'AI Act non fa venire meno il ruolo delle autorità settoriali: la vigilanza e la regolazione dei mercati seguono una organizzazione a matrice con autorità incaricata di vigilare e (molte di esse anche) regolare mercati "verticali" (banche, assicurazioni, mercati finanziari, energia/gas/rifiuti, comunicazioni elettroniche, trasporti) e autorità competenti a vigilare "orizzontalmente" su tutti i mercati, in modo trasversale (concorrenza, privacy, tutela dei consumatori); queste autorità operano in mercati nei quali i sistemi di intelligenza artificiale stanno modificando velocemente parametri di riferimento e rapporti di forza tra gli operatori ed ognuna di esse, in ragione dell'*expertise* acquisita nel settore, è la più idonea a intervenire sul mercato che già vigila e come nel caso del settore bancario, regola²¹⁹.

Si comprende perciò perché l'art. 74 del Regolamento²²⁰ espressamente affermi che «6. *Per i sistemi di IA ad alto rischio immessi sul mercato, messi in servizio o usati da istituti finanziari disciplinati dal diritto dell'Unione in materia di servizi finanziari, l'autorità di vigilanza del mercato ai fini del presente regolamento è l'autorità nazionale pertinente responsabile della vigilanza finanziaria di tali enti ai sensi di tale diritto, nella misura in cui l'immissione sul mercato, la messa in servizio o l'uso del sistema di IA siano direttamente collegati alla fornitura di tali servizi finanziari*».

Il legislatore europeo consapevole della necessità di un coordinamento in vista delle interrelazioni fra silos verticali e orizzontali ha agito su due fronti: da un

²¹⁹ M. Rabitti, F. Bassan, *L'applicazione dell'AI Act in Italia e la tutela del consumatore. Il ruolo delle autorità indipendenti*, <https://romatrepress.uniroma3.it/wp-content/uploads/2025/07/DECA-n.-10-2025-ebook.pdf> p.20

²²⁰ Il regolamento permette, in casi motivati, al legislatore nazionale di attribuire tale compito alla autorità "generale" sull'intelligenza artificiale fermo restando la necessità di garantire il coordinamento e di comunicare eventuali elementi rilevanti per la vigilanza prudenziale alle autorità competenti ai sensi del Meccanismo di Vigilanza Unico.

lato ha inserito disposizioni di coordinamento in capo ai soggetti vigilati; in tal senso rilevano le previsioni contenute negli articoli 17 e 26 del Regolamento che considerano assolto, dagli enti creditizi fornitori e/o utilizzatori di sistemi di IA ad alto rischio, l'obbligo di istituire un sistema di gestione della qualità e di monitorare il sistema *high risk*, laddove tali enti rispettino le norme sui dispositivi, i processi e i meccanismi di governance interna stabiliti nelle rigorose disposizioni di vigilanza (es. CRD e CRR). Dall'altro lato ha demandato agli Stati membri il compito di agevolare il coordinamento fra le varie autorità.

Il Parlamento italiano, per dare corretta applicazione al Regolamento Europeo e promuovere «*un utilizzo corretto, trasparente e responsabile, in una dimensione antropocentrica, dell'intelligenza artificiale, volto a coglierne le opportunità*²²¹» ha approvato la Legge n. 132/2025 (c.d. Legge sull'Intelligenza Artificiale) con la quale sono state designate, quali autorità nazionali per l'intelligenza artificiale, l'Agenzia per l'Italia Digitale (autorità di notifica) e l'Agenzia per la Cybersicurezza Nazionale (autorità di controllo). L'articolo 20 nell'attribuire i compiti alle due autorità menzionate accoglie la deroga funzionale, di cui all'art. 74 dell'AI Act per il settore finanziario affermando esplicitamente che è fatta salva «*l'attribuzione alla Banca d'Italia, alla CONSOB e all'IVASS del ruolo di autorità di vigilanza del mercato*».

Per evitare frammentazioni nella prassi applicativa l'articolo 20 demanda altresì alle Autorità nazionali per l'intelligenza artificiale il compito di assicurare il coordinamento con le autorità indipendenti ed il raccordo tra le funzioni; per la realizzazione di tale compiuto viene istituito, presso la Presidenza del Consiglio dei ministri un Comitato di coordinamento, composto dai direttori generali delle due Agenzie e dal capo del Dipartimento per la trasformazione digitale della Presidenza del Consiglio dei ministri cui partecipano, quando si trattano

²²¹ Legge 23 settembre 2025, n. 132, Art. 1

questioni di rispettiva competenza, rappresentanti di vertice della Banca d'Italia, della CONSOB e dell'IVASS.

È inoltre attribuita al governo, dall'art. 24, la delega ad «*apportare alla normativa vigente, ivi inclusa quella in materia di servizi bancari, finanziari, assicurativi e di pagamento, le modifiche, le integrazioni e le abrogazioni necessarie al corretto e integrale adeguamento al regolamento (UE) 2024/1689*».

L'utilizzo dell'AI nell'ambito del settore finanziario viene dunque esclusa dal mandato di sorveglianza generale di ACN e AgID a favore di un regime di "specialità" in cui è l'autorità che vigila sull'istituzione finanziaria ad essere competente anche sul rispetto della normativa in materia di intelligenza artificiale con riguardo ai sistemi che questi usa.

L'approccio seguito dal legislatore italiano è pienamente conforme con quanto affermato dal considerando 158 dell'AI Act laddove viene affermato come sia opportuno che le autorità competenti del controllo e dell'esecuzione della *Capital Requirements Regulation*²²², della *Capital Requirements Directive*²²³, della *Consumer Credit Directive*²²⁴ e della *Mortgage Credit Directive*²²⁵ siano designate, nell'ambito delle rispettive competenze, quali autorità competenti ai fini del controllo dell'attuazione del regolamento IA stesso, anche in relazione alle attività di vigilanza del mercato, per quanto riguarda i sistemi di IA ad alto rischio forniti o utilizzati da istituti finanziari regolamentati e sottoposti a vigilanza.

²²² Regolamento (UE) n. 575/2013 del Parlamento europeo e del Consiglio, del 26 giugno 2013, relativo ai requisiti prudenziali per gli enti creditizi e le imprese di investimento e che modifica il regolamento (UE) n. 648/2012

²²³ Direttiva 2013/36/UE del Parlamento Europeo e del Consiglio del 26 giugno 2013 sull'accesso all'attività degli enti creditizi e sulla vigilanza prudenziale sugli enti creditizi e sulle imprese di investimento, che modifica la direttiva 2002/87/CE e abroga le direttive 2006/48/CE e 2006/49/CE

²²⁴ Direttiva 2008/48/CE del Parlamento Europeo e del Consiglio del 23 aprile 2008 relativa ai contratti di credito ai consumatori e che abroga la Direttiva 87/102/CEE

²²⁵ Direttiva 2014/17/UE del Parlamento Europeo e Del Consiglio del 4 febbraio 2014 in merito ai contratti di credito ai consumatori relativi a beni immobili residenziali e recante modifica delle direttive 2008/48/CE e 2013/36/UE e del regolamento (UE) n. 1093/2010

Tale scelta è giustificata sulla base del fatto che gli intermediari finanziari soggiacciono già a significativi obblighi di governance (sana e prudente gestione, trasparenza...) i quali si applicano sempre indipendentemente dagli strumenti cui si fa ricorso; in altre parole, è indifferente che la banca usi un pallottoliere od un algoritmo di machine learning: gli "elementi sostanziali" (*rectius* la necessità di proteggere il risparmiatore e la stabilità del sistema) non cambiano. L'IA è solo un nuovo strumento utilizzabile per svolgere le stesse attività motivo per il quale non vi è ragione di sostituire il controllore ma solo di integrarne i compiti. Ma non solo, non può ignorarsi come solo l'autorità settoriale possa essere in grado di meglio comprendere modelli applicati ai mercati finanziari i quali necessitano di una *expertise* tecnico-economica difficilmente replicabile da una autorità "generalista". Inoltre, dal punto di vista della vigilanza prudenziale, separare la vigilanza sull'algoritmo da quella sul rischio di credito avrebbe rischiato di creare zone d'ombra nella valutazione della stabilità dell'intermediario.

Per tali motivi una applicazione coerente e armonizzata degli obblighi dell'AI Act con la normativa prevista per le banche può realizzarsi solo qualora la stessa venga supervisionata dalla autorità di vigilanza bancaria.

Assume dunque assoluta logicità, nell'ambito del credit scoring algoritmico, il fatto che sia la Banca d'Italia a dover assicurare da un lato che i sistemi non siano discriminatori, escludendo ad esempio dall'accesso al credito alcuni richiedenti sulla scorta di *bias* storici, magari incentrati su genere o razza, e dall'altro lato che non siano lesi i principi di sana e prudente gestione dell'intermediario finanziario, primo fra tutti quello secondo il quale il credito non va concesso a tutti coloro che lo richiedono ma solo a coloro che meritano di averlo in base alle capacità di rimborso.²²⁶

²²⁶ L. Ammannati, G. Grieco, *Il credit scoring "intelligente": esperienze, rischi e nuove regole*, in *Rivista di diritto bancario*, luglio/settembre 2023 p. 508

3. RISCHIO DI CREDITO E MACHINE LEARNING

1. Il rischio di credito nell'ambito della disciplina prudenziale delle banche.

Nel precedente capitolo si è fatto riferimento al principio di sana e prudente gestione, ex art. 5 del Tub, quale bussola che deve guidare sia l'attività d'impresa delle banche che l'attività di vigilanza sulle stesse. Il rispetto di tale principio viene realizzato attraverso regole e meccanismi di controllo dell'attività delle bancaria che prendono il nome di vigilanza prudenziale.

La disciplina prudenziale si fonda sulla impostazione tracciata dagli Accordi di Basilea ed in particolare sui tre pilastri enucleati da Basilea 2²²⁷:

- 1) possesso di requisiti patrimoniali minimi per fronteggiare i rischi tipici dell'attività bancaria (credito, mercato, controparte, operativi);
- 2) adozione, da parte delle banche, di una strategia e di un processo di controllo prudenziale della adeguatezza patrimoniale rispetto al quale sussiste un corrispettivo compito, in capo all'autorità di vigilanza, di verificare l'affidabilità e la coerenza dei risultati e di adottare, ove necessario, misure correttive;
- 3) il rispetto di obblighi di informativa nei confronti del pubblico in merito alla adeguatezza patrimoniale, l'esposizione ai rischi e le caratteristiche generali dei sistemi di gestione e controllo²²⁸.

²²⁷ Le previsioni di Basilea 2 erano state recepite nell'Unione Europea dalla Direttiva 2006/48/CE, sull'accesso all'attività degli enti creditizi, e dalla Direttiva 2006/49/CE, sulla adeguatezza patrimoniale, note congiuntamente come pacchetto CRD (Capital Requirements Directives).

Le "Direttive" avevano tuttavia concesso ai legislatori nazionali ampia flessibilità nella loro applicazione il che aveva portato a un quadro frammentato a livello unionale. La crisi finanziaria del 2008 e la successiva crisi del debito sovrano europeo hanno fatto emergere la necessità di una risposta normativa uniforme, capace di garantire che il capitale di una banca avesse la medesima consistenza e qualità in ogni giurisdizione dell'Unione; cosicché nel recepire gli accordi di Basilea 3, l'Unione Europea ha optato per utilizzare maggiormente lo strumento del regolamento ed ha pertanto enucleato le norme prettamente tecniche e quantitative, quali il calcolo dei requisiti patrimoniali, dello standard minimo di liquidità e della leva finanziaria, nel Regolamento UE n. 575/2013 (c.d. Capital Requirements Regulation). La Direttiva 2013/36/UE (c.d. Capital Requirements Directive IV) ha invece regolato la materia della governance bancaria, delle politiche di remunerazione, delle sanzioni amministrative e delle riserve di capitale anticicliche così da potere mantenere un necessario grado di adattamento alle specificità del diritto nazionale degli stati.

²²⁸ R. Costi, L'ordinamento bancario, Il Mulino, Quinta Edizione 2012, pp. 582 ss.

Come detto, il primo pilastro di Basilea richiede che la banca sia in possesso di risorse, (la Circolare della Banca d'Italia n. 285/2013 le denomina "fondi propri") finalizzate di fronteggiare i rischi e coprire le perdite di varia natura; ai sensi dell'art. 92 del Regolamento UE n. 575/2013 (c.d. CRR), che ha trasposto le norme di Basilea 3, lasciando immutata la previsione di Basilea 2²²⁹, il requisito patrimoniale minimo di fondi propri che le banche devono detenere come riserva è pari all'8% delle attività ponderate per il rischio. Tale attività di ponderazione rispetto alle diverse tipologie di rischio viene svolta applicando metodi di calcolo dettagliati nella CRR e nelle disposizioni di vigilanza²³⁰.

Soffermandoci sulla ponderazione delle attività rispetto al rischio di credito (la possibilità che il debitore non adempia alla propria obbligazione pecuniaria nei tempi contrattualmente previsti) va rilevato come esso possa calcolarsi tramite due differenti sistemi: il metodo standard ed il metodo dei rating interni (*Internal Risk Based*, in seguito anche IRB).

Nel metodo standard le esposizioni vengono suddivise in classi di rischio in base alla natura della controparte, delle caratteristiche tecniche del rapporto nonché delle modalità di svolgimento di quest'ultimo: l'art. 112 della CRR elenca diciassette classi di esposizioni, tra le quali si annoverano amministrazioni e banche centrali, organismi del settore pubblico, organizzazioni internazionali, imprese, clientela al dettaglio, ecc.... L'inserimento di un finanziamento in una determinata classe di esposizione comporta la moltiplicazione del suo ammontare per un determinato coefficiente di rischio, dettagliatamente esplicitato dagli articoli 114 e seguenti della CRR. Tuttavia, il coefficiente opera solo qualora non sussista, per quel dato finanziamento una valutazione del merito creditizio effettuata da una *External Credit Assessment Institution* (c.d.

²²⁹ Basilea 3 ha lasciato invariata il requisito complessivo dell'8% ma ha previsto che esso vada soddisfatto per più della metà con il *Common equity tier 1*, composto dalla somma algebrica di azioni ordinarie emessa dalla banca, sovrapprezzo delle azioni, riserve di utili, riserve da valutazione e altre riserve palesi.

²³⁰ C. Brescia Morra, *Il diritto delle banche*, Il Mulino, Seconda Edizione, p. 212 ss.

ECAI) ossia una agenzia di rating del credito registrata o certificata conformemente al regolamento (CE) n. 1060/2009; le agenzie di rating svolgono una attività assimilabile a quella di “scoring” di società come la “Schufa” ma se ne differenziano per la maggiore analiticità della valutazione e per i soggetti che sono valutati: se il credit scoring si rivolge normalmente a consumatori, micro, piccole e medie imprese e, compie i propri giudizi in modo automatizzato usando un metodo statistico od il ML, il *credit rating* viene condotto, di norma, nei confronti di grandi società e stati sovrani e prevede un’analisi approfondita da parte di esperti del settore. Oltre a valutare i dati quantitativi, gli analisti esaminano anche dati qualitativi, quali l’affidabilità e le competenze del management aziendale, nonché la credibilità dei progetti e degli obiettivi prefissati dalla società oggetto di valutazione²³¹.

In alternativa al metodo standard, le banche possono chiedere all’autorità di vigilanza di essere autorizzate ad usare i loro rating interni.

Nel metodo IRB l’intermediario calcola la rischiosità di ciascun credito concesso tramite modelli interni, elaborati dallo stesso ed aggiornati periodicamente (cosicché la ponderazione del credito ed il correlato ammontare di fondi propri che deve detenere la banca per far fronte alla sua rischiosità può modificarsi con il passare del tempo)²³².

Il metodo IRB si basa su quattro componenti di rischio²³³:

- i. **Probability of Default (PD):** la probabilità che una controparte diventi insolvente entro un orizzonte temporale di un anno;

²³¹ World Bank, Credit scoring approaches guidelines, <https://thedocs.worldbank.org/en/doc/935891585869698451-0130022020/original/CREDITSCORINGAPPROACHESGUIDELINESFINALWEB.pdf> p. 3

²³² R. Gallo, The IRB approach and bank lending to firms, in Temi di discussione della Banca d’Italia n. 1347, 2021 <https://www.bancaditalia.it/pubblicazioni/temi-discussione/2021/2021-1347/index.html?com.dotmarketing.htmlpage.language=1> p. 8

²³³ R. Costi, L’ordinamento bancario, Il Mulino, Quinta Edizione 2012, pp. 592 ss.

- ii. **Loss Given Default (LGD):** la porzione di credito che la banca prevede di perdere in caso di insolvenza della controparte, al netto dei recuperi;
- iii. **Exposure at Default (EAD):** la stima dell'ammontare complessivo delle somme che la banca vanta nei confronti di un debitore al momento in cui diventa insolvente; si tratta di una previsione futura composta dalla esposizione per cassa (*drawn*) ossia la parte di credito già utilizzata dal cliente e dall'esposizione fuori bilancio (*undrawn*) ossia la parte di credito concessa ma non ancora utilizzata che il cliente potrebbe utilizzare prima del default
- iv. **Maturity, (M):** la media, per una data esposizione, delle durate residue contrattuali dei pagamenti, ciascuna ponderata per il relativo importo²³⁴.

A seconda di quali variabili la banca è autorizzata a calcolare autonomamente il metodo IRB si distingue fra base e avanzato: nel metodo di base, la banca calcola autonomamente la PD dei debitori mentre si avvale di parametri fissati dalle autorità per stimare le altre componenti del rischio; di converso in quello avanzato anche le altre variabili sono calcolate autonomamente dall'intermediario²³⁵.

Entrando maggiormente nel dettaglio del funzionamento dei modelli IRB va notato come il suo funzionamento sia sussumibile in tre elementi principali²³⁶:

- un punteggio di rischio continuo: ottenuto tramite una regressione logistica in cui si regredisce una variabile dummy²³⁷ (pari a uno se l'impresa va in default entro un orizzonte di un anno) su un numero

²³⁴ Banca d'Italia, *Recepimento della Nuova Regolamentazione Prudenziale Internazionale Metodo Dei Rating Interni per il Calcolo del Requisito Patrimoniale a Fronte del Rischio di Credito*, https://www.bancaditalia.it/compiti/vigilanza/normativa/consultazioni/2006/basilea2/Doc_Cons_IRB.pdf P. 5

²³⁵ F. Vella, G. Bosi, *Diritto ed Economia di Banche e Mercati Finanziari*, Il Mulino, 2023 Seconda Edizione p. 268

²³⁶ H. Fraisse, M. Laporte, *Return on Investment on AI: The Case of Capital Requirement*, Banque de France Working Paper n. 809

²³⁷ Variabile numerica costruita in modo da poter assumere solo valore 1 o 0 qualora una particolare caratteristica sia presente o meno.

limitato di variabili esplicative che presumibilmente determinano la salute finanziaria dell'impresa;

- una scala di rischio discreta²³⁸: le imprese vengono raggruppate in N classi di rating in base a vincoli di omogeneità del punteggio di rischio all'interno di ciascuna classe e di eterogeneità del punteggio di rischio tra le diverse classi;
- l'associazione della probabilità di default a ciascuna classe: tipicamente la PD rappresenta una media di lungo periodo del tasso di default delle esposizioni appartenenti a quella classe.

La realizzazione del modello e l'autorizzazione al suo utilizzo richiedono un processo lungo e complesso che può essere diviso in tre macro-fasi ciascuna scandita da vari momenti:

1. **Fase di sviluppo del modello:** è la fase in cui il modello viene ideato e strutturato dopo una pianificazione preliminare che, si focalizza sulla definizione degli obiettivi e delle modalità con cui raggiungerli, nonché sull'identificazione dei rischi anche connessi alla qualità e disponibilità dei dati precedendo e poi sull'analisi statistica. In questo stadio, il progetto viene configurato come una soluzione decisionale integrata si stabiliscono i ruoli dei partecipanti, si determina la strategia di implementazione più idonea e si procede poi ad attuarla²³⁹. Fondamentali in tale fase sono i momenti di:

- a) **Data preparation e feature selection:** la *data preparation* ha ad oggetto la raccolta, la scelta e la preparazione dei dati che saranno alla base della serie storica utilizzata per la stima. Conformemente alle previsioni di cui agli articoli 174, lett b) e 180 c.1. lett h della CRR i dati

²³⁸ Strumento qualitativo utilizzato per classificare il rischio in livelli distinti (ad es. basso, medio, alto) in base alla probabilità che avvenga un certo evento.

²³⁹ N. Siddiqi, *Intelligent Credit Scoring: Building and Implementing Better Credit Risk Scorecards*, Second Edition, Newark: John Wiley & Sons, Incorporated, 2017, pp. 55 ss

dovranno essere accurati, robusti, pertinenti e coprire un intero ciclo economico fermo restando l'obbligo di ricomprendere, in ogni caso, un orizzonte temporale inferiore a 5 anni per la PD e a 7 anni per la LGD e la EAD.

La *feature selection* si sostanzia nella selezione e nella trasformazione delle variabili al fine di ottenere un sottoinsieme ottimale di indicatori dal set originale di dimensioni ridotte al fine di diminuire la complessità computazionale dei modelli e migliorarne l'accuratezza riflettendo al contempo in maniera completa le informazioni del set originale²⁴⁰.

- b) Risk Differentiation:** è la capacità del modello di ordinare i debitori in base alla loro rischiosità, facendo corrispondere a classi di rating peggiori valori attesi di default o perdita progressivamente superiori²⁴¹. Scopo dei modelli di PD, come visto pocanzi, è quello di stimare un punteggio basato sui fattori di rischio ed in base a questo classificare i debitori o le esposizioni di credito in diversi segmenti o classi di rischio omogenee, garantendo l'uniformità del rischio di default all'interno di ciascuna classe e l'eterogeneità di questo rischio tra le diverse classi²⁴².

Per riuscire in tale obiettivo il modello deve essere in grado di riflettere le caratteristiche rilevanti del debitore e delle transazioni e di garantire una buona differenziazione del rischio. Al fine di verificare la capacità del modello di differenziare il rischio, la banca compara le performance

²⁴⁰ Y. Lu, L. Yang, B. Shi, et al., *A novel framework of credit risk feature selection for SMEs during industry 4.0 in Ann Oper Res* 350, 425–452 (2025). <https://doi.org/10.1007/s10479-022-04849-3> p. 428

²⁴¹ Regolamento Delegato (UE) 2022/439 della commissione del 20 ottobre 2021 che integra il regolamento (UE) n. 575/2013 del parlamento europeo e del consiglio per quanto riguarda le norme tecniche di regolamentazione che specificano la metodologia che l'autorità competente deve seguire nel valutare la conformità degli enti creditizi e delle imprese d'investimento ai requisiti relativi all'uso del metodo basato sui rating interni, art. 23 ss.

²⁴² C. Hurlin, C. Pérignon, *Machine Learning and IRB Capital Requirements: Advantages, Risks, and Recommendations*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4483793 p.9

misurate nel campione di stima del modello con quelle ottenute su portafogli più recenti²⁴³.

c) **Risk Quantification:** le probabilità di default sono i dati chiave per il calcolo dei requisiti patrimoniali; questa fase concerne l'assegnazione della probabilità di default a ciascuna classe di rischio omogenea. Le probabilità vengono generalmente stimate sulla base della frequenza empirica dei default osservata su un lungo periodo all'interno di ciascuna classe.

2. **Fase di autorizzazione:** affinché il modello sviluppato possa essere utilizzato dalla banca per stimare il proprio rischio di credito, deve essere sottoposto ad una complessa procedura autorizzatoria composta da due fasi, una svolta dall'intermediario e l'altra dall'autorità di vigilanza.

a) **Validazione interna:** in tale fase la banca valida l'accuratezza e la coerenza dei sistemi e dei processi di rating, nonché delle stime di tutte le componenti rilevanti di rischio; a tal fine vengono comparate l'incidenza effettiva delle inadempienze con le stime della PD per ciascuna classe di rischio mediante serie storiche quanto più possibile estese nel tempo così da dimostrare che tale incidenza rientri nell'intervallo atteso di valori per la classe in questione²⁴⁴.

Come visibile nella "Figura 1" la performance statistica del modello viene valutata in due fasi: nella prima fase si verifica che il modello identifichi correttamente il rischio in termini di capacità di classificazione, assegnando ad ogni debitore una classe di rischio

²⁴³ European Banking Authority, *Orientamenti sulla stima della probabilità di default (PD) e delle perdite in caso di default (LGD) e sul trattamento delle esposizioni in stato di default* (EBA-GL-2017-16), <https://www.eba.europa.eu/sites/default/files/documents/10180/2192133/f086e543-dad8-404e-83d8-45e49d3ccb59/Guidelines%20on%20PD%20and%20LGD%20estimation%20%28EBA-GL-2017-16%29-IT.pdf> par. 218

²⁴⁴ E. Bernardini, *La Validazione Qualitativa Dei Sistemi Di Rating*, in *Rivista bancaria*, giugno 2008, pp. 6 ss.

specifica; in secondo luogo, il modello viene adeguatamente calibrato, affinché i debitori siano classificati in modo che la probabilità di default stimata per ciascuna classe rappresenti in modo realistico il loro profilo di rischio²⁴⁵.

Nella validazione delle stime dei modelli statistici, l'approccio IRB prevede due principali meccanismi: *backtesting* e *benchmarking*; la validazione deve passare il *backtesting* mediante il quale si misura il tasso di errore dei modelli quando vengono utilizzati dati non inclusi nel campione.

Successivamente al *backtesting*, per ottenere un confronto più accurato, viene effettuato il *benchmarking* ossia il confronto delle stime con quelle ottenute da altre istituzioni comparabili e/o da fornitori esterni di dati (ad esempio, agenzie di rating) e con i risultati di un eventuale modello di riferimento predisposto dal supervisore²⁴⁶.

²⁴⁵ European Banking Authority, *Orientamenti sulla stima della probabilità di default (PD) e delle perdite in caso di default (LGD) e sul trattamento delle esposizioni in stato di default* (EBA-GL-2017-16), <https://www.eba.europa.eu/sites/default/files/documents/10180/2192133/f086e543-dad8-404e-83d8-45e49d3ccb59/Guidelines%20on%20PD%20and%20LGD%20estimation%20%28EBA-GL-2017-16%29-IT.pdf> p. 10

²⁴⁶ A. Alonso, J. M. Carbó, *Machine Learning in Credit Risk: Measuring The Dilemma Between Prediction and Supervisory Cost*, <https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosTrabajo/20/Files/dt2032e.pdf> p.27

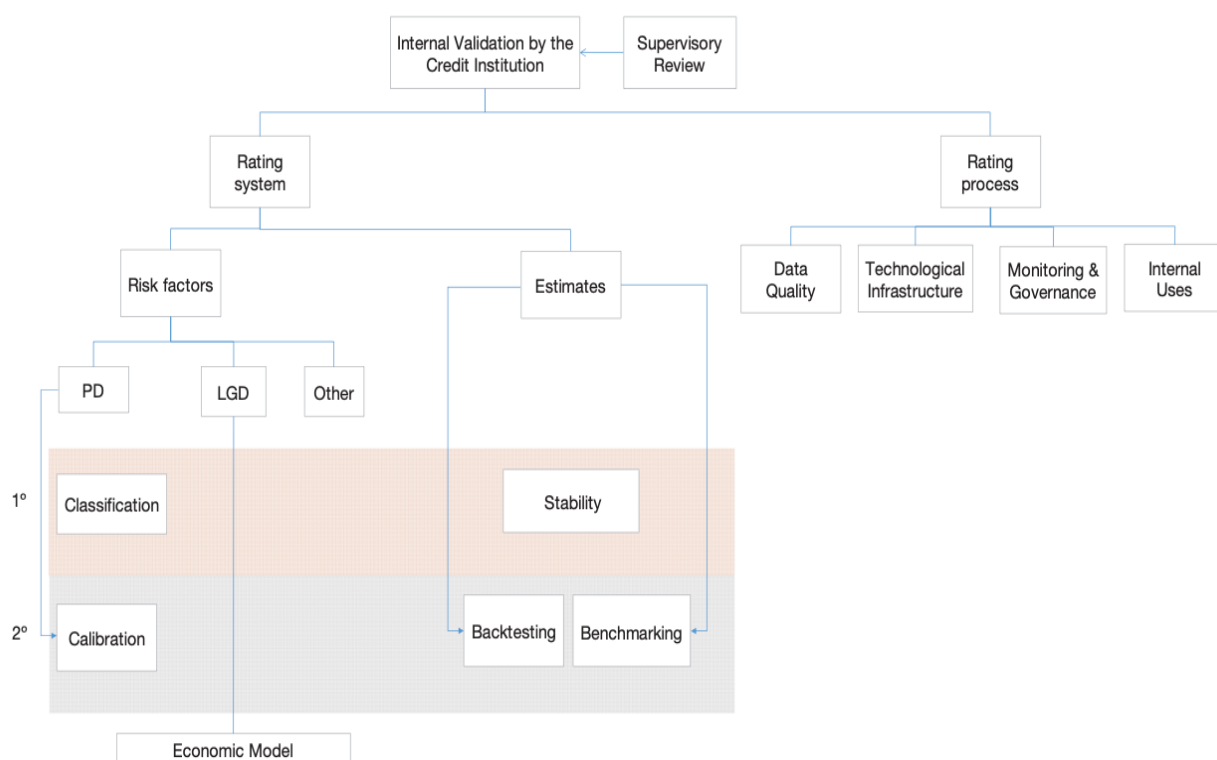


Figura 1. Schema del processo di validazione di un modello IRB

Fonte: A. Alonso, J. M. Carbó, Machine Learning in Credit Risk

b) **Approvazione regolamentare:** l'autorità di vigilanza competente²⁴⁷, al fine di autorizzare all'uso del modello, revisiona il processo di validazione interna compiuto dalla banca e verifica che questa si sia conformata ai requisiti di carattere qualitativo e quantitativo nella creazione e gestione dei modelli interni, stabiliti dalla normativa prudenziale.

Tra gli aspetti **qualitativi** è ricompreso il rispetto di previsioni in materia di documentazione, conservazione e gestione dei dati nonché

²⁴⁷ Conformemente alle previsioni del Meccanismo di Vigilanza Unico di cui al Regolamento UE 1024/2013, la competenza viene ripartita, a seconda che la banca sia o meno significativa, tra Banca Centrale Europea e l'Autorità Nazionale Competente (nel nostro paese Banca d'Italia)

sulla presenza di un appropriato assetto di governance e controlli interni:

- i. **Documentazione e Data Governance** (art. 175-176 CRR): l'impiego di modelli statistici deve essere supportato da adeguata documentazione che illustri la teoria, le ipotesi, le basi matematiche e le fonti dei dati utilizzati per costruire il modello, nonché il procedimento statistico di validazione. La documentazione permette, a terze parti qualificate, di comprendere in modo indipendente la metodologia, le ipotesi, le limitazioni e l'utilizzo del modello, nonché di replicarne lo sviluppo e l'implementazione. La banca è tenuta a dimostrare come la propria documentazione e il registro dei modelli interni facilitino la comprensione dei modelli sia internamente che esternamente²⁴⁸.

Tra le informazioni che le banche devono conservare figurano quelle sui dettagli del debitore, sull'operazione finanziata, il rating assegnato, la metodologia e i parametri impiegati per assegnare quel dato rating, nonché informazioni generali su tutte le inadempienze (generalità dei debitori e delle operazioni, i tempi, le circostanze, PD, LGD, EAD, i costi e gli importi dei recuperi) anche al fine di determinare i tassi effettivi di insolvenza per classe e di consentire una riclassificazione dei rating assegnati (ad esempio, in caso di modifiche al modello)²⁴⁹.

Nel gestire e conservare i dati la banca deve adottare solide pratiche di governance che ricomprendano adeguate misure organizzative per la gestione, il monitoraggio della sicurezza dei dati e la reazione tempestiva ad eventuali violazioni.

²⁴⁸ European Central Bank, *Guide to internal models*, https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.supervisory_guide202507.en.pdf p. 9

²⁴⁹ E. Bernardini, La Validazione Qualitativa Dei Sistemi Di Rating, in *Rivista bancaria*, giugno 2008, pp 6 ss.

ii. **Appropriato assetto di governance e controlli interni:** le banche devono definire chiaramente i ruoli e le responsabilità in materia IRB nell'ambito dei loro modelli di corporate governance²⁵⁰. Conformemente alle previsioni di cui all'art. 189 CRR gli aspetti sostanziali del processo di stima devono essere approvati dall'organo di gestione e dall'alta dirigenza della banca, la quale ha il compito di assicurare la corretta operatività dei sistemi di rating ed è tenuta ad informare l'organo di gestione quando si verificano cambiamenti che possono influire in modo sostanziale sui sistemi dell'ente.

Per ridurre il rischio di perdite e sottostime dei requisiti di fondi propri, che possono venire causati da difetti nello sviluppo, nell'implementazione o nell'utilizzo dei modelli, le banche devono avere un quadro di gestione del rischio da modello che permetta di identificare, comprendere e gestire i rischi dei modelli interni a livello di gruppo; in tale quadro devono essere incluse linee guida per identificare e mitigare le incertezze di misurazione e le carenze dei modelli; vanno enucleate le procedure per la comunicazione e la rendicontazione del rischio di modello; definiti i ruoli e le responsabilità all'interno del quadro di gestione del rischio di modello nonché esplicitati i livelli appropriati di competenze, capacità ed esperienza richiesti per ciascun gruppo di stakeholder, proporzionati al rischio di modello e ai compiti di ciascun portatore di interesse²⁵¹.

Nell'ambito del sistema dei controlli, inoltre deve essere prevista una unità di controllo del rischio²⁵², indipendente dal personale responsabile della concessione dei fidi, che ha il compito di effettuare

²⁵⁰ European Central Bank, *Guide to internal models*, https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.supervisory_guide202507.en.pdf p. 12

²⁵¹ Ivi, pp. 10 ss.

²⁵² L'unità di controllo partecipa anche alla fase di validazione dei modelli.

la verifica e il monitoraggio dei sistemi di rating, di accertare che i criteri di valutazione siano omogeneamente applicati.

Infine, si deve prevedere una funzione di *audit interno* che svolgendo, conformemente a quanto previsto dalle linee guida EBA in materia di governance interna²⁵³, le revisioni in maniera efficiente ed effettiva²⁵⁴, annualmente riveda il modello e il suo funzionamento, al fine di accertarne la conformità con i requisiti minimi, documentando i risultati degli accertamenti²⁵⁵.

Tra gli aspetti **quantitativi** dello sviluppo dei modelli (concernenti la determinazione delle procedure volte alla discriminazione del rischio tra le diverse controparti) sono ricompresi²⁵⁶:

- i. **Qualità e rappresentatività dei dati di input dei modelli:** dalla qualità e dall'affidabilità dei dati dipende ampiamente il sistema di misurazione²⁵⁷; in ragione di ciò le previsioni di cui all'art. 174 lett. b) e c) della CRR richiedono che il modello sia costruito utilizzando dati pertinenti, accurati, completi e rappresentativi dell'effettiva popolazione dei debitori e delle esposizioni dell'ente.

Per realizzare tale obiettivo è necessario in primo luogo “pulire” i dati storici togliendo anomalie, stagionalità o segmenti di clientela non più

²⁵³ European Banking Authority, Orientamenti in materia di governance interna ai sensi della direttiva (UE) 2019/2034 (EBA/GL/2021/14), https://www.eba.europa.eu/sites/default/files/document_library/Publications/Guidelines/2021/EBA-GL-2021-14%20Guidelines%20on%20internal%20governance%20under%20IFD/translations/1028039/GL%20on%20internal%20governance%20under%20IFD_IT_COR.pdf

²⁵⁴ European Central Bank, *Guide to internal models*, https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.supervisory_guide202507.en.pdf p. 14

²⁵⁵ E. Bernardini, La Validazione Qualitativa Dei Sistemi Di Rating, in *Rivista bancaria*, giugno 2008, pp 6 ss.

²⁵⁶ E. Bonaccorsi di Patti, F. Calabresi, B. De Varti, F. Federico, M. Affinito, M. Antolini, F. Lorzio, S. Marchetti, I. Masiani, M. Moscatelli, F. Privitera e G. Rinna, *Intelligenza artificiale nel credit scoring. Analisi di alcune esperienze nel sistema finanziario italiano*, https://www.bancaditalia.it/pubblicazioni/qef/2022-0721/QEF_721_IT.pdf pp. 22 ss.

²⁵⁷ Comitato di Basilea per la vigilanza bancaria, *Convergenza internazionale della misurazione del capitale e dei coefficienti patrimoniali*, 2004, <https://www.bis.org/publ/bcbs107ita.pdf> p. 141

pertinenti attraverso rigorose procedure di esclusione; la solidità dei parametri di rischio dipende non solo dalla "pulizia" del dato, ma anche dalla capacità del database di garantire stabilità e significatività statistica nel tempo, evitando che la capacità predittiva del modello venga compromessa da bias informativi o da finestre temporali di osservazione incongrue²⁵⁸.

- ii. **Appropriatezza delle determinanti del rischio:** conformemente alle previsioni di cui all'art. 179 c.1 lett. a), i valori dei parametri di rischio devono essere plausibili, fondarsi su variabili rilevanti; inoltre, le banche devono verificare che i parametri così ottenuti siano in linea con il loro significato economico²⁵⁹.
- iii. **Verifica della filosofia di rating:** la banca può adottare due differenti approcci/filosofie nel calcolo delle stime di rischio il "*point in time*" ("PIT") e il *through the cycle* ("TTC"); il primo produce una probabilità di default (PD) del debitore sensibile alle variazioni macroeconomiche di breve periodo; l'approccio TTC invece, esclude i fattori ciclici dalla PD producendo stime del rischio di credito meno volatili, che permettono di cogliere meglio la componente di medio-lungo periodo del merito di credito dei soggetti debitori a discapito di quella di breve periodo²⁶⁰. Scelto l'approccio per ciascun modello che utilizzerà la

²⁵⁸ European Banking Authority, *Orientamenti sulla stima della probabilità di default (PD) e delle perdite in caso di default (LGD) e sul trattamento delle esposizioni in stato di default* (EBA-GL-2017-16), <https://www.eba.europa.eu/sites/default/files/documents/10180/2192133/f086e543-dad8-404e-83d8-45e49d3ccb59/Guidelines%20on%20PD%20and%20LGD%20estimation%20%28EBA-GL-2017-16%29-IT.pdf> par. 36-40; 81-99

²⁵⁹ E. Bonaccorsi di Patti, F. Calabresi, B. De Varti, F. Federico, M. Affinito, M. Antolini, F. Lorzio, S. Marchetti, I. Masiani, M. Moscatelli, F. Privitera e G. Rinna, *Intelligenza artificiale nel credit scoring. Analisi di alcune esperienze nel sistema finanziario italiano*, https://www.bancaditalia.it/pubblicazioni/qef/2022-0721/QEF_721-IT.pdf p. 22 ss.

²⁶⁰ AIFIRM, *Il principio contabile IFRS 9 in banca: la prospettiva del Risk Manager*, 2016 <https://www.aifirm.it/wp-content/uploads/2016/12/2016-Position-Paper-8-Il-principio-contabile-IFRS-9-in-banca-la-prospettiva-del-Risk-Manager.pdf> pp 29 ss.

banca, deve valutare che i modelli statistici utilizzati consentano di rispettare la filosofia individuata²⁶¹.

3) Fase di monitoraggio e revisione periodica: una volta che il modello statistico è stato approvato le stime ed i risultati vengono usati per il calcolo del capitale regolamentare. La banca è tenuta a monitorare il modello periodicamente per verificare se sia ancora solido e non vi sia stato un deterioramento delle sue performance e del suo potere discriminatorio²⁶². Conformemente alle previsioni di cui all'art 173 e 174 della CRR con cadenza almeno annuale deve essere effettuata la revisione del rating, da parte di un'unità indipendente nonché ogni qualvolta si verifichino eventi significativi quali:

i) venga svolta una valutazione dell'indipendenza della funzione di convalida; ii) vi siano cambiamenti nel contesto giuridico; (iii) la funzione di audit interno effettui un controllo; (iv) avvenga una revisione da parte dell'autorità competente²⁶³.

L'utilizzo dei metodi IRB, rispetto al metodo standard consente una maggiore flessibilità nello stimare i principali parametri di rischio e permette ai requisiti patrimoniali di essere maggiormente allineati ai profili effettivi di rischio delle esposizioni con conseguente ottimizzazione dell'allocazione del capitale.

Le banche con solidi quadri di gestione del rischio e portafogli di alta qualità usando il metodo IRB possono detenere minori fondi propri dal momento che la diminuzione della probabilità di insolvenza del debitore (PD), della vita residua del debito (M), dell'esposizione al momento di insolvenza (EAD) e della perdita

²⁶¹ European Banking Authority, *Orientamenti sulla stima della probabilità di default (PD) e delle perdite in caso di default (LGD) e sul trattamento delle esposizioni in stato di default* (EBA-GL-2017-16), <https://www.eba.europa.eu/sites/default/files/documents/10180/2192133/f086e543-dad8-404e-83d8-45e49d3ccb59/Guidelines%20on%20PD%20and%20LGD%20estimation%20%28EBA-GL-2017-16%29-IT.pdf> par. 66-68

²⁶² Ivi, parr. 218 ss.

²⁶³ Ivi, par. 215 ss.

attesa nel caso di insolvenza (LGD) comporta la diminuzione dei fondi propri che la banca deve accantonare e viceversa²⁶⁴.

Ma tale autonomia e benefici portano con sé oneri regolamentari più stringenti, per gli istituti di credito che devono dimostrare la solidità e l'accuratezza dei propri modelli interni alla autorità di vigilanza e la loro capacità di assicurare rigorosi sistemi di governo, gestione e controllo dei rischi e dei dati²⁶⁵ attuando valutazioni periodiche della complessità dei modelli e processi di validazione.

2. Modelli IRB basati sul Machine Learning.

L'ammontare dei fondi propri che devono essere detenuti è uno degli elementi che influenzano il costo del finanziamento assieme ad altri fattori quali il *funding* (costo che la banca paga per reperire le risorse con cui erogare i finanziamenti) i costi interni di gestione, valutazioni di carattere commerciale ed i fattori dipendenti dalle componenti di rischio come il costo della perdita attesa (EL)²⁶⁶ ossia la perdita media che si manifesta entro un intervallo temporale di un anno su ogni esposizione esistente in portafoglio e che viene ricavata moltiplicando la PD per la LGD.

Viene da chiedersi se applicando il *machine learning* (ML) ai modelli IRB si possano ottenere dal punto di vista del *rating* benefici predittivi simili a quelli riscontrati per lo *scoring* con miglioramenti del potere discriminante dei modelli e nella precisione del calcolo della PD e della LGD e conseguente ottimizzazione

²⁶⁴ D. Cucinelli, M. L. Di Battista, M. Marchese, L. Nieri, *Credit Risk in European Banks: The Bright Side of the Internal Ratings Based Approach*, in *Journal of Banking & Finance*, 93, 2018, pp. 213-229

²⁶⁵ Banca d'Italia, Raccomandazione in materia di modelli IRB ai fini della misurazione del rischio di credito, <https://www.bancaditalia.it/compiti/vigilanza/normativa/orientamenti-vigilanza/Raccomandazione-IRB-misurazione-rischio-di-credito.pdf>

²⁶⁶ Associazione Bancaria Italiana, Come viene valutata l'affidabilità delle imprese con l'Accordo di Basilea, https://bancaria.it/pozzo/Conoscere_il_rating.pdf pp 7-8

dell'attività di ponderazione per il rischio idonea a riflettersi sia sulla quantità che sul costo del credito.

A supporto di tale tesi vi sono su due assunti: in primo luogo i modelli IRB di PD e LGD basati su tecniche di ML sono in grado di sovraperformare i modelli statistici tradizionali (ad esempio, il ML può prevedere meglio i default su alcune sottopopolazioni del portafoglio di credito, cogliendo segnali deboli che i modelli standard non riescono a rilevare); in secondo luogo i miglioramenti in accuratezza sono idonei a tradursi in vantaggi in termini di capitale regolamentare²⁶⁷.

Il tema è stato attenzionato dalla European Banking Authority che ha dedicato un *discussion paper* all'uso in tale ambito²⁶⁸ ed un successivo *follow-up report*²⁶⁹ che ha raccolto le opinioni degli *stakeholder*.

È stato osservato come, nell'ambito dei modelli IRB il ML possa essere utilizzato per stimare diversi parametri di rischio²⁷⁰, tra cui la PD e la LGD, in diverse fasi della loro stima: preparazione dei dati, selezione delle variabili, differenziazione e quantificazione del rischio, validazione del modello²⁷¹.

Focalizzando sulla PD va notato come il *machine learning* possa applicarsi in varie fasi:

1. **Fase di data preparation e feature engineering**²⁷²: gli algoritmi di machine learning possono essere utilizzati a monte della modellazione, nelle fasi di

²⁶⁷ C. Hurlin, C. Pérignon, Machine Learning and IRB Capital Requirements: Advantages, Risks, and Recommendations, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4483793 p.21.

Vantaggi dal punto di vista macrofinanziario sono stati altresì evidenziati dal Financial Stability Board in *Artificial intelligence and machine learning in financial services, market developments and financial stability implications*, <https://www.fsb.org/uploads/P011117.pdf>

²⁶⁸ European Banking Authority, *Discussion Paper on Machine Learning for IRB Models* (EBA/DP/2021/04),

²⁶⁹ European Banking Authority, *Machine Learning for IRB Models Follow-Up Report from the Consultation on the Discussion Paper on Machine Learning for IRB Models* (EBA/REP/2023/28)

²⁷⁰ Ivi, pp 6 ss

²⁷¹ C. Hurlin, C. Pérignon, Machine Learning and IRB Capital Requirements: Advantages, Risks, and Recommendations, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4483793 p.4

²⁷² Fase di preelaborazione che trasforma i dati grezzi in un insieme di input più efficaci. Ogni input è costituito da diversi attributi noti come caratteristiche (feature); fornendo ai modelli informazioni pertinenti, il feature engineering ne migliora sensibilmente la precisione predittiva e la capacità di prendere decisioni.

preprocessing per la preparazione dei dati e la predisposizione degli input. Tali attività possono aumentare la robustezza del modello (ad es. gestendo valori mancanti e anomalie), facilitare l'interpretazione dei risultati (discretizzando²⁷³ le variabili continue), eliminare le informazioni ridondanti, preselezionare le variabili candidate per la modellazione (analizzando la correlazione con la variabile target) ed incrementare le prestazioni predittive del modello (ottimizzando i metodi di binning²⁷⁴ per discretizzare le caratteristiche continue)²⁷⁵.

Ma non solo, i modelli di ML potrebbero consentire alle banche di utilizzare dati non strutturati (ad esempio dati qualitativi come i rapporti aziendali), ampliando i data set utilizzabili per la stima dei parametri²⁷⁶.

2. **Fase di risk differentiation:** il *machine learning* potrebbe aumentare il potere discriminante del modello e fornire strumenti utili per identificare i fattori di rischio rilevanti e le relazioni tra di essi; i modelli di ML possono essere utilizzati per ottimizzare la segmentazione del portafoglio, costruire modelli solidi nonché per prendere decisioni guidate dai dati che bilancino la disponibilità dei dati con la granularità richiesta dal modello. Inoltre, essi potrebbero aiutare a confermare le *data features* selezionate dagli sviluppatori nei modelli statistici tradizionali, offrendo una prospettiva basata sui dati al processo di selezione delle caratteristiche²⁷⁷.
3. **Fase di risk quantification:** il ML può migliorare la capacità predittiva del modello di prevedere la probabilità di default nonché supportare il calcolo

²⁷³ Processo matematico che trasforma una variabile continua (che può assumere infiniti valori) in una variabile discreta (costituita da un numero finito di categorie o intervalli).

²⁷⁴ Tecnica di elaborazione dei dati che consiste nell'aggregare valori continui in un numero ridotto di intervalli, c.d. "bin" (*cestini*) al fine di raggruppare una grande quantità di dati in categorie generali per semplificarne l'analisi.

²⁷⁵ C. Hurlin, C. Pérignon, Machine Learning and IRB Capital Requirements: Advantages, Risks, and Recommendations, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4483793 P.9 ss

²⁷⁶ European Banking Authority, *Discussion Paper on Machine Learning for IRB Models* (EBA/DP/2021/04), p. 21

²⁷⁷ European Banking Authority, *Discussion Paper on Machine Learning for IRB Models* (EBA/DP/2021/04) p. 20

degli adeguamenti necessari alle stime di rischio (di cui agli artt. 180 c.1 lett d e 181, c. 1 lett b) per renderle più affidabili anche ricorrendo all'analisi testuale di documenti quali i bilanci²⁷⁸.

4. **Fase di model validation:** uno degli utilizzi più comuni del ML, in ambito IRB è stato riscontrato nella fase di validazione dei modelli laddove tali tecniche sono impiegate per sviluppare “*modelli challenger*” per testare il “modello ufficiale” utilizzato per il calcolo interno dei requisiti patrimoniali²⁷⁹. L'ampio utilizzo di tale tecnica è dovuto al fatto che i modelli challenger, non essendo utilizzati direttamente per il calcolo dei requisiti patrimoniali regolamentari, non soggiacciono ai vincoli imposti dal CRR.

Sebbene il ML possa aiutare a stimare i parametri di rischio in modo più preciso, l'aumento della capacità predittiva porta con sé una maggiore complessità del modello, la quale può ostacolare una chiara illustrazione della teoria, delle ipotesi e della base matematica dell'assegnazione finale delle stime a classi, singoli debitori, esposizioni o gruppi²⁸⁰.

Il *trade-off* fra potere predittivo e interpretabilità è un elemento critico specie considerando che i modelli IRB incidono sulla stabilità patrimoniale delle banche che li utilizzano ed erranee modalità di calcolo potrebbero causare una situazione di *distress*.

In considerazione di ciò, vi è in dottrina, chi ritiene che il machine learning non dovrebbe essere utilizzato per il calcolo dei requisiti patrimoniali poiché sarebbe improbabile che gli algoritmi di apprendimento automatico possano migliorare la predittività al punto da giustificare la loro inclusione nei modelli IRB senza

²⁷⁸ Ibid.

²⁷⁹ Institute of International Finance, Machine learning in credit risk, https://www.iif.com/Portals/0/Files/content/Research/iif_mlcr_2nd_8_15_19.pdf

²⁸⁰ European Banking Authority, *Discussion Paper on Machine Learning for IRB Models* (EBA/DP/2021/04) pp. 13-14

contare che la loro intrinseca complessità e imperscrutabilità può essere foriera di nuove forme di bias e di arbitraggio regolamentare nella normativa prudenziale; addestrare gli algoritmi di ML non è infatti un processo passivo: la selezione dell'algoritmo, dei dati, nonché le fasi di etichettatura, calibrazione (tuning) e test, costituiscono ambiti in cui l'intervento umano può influenzare e manipolare il funzionamento del sistema.

Nonostante l'apporto umano, i risultati generati dai computer vengono spesso percepiti come dotati di una legittimità intrinseca cosicché le autorità di vigilanza potrebbero finire per rimettersi acriticamente a essi (c.d. *automation bias*) e le banche potrebbero quindi sfruttare tale distorsione selezionando deliberatamente le forme di algoritmi più complesse e meno spiegabili e con la loro opacità celare ai supervisor le numerose opportunità che l'istituto ha, durante le fasi di selezione e test, di accentuare la tendenza degli algoritmi a ignorare i c.d. *tail events*²⁸¹ e a sottostimare la ponderazione del rischio, con l'obiettivo ultimo di minimizzare i requisiti patrimoniali obbligatori²⁸².

Ci sentiamo tuttavia di dissentire da tale approccio radicale! Come hanno dimostrato alcuni studi²⁸³ i benefici derivanti dall'integrazione del ML nei modelli IRB assumono valori significativi basti pensare che questi non solo possono offrire un miglioramento della *risk differentiation* anche del 20%²⁸⁴ ma anche ridurre il valore della ponderazione media delle attività (RWA Density) in

²⁸¹ Eventi rari e imprevedibili che si collocano nelle estremità di una distribuzione statistica (c.d. "code"); seppure tali eventi abbiano una probabilità molto bassa di verificarsi il loro impatto può essere estremamente significativo quando si manifestano.

²⁸² H. J. Allen, 'Just Say No' to Machine Learning in IRB Models, in Oxford Business Law Blog, 2021

²⁸³ - P. Di Biasi, R. Gnutti, A. Resti, D. Vergari, *Machine Learning for Credit risk: three successful Case Histories*, in *Risk Management Magazine*, Vol. 17, Issue 2, 2022 pp. 7 ss

- M. Wallis, K. Kumar, A. Gepp, *Credit Rating Forecasting Using Machine Learning Techniques*, in Z. Sun (Ed.), *Managerial Perspectives on Intelligent Big Data Analytics* (pp. 180-198). IGI Global (2019)

²⁸⁴ A. Alonso, J. M. Carbó, *Machine Learning in Credit Risk: Measuring the Dilemma Between Prediction and Supervisory Cost*, <https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosTrabajo/20/Files/dt2032e.pdf> p.9

una percentuale variabile tra i 2 ed i 27 punti percentuali rispetto ai modelli Logit²⁸⁵.

Per quanto concerne i rischi derivanti dai modelli essi sono mitigabili dall'applicazione della legislazione in materia.

3. Il quadro regolatorio dei modelli IRB-ML

Nell'esaminare le fonti che disciplinano la materia è necessario partire da un assunto: gli strumenti di intelligenza artificiale applicati ai modelli IRB, non rientrano nell'ambito dei sistemi ad alto rischio di cui agli allegati I e III del Regolamento Ue 2024/1689²⁸⁶!

Essi, infatti, né vengono menzionati espressamente dall'AI Act come modelli ad alto rischio né possono essere assimilati ai sistemi di credit scoring finalizzati a valutare «l'affidabilità creditizia delle persone fisiche o [a] stabilire il loro merito di credito» di cui all'allegato III, punto 5, lett. b.: se la finalità dell'AI Act è, come espresso nel considerando 1 del Regolamento, quella di garantire che l'intelligenza artificiale venga utilizzata rispettando la salute, la sicurezza e i diritti fondamentali dei cittadini (tra cui rientra anche il diritto di non essere esclusi dall'accesso al credito) ben diversa è la finalità della normativa in materia di modelli IRB che si sostanzia nel garantire la sana e prudente gestione degli enti creditizi. È dunque con tale ottica che gli strumenti per calcolare e verificare i coefficienti patrimoniali delle banche vanno analizzati non essendo questi idonei ad incidere in maniera diretta sui diritti fondamentali delle persone bensì sulla stabilità finanziaria dell'intermediario²⁸⁷.

²⁸⁵ H. Fraisse, M. Laporte, *Return on Investment on AI: The Case of Capital Requirement*, in Banque de France Working Paper n. 809

²⁸⁶ In tal senso si è espressa anche l'EBA in EBA/REP/2023/28 par. 48

²⁸⁷ Banca Centrale Europea, *Parere della Banca Centrale Europea del 29 dicembre 2021 relativo a una proposta di regolamento che stabilisce regole armonizzate sull'intelligenza artificiale*

Se dunque i modelli IRB sono soggetti solo alla disciplina prudenziale delle banche a ben vedere la normativa settoriale contiene norme per molti aspetti sovrapponibili a quelle dell'AI Act, il quale detta norme che non sono antitetiche ma complementari rispetto alla normativa bancaria²⁸⁸.

A riprova di ciò si possono annoverare alcune previsioni della CRR:

- l'art. 171, comma 1, lett. a) e b): il quale richiede che i criteri di assegnazione delle esposizioni alle classi di rating siano sufficientemente trasparenti da consentire a terzi di comprendere, replicare e valutare l'adeguatezza del processo di classificazione.

La disposizione in oggetto presenta forti analogie con l'art. 13 dell'AI Act, ai sensi del quale i sistemi di IA ad alto rischio devono essere progettati e sviluppati in modo che il loro funzionamento sia *«sufficientemente trasparente da consentire ai deployer di interpretare l'output del sistema e utilizzarlo adeguatamente»*;

- l'art. 175: il quale impone alla banca che utilizzi un modello statistico nel processo di assegnazione dei rating, di documentare le ipotesi di modellizzazione e la teoria alla base del modello al fine di comprovare l'osservanza dei requisiti del regolamento. Evidente è qui il parallelismo con l'art. 11 del Regolamento sull'IA che impone ai *developers* la redazione di una documentazione tecnica che deve essere tenuta a disposizione delle autorità e periodicamente aggiornata, al fine di fornire *«in forma chiara e comprensibile, le informazioni necessarie per valutare la conformità del sistema di IA»* ai requisiti del regolamento;
- l'art. 174, lett. e) il quale stabilisce che il modello statistico vada combinato *«con la valutazione e la revisione umana in modo da verificare le assegnazioni*

²⁸⁸ European Banking Authority, *AI Act: implications for the EU banking and payments sector*, <https://www.eba.europa.eu/sites/default/files/2025-11/d8b999ce-a1d9-4964-9606-971bbc2aaf89/AI%20Act%20implications%20for%20the%20EU%20banking%20sector.pdf>

effettuate in base al modello e da assicurare che i modelli siano utilizzati in modo appropriato». Non si può non riscontrare una somiglianza con l'art. 14 dell'AI Act rubricato *“sorveglianza umana”* il quale dispone che i sistemi ad alto rischio devono essere progettati con modalità che permettano una efficace supervisione da parte di persone fisiche che devono essere in grado di *“comprendere correttamente le capacità e i limiti pertinenti del sistema”*, *“monitorarne debitamente il funzionamento”* e se necessario di *“intervenire sul funzionamento”* del sistema; essi inoltre devono avere la possibilità di decidere di non usare il sistema o di ignorarne, annullarne o ribaltarne l'output generato,

- l'art. 179: il quale richiede che le stime dei parametri di rischio, da associare alle classi di rating siano plausibili e intuitive. Tali requisiti a ben vedere si ritrovano anche nell'AI Act laddove il concetto di *“intuitività”* è assimilabile al requisito di trasparenza dell'Articolo 13, che mira a far sì che l'utente possa *«interpretare l'output del sistema»* mentre la *“plausibilità”* è assimilabile alla previsione per cui i sistemi devono avere prestazioni coerenti durante tutto il ciclo di vita del sistema e garantire un livello adeguato di *“accuratezza”* e *“robustezza”*, ex art. 15 dell'AI Act.

Ulteriori similitudini sono riscontrabili guardando alla disciplina prudenziale di livello secondario ed in particolare sugli *“Orientamenti in materia di concessione e monitoraggio dei prestiti”* dell'EBA, recepiti dalla Banca d'Italia con Nota n. 13 del 20 luglio 2021²⁸⁹. Tali orientamenti prevedono una serie di requisiti che i modelli *“automatizzati”* delle banche per la valutazione del

²⁸⁹ Banca d'Italia, *Attuazione degli Orientamenti dell'Autorità bancaria europea in materia di concessione e monitoraggio dei prestiti* (EBA/GL/2020/06), <https://www.bancaditalia.it/compiti/vigilanza/normativa/orientamenti-vigilanza/elenco-esa/note/Nota-n-13-del-20-luglio-2021.pdf>

merito di credito debbono avere, affinché possano essere validati dall'autorità di vigilanza, tra i quali figura la necessità di²⁹⁰:

- i) comprendere adeguatamente i modelli utilizzati con particolare riguardo alla loro metodologia, ai dati inseriti, alle ipotesi, ai limiti e ai risultati;
- ii) adottare politiche e procedure interne finalizzate a individuare e prevenire le distorsioni e garantire la qualità dei dati inseriti;
- iii) valutare l'adeguatezza (in termini di tracciabilità, verificabilità, robustezza e resilienza) degli *input* e degli *output* dei modelli;
- iv) dotarsi di policy che assicurino la verifica regolare della qualità dei risultati dei modelli;
- v) prevedere meccanismi di controllo del risultato prodotto che integrino il giudizio di esperti in materia;
- vi) disporre di un'adeguata documentazione relativa alla progettazione e all'utilizzo dei modelli.

Ebbene, oltre alla somiglianza con gli articoli 11, 13, 14 e 15, pocanzi visti, si trovano similitudini in materia di *data governance*, con l'art. 10 dell'AI Act, il quale richiede che i dati di addestramento, convalida e prova siano «*pertinenti, sufficientemente rappresentativi e, nella misura del possibile, esenti da errori e completi*» e che vadano esaminati per valutare eventuali distorsioni che renderebbero le stime non plausibili.

Ma non solo, somiglianze vi sono anche tra la verifica regolare della qualità dei modelli (iv) e gli articoli 26 e 72 dall'AI Act laddove il primo impone ai *deployer* di «*monitorare il funzionamento del sistema di IA ad alto rischio sulla base delle*

²⁹⁰ European Banking Authority, Orientamenti in Materia di Concessione e Monitoraggio dei Prestiti (EBA-GL-2020-06)

https://www.eba.europa.eu/sites/default/files/document_library/Publications/Guidelines/2020/Guidelines%20on%20loan%20origination%20and%20monitoring/Translations/886685/Final%20Report%20on%20GL%20on%20loan%20origination%20and%20monitoring_COR_IT.pdf parr. 53, 54 e 55.

istruzioni per l'uso» ed il secondo ai fornitori di istituire un sistema per raccogliere e analizzare attivamente i dati sulle prestazioni del sistema di IA «*per tutta la durata del loro ciclo di vita*» così da valutare la costante conformità e qualità dei risultati nel tempo.

Da ultimo, l'avvicinamento tra le normative è avvenuto ad opera della Banca Centrale Europea che ha, nel luglio del 2025, aggiornato la sua "*Guide to internal models*²⁹¹" ossia il documento nel quale esplicita le modalità con cui intende applicare le disposizioni prudenziali in materia di modelli interni nei confronti dei soggetti sottoposti alla sua vigilanza.

Nel nuovo aggiornamento della guida è ora presente un nuovo capitolo, il nove, dedicato all'uso delle tecniche di machine learning nei modelli interni e che elenca principi, per l'utilizzo di tali strumenti, conformi alle previsioni dell'AI Act:

- i. **Governance²⁹² e competenze tecniche (parr. 35-37-68):** le banche devono integrare i rischi del ML nei loro modelli e processi di governance; le loro politiche interne dovrebbero definire in maniera chiara l'ambito e lo scopo dell'uso dei modelli IRB basati su ML, nei loro processi di gestione del rischio e di decisione, nonché nelle funzioni di approvazione del credito, allocazione interna del capitale e governance aziendale. Tale definizione dovrebbe contenere la precisa delimitazione delle aree e dei processi in cui i modelli saranno applicati, descrivere i compiti e delle funzioni specifiche per cui i modelli sono progettati (es. analisi predittiva, riconoscimento di schemi, ottimizzazioni) e delle limitazioni di tali modelli ivi inclusi i

²⁹¹ European Central Bank, *Guide to internal models*, https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.supervisory_guide202507.en.pdf

²⁹² Parallelismo in tema di governance si ritrova con l'art. 9 dell'AI Act, il quale impone l'istituzione di un "sistema di gestione dei rischi" pensato per essere un processo iterativo continuo durante l'intero ciclo di vita del sistema, integrato nei processi decisionali.

potenziali bias, i livelli di accuratezza, le dipendenze dai dati e gli scenari in cui potrebbe essere necessario l'intervento umano.

Viene altresì richiesto a tutti i principali *stakeholder* che interagiscono con i modelli di ML di possedere competenze e conoscenze specifiche sulle tecniche di ML utilizzate tali da permettere loro di svolgere consapevolmente i loro compiti; in ragione di ciò le banche dovrebbero assicurare che il personale addetto allo sviluppo, al controllo ed alla validazione dei modelli abbia competenze adeguate a comprendere compiutamente gli esiti delle proprie azioni, per valutare la pertinenza e l'appropriatezza dei driver di rischio utilizzati così come la solidità della logica economica sottostante dell'intero modello. Ma non solo, gli intermediari devono garantire che l'organo di gestione e il senior management della banca siano messi nelle condizioni di comprendere il modello fornendo loro una documentazione adeguata che gli permetta di capire quali indicatori o variabili sono i principali driver per l'assegnazione delle esposizioni a classi o pool, nonché come i modelli di ML incidano sulla quantificazione del rischio.

- ii. **Validazione rigorosa e internal audit (parr 43-49):** le banche devono garantire che le procedure di validazione interna dei modelli tengano esplicitamente in conto dell'uso di tecniche di ML nei modelli interni e nello stesso processo di validazione. L'attività di validazione deve essere commisurata alla complessità e al rischio del modello basato sul ML e, se necessario, essere rivista anche a cadenze ravvicinate; essa dovrebbe "sfidare" in maniera effettiva le decisioni di modellizzazione adottate riguardo ai modelli interni basati sul ML, valutando se la loro complessità sia giustificata ed esaminando la stabilità e la robustezza dei risultati. Inoltre, tale attività dovrebbe valutare le implicazioni

legate alla capacità delle tecniche di ML utilizzate nei modelli interni di adattarsi al set di dati di addestramento oltre le proprie capacità di generalizzazione (incluso l'*overfitting*) e dovrebbe prestare particolare attenzione alle prestazioni dei modelli basati su ML misurate utilizzando dati *out-of-sample* (fuori campione) e *out-of-time* (fuori intervallo temporale).

L'*audit* interno deve valutare l'adeguatezza delle risorse e dei processi per essere svolto su modelli basati su ML, considerando i rischi specifici associati a queste tecniche; l'intensità e la frequenza di tale attività devono essere proporzionate alla complessità e dinamismo dei modelli, prevedendo audit aggiuntivi, se necessario, in caso di esigenze informative speciali o eventi imprevisti; l'*audit* dovrebbe inoltre valutare anche aspetti rilevanti specifici per i modelli ML e includere la propria opinione nella valutazione del rischio.

- iii. **Data governance (parr 50-51):** Le banche devono definire standard per valutare l'adeguatezza dei dati utilizzati nei modelli di ML, considerando anche dati sintetici (cioè generati artificialmente) e non strutturati (documenti di testo, interazioni sui social media, registrazioni video...). Inoltre, i dati di input per le tecniche di ML nei modelli interni devono essere analizzati per comprendere formati, valori mancanti e potenziali bias.
- iv. **Spiegabilità:** La BCE si aspetta che gli intermediari evitino il ricorso ad approcci di modellizzazione inutilmente complessi e garantiscano un bilanciamento tra la complessità delle tecniche di modellizzazione e l'incremento delle performance dei modelli interni basati sul machine learning. La necessità di evitare una eccessiva complessità trova fondamento nella CRR agli articoli 179 (il quale richiede che le stime

dei parametri di rischio siano plausibili ed intuitive e basate su risk drivers sostanziali) e 174 comma 1, lett a (il quale esige che le variabili di *input* costituiscano una base ragionevole ed efficace per la previsione risultante). Inoltre, osserva la BCE, come un eccessivo numero di fattori di rischio e di variabili di input potrebbe essere di ostacolo al rispetto delle previsioni degli articoli 171 comma1 lett. b e art. 175 c.2 per le quali gli enti devono documentare (*rectius* motivare) le ragioni alla base della scelta dei criteri di rating in modo tale da consentire ai terzi di comprendere l'assegnazione delle esposizioni alle classi, di replicarle e di valutarne l'appropriatezza²⁹³.

La necessità di garantire la spiegabilità del modello è la chiave di lettura che permette di comprendere le similitudini fra AI Act e normativa prudenziale: lo scopo dell'AI Act è garantire la diffusione di «*un'intelligenza artificiale antropocentrica ed affidabile*», la disciplina dei modelli IRB si prefigge di garantire che sia preservata la stabilità finanziaria dell'intermediario e più in generale dell'intero sistema bancario. Il conseguimento di tali obiettivi richiede un presupposto comune: che i modelli vengano sviluppati in modo tale da essere non solo affidabili ma anche spiegabili (*rectius* trasparenti).

L'effettiva trasparenza del modello IRB va assicurata, *by design*, fin dalla fase di progettazione, evitando di predisporre sistemi eccessivamente complessi se non quando tale complessità venga compensata e giustificata da un significativo miglioramento nelle capacità predittive del modello stesso²⁹⁴; le banche devono pertanto evitare di adottare scelte di modellizzazione eccessivamente complesse se sono disponibili approcci più semplici in grado di fornire risultati analoghi e che i modelli non includano un numero eccessivo di variabili esplicative né

²⁹³ European Central Bank, *Guide to internal models*, https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.supervisory_guide202507.en.pdf pp 18 ss

²⁹⁴ European Central Bank, *Press Release ECB publishes revised guide to internal models*, <https://www.bankingsupervision.europa.eu/press/pr/date/2025/html/ssm.pr250728~2b36305822.it.html>

utilizzino dati non strutturati qualora siano disponibili dati, maggiormente convenzionali, che permettano di ottenere capacità predittive simili²⁹⁵.

Quando la comprensibilità dei modelli basati sul ML non è realizzabile mediante una interpretabilità *by design*²⁹⁶ poiché vengono scelti algoritmi di ML che, pur fornendo migliori performance, non permettono di delineare un nesso causale ed economico univoco tra le variabili di input e gli output prodotti (si pensi ad es. alle reti neurali²⁹⁷) individuare la teoria economica sottostante al modello, e renderlo compatibile con le previsioni di spiegabilità e trasparenza, sancite dalla CRR, può rappresentare una sfida che tuttavia può essere superata affidandosi alla tecnologia ed in particolare a tecniche e strumenti avanzati di interpretazione.

4. La Tecnologia come strumento di trasparenza: eXAI

L'assenza di trasparenza è una caratteristica intrinseca a molti algoritmi di *machine learning* il cui grado di spiegabilità intrinseca tende a diminuire con l'aumento della capacità previsiva, in un rapporto di proporzionalità inversa: se ad esempio le reti neurali offrono performance molto elevate e possono essere rese in qualche misura spiegabili solo adottando tecniche avanzate di

²⁹⁵ European Banking Authority, *Discussion Paper on Machine Learning for IRB Models* (EBA/DP/2021/04) pp. 23 ss

²⁹⁶ Come visto in Cap. II, par. 2, del presente elaborato, *l'interpretabilità* ha ad oggetto la capacità intrinseca del modello di permettere all'essere umano di comprenderne la logica decisionale esaminandone direttamente i componenti matematici (es. i nodi di un albero decisionale o i coefficienti di una regressione lineare), senza necessità di strumenti esterni;

²⁹⁷ G. Ziccardi, *Una lettura dell'Artificial Intelligence Act: Norme, Etica, Adempimenti, Attuazione*, in *Intelligenza Artificiale Diritto, Giustizia, Economia ed Etica*, a cura di F. Basile, M. Biasi, L. Camaldo, G. Caneschi, B. Fragasso, D. Milani, G. Giappichelli Editore, p.22 https://www.giappichelli.it/media/catalog/product/openaccess/9791221162905.pdf?srsId=AfmBOorGBYZcEsgMRG18xXnnAkKNZahCiLf_8M6dmkntwrwhcDSOmmUC1

interpretazione dall'altro lato un modello intrinsecamente spiegabile, quale un *decision tree*, presenta di norma minore capacità previsiva²⁹⁸.

Per evitare risultati opachi è necessario promuovere strumenti tecnici che permettano di ottenere una valutazione ed una quantificazione del contributo apportato dai singoli fattori di rischio alla previsione del modello dal punto di vista globale nonché di conoscere il contributo dei singoli fattori di rischio chiave a ogni singola previsione.

Questi strumenti devono essere utilizzati nell'ambito dei controlli interni della banca al fine di identificare in modo attendibile le circostanze e i casi specifici in cui un modello basato sul ML non funzioni efficacemente (art. 175, c. 4, lett. c CRR), le cause di un deterioramento delle prestazioni e della stabilità del modello (art 174, lett. d, CRR), le ragioni dello scostamento tra le stime di rischio e i valori realizzati (art. 185, lett. e, CRR).

Per realizzare siffatti obiettivi tali mezzi devono poter fornire una spiegazione robusta ed accurata che consenta una supervisione umana attendibile e proattiva e permetta di avere un giudizio umano informato rispetto agli output del modello; inoltre, per garantire tale risultato è necessario che le tecniche e gli strumenti di spiegabilità siano anch'essi oggetto della attività di validazione interna per verificarne l'adeguatezza ed appropriatezza.

Le tecniche di interpretazione rilevano altresì dal punto di vista delle autorità di vigilanza le quali, nella loro attività autorizzatoria e di verifica periodica devono essere nelle condizioni di comprendere appieno non solo gli output generati dal modello interno, ma anche le implicazioni e il funzionamento dello stesso.

²⁹⁸ E. Bonaccorsi di Patti, F. Calabresi, B. De Varti, F. Federico, M. Affinito, M. Antolini, F. Lorzio, S. Marchetti, I. Masiani, M. Moscatelli, F. Privitera e G. Rinna, *Intelligenza artificiale nel credit scoring. Analisi di alcune esperienze nel sistema finanziario italiano*, https://www.bancaditalia.it/pubblicazioni/qef/2022-0721/QEF_721_IT.pdf p. 12

È in tale quadro che assumono un ruolo fondamentale le tecnologie di eXplainable AI (eXAI), termine con il quale si indica l'insieme di strumenti applicati dall'analista volti ad integrare il risultato primario dell'algoritmo di ML con un insieme di interpretazioni e spiegazioni circa il funzionamento del modello, il cui uso permette di ottenere valutazioni qualitative e giustificazioni comprensibili in merito al funzionamento ed ai risultati prodotti dagli algoritmi di intelligenza artificiale.

Più nello specifico l'eXAI si riferisce a due approcci principali²⁹⁹:

- i. *interpretability*: riguarda la capacità intrinseca del modello di essere comprensibile, *by design*, all'essere umano; tale approccio privilegia la trasparenza strutturale permettendo di tracciare quantitativamente i meccanismi interni che trasformano gli input in output. Un modello è interpretabile se un osservatore può comprenderne la logica decisionale esaminandone direttamente i componenti matematici (es. i nodi di un albero decisionale o i coefficienti di una regressione lineare), senza necessità di strumenti esterni;
- ii. *explainability*: riguarda la capacità di fornire valutazioni qualitative e giustificazioni comprensibili in merito ai risultati prodotti. L'approccio non è quello di creare un modello che sia intrinsecamente interpretabile ma costruire un ulteriore modello che permetta di spiegare a posteriori (*post-hoc*) il modello originario chiarendo il perché di una specifica decisione.

L'implementazione di tecniche di eXAI può riguardare diverse fasi del ciclo di vita dei modelli³⁰⁰: prima dello sviluppo, al fine di comprendere i dati utilizzati

²⁹⁹ C. Rudin, *Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead*, in *Nature Machine Intelligence* 1, (2019), <https://doi.org/10.1038/s42256-019-0048-x> pp. 206–215

³⁰⁰ E. Bonaccorsi di Patti, F. Calabresi, B. De Varti, F. Federico, M. Affinito, M. Antolini, F. Lorizzo, S. Marchetti, I. Masiani, M. Moscatelli, F. Privitera e G. Rinna, *Intelligenza artificiale nel credit scoring. Analisi di alcune esperienze nel sistema finanziario italiano*, https://www.bancaditalia.it/pubblicazioni/qef/2022-0721/QEF_721_IT.pdf pp 12 ss.

per l'addestramento (*data explainability*); durante lo sviluppo, per migliorarne l'interpretabilità (*explainable modeling*); a valle dello sviluppo, per fornire spiegazioni a posteriori sulla logica decisionale (*post-modeling explainability*). Soffermandoci sulle tecniche maggiormente utilizzate e applicabili su qualunque tipo di algoritmo di ML (c.d. *model-agnostic explainability*³⁰¹) ne vanno evidenziate in particolare due:

1. Shapley Additive Explanation (SHAP): questo metodo trova il suo fondamento nella teoria dei giochi cooperativi, formulata da Lloyd Shapley, nel 1953³⁰², al fine di calcolare come distribuire il profitto realizzato tra un gruppo di giocatori sulla base del contributo che questi hanno fornito all'output generato nel complesso dal gruppo stesso. Traslata nel contesto del machine learning, i "giocatori" sono le variabili esplicative (*feature*) e il "profitto" corrisponde alla previsione del modello cosicché tale metodologia permette di determinare, anche nell'ambito delle reti neurali, quanto ciascuna *feature* influisca sull'output generato dal modello³⁰³ (Figura 2).

Il metodo SHAP è considerato essere, dalle banche, uno dei più efficaci tanto che, il 40% dei rispondenti alla call dell'EBA ha dichiarato di utilizzarlo³⁰⁴. I suoi pregi sono stati ben riassunti nella risposta trasmessa all'EBA da Intesa Sanpaolo³⁰⁵ che ha evidenziato come SHAP fornendo un metodo per analizzare il contributo marginale di ciascuna variabile e per considerare ogni

³⁰¹ S. Ali, T. Abuhmed, S. El-Sappagh, K. Muhammad, J. M. Alonso-Moral, R. Confalonieri, R. Guidotti, J. Del Ser, N. Díaz-Rodríguez, F. Herrera, *Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence*, in *Information Fusion*, Volume 99, 2023, p. 11

³⁰² L. Shapley, *A Value for n-Person Games*, in Kuhn, H. and Tucker, A., Eds., *Contributions to the Theory of Games II*, Princeton University Press, Princeton, (1953) pp. 307-317

³⁰³ P. Giudici, E. Raffinetti, *Shapley-Lorenz eXplainable Artificial Intelligence*, in *Expert Systems with Applications*, Vol. 167, 2021

³⁰⁴ European Banking Authority, *Machine Learning for IRB Models Follow-Up Report from the Consultation on the Discussion Paper on Machine Learning for IRB Models* (EBA/REP/2023/28), p. 11

³⁰⁵ Intesa San Paolo, *Response to discussion on machine learning for IRB models*, <https://www.eba.europa.eu/eba-response/31935?destination=/discussion-paper-machine-learning-irb-models>

possibile combinazione con le altre (tramite l'osservazione della variazione del punteggio per ciascuna combinazione) permette alla banca di valutare quali caratteristiche influenzano maggiormente l'output del modello.

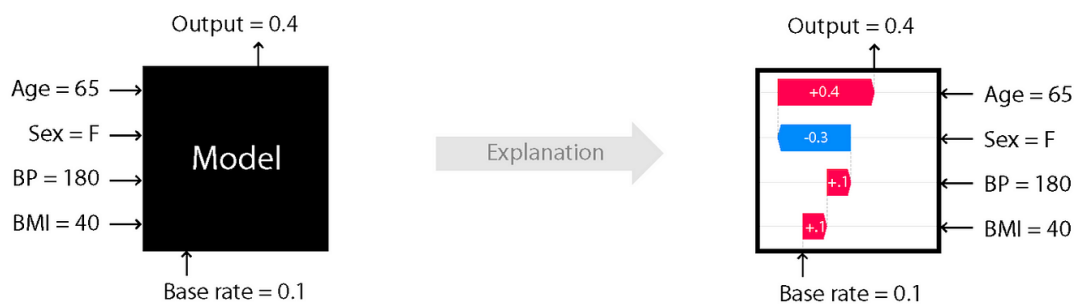


Figura 2: funzionamento di un modello SHAP

Fonte: E. Bonaccorsi di Patti Et al., "Intelligenza artificiale nel credit scoring. Analisi di alcune esperienze nel sistema finanziario italiano".

2. Local Interpretable Model-Agnostic Explanations (LIME): questo metodo fornisce una spiegazione sul comportamento del modello nei confronti di una specifica previsione mediante la variazione dei valori delle features e la successiva osservazione e spiegazione degli impatti che tale variazione ha sull'output grazie all'ausilio di un modello lineare³⁰⁶.

Il funzionamento di un modello LIME può essere meglio compreso osservando la "Figura 3"; dato un modello di classificazione che genera un output binario (lo sfondo blu e rosa) al fine di spiegare il contributo di una *feature* all'ottenimento dell'*output* (la croce rossa in grassetto) sarà necessario compiere i seguenti passaggi:

³⁰⁶ M. T. Ribeiro, S. Singh, C. Guestrin, *Model-Agnostic Interpretability of Machine Learning*, (2016) DOI:[10.48550/arXiv.1606.05386](https://doi.org/10.48550/arXiv.1606.05386)

- Campionare l'istanza: individuato un dato da spiegare (la croce rossa in grassetto) il modello LIME la utilizza per creare delle varianti con dati molto simili ma leggermente modificati (le altre croci e i cerchi blu).
- Ponderare i dati: a ogni variante viene assegnato un peso in base alla sua somiglianza con l'originale; i campioni più simili (graficamente più grandi) hanno un'influenza maggiore, poiché descrivono meglio il comportamento del modello nelle immediate vicinanze dell'istanza analizzata.
- Addestrare il modello surrogato: sulla base dei campioni ponderati, viene addestrato un modello semplice e interpretabile (di norma una regressione lineare); il modello (graficamente rappresentato dalla linea tratteggiata) non cerca di spiegare l'intero funzionamento del modello, ma solo come esso si comporta in quel punto specifico.
- Generazione della spiegazione: la linea tratteggiata ottenuta rappresenta la spiegazione locale; i coefficienti del modello lineare indicano il contributo e l'importanza di ciascuna *feature* nel determinare l'output specifico del modello per l'istanza analizzata. Il risultato finale è una spiegazione "*localmente fedele*", capace di rendere comprensibile una decisione complessa attraverso pochi elementi chiave.

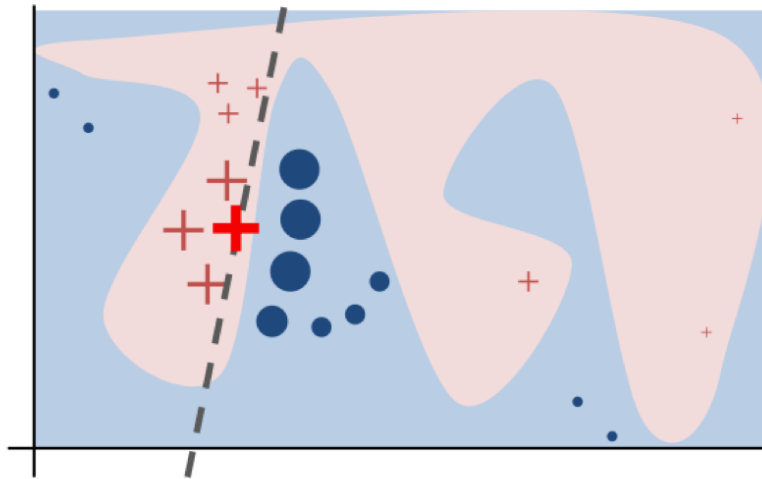


Figura 3: funzionamento di un modello LIME

Fonte M. T. Ribeiro et Al, "Model-Agnostic Interpretability of Machine Learning".

Un esempio pratico può aiutare a meglio chiarire il procedimento illustrato³⁰⁷: supponiamo che ad un modello venga chiesto di classificare (*rectius* determinare) degli animali sulla base delle foto fornitegli ma il modello classifica erroneamente un cane (Figura 4) come un lupo. Per comprendere il motivo dell'errore un modello LIME opererà creando delle versioni della foto oscurando diverse parti dell'immagine ed osservando la probabilità che il modello classifichi l'immagine come "lupo" al variare delle parti visibili. Sulla base di tali dati verrà creato un modello *logit* tramite il quale si evince che il modello di classificazione basa la sua predizione non sui tratti somatici dell'animale bensì sulla presenza o meno della neve nello sfondo.

³⁰⁷ M. T. Ribeiro, S. Singh, C. Guestrin, "Why Should I Trust You?": Explaining the Predictions of Any Classifier in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)* Association for Computing Machinery, New York, USA, 1135–1144 2016

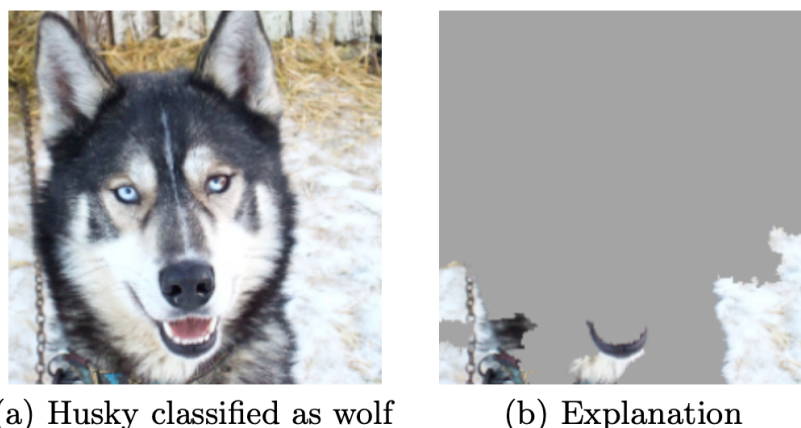


Figura 4: esempio di spiegazione di un modello LIME

Fonte M. T. Ribeiro et Al, Why Should I Trust You? Explaining the Predictions of Any Classifier

Sia LIME che SHAP si basano sull'analisi della relazione tra gli input e gli output del modello piuttosto che sull'esame del funzionamento interno del modello stesso ma si differenziano tra di loro perché LIME spiega le predizioni del modello, approssimandolo localmente attraverso un modello surrogato interpretabile costruito a partire da perturbazioni delle caratteristiche in input, mentre SHAP quantifica il contributo di ciascuna caratteristica a una predizione utilizzando i valori di Shapley derivati dalla teoria dei giochi cooperativi³⁰⁸.

Le tecniche illustrate non sono le uniche utilizzabili, moltissimi sono infatti gli strumenti di eXAI sviluppati³⁰⁹ tra cui, sempre in ambito dei modelli IRB³¹⁰, si possono citare i *partial dependence plots* (PDP³¹¹) e gli *accumulated local effects*

³⁰⁸ P. Knab, S. Marton, U. Schlegel, C. Bartelt, *Which LIME should I trust? Concepts, Challenges, and Solutions*, in: *Explainable Artificial Intelligence. xAI 2025. Communications in Computer and Information Science*, a cura di R. Guidotti, U. Schmid, L. Longo, (eds), vol 2577. Springer, Cham <https://arxiv.org/pdf/2503.24365v1> p.2.

³⁰⁹ Per un elenco dettagliato si rinvia a S. Ali, T. Abuhmed, S. El-Sappagh, K. Muhammad, J. M. Alonso-Moral, R. Confalonieri, R. Guidotti, J. Del Ser, N. Díaz-Rodríguez, F. Herrera, *Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence*, in *Information Fusion*, Volume 99, 2023

³¹⁰ French Banking Federation, *Response to discussion on machine learning for IRB models*, 15. <https://www.eba.europa.eu/eba-response/31906?destination=/discussion-paper-machine-learning-irb-models>

³¹¹ Metodo grafico di interpretabilità che permette di mostrare l'effetto marginale medio che una variabile ha sulla previsione finale del modello.

(ALE³¹²). Nel complesso queste tecniche forniscono informazioni su come i sistemi di intelligenza artificiale elaborino i dati e generino il loro output, offrendo una comprensione del “ragionamento” che ha portato a tali conclusioni o decisioni dando una maggiore comprensione ai *deployer* e alle autorità di vigilanza in merito al funzionamento dei sistemi e fornendo agli individui interessati dalla decisione algoritmica la possibilità di comprendere e interagire con il processo decisionale³¹³.

5. eXAI, regolamentazione e neutralità tecnologica.

Le tecniche di eXAI per la loro capacità di spiegare le decisioni opache della macchina non si limitano a permettere il controllo del rispetto della normativa prudenziale da parte delle banche che vogliono utilizzare i modelli IRB ma possono aiutare a garantire la conformità dei modelli alle previsioni del GDPR e dell'AI Act: il combinato disposto degli artt. 13, 14 e 15 del GDPR, così come interpretato alla luce della giurisprudenza comunitaria prevede che l'interessato abbia diritto ad ottenere una spiegazione da parte del titolare del trattamento dei motivi sottostanti alla decisione algoritmica che lo ha riguardato; i sistemi *high risk* potranno essere messi in commercio e utilizzati, a condizione che (art. 13 AI Act) siano progettati e sviluppati in modo da garantire agli utilizzatori di interpretare e utilizzare l'output del sistema in modo appropriato garantendone una sufficiente trasparenza (il considerando 72 la declina come necessità di «comprendere il funzionamento del sistema di IA, valutarne la funzionalità e comprenderne i punti di forza e i limiti»).

³¹² Metodo grafico di interpretabilità che permette di visualizzare l'impatto medio di una *feature* sulla previsione di un modello; si differenzia dal PDP perché mentre questi fa una media globale, l'ALE calcola come cambia la previsione "localmente" (in piccoli intervalli della *feature*) ed accumula questi cambiamenti incrementali.

³¹³ European Data Protection Supervisor, *TechDispatch on Explainable Artificial Intelligence*, https://www.edps.europa.eu/system/files/2023-11/23-11-16_techdispatch_xai_en.pdf p.9

Ora l'uso di tecniche di eXAI non viene previsto specificatamente da alcuna disposizione; né dal GDPR che impone di fornire «*informazioni significative sulla logica utilizzata, nonché l'importanza e le conseguenze previste di tale trattamento per l'interessato*» né dall'AI Act, il quale ricollega la trasparenza solo all'obbligo di fornire documentazione pertinente e istruzioni per l'uso riguardo il sistema di intelligenza artificiale³¹⁴ né tanto meno dalla normativa prudenziale.

Questo approccio è certamente conforme al principio di neutralità tecnologica in virtù del quale i *policy maker* non devono privilegiare un tipo particolare di tecnologia ad un'altra poiché così facendo la regolazione potrebbe pregiudicare la libera concorrenza³¹⁵.

Viene tuttavia da chiedersi se applicare pedissequamente tale principio in ambito finanziario sia la scelta più opportuna da seguire: quando un modello di machine learning di tipo avanzato, quale ad esempio una rete neurale, apprende dai dati e costruisce il proprio schema decisionale lo fa senza considerare la necessità di comprensione da parte dell'uomo; in altre parole, l'autoapprendimento della macchina non si accorda per sua natura con l'esigenza di fornire spiegazioni comprensibili all'uomo³¹⁶; il fenomeno della *black box*, rendendo difficile spiegare perché un modello abbia generato un particolare risultato o decisione e quale combinazione di *features* vi abbia contribuito³¹⁷, evidenzia la necessità di doversi

³¹⁴ C. Panigutti, R. Hamon, I. Hupont, D. F. Llorca, D. F. Yela, H. Junklewitz, S. Scalzo, G. Mazzini, I. Sanchez, J. S. Garrido, E. Gomez, *The role of explainable AI in the context of the AI Act*, in *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)*, Association for Computing Machinery, New York, NY, USA, 1139–1150 2023 il quale osserva come i *deployer* possano garantire la trasparenza dei modelli ad esempio progettando l'interfaccia uomo-macchina, di cui all'art. 14 dell'AI Act in modo da rendere facile per un professionista qualificato comprendere l'output del sistema.

³¹⁵ M. Rabitti, *Le Regole di Supervisione nel Mercato Digitale: Considerazioni Intorno alla Comunicazione Banca d'Italia in Materia di Tecnologie Decentralizzate nella Finanza e Cripto-Attività*, in *Rivista Trimestrale di Diritto dell'economia*, Supplemento n. 3 al n. 1/2022, p. 126

³¹⁶ J. Burrell, *How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms*, in *Big Data & Society*, 1/2016, p.10

³¹⁷ E. Troisi, *Decisione Algoritmica, Black-Box e Ai Etica: il Diritto di Accesso Come Diritto a Ottenere una Spiegazione*, in *Jus Civile*, Fascicolo 4, 2022, G. Giappichelli p. 955

servire delle tecniche di eXAI per rispettare il principio di *explainability* richiesto per l'uso dei modelli AI³¹⁸.

E allora bisogna ammettere i limiti del principio di neutralità (peraltro messo in dubbio dalla dottrina che sottolinea come nella prassi sia spesso smentito) e riconoscere che determinati modelli di intelligenza artificiale necessitano di essere affiancati da tecnologie per la spiegazione degli output non potendosi sostenere che la regolamentazione e l'attività di vigilanza debbano mantenersi equidistanti sia da soluzioni efficienti che inefficienti³¹⁹.

La rapidità con cui si evolve la tecnologia osta ad imporre unilateralmente una specifica soluzione tecnologica (ad es. la tecnica di eXAI cui fare ricorso) attraverso fonti di *hard law* per loro natura più lunghe da modificare per adattarle al rapido cambio normativo. Tuttavia, se una soluzione tecnologica riesce ad imporsi rispetto alle altre (come nel caso della tecnica SHAP) nel garantire la spiegabilità del modello è lecito ritenere che l'autorità di vigilanza possa raccomandare quella certa tecnologia tramite strumenti di *soft law*³²⁰, i quali basandosi su procedure maggiormente rapide permettono modifiche più agevoli delle previsioni così da riuscire a mantenere il passo con il rapido progresso dell'innovazione tecnologica³²¹.

Ma non solo, visti gli interessi in gioco sarebbe opportuno che siano le stesse autorità di vigilanza bancaria e dell'intelligenza artificiale a farsi driver dello sviluppo di soluzioni tecnologiche efficienti cooperando con l'industria e la

³¹⁸ N. Bussmann, P. Giudici, D. Marinelli, J. Papenbrock, . Explainable Machine Learning in Credit Risk Management in *Computational Economics* vol. 57 pp. 203–216 (2021).

³¹⁹ M. Rabitti, *Le Regole di Supervisione nel Mercato Digitale: Considerazioni Intorno alla Comunicazione Banca d'Italia in Materia di Tecnologie Decentralizzate nella Finanza e Cripto-Attività*, in *Rivista Trimestrale di Diritto dell'economia*, Supplemento n. 3 al n. 1/2022

³²⁰ Come osservato da M. Ramajoli, Self regulation, soft regulation e hard regulation nei mercati finanziari, in *Rivista della Regolazione dei mercati*, Fascicolo 2 (2016) p. 65 ³²⁰ tali strumenti trovano nei mercati finanziari un terreno molto fertile con autorità di vigilanza che fanno ampio ricorso a raccomandazioni, orientamenti o altri documenti di *soft regulation* emanati per istituire prassi di vigilanza uniformi, efficienti ed efficaci e assicurare l'applicazione comune, uniforme e coerente del diritto UE.

³²¹ G. Morbidelli, *Degli Effetti Giuridici della Soft Law*, in *Rivista della Regolazione dei mercati*, Fascicolo 2 (2016)

ricerca. Così facendo gli operatori avrebbero davanti a loro una maggiore certezza rispetto al quadro normativo e le autorità riceverebbero spunti e indicazioni per regole efficienti e best practices, che potrebbero portare anche a miglioramenti dell'attività di controllo³²².

In conclusione, l'intrinseca opacità dei modelli avanzati di ML richiede di superare il principio di neutralità a favore di previsioni focalizzate sulla tecnologia stessa. L'integrazione tra strumenti di *soft law* e tecniche di eXAI non rappresenta solo una soluzione tecnica, ma un necessario cambio di paradigma nella governance finanziaria: un approccio in cui l'autorità di vigilanza non è semplicemente il controllore ma un partner strategico nella definizione delle tecniche di interpretabilità. Solo attraverso questa sinergia tra regolatore e industria sarà possibile cogliere i benefici dell'innovazione algoritmica senza sacrificare la comprensibilità dei modelli e la sana e prudente gestione delle banche pilastri imprescindibili per la stabilità del sistema.

³²² A. Perrone, *La nuova vigilanza. Regtech e capitale umano*, in *Banca Borsa Titoli di Credito*, p. 516

CONCLUSIONI

Nel corso del lavoro si è cercato di identificare i benefici che l'intelligenza artificiale e più nel dettaglio le tecniche di *machine learning* possano apportare al mercato del credito sia dal punto di vista economico che sociale.

Si è riscontrato come tali algoritmi riescano a sovraperformare i modelli statistici tradizionali nel predire i *default*³²³, sia in accuratezza³²⁴ che in potere discriminatorio³²⁵ grazie alla loro capacità di elaborare grandi quantità di dati in volume e varietà e di considerare variabili prive di una chiara relazione economica, permettendo così di cogliere schemi complessi e interazioni non lineari tra le variabili che i modelli *logit* non riescono a rilevare³²⁶.

I benefici economici connessi si manifestano in primo luogo nella riduzione della asimmetria informativa e dei consequenziali problemi di *adverse selection* e moral hazard permettendo così alle banche di assumere decisioni maggiormente consapevoli e di ridurre il razionamento del credito rispetto a progetti maggiormente rischiosi e/o innovativi³²⁷ senza dover far ricorso alla richiesta di collaterali, i quali sono idonei a generare allocazioni subottimali del capitale favorendo soggetti già "ricchi" a discapito di soggetti i cui progetti, seppur

³²³ Tali considerazioni sono state nuovamente riaffermate in uno studio della Banca d'Italia pubblicato all'inizio di quest'anno: F. Columba, M. Cugliari, S. Di Virgilio, Credit Risk Assessment with Stacked Machine Learning in Mercati, infrastrutture, sistemi di pagamento n.73, 2026

³²⁴ Proporzioni di previsioni corrette (sia positive che negative) sul totale delle osservazioni considerate in un modello di classificazione.

³²⁵ Capacità di un modello di distinguere correttamente tra casi positivi e negativi, ordinandoli in base alla probabilità stimata di appartenenza alla classe positiva.

³²⁶ H. Ayari, Ramzi Guetari, N. Kraïem, *Machine learning powered financial credit scoring: a systematic literature review*, in *Artificial Intelligence Review* Vol 59, 2025

³²⁷ Per una sintesi efficace delle risultanze riscontrate si rinvia a:

T. Berg, V. Burg, A. Gombović, M. Puri, *On the Rise of FinTechs: Credit Scoring Using Digital Footprints in The Review of Financial Studies*, Volume 33, Issue 7, July 2020, Pages 2845–2897

Bonaccorsi di Patti, F. Calabresi, B. De Varti, F. Federico, M. Affinito, M. Antolini, F. Loriglio, S. Marchetti, I. Masiani, M. Moscatelli, F. Privitera e G. Rinna, *Intelligenza artificiale nel credit scoring. Analisi di alcune esperienze nel sistema finanziario italiano*, https://www.bancaditalia.it/pubblicazioni/qef/2022-0721/QEF_721_IT.pdf

economicamente validi vedono denegato il finanziamento a causa della assenza di garanzie³²⁸.

In secondo luogo, come è stato dimostrato in uno studio che ha analizzato i dati delle banche italiane prima e durante la crisi del COVID-19, la maggiore accuratezza predittiva ottenuta grazie a tali strumenti riesce a mitigare gli effetti delle fluttuazioni macroeconomiche smorzando gli effetti contro ciclici e stabilizzando l'offerta di credito, con impatti positivi su investimenti e occupazione delle imprese clienti³²⁹.

Infine, tali sistemi possono ridurre le riserve di capitale che le banche devono detenere per far fronte al rischio di credito (che possono così essere utilizzati per erogare ulteriori prestiti) dal momento che l'utilizzo del *machine learning* nei modelli IRB si è dimostrato idoneo a rendere più accurata la *risk differentiation* anche del 20%³³⁰ e a ridurre la ponderazione media delle attività rispetto al rischio tra il 2% ed il 27% rispetto ai modelli Logit³³¹.

Dal punto di vista "sociale" l'uso di dati alternativi, complementari a quelli impiegati nei modelli tradizionali, e di algoritmi di ML per elaborarli sono idonei a migliorare l'inclusione finanziaria ampliando la capacità di valutazione del merito creditizio nei confronti di soggetti che altrimenti sarebbero esclusi o penalizzati nell'accesso al credito. Tali tecniche permettono infatti di erogare risorse a richiedenti che i sistemi tradizionali classificherebbero ad alto rischio di

³²⁸ W. G. Gale, *Collateral, Rationing, and Government Intervention in Credit Markets*, in *Asymmetric Information, Corporate Finance, and Investment* a cura di R. G. Hubbard, University of Chicago Press (1990) pp 43-62

³²⁹ L. Gambacorta, F. Sabatini, S. Schiaffi, *Artificial intelligence and relationship lending*, https://www.bancaditalia.it/pubblicazioni/temi-discussione/2025/2025-1476/en_tema_1476.pdf?language_id=1

³³⁰ A. Alonso, J. M. Carbó, *Machine Learning in Credit Risk: Measuring the Dilemma Between Prediction and Supervisory Cost*, <https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosTrabajo/20/Files/dt2032e.pdf> p.9

³³¹ H. Fraisse, M. Laporte, *Return on Investment on AI: The Case of Capital Requirement*, in Banque de France Working Paper n. 809

insolvenza³³² nonché a soggetti con limitato storico creditizio ("*thin-file*" e "*no-file*")³³³ e a categorie vulnerabili, quali, ad esempio, le minoranze etniche³³⁴.

I benefici economici e sociali vengono tuttavia ottenuti in cambio di una contropartita: la mancanza di trasparenza degli output, dovuta alla natura intrinseca di *black box*³³⁵ di alcune tipologie di modelli di *machine learning*³³⁶ la quale comporta che «*we can observe its inputs and outputs, but we cannot tell how one becomes the other*»³³⁷.

Sussiste un rapporto di proporzionalità inversa fra l'aumentare della capacità previsiva e l'interpretabilità della decisione: un modello intrinsecamente interpretabile come un *decision tree*, presenta infatti una minore capacità previsiva rispetto ad una rete neurale la cui spiegabilità può essere realizzata solo utilizzando algoritmi che possano rendere trasparente la decisione³³⁸.

Il tema della trasparenza della decisione algoritmica e della conseguente necessità di "aprire la *black box*" e rendere comprensibile il sillogismo tra *input* e *output* è un principio cardine, valido per tutte le applicazioni dell'intelligenza artificiale, che è stato dettagliato, dalla giurisprudenza europea, proprio in

³³² J. Jagtiani, C. Lemieux, *The roles of alternative data and machine learning in fintech lending: Evidence from the LendingClub consumer platform*, <https://www.philadelphiafed.org/-/media/frbp/assets/working-papers/2018/wp18-15r.pdf>

³³³ L. Ammannati, G. L. Greco, *Il credit scoring "intelligente": esperienze, rischi e nuove regole*, in *Rivista di diritto bancario* p. 465 ss.

³³⁴ Per un approfondimento su come il *credit scoring* algoritmico abbia ridotto le disparità di accesso al credito delle comunità ispanica e afroamericana nel mercato USA si rinvia a R. Bartlett, A. Morse, R. H Stanton, N. Wallace, *Consumer-Lending Discrimination in the Fintech Era*, in NBER Working Paper No. w25943, (June 2019).

³³⁵ E. Troisi, *Decisione Algoritmica, Black-Box e Ai Etica: il Diritto di Accesso Come Diritto a Ottenere una Spiegazione*, in *Jus Civile*, Fascicolo 4, 2022, G. Giappichelli p. 955

³³⁶ G. Cascarino, M. Moscatelli, F. Parlapiano, *Explainable Artificial Intelligence: Interpreting Default Forecasting Models Based on Machine Learning*, in *Questioni di Economia e Finanza della Banca d'Italia* n. 674 (2022)

³³⁷ F. Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information*, Harvard University Press, 2015, p.3.

³³⁸ E. Bonaccorsi di Patti, F. Calabresi, B. De Varti, F. Federico, M. Affinito, M. Antolini, F. Lorzio, S. Marchetti, I. Masiani, M. Moscatelli, F. Privitera e G. Rinna, *Intelligenza artificiale nel credit scoring. Analisi di alcune esperienze nel sistema finanziario italiano*, https://www.bancaditalia.it/pubblicazioni/qef/2022-0721/QEF_721_IT.pdf p. 12

materia di *credit scoring*³³⁹: come si ha avuto modo di commentare nel secondo capitolo, la Corte di Giustizia nelle sentenze Schufa³⁴⁰ e D&B³⁴¹, richiamandosi agli articoli 15, lett. h) e 22 del GDPR, in materia di processo decisionale automatizzato, ha infatti affermato il diritto, del soggetto interessato da una valutazione di *credit scoring* algoritmico, di ricevere una spiegazione sufficientemente trasparente e intellegibile in merito alle ragioni a fondamento del diniego della richiesta di credito; in altre parole non basta infatti comunicare l'esito della richiesta bensì è necessario esplicitare il percorso logico che ha condotto a tale risultato al fine di garantire la contestabilità della decisione.

Tale diritto è stato trasposto nella nuova Direttiva sul Credito al Consumo dagli articoli 18 e 19, che legittimano il ricorso agli strumenti di *credit scoring* algoritmico a condizione che il destinatario abbia il diritto di «*di chiedere ed ottenere dal creditore una spiegazione chiara e comprensibile della valutazione del merito creditizio, compresi la logica e i rischi derivanti dal trattamento automatizzato dei dati personali nonché la rilevanza e gli effetti sulla decisione*» e che per assumere tale decisione non siano stati usati come input né le categorie di dati particolari di cui all'art. 9, par. 1 del GDPR (origine razziale, opinioni politiche, convinzioni

³³⁹ A ben vedere, ancor prima delle pronunce della Corte di Giustizia dell'Unione Europea il principio di trasparenza algoritmica era stato affermato dal Consiglio di Stato con "Sentenza 8472 del 2019" nella quale è stata ammessa la possibilità per la Pubblica Amministrazione di adottare decisioni algoritmiche, nei provvedimenti caratterizzati da discrezionalità tecnica, a patto che queste rispettassero un "decalogo della legalità algoritmica" sintetizzabile in tre principi che sono stati desunti dagli articoli 13, 14, 15 e 22 del GDPR:

- Trasparenza dell'algoritmo: per il quale cui ognuno *"ha diritto a conoscere l'esistenza di processi decisionali automatizzati che lo riguardano ed in questo caso a ricevere informazioni significative sulla logica utilizzata"*; l'Amministrazione è tenuta cioè a mettere il cittadino nelle condizioni di avere ragionevolmente chiaro il senso della decisione che lo interessa e quindi la logica sottesa all'algoritmo.

- Non esclusività algoritmica: l'interessato ha il diritto di non essere sottoposto a una decisione basata unicamente su di un trattamento automatizzato.

- Non discriminazione algoritmica: la legittimità dell'azione pubblica non è garantita dalla sola presenza di un algoritmo conoscibile e comprensibile, oggetto di controllo e validazione da parte di un funzionario, ma *"occorre che lo stesso non assuma carattere discriminatorio"*; il che potrà essere valutato conoscendo l'algoritmo, ma soprattutto, attraverso il sindacato del giudice tramite le categorie della violazione di legge e dell'eccesso di potere.

³⁴⁰ Corte di Giustizia dell'Unione Europea, Sentenza, SCHUFA Holding (Scoring), Causa C-634/21 <https://eur-lex.europa.eu/legal-content/IT/ALL/?uri=CELEX%3A62021CA0634>

³⁴¹ Corte di Giustizia dell'Unione Europea, Sentenza - 27/02/2025 - Dun & Bradstreet Austria Causa C-203/22, <https://eur-lex.europa.eu/legal-content/IT/ALL/?uri=CELEX:62022CJ0203>

religiose, salute, l'orientamento sessuale ecc.) né informazioni personali provenienti dai social network.

L'intellegibilità dei modelli è un requisito richiesto altresì dalla normativa prudenziale, per tutti i modelli IRB, a prescindere che si basino o meno sull'intelligenza artificiale. In tal senso depongono le previsioni di cui agli articoli 179 (in base al quale le stime dei parametri di rischio devono essere plausibili ed intuitive) e 174 comma 1, lett a (secondo cui le variabili di *input* devono essere una base ragionevole ed efficace per la previsione risultante) del Regolamento 575/2013. La Banca Centrale Europea, nell'aggiornare, nel luglio 2025, la sua guida ai modelli IRB ha declinato la sue aspettative in merito all'uso dell'AI in tale ambito, invitando le banche ad evitare di ricorrere ad approcci di modellizzazione inutilmente complessi; a garantire un bilanciamento tra la complessità delle tecniche di modellizzazione e l'incremento delle performance dei modelli interni basati sul machine learning; nonché a documentare (*rectius* motivare) le ragioni alla base della scelta dei criteri di *rating* in modo tale da consentire ai terzi di comprendere l'assegnazione delle esposizioni alle classi, di replicarle e di valutarne l'appropriatezza³⁴².

La trasparenza della decisione algoritmica non si limita ad applicarsi alla elaborazione di dati personali ed ai modelli prudenziali delle banche, ma è stato elevato dall'AI Act a principio generale di mitigazione del rischio³⁴³ rispetto a quei sistemi di intelligenza artificiale che possono impattare significativamente sulla salute la sicurezza o i diritti fondamentali delle persone. Il principio di trasparenza, il cui rispetto è presupposto per la messa in commercio dei sistemi *high risk* (fra cui rientra il credit scoring algoritmico) richiede infatti che i sistemi siano sorvegliabili, tracciabili e intellegibili ossia sufficientemente trasparenti da

³⁴² European Central Bank, *Guide to internal models*, https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.supervisory_guide202507.en.pdf pp 18 ss

³⁴³ B. Parenzo, "Informazioni significative" nei processi decisionali automatizzati: parla la Corte di Giustizia, in *La nuova giurisprudenza civile commentata*, 4/2025, Wolters Kluwer, pp 974 ss.

consentire ai deployer di «interpretare l'output del sistema e utilizzarlo adeguatamente (art. 13 AI Act)»

Affinché i modelli siano trasparenti è necessario che questi o siano progettati in maniera da essere interpretabili *by default* (si pensi, ad esempio all'uso di un *decision tree*) o, qualora ciò non sia tecnicamente possibile in ragione della tipologia di algoritmo necessaria per ottenere una *performance* migliore, che siano adottati strumenti per rendere spiegabile l'output generato.

Tali strumenti di eXplainable AI (eXAI), fornendo informazioni su come un determinato sistema di intelligenza artificiale abbia elaborato i dati e generato un output, offrono a sviluppatori, utilizzatori, autorità di vigilanza, nonché ai soggetti interessati da una decisione algoritmica, la possibilità di comprendere il “ragionamento” della macchina così da valutarne la correttezza³⁴⁴.

Dal punto di vista informatico, l'analisi condotta in queste pagine, ha rilevato come allo stato della tecnica attuale, il metodo *Shapley Additive Explanation* (SHAP) sia quello ritenuto dalle banche maggiormente efficace per interpretare gli output³⁴⁵ grazie alla sua capacità di analizzare il contributo marginale di ciascuna variabile e di considerare ogni possibile combinazione con le altre così da consentire all'intermediario di valutare quali tipologie di dati influenzino maggiormente l'output del modello³⁴⁶.

Ora va notato come, seppure le normative esaminate, nel richiedere la trasparenza delle decisioni, non impongano per la loro *compliance* l'utilizzo di tecniche di eXAI³⁴⁷ (conformemente al principio di neutralità tecnologica),

³⁴⁴ European Data Protection Supervisor, *TechDispatch on Explainable Artificial Intelligence*, https://www.edps.europa.eu/system/files/2023-11/23-11-16_techdispatch_xai_en.pdf p.9

³⁴⁵ European Banking Authority, *Machine Learning for IRB Models Follow-Up Report from the Consultation on the Discussion Paper on Machine Learning for IRB Models* (EBA/REP/2023/28), p. 11

³⁴⁶ Intesa San Paolo, *Response to discussion on machine learning for IRB models*, <https://www.eba.europa.eu/eba-response/31935?destination=/discussion-paper-machine-learning-irb-models>

³⁴⁷ Il GDPR chiede di fornire «informazioni significative sulla logica utilizzata, nonché l'importanza e le conseguenze previste di tale trattamento per l'interessato»; l'AI Act ricollega la trasparenza all'obbligo di fornire documentazione pertinente e istruzioni per l'uso riguardo il sistema di intelligenza artificiale.

l'interpretabilità dei modelli di intelligenza artificiale più complessi non può prescindere dall'utilizzo di tali strumenti che sono gli unici in grado di garantirne la spiegabilità. L'autoapprendimento degli algoritmi risulta, infatti, incompatibile per sua natura, con la capacità di fornire spiegazioni intelleggibili dall'essere umano³⁴⁸. Per tale motivo dovrebbero essere ammessi i limiti del principio di neutralità e andrebbe riconosciuto che determinati modelli di intelligenza artificiale necessitano di essere affiancati da tecnologie per la spiegazione degli output; non si può infatti sostenere che la regolamentazione e l'attività di vigilanza debbano mantenersi equidistanti sia da soluzioni efficienti che inefficienti³⁴⁹.

Il ripensamento di tale principio non deve tuttavia avere come precipitato l'imposizione di una specifica soluzione tecnologica (ad es. il fare ricorso ad una determinata tecnica di eXAI) attraverso strumenti di *hard law* in quanto il loro complesso iter di formazione e revisione li rende inadatti a permettere un rapido adeguamento per stare al passo con l'innovazione tecnologica. Tuttavia, se una soluzione tecnologica riesce ad imporsi rispetto alle altre per la sua capacità di garantire la spiegabilità del modello (come nel caso di SHAP) l'autorità di vigilanza dovrebbe promuovere tali standard tecnologici mediante strumenti di *soft law* che garantiscono quella flessibilità necessaria ad intercettare le rapide evoluzioni dell'innovazione³⁵⁰ e che proprio nel settore finanziario hanno dato prova di riuscire ad assicurare uniformità e coerenza nella applicazione del diritto UE³⁵¹.

³⁴⁸ J. Burrell, *How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms*, in *Big Data & Society*, 1/2016, p.10

³⁴⁹ M. Rabitti, *Le Regole di Supervisione nel Mercato Digitale: Considerazioni Intorno alla Comunicazione Banca d'Italia in Materia di Tecnologie Decentralizzate nella Finanza e Cripto-Attività*, in *Rivista Trimestrale di Diritto dell'economia*, Supplemento n. 3 al n. 1/2022

³⁵⁰ G. Morbidelli, *Degli Effetti Giuridici della Soft Law*, in *Rivista della Regolazione dei mercati*, Fascicolo 2 (2016)

³⁵¹ M. Ramajoli, *Selfregulation, soft regulation e hard regulation nei mercati finanziari*, in *Rivista della Regolazione dei mercati*, Fascicolo 2 (2016) p. 65

Ma non solo, visti gli interessi in gioco sarebbe opportuno fare un passo ulteriore ed ammettere che siano le stesse autorità che vigilano sulle banche e sull'intelligenza artificiale a farsi *driver* dello sviluppo di soluzioni tecnologiche efficienti, cooperando con l'industria e la ricerca al fine di dare, da un lato, agli operatori una maggiore certezza del quadro normativo e, dall'altro lato, fornire alle autorità spunti e indicazioni per regole efficienti e best practices, che potrebbero migliorare l'attività di controllo³⁵².

In tal senso i "facilitatori dell'innovazione" quali *sandbox* regolamentari ed *innovation hub*, come quello della Banca d'Italia a Milano possono rivelarsi fondamentali per supportare l'adozione delle tecnologie AI da parte del mercato finanziario, facilitando l'adozione di soluzioni in grado di presidiare i rischi e permettere al contempo alle autorità di vigilanza di restare al passo con le innovazioni e di identificare meglio i casi in cui potrebbe essere necessaria una rivalutazione dei perimetri regolamentari³⁵³. Su tale scia si sono peraltro orientate le autorità di vigilanza europee (EBA, EIOPA ed ESMA) che nel loro report congiunto sugli "innovation facilitators"³⁵⁴ hanno raccomandato che gli hub di innovazione e i sandbox normativi siano preferiti dalle autorità di regolamentazione e vigilanza quali soluzioni per incoraggiare l'innovazione nel settore finanziario nell'area economica europea³⁵⁵. In conclusione, dallo studio condotto emerge come l'intelligenza artificiale apporti significativi benefici nel settore della concessione del credito pur portando con sé nuovi rischi, in particolare per gli individui. Tuttavia, nella comparazione fra benefici e rischi emerge come i primi abbiano delle dimensioni tali da compensare i secondi. I

³⁵² A. Perrone, *La nuova vigilanza. Regtech e capitale umano*, in *Banca Borsa Titoli di Credito*, p. 516

³⁵³ <https://www.bancaditalia.it/focus/milano-hub/index.html>

³⁵⁴ European Supervisory Authorities, *Report Update on the functioning of innovation facilitators – innovation hubs and regulatory sandboxes* (ESA 2023-27), <https://www.esma.europa.eu/document/joint-esas-report-innovation-facilitators-2023>

³⁵⁵ EIOPA, *Details on Joint ESAs Report on Innovation Facilitators: innovation hubs and regulatory sandboxes* https://www.eiopa.europa.eu/publications/joint-esas-report-innovation-facilitators-innovation-hubs-and-regulatory-sandboxes_en

rischi non vanno, tuttavia, sottovalutati bensì mitigati; in tale processo di mitigazione assume una fondamentale importanza l'assicurare la trasparenza dei modelli che può realizzarsi o utilizzando modelli interpretabili *by default*, come i *decision trees*, o nel caso in cui l'ottenimento di performance maggiori giustifichi la scelta di modelli non intrinsecamente comprensibili, adottando tecniche di eXAI. Alla luce di quanto esposto, tale esigenza di trasparenza non può essere rimessa esclusivamente alla discrezionalità degli operatori economici, ma deve essere incorporata in un assetto di vigilanza che sia capace di orientare il mercato verso soluzioni efficaci ed efficienti; in questa prospettiva, la neutralità tecnologica non può tradursi in un'indifferenza regolatoria rispetto alla qualità degli strumenti impiegati bensì è necessario che la supervisione privilegi un approccio funzionale e risk-based, fondato su standard flessibili basati sulla soft law e su percorsi cooperativi con ricerca e industria, così da assicurare un aggiornamento continuo delle *best practices* e, al contempo, una maggiore certezza applicativa per gli intermediari.

Solo grazie ad una coregolazione consapevole dei limiti della neutralità e che sia in grado di promuovere trasparenza e accountability, senza cristallizzare specifiche tecnologie, si potranno valorizzare i benefici dell'IA nella concessione del credito rendendoli compatibili con la tutela degli individui e con la stabilità complessiva del sistema finanziario.

Bibliografia

- Agenzia per la Cybersicurezza Nazionale, *Operational Summary 1° semestre 2025*, https://www.acn.gov.it/portale/documents/20119/989383/Operational_Summary_H1_2025_CLEAR.pdf/93dfccd6-695f-5e5c-f166-03f7550b2430?t=1754322719758
- AIFIRM, *Il principio contabile IFRS 9 in banca: la prospettiva del Risk Manager*, Position Paper n. 8, 2016.
- G. A. Akerlof, *The Market for “Lemons”: Quality Uncertainty and the Market Mechanism*, in *The Quarterly Journal of Economics*, 84(3), 1970, pp. 488–500.
- S. Ali, T. Abuhmed, S. El-Sappagh, K. Muhammad, J. M. Alonso-Moral, R. Confalonieri, R. Guidotti, J. Del Ser, N. Díaz-Rodríguez, F. Herrera, *Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence*, in *Information Fusion*, Volume 99, 2023.
- H. J. Allen, *‘Just Say No’ to Machine Learning in IRB Models*, in *Oxford Business Law Blog*, 2021.
- A. Alonso, J. M. Carbó, *Machine Learning in Credit Risk: Measuring The Dilemma Between Prediction and Supervisory Cost*, Banco de España Working Paper, <https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosTrabajo/20/Files/dt2032e.pdf>
- L. Ammannati, G. L. Grieco, *Il credit scoring “intelligente”: esperienze, rischi e nuove regole*, in *Rivista di diritto bancario*, luglio/settembre 2023.
- Associazione Bancaria Italiana, *Come viene valutata l’affidabilità delle imprese con l’Accordo di Basilea*, https://bancaria.it/pozzo/Conoscere_il_rating.pdf
- H. Ayari, R. Guetari, N. Kraïem, *Machine learning powered financial credit scoring: a systematic literature review*, in *Artificial Intelligence Review*, Vol. 59, 2025.

- B. Baesens, *Analytics in a big data world: the essential guide to data science and its Applications*, John Wiley & Sons, 2014.
- Banca Centrale Europea, *Parere della Banca Centrale Europea del 29 dicembre 2021 relativo a una proposta di regolamento che stabilisce regole armonizzate sull'intelligenza artificiale*.
- Banca d'Italia, *Attuazione degli Orientamenti dell'Autorità bancaria europea in materia di concessione e monitoraggio dei prestiti (EBA/GL/2020/06)*, Nota n. 13 del 20 luglio 2021.
- Banca d'Italia, *Milano Hub*, <https://www.bancaditalia.it/focus/milano-hub/index.html>
- Banca d'Italia, *Raccomandazione in materia di modelli IRB ai fini della misurazione del rischio di credito*, <https://www.bancaditalia.it/compiti/vigilanza/normativa/orientamenti-vigilanza/Raccomandazione-IRB-misurazione-rischio-di-credito.pdf>
- Banca d'Italia, *Recepimento della Nuova Regolamentazione Prudenziale Internazionale Metodo Dei Rating Interni per il Calcolo del Requisito Patrimoniale a Fronte del Rischio di Credito*, Documento di Consultazione, 2006.
- V. Barone, *La discriminazione ai tempi dell'intelligenza artificiale: la questione algoritmica*, in T. Casadei, S. Pietropaoli (a cura di), *Diritto e Tecnologie Informatiche*, Wolters Kluwer, 2024.
- R. Bartlett, A. Morse, R. H. Stanton, N. Wallace, *Consumer-Lending Discrimination in the Fintech Era*, NBER Working Paper No. w25943, June 2019.
- M. Bazarbash, *FinTech in Financial Inclusion: Machine Learning Applications in Assessing Credit Risk*, in IMF Working Paper (WP/19/109), 2019, <https://www.imf.org/-/media/files/publications/wp/2019/wpiea2019109.pdf>

- T. Berg, V. Burg, A. Gombović, M. Puri, *On the Rise of FinTechs: Credit Scoring Using Digital Footprints*, in *The Review of Financial Studies*, Volume 33, Issue 7, 2020, pp. 2845–2897.
- E. Bernardini, *La Validazione Qualitativa Dei Sistemi Di Rating*, in *Rivista bancaria*, giugno 2008.
- D. Besanko, A. V. Thakor, *Collateral and Rationing: Sorting Equilibria in Monopolistic and Competitive Credit Markets*, in *International economic review*, vol 28 n.3, 1987, pp. 671–689.
- M. Biasin, S. Cosma, M. Oriani, *La Banca. Istituzione e impresa*, Isedi, 2013.
- M. Bianco, *Intelligenza artificiale nel credit scoring: analisi di alcune esperienze nel sistema finanziario italiano*, Presentazione del QEF in pubblicazione.
- C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
- E. Bonaccorsi di Patti, F. Calabresi, B. De Varti, F. Federico, M. Affinito, M. Antolini, F. Loriggio, S. Marchetti, I. Masiani, M. Moscatelli, F. Privitera e G. Rinna, *Intelligenza artificiale nel credit scoring. Analisi di alcune esperienze nel sistema finanziario italiano*, Questioni di Economia e Finanza (Occasional Papers) n. 721, 2022, https://www.bancaditalia.it/pubblicazioni/qef/2022-0721/QEF_721_IT.pdf
- F. Z. Borgesius, *Discrimination, Artificial Intelligence and Algorithmic*, Directorate General of Democracy, Council of Europe (2018).
- C. Brescia Morra, *Il diritto delle banche*, Il Mulino, Seconda Edizione.
- J. Bridges, D. Gregory, M. Nielsen, S. Pezzini, A. Radia, M. Spaltro, *The impact of capital requirements on bank lending*, Working Paper No. 486, Bank of England, 2014, <https://www.bankofengland.co.uk/-/media/boe/files/working-paper/2014/the-impact-of-capital-requirements-on-bank-lending.pdf>
- Bundesverwaltungsgericht, Sentenza W256 2217011-1 del 23/10/2019.

- J. Burrell, *How the Machine “Thinks”: Understanding Opacity in Machine Learning Algorithms*, in *Big Data & Society*, 1/2016.
- N. Bussmann, P. Giudici, D. Marinelli, J. Papenbrock, *Explainable Machine Learning in Credit Risk Management*, in *Computational Economics* vol. 57, 2021, pp. 203–216.
- V. Caforio, F. Paolucci, *Processi decisionali automatizzati, tutela dei dati personali e del segreto commerciale nell’Unione europea: spunti dal caso “SCHUFA” (causa C-634/21)*, in *Il Diritto industriale*, 3/2025.
- Camera dei Deputati, *Relazione Illustrativa allo Schema di decreto legislativo recante recepimento della direttiva (UE) 2023/2225*, https://documenti.camera.it/apps/nuovoSito/attiGoverno/schedaLavori/getTesto.ashx?leg=XIX&file=0321_F001.pdf
- E. Capobianco, *Intelligenza artificiale e rapporti bancari*, in *Rivista di diritto bancario*, 2024.
- J. A. Carrillo Donaire, *Notifying Authorities and Notified Bodies (Chapter III, Section 4). Standards, Conformity Assessment, Certificates, Registration (Chapter III, Section 5)*, in *The EU regulation on Artificial Intelligence: a commentary* a cura di G. M. Díaz González; A. J. Huergo Lora, Wolters Kluwer Italia, 2025
- G. Cascarino, M. Moscatelli, F. Parlapiano, *Explainable Artificial Intelligence: Interpreting Default Forecasting Models Based on Machine Learning*, in *Questioni di Economia e Finanza della Banca d’Italia* n. 674, 2022.
- F. Ciruolo, *Le valutazioni automatizzate del merito creditizio nel quadro regolatorio europeo. Quale futuro per il credit scoring algoritmico?*, in *Rivista di Diritto Bancario*, Fascicolo I, Anno 2025, Sezione Prima.

- F. Columba, M. Cugliari, S. Di Virgilio, *Credit Risk Assessment with Stacked Machine Learning*, in *Mercati, infrastrutture, sistemi di pagamento* n. 73, Banca d'Italia, 2026.
- Comitato di Basilea per la vigilanza bancaria, *Basilea 3 – Schema di regolamentazione internazionale per il rafforzamento delle banche e dei sistemi bancari*, 2011, https://www.bis.org/publ/bcbs189_it.pdf
- Comitato di Basilea per la vigilanza bancaria, *Convergenza internazionale della misurazione del capitale e dei coefficienti patrimoniali*, 2004, <https://www.bis.org/publ/bcbs107ita.pdf>
- Commissione Europea, *Comunicazione della Commissione: Piano Coordinato sull'Intelligenza Artificiale* (COM 2018/795), 2018.
- Commissione Europea, *Explanatory Memorandum*, <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021PC0206>
- Commissione Europea, *Libro bianco sull'intelligenza artificiale: un approccio europeo all'eccellenza e alla fiducia*, <https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:52020DC0065>
- Consumer Financial Protection Bureau, *Data Point: Credit Invisibles*, 2015, https://files.consumerfinance.gov/f/201505_cfpb_data-point-credit-invisibles.pdf.
- R. Corigliano, *Banca e impresa in Italia: caratteri evolutivi del relationship lending e sostegno dello sviluppo* in Banca Impresa Società, Fascicolo 1, Anno 2007, Il Mulino.
- Corte di Giustizia dell'Unione Europea, Sentenza del 7/12/2023, *SCHUFA Holding (Scoring)*, Causa C-634/21.
- Corte di Giustizia dell'Unione Europea, Sentenza del 27/02/2025, *Dun & Bradstreet Austria*, Causa C-203/22.

- A. Cosentini, O. Pollicino, G. De Gregorio, A. Ermellino, D. Fontanella, I. Nicole, F. Paolucci, I. G. Penco, D. Regoli, S. Tessaro Trapani, *Assessing the Impact of Artificial Intelligence Systems on Fundamental Rights Available*, <https://ssrn.com/abstract=5168579>.
- R. Costi, *L'ordinamento bancario*, Il Mulino, Quinta Edizione, 2012.
- G. S. Crawford, N. Pavanini, F. Schivardi, *Asymmetric Information and Imperfect Competition in Lending Markets*, in *American Economic Review*, vol. 108, no. 7, 2018, pp. 1659–1701.
- K. Crawford, *Atlas of AI: power, politics, and the planetary costs of artificial intelligence*, Yale University Press, 2021.
- D. Cucinelli, M. L. Di Battista, M. Marchese, L. Nieri, *Credit Risk in European Banks: The Bright Side of the Internal Ratings Based Approach*, in *Journal of Banking & Finance*, 93, 2018.
- A. Davola, *Algoritmi decisionali e trasparenza Bancaria*, Utet Giuridica 2020
- D. De Meza, D. C. Webb, *Too Much Investment: A Problem of Asymmetric Information*, in *The Quarterly Journal of Economics*, Volume 102, Issue 2, 1987, pp. 281–292.
- M. De Poli, *Spiegazioni Adeguate*, in *Rivista Trimestrale di Diritto dell'Economia*, Supplemento al n. 4/2024, pp. 201-214
- P. Deriu, S. Racioppi, *Riflessioni in tema di intelligenza artificiale e attività di vigilanza*, in *Quaderni Fintech Consob*, n. 15.
- P. Di Biasi, R. Gnutti, A. Resti, D. Vergari, *Machine Learning for Credit risk: three successful Case Histories*, in *Risk Management Magazine*, Vol. 17, Issue 2, 2022.
- L. Di Brina, *L'esercizio del credito: profili generali*, in U. Urbani (a cura di), *L'attività delle Banche*, Wolters Kluwer Italia, Seconda Edizione, 2019.
- D. W. Diamond, *Financial Intermediation and Delegated Monitoring*, in *The Review of Economic Studies*, 51, no. 3, 1984, pp. 393–414.

- D. W. Diamond, *Financial Intermediation as Delegated Monitoring: A Simple Example*, in *Economic Quarterly*, Summer 1996.
- I. D. Dichev, D. J. Skinner, *Large-Sample Evidence on the Debt Covenant Hypothesis*, in *Journal of Accounting Research*, Vol. 40, No. 4, 2002.
- Dipartimento per il programma di Governo, *Analisi del quadro normativo in materia di Intelligenza artificiale*, <https://www.programmagoverno.gov.it/media/je0lo4i0/focus-ia.pdf>
- F. D’Orazio, *Il credit scoring e l’art. 22 del GDPR al vaglio della Corte di giustizia*, in *La nuova giurisprudenza civile commentata* 2024, n° 2, p. 410-421.
- EIOPA, *Details on Joint ESAs Report on Innovation Facilitators: innovation hubs and regulatory sandboxes*, https://www.eiopa.europa.eu/publications/joint-esas-report-innovation-facilitators-innovation-hubs-and-regulatory-sandboxes_en
- L. Enqvist, *‘Human oversight’ in the EU artificial intelligence act: what, when and by whom?*, in *Law, Innovation and Technology*, 15(2), 2023.
- European Banking Authority, *AI Act: implications for the EU banking and payments sector*, <https://www.eba.europa.eu/sites/default/files/2025-11/d8b999ce-a1d9-4964-9606-971bbc2aaf89/AI%20Act%20implications%20for%20the%20EU%20banking%20sector.pdf>
- European Banking Authority, *Discussion Paper on Machine Learning for IRB Models* (EBA/DP/2021/04).
- European Banking Authority, *Guidelines on Loan Origination and Monitoring*, 2020, <https://www.eba.europa.eu/activities/single-rulebook/regulatory-activities/credit-risk/guidelines-loan-origination-and-monitoring>

- European Banking Authority, *Machine Learning for IRB Models Follow-Up Report from the Consultation on the Discussion Paper on Machine Learning for IRB Models* (EBA/REP/2023/28).
- European Banking Authority, *Orientamenti in materia di governance interna ai sensi della direttiva (UE) 2019/2034*(EBA/GL/2021/14).
- European Banking Authority, *Orientamenti sulla stima della probabilità di default (PD) e delle perdite in caso di default (LGD) e sul trattamento delle esposizioni in stato di default* (EBA-GL-2017-16).
- European Banking Authority, *Report On Big Data And Advanced Analytics*, https://www.eba.europa.eu/sites/default/files/document_library/Final%20Report%20on%20Big%20Data%20and%20Advanced%20Analytics.pdf
- European Central Bank, *Guide to internal models*, https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.supervisory_guide202507.en.pdf
- European Central Bank, *Press Release ECB publishes revised guide to internal models*, <https://www.bankingsupervision.europa.eu/press/pr/date/2025/html/ssm.pr250728~2b36305822.it.html>
- European Data Protection Supervisor, *TechDispatch on Explainable Artificial Intelligence*, https://www.edps.europa.eu/system/files/2023-11/23-11-16_techdispatch_xai_en.pdf
- European Supervisory Authorities (ESA), *Report Update on the functioning of innovation facilitators – innovation hubs and regulatory sandboxes* (ESA 2023-27), <https://www.esma.europa.eu/document/joint-esas-report-innovation-facilitators-2023>
- European Union Agency for Fundamental Rights, *Bias in Algorithms – Artificial Intelligence and*

Discrimination, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2022-bias-in-algorithms_en.pdf

- E. Falletti, C. Gallese, *Credit scoring and transparency between the AI Act and the Court of Justice of the European Union*, AIMMES, 2024.
- Federation of European Risk Management Associations, *The EU's Artificial Intelligence Act*, https://www.ferma.eu/wp-content/uploads/2024/10/FERMA-Policy-Note-EU-AI-Act_v0.2.pdf?utm_source=chatgpt.com
- Financial Stability Board, *Artificial intelligence and machine learning in financial services, market developments and financial stability implications*, 2017, <https://www.fsb.org/uploads/P011117.pdf>
- Financial Stability Board, *The Financial Stability Implications of Artificial Intelligence*, 2024, <https://www.fsb.org/uploads/P14112024.pdf>
- B. Fragasso, *Il regolamento europeo sull'I.A. (AI ACT) e i primi tentativi di regolazione in Italia e all'estero*, in *Intelligenza artificiale e responsabilità penale*, Giappichelli Editore, 2025.
- H. Fraisse, M. Laporte, *Return on Investment on AI: The Case of Capital Requirement*, Banque de France Working Paper n. 809.
- French Banking Federation, *Response to discussion on machine learning for IRB models*, <https://www.eba.europa.eu/eba-response/31906?destination=/discussion-paper-machine-learning-irb-models>
- X. Freixas, J.C. Rochet, *Microeconomics of Banking*, MIT Press, I edizione, 1997.
- G7 Cyber Expert Group, *Statement on AI and Cybersecurity*, 2025, <https://home.treasury.gov/system/files/136/G7-Cyber-Expert-Group-Statement-AI-and-Cybersecurity-2025.pdf>
- F. M. Gabbricci, *La vigilanza sull'intelligenza artificiale in ambito bancario*, in M. Rabitti, F. Bassan (a cura di), *L'applicazione dell'AI Act in Italia e la tutela del*

consumatore. Il ruolo delle autorità indipendenti, <https://romatrepress.uniroma3.it/wp-content/uploads/2025/07/DECA-n.-10-2025-ebook.pdf>

- P. Gaggero, C. A. Valenza, *Le moderne tecniche di credit scoring tra GDPR, disciplina di settore e AI Act*, in *Rivista di Diritto Bancario*, 2024.
- W. G. Gale, *Collateral, Rationing, and Government Intervention in Credit Markets*, in R. G. Hubbard (a cura di), *Asymmetric Information, Corporate Finance, and Investment*, University of Chicago Press, 1990.
- W. G. Gale, *Federal Lending and the Market for Credit*, in *Journal of Public Economics*, 42 (2), 1990, pp. 177–193.
- R. Gallo, *The IRB approach and bank lending to firms*, in *Temi di discussione della Banca d'Italia* n. 1347, 2021.
- L. Gambacorta, Y. Huang, Z. Li, H. Qiu, S. Chen, *Data versus Collateral*, in *Review of Finance*, Volume 27, Issue 2, 2023, pp. 369–398.
- L. Gambacorta, F. Sabatini, S. Schiaffi, *Artificial intelligence and relationship lending*, *Temi di Discussione Banca d'Italia* n. 1476, 2025.
- P. Giudici, E. Raffinetti, *Shapley-Lorenz eXplainable Artificial Intelligence*, in *Expert Systems with Applications*, Vol. 167, 2021.
- Google Cloud, *Che cos'è l'apprendimento non supervisionato*, <https://cloud.google.com/discover/what-is-unsupervised-learning>
- M. Hall, L. van der Maaten, L. Gustafson, M. Jones, A. Adcock, *A Systematic Study of Bias Amplification*, <https://arxiv.org/abs/2201.11706>
- C. Hurlin, C. Pérignon, *Machine Learning and IRB Capital Requirements: Advantages, Risks, and Recommendations*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=44837

- IBM, *Data Poisoning*, <https://www.ibm.com/it-it/think/topics/data-poisoning>
- B. Indovina, *Informatica, diritto, intelligenza artificiale*, Egea S.p.a., 2024.
- Institute of International Finance, *Machine learning in credit risk*, https://www.iif.com/Portals/0/Files/content/Research/iif_mlcr_2nd_8_15_19.pdf
- Intesa San Paolo, *Response to discussion on machine learning for IRB models*, <https://www.eba.europa.eu/eba-response/31935?destination=/discussion-paper-machine-learning-irb-models>
- International Monetary Fund, *Fintech: The Experience So Far*, Working Paper, 2019, <https://www.imf.org/-/media/files/publications/wp/2019/wpiea2019109.pdf>
- J. Jagtiani, C. Lemieux, *The roles of alternative data and machine learning in fintech lending: Evidence from the LendingClub consumer platform*, Federal Reserve Bank of Philadelphia Working Papers.
- P. Knab, S. Marton, U. Schlegel, C. Bartelt, *Which LIME should I trust? Concepts, Challenges, and Solutions*, in R. Guidotti, U. Schmid, L. Longo (eds), *Explainable Artificial Intelligence. xAI 2025*, Springer, Cham.
- M. Lembo, *L'intelligenza artificiale e i contratti bancari: il caso del credit scoring*, in *I contratti: rivista di dottrina e giurisprudenza*, n. 4/2025, Wolters Kluwer.
- Y. Li, W. Chen, *A Comparative Performance Assessment of Ensemble Learning for Credit Scoring*, in *Mathematics*, 8(10), 2020.
- Z. C. Lipton, *The Mythos of Model Interpretability*, in *Communications of the ACM*, Vol. 61, Issue 10, 2018
- Y. Lu, L. Yang, B. Shi, et al., *A novel framework of credit risk feature selection for SMEs during industry 4.0*, in *Ann Oper Res* 350, 2025.

- G. Malgieri, G. Comandé, *Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation*, in *International Data Privacy Law*, vol. 7, Issue 3, (2017).
- N. G. Mankiw, *The Allocation of Credit and Financial Collapse*, in *Quarterly Journal of Economics*, 101 (3), 1986, pp. 455–470.
- A. Mantelero, G. Resta, G. M. Riccio, *Intelligenza Artificiale Commentario*, Wolters Kluwer, I Edizione, settembre 2025.
- P. Marano, *L'Intelligenza Artificiale nella Realizzazione e Distribuzione dei Prodotti Assicurativi: Verso una Compliance Integrata tra AI Act, IDD e POG*, in *Rivista di Diritto Bancario*, Fascicolo II, 2025, 275-309
- F. Mattassoglio, *Algoritmi e Regolazione, Circa i Limiti del Principio di Neutralità Tecnologica*, in *Rivista della Regolazione dei mercati*, Fascicolo 2, 2018.
- F. Mattassoglio, *La Corte di giustizia europea, algoritmi e credit scoring. L'apertura del vaso di Pandora delle società che si "limitano" a elaborare gli scoring*, in *Dialoghi di diritto dell'Economia*, gennaio 2025.
- M. McFadden, K. Jones, E. Taylor, G. Osbornpp, *Harmonising Artificial Intelligence: The Role of Standards in the EU AI Regulation*, 2021, Oxford Commission on AI and Good Governance, 2021
- F. S. Mishkin, S. G. Eakins, E. Beccalli, *Istituzioni e mercati finanziari*, Pearson, 9 edizione, (2019)
- F. S. Mishkin, *The Economics of Money, Banking, and Financial Markets*, Pearson Education Limited, XIII Edizione Globale, 2022.
- G. Morbidelli, *Degli Effetti Giuridici della Soft Law*, in *Rivista della Regolazione dei mercati*, Fascicolo 2, 2016.
- G. Nicodano, *Asimmetria Informativa*, in *Dizionario di Economia e Finanza*, Treccani, 2012.

- OECD, *AI in Business and Finance 2021: OECD Business and Finance Outlook*, 2021, https://www.oecd.org/content/dam/oecd/en/publications/reports/2021/09/oecd-business-and-finance-outlook-2021_377c2c18/ba682899-en.pdf
- O. Fernández, *Accuracy, Robustness and Cybersecurity (Article 15)*, in *The EU regulation on Artificial Intelligence: A commentary*, a cura di A. J. Huergo Lora (ed.), G. M. Díaz González, (pp. 267-288), Wolters Kluwers Italia
- D. Palazzo, *Il diritto alla spiegazione delle decisioni automatizzate. Trasparenza e tutela del segreto commerciale alla luce del GDPR e dell'AI Act*, in *Rivista Diritto e Società*.
- F. Palmiotto, *When Is a Decision Automated? A Taxonomy for a Fundamental Rights Analysis*, in *German Law Journal*, 2024 pp. 1–27.
- F. Panetta, P. Angelini, et al., *La prociclicità del sistema finanziario: lezioni dalla crisi*, *Questioni di Economia e Finanza* n. 44, 2009.
- C. Panigutti, R. Hamon, I. Hupont, D. F. Llorca, D. F. Yela, H. Junklewitz, S. Scalzo, G. Mazzini, I. Sanchez, J. S. Garrido, E. Gomez, *The role of explainable AI in the context of the AI Act*, in *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)*.
- F. Paolucci, O. Pollicino, *Diritti fondamentali Valutazione d'impatto per banche, assicurazioni, enti o servizi pubblici*, in *Norme & Tributi, Il Sole 24 Ore*, 30 ottobre 2025.
- B. Parenzo, *“Informazioni significative” nei processi decisionali automatizzati: parla la Corte di Giustizia*, in *La nuova giurisprudenza civile commentata*, 4/2025, Wolters Kluwer.
- F. Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information*, Harvard University Press, 2015.
- U. Patroni Griffi, *I covenants finanziari. Note introduttive*, in *Rivista di Diritto Societario*, Fascicolo 3, 2009.

- S. Pellegrini, R. Pupella, *Competenze suddivise tra le Ue e le autorità degli Stati Membri*, in *Norme & Tributi*, Il Sole 24 Ore, 30 ottobre 2025.
- A. Perrone, *La nuova vigilanza. Regtech e capitale umano*, in *Banca Borsa Titoli di Credito*, p. 516.
- V. Pietrella, S. Racioppi, *Il credit scoring e la protezione dei dati personali: commento alle sentenze della Corte di giustizia dell'Unione europea del 7 dicembre 2023*, in *Rivista Italiana di Informatica e Diritto*, Fascicolo 1, 2024.
- P. Pikamäe, *Conclusioni dell'Avvocato Generale nella Causa C-634/21*, <https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:62021CC0634>
- M. Rabitti, *Credit scoring via machine learning e prestito responsabile*, in *Rivista di Diritto Bancario*, Fascicolo I, Anno 2023.
- M. Rabitti, *Le Regole di Supervisione nel Mercato Digitale: Considerazioni Intorno alla Comunicazione Banca d'Italia in Materia di Tecnologie Decentralizzate nella Finanza e Cripto-Attività*, in *Rivista Trimestrale di Diritto dell'economia*, Supplemento n. 3 al n. 1/2022.
- M. Rabitti, F. Bassan, *L'applicazione dell'AI Act in Italia e la tutela del consumatore. Il ruolo delle autorità indipendenti*, Roma TrE-Press, 2025.
- M. Ramajoli, *Selfregulation, soft regulation e hard regulation nei mercati finanziari*, in *Rivista della Regolazione dei mercati*, Fascicolo 2, 2016.
- M. T. Ribeiro, S. Singh, C. Guestrin, *Model-Agnostic Interpretability of Machine Learning*, (2016) DOI:10.48550/arXiv.1606.05386
- M. T. Ribeiro, S. Singh, C. Guestrin, *"Why Should I Trust You?": Explaining the Predictions of Any Classifier*, in *Proceedings of the 22nd ACM SIGKDD*, 2016.
- C. Rudin, *Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead*, in *Nature Machine Intelligence* 1, 2019.

- H. Sadok, F. Sakka, M. E. H. El Maknouzi, *Artificial intelligence and bank credit analysis: A review*, in *Cogent Economics & Finance*, 2022.
- D. U. Santosuosso, *Covenants finanziari: rischio d'impresa e responsabilità gestionali*, in *Rivista di Diritto Societario*, Fascicolo 3, 2009.
- S. Sapienza, M. Palmirani, Algorithmic Knowability: A Unified Approach to Explanations in the AI Act, in R. Guidotti, U. Schmid, L. Longo, *Explainable Artificial Intelligence, Communications, Computer and Information Science*, vol 2576. Springer 185-209
- Servizio Studi della Camera dei deputati, *Dossier relativo al recepimento della direttiva (UE) 2023/2225, relativa ai contratti di credito ai consumatori*.
- L. Shapley, *A Value for n-Person Games*, in Kuhn, H. and Tucker, A. (Eds.), *Contributions to the Theory of Games II*, Princeton University Press, 1953.
- D. Schnurr, *Effective Implementation of Requirements for High-Risk AI Systems Under the AI Act: Transparency and Appropriate Accuracy*, Centro on Regulation in Europe (2025).
- K. Shrishak, *AI-Complex Algorithms and effective Data Protection Supervision - Bias evaluation pp. 5 ss.*, https://www.edpb.europa.eu/system/files/2025-01/d1-ai-bias-evaluation_en.pdf
- G. Siani, *Cyber Risk e Sistema finanziario*, Intervento al convegno "Sadibaquarantotto", Milano, 1° aprile 2025.
- D. Siclari, R. Lo Conte, *Informazioni da Inserire nel Contratto di Credito*, in *Rivista Trimestrale di Diritto dell'Economia*, Supplemento al n. 4 /2024.
- N. Siddiqi, *Intelligent Credit Scoring: Building and Implementing Better Credit Risk Scorecards*, Second Edition, John Wiley & Sons, 2017.
- E. Stanghellini, *Introduzione ai Metodi statistici per il credit scoring*, Springer, 2009.

- J. E. Stiglitz, A. Weiss, *Credit Rationing in Markets with Imperfect Information*, in *The American Economic Review*, Vol. 71, No. 3, 1981.
- H. Suresh, J.V. Guttag, *A Framework for Understanding Unintended Consequences of Machine Learning*, 2021, <https://dl.acm.org/doi/10.1145/3465416.3483305>
- R. S. Sutton, A. G. Barto, *Reinforcement Learning: An Introduction*, The MIT Press, Second Edition, 2018.
- The Guardian, *Demis Hassabis: AI future 10 times bigger than industrial revolution and 10 times faster*, 4 agosto 2025, <https://www.theguardian.com/technology/2025/aug/04/demis-hassabis-ai-future-10-times-bigger-than-industrial-revolution-and-10-times-faster>
- L. Thomas, J. Crook, D. Edelman, *Credit Scoring and Its Applications*, SIAM, Second Edition, 2017.
- Treccani, *Prompt*, Vocabolario on line, <https://www.treccani.it/vocabolario/prompt/>
- E. Troisi, *Decisione Algoritmica, Black-Box e Ai Etica: il Diritto di Accesso Come Diritto a Ottenere una Spiegazione*, in *Jus Civile*, Fascicolo 4, 2022.
- Ufficio Legislativo Economia, *Analisi Tecnico Normativa allo schema di decreto legislativo recante il recepimento della direttiva (UE) 2023/2225*, https://documenti.camera.it/apps/nuovoSito/attiGoverno/schedaLavori/getTesto.ashx?leg=XIX&file=0321_F001.pdf
- C. A. Valenza, *Obbligo di Valutazione del Merito Creditizio del Consumatore e Regole in Materia di Banche Dati*, in *Rivista Trimestrale di Diritto dell'Economia*, Supplemento al n. 4 /2024.
- F. Vella, G. Bosi, *Diritto ed Economia di Banche e Mercati Finanziari*, Il Mulino, Seconda Edizione, 2023.

- Verwaltungsgericht Wien, Sentenza VGW-101/042/791/2020-44 dell'11/02/2022.
- S. Wachter, B. Mittelstadt, L. Floridi, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, in *International Data Privacy Law*, Volume 7, Issue 2, 2017.
- M. Wallis, K. Kumar, A. Gepp, *Credit Rating Forecasting Using Machine Learning Techniques*, in Z. Sun (Ed.), *Managerial Perspectives on Intelligent Big Data Analytics*, IGI Global, 2019.
- World Bank, *Credit scoring approaches guidelines*, <https://thedocs.worldbank.org/en/doc/935891585869698451-0130022020/original/CREDITSCORINGAPPROACHESGUIDELINESFINALWEB.pdf>
- C. Yuk-Shee, G. Kanatas, *Asymmetric Valuations and the Role of Collateral in Loan Agreements*, in *Journal of Money, Credit and Banking*, vol. 17, no. 1, 1985.
- S. Zedda, *Credit scoring: Does XGboost outperform logistic regression? A test on Italian SMEs*, in *Research in International Business and Finance*, Vol 70, Part B, 2024.
- G. Ziccardi, *Una lettura dell'Artificial Intelligence Act: Norme, Etica, Adempimenti, Attuazione*, in *Intelligenza Artificiale Diritto, Giustizia, Economia ed Etica*, a cura di F. Basile, et al., G. Giappichelli Editore.