

Machine learning per il rischio di credito: quale ruolo nei modelli regolamentari?

Machine learning for credit risk: what role in regulatory models?

Paolo Di Biasi, Rita Gnutti, Intesa Sanpaolo | Andrea Resti, Università Bocconi e Crif | Daniele Vergari, Crif

Keywords

Machine learning, rischio di credito, banche, modelli regolamentari

Jel codes

G21, G28, G32, O33

L'uso del machine learning per l'analisi del rischio di credito ha suscitato un vivace dibattito. Tale etichetta racchiude tecniche estremamente diverse, che non dovrebbero essere trattate allo stesso modo: gli approcci più consolidati e maturi garantiscono infatti un'elevata trasparenza e bassi rischi di eccessivo adattamento (overfitting). Questi modelli possono essere utilizzati per finalità di calcolo del capitale regolamentare, ma rispondono anche a finalità più ampie: ridurre i costi, accorciare i tempi di risposta, rafforzare il monitoraggio, migliorare i canali distributivi e rafforzare le relazioni con i clienti. Scoraggiarne l'uso può porre le banche in condizione di svantaggio rispetto agli operatori innovativi, a tutto detrimento della reale efficacia della vigilanza.

The use of machine learning (ML) for credit risk analysis has sparked a lively debate. ML refers to a set of heterogeneous techniques, and one should not treat them all in the same way, since the most consolidated and stable approaches guarantee a high level of transparency and a low risk of overfitting. They can be used for regulatory capital calculation purposes, but may also serve broader purposes: reducing costs, shortening response times, strengthening credit monitoring, improving distribution channels and cementing customer relationships. Discouraging their use can put banks at a disadvantage vis à vis their most innovative competitors, making supervision less effective.

I. Premessa

L'uso dei modelli di machine learning (ML) nelle banche ha suscitato un vivace dibattito tra studiosi e practitioner; nei mesi scorsi, anche l'Autorità bancaria europea (Eba) ha pubblicato un documento di discussione sull'uso del ML per i modelli basati sui rating interni (Irb). Il documento (European Banking Authority, 2021) osserva che, nonostante il loro potere predittivo, i modelli di ML sono stati incorporati nei modelli Irb in modo lento e graduale, a causa della loro complessità e della mancanza di una comprensione comune delle loro principali caratteristiche e delle sfide insite nel loro utilizzo per finalità normative.

Ciò ha portato l'Eba a valutare una serie di raccomandazioni per garantire che i modelli di ML siano conformi alle disposizioni prudenziali, e in particolare alcuni requisiti di

alto livello sulla qualità dei dati e sull'interpretazione del modello. Con riferimento al primo tema, è richiesto che l'utilizzo di dati non strutturati non indebolisca la capacità degli enti di garantire accuratezza, completezza e rappresentatività delle informazioni utilizzate per alimentare i modelli. Per quanto riguarda l'interpretabilità, l'organo di gestione e l'alta dirigenza dovrebbero avere una buona conoscenza dei modelli, compresi i loro fattori chiave e il loro impatto sui parametri di rischio; per raggiungere questo obiettivo, si dovrebbe evitare la complessità, quando non necessaria, ricercando un equilibrio tra le prestazioni di un modello e la sua «spiegabilità» (explainability).

Avendo implementato modelli di ML nelle banche per diversi anni, proviamo a inserirci nel dibattito con questo scritto, che riporta i principali risultati di un lavoro più ampio redatto in lingua inglese (Di Biasi et al., 2022). A tal

Questo articolo riprende alcuni temi sviluppati nel paper *Machine Learning for Credit Risk Management and Irb Models: Lessons from success Case Histories*, scritto con Angelo Basile, Fiorella Bernabei, Cristina Caprara, Dario Cavarero, Mattia Marigliano, Roberta Ranaldi e Marco Vignolo (<https://www.crif.it/ricerche-e-pubblicazioni/altre-risorse-e-ricerche/2022/marzo/l-applicazione-di-tecniche-di-ml-al-credit-risk-management-e-ai-modelli-irb/>). Siamo grati per i commenti a Simone Casellina (European Banking Authority), Francesco Cannata (Banca d'Italia) e a tutti i partecipanti a un seminario di presentazione del paper all'Eba tenuto il 17 marzo 2022.

fine, nel paragrafo 2 discutiamo delle principali opportunità e rischi associati ai modelli di ML, fornendone una sommaria classificazione e illustrandone i possibili benefici e le potenziali insidie; esaminiamo quindi nel paragrafo 3 i principali vincoli normativi al loro utilizzo nei sistemi di rating Irb. Il paragrafo 4 conclude.

2. Modelli di ML: opportunità e svantaggi

Una rapida foto di gruppo. Come notato da Nagy (2018), il ML riguarda tecniche per fornire ai computer la capacità di apprendere¹ come risolvere un problema senza essere preventivamente stati programmati in modo specifico. Il ML viene talvolta confuso con l'intelligenza artificiale (AI), ovvero lo sviluppo di sistemi informatici per svolgere compiti che normalmente richiedono l'intelligenza umana; quest'ultima, tuttavia, si riferisce a un concetto più ampio e include tecniche che non si qualificano come machine learning (si pensi ad esempio a un programma per computer che gioca a scacchi perché le regole sono state codificate da un esperto e non già apprese dalla macchina osservando una serie di partite a scacchi).

Il ML può essere supervisionato (quando un esperto umano «etichetta» i dati, arricchendo ogni osservazione con informazioni su una variabile-obiettivo) o non supervisionato (quando la macchina analizza e raggruppa set di dati non etichettati, mettendo a punto i propri schemi di classificazione senza l'intervento umano). Comprende una serie di diverse tecniche², tra cui principalmente le seguenti:

- analisi di regressione. Può essere vista come un semplice modello di ML e arricchita con varianti (ad esempio, la regressione «ridge» e «lasso») per migliorare la capacità del modello di gestire dati esterni al campione;
- alberi decisionali. Combinano una sequenza di test sugli attributi del campione (rami) al fine di etichettare ogni osservazione attraverso una regola di classificazione composita;
- random forest. Algoritmi che uniscono i risultati di più alberi decisionali (ciascuno basato su un sottoinsieme casuale di caratteristiche e/o osservazioni) per ottenere una strategia di classificazione più solida. Sebbene siano rela-

tivamente semplici, si è riscontrato che funzionano in modo soddisfacente in un'ampia gamma di situazioni (Fernández-Delgado et al., 2014);

- nearest neighbour. Algoritmi che classificano ogni osservazione (o forniscono una stima quantitativa per qualche variabile-obiettivo) osservando i k individui a essa più vicini (cioè più simili) all'interno di uno o più campioni di addestramento³;
- k-means. Tecnica di clustering in base alla quale le osservazioni sono suddivise in k «grappoli» (cluster), in modo che ogni osservazione appartenga al cluster con la media più vicina;
- algoritmi di boosting. Sono meta-algoritmi utilizzati per migliorare un modello ML basati su alcune «funzioni di perdita» che misurano il costo degli errori commessi dalle precedenti soluzioni candidate. Includono diverse varianti, come gradient boosting e XGBoost⁴;
- reti neurali artificiali. Sistemi di unità connesse (nodi) organizzate in livelli (uno strato di input che riceve informazioni dall'esterno, uno strato di output che fornisce risultati all'utente e uno o più strati intermedi). A seconda delle informazioni ricevute, ogni nodo trasmette segnali ad altri nodi negli strati successivi, sulla base di funzioni calibrate in modo iterativo al fine di fornire in output una classificazione corretta;
- deep learning. Modelli costituiti principalmente da reti neurali artificiali con più livelli intermedi, nonché da altri algoritmi complessi come «deep belief networks» e «deep reinforcement learning».

Questo rapido elenco (necessariamente incompleto) mostra come i modelli di ML siano altamente eterogenei tra loro: alcuni molto vicini (o addirittura sovrapponibili) alle tecniche statistiche tradizionali, altri che consentono soluzioni più flessibili e altrettanto complesse (figura 1). Ne consegue che le linee guida emanate dalla vigilanza dovrebbero rifuggire un approccio «a taglia unica» (one size fits all) che impone gli stessi vincoli ad algoritmi molto diversi.

Machine learning: vantaggi per le banche. Il settore dei servizi finanziari sta cambiando rapidamente (figura 2) a causa, tra l'altro dei seguenti fenomeni:

¹ Un programma per computer apprende da un'esperienza (E), dati un compito (T) e una misura delle sue prestazioni (P), se le sue prestazioni su T, misurate da P, migliorano con l'esperienza E (Mitchell, 1997).

² Cfr: Eba (2020) per una breve discussione e Breeden (2021) per una tassonomia degli algoritmi ML applicati al rischio di credito.

³ Cfr: Cover e Hart (1967). Non esistono metodi statistici predefiniti per trovare il valore ottimale di k . La scelta può dipendere dalla dimensione del campione: all'aumentare della dimensione del campione, possono essere scelti valori di k più elevati. In generale, la scelta di un valore basso per k porta a limiti decisionali instabili, mentre un valore k elevato porta a previsioni più accurate e stabili.

⁴ Cfr: ad esempio, Chen et al. (2019).

Figura 1

Esempi di modelli MI con diversi mix tra interpretabilità e complessità/prestazioni

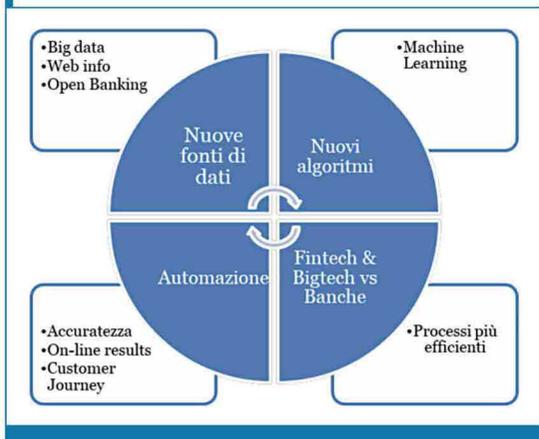


- digitalizzazione dei processi creditizi, dalla «presa a bordo» del cliente fino al recupero crediti;
- normative orientate all'open banking, che offrono accesso a una gamma di dati senza precedenti;

- crescente concorrenza da parte di operatori non bancari che sfidano gli operatori storici e promuovono l'innovazione;
- normative più prescrittive (che incrementano gli oneri legati alla compliance) e requisiti patrimoniali più severi, che spingono le banche ad aumentare il reddito operativo per sostenere costi più elevati.

Figura 2

Tendenze recenti nel settore dei servizi finanziari



Di fronte a tali minacce e opportunità, le tecniche tradizionali potrebbero non fornire risposte adeguate. Il ML può invece essere utilizzato dalle banche per:

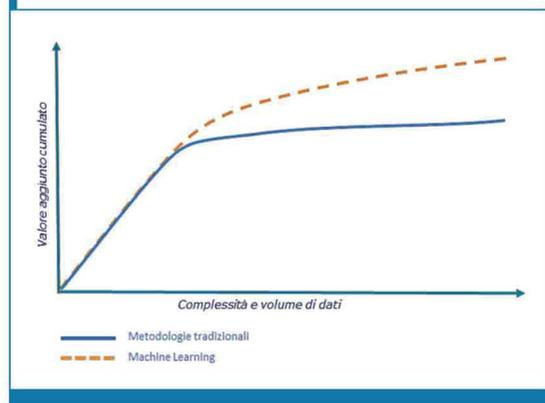
- espandere i propri database attraverso fonti alternative che forniscono informazioni non strutturate e semi-strutturate, come i dati sulle transazioni (comprese le causali descrittive associate ai singoli pagamenti) e la digital footprint dei clienti (nel rispetto delle norme sulla privacy);
- gestire i dati multi-dimensionali selezionando automaticamente le caratteristiche più significative o creando meta-variabili che riassumono le informazioni più rilevanti (garantendo nel contempo qualità e tracciabilità);
- migliorare le prestazioni dei sistemi interni di scoring/ra-

- ting attraverso questi dati e tramite algoritmi non standard capaci di catturare relazioni altamente non lineari;
- sfruttare l'interazione tra nuovi dati e modelli innovativi che imparano dai propri errori, garantendo nel contempo che le stime del rischio di credito rimangano stabili, solide e interpretabili;
 - conseguire così un minor costo del rischio lungo tutto il ciclo di vita del credito, migliorando la stabilità finanziaria;
 - ampliare e diversificare la propria rete distributiva grazie alla capacità di fornire risposte in tempo reale alle richieste di credito ricevute attraverso canali virtuali a basso costo (ad esempio, siti di e-commerce). Così facendo, accrescere i volumi e ridurre le spese operative, aumentando la redditività e rimanendo competitive di fronte alla concorrenza di operatori non bancari;
 - fornire una pronta risposta a scenari imprevisti (ad esempio, pandemia di Covid-19) che alterano i driver del rischio di credito e richiedono un cambio di paradigma al fine di identificare e supportare i debitori migliori.

In poche parole (figura 3), il ML può efficacemente innescare un aumento del potere esplicativo e dell'efficacia dei modelli di una banca in presenza di informazioni complesse e non strutturate (Jones et al., 2015); inoltre, fornisce prestazioni superiori se utilizzato su dati altamente granulari, la cui ric-

Figura 3

Profilo a valore aggiunto dei modelli tradizionali e dei modelli basati su ML



chezza potrebbe non essere completamente sfruttata dai modelli tradizionali. In tal modo, le decisioni sui prestiti possono diventare più rapide, accurate e automatizzate, consentendo lo sviluppo di nuove attività a basso costo, la riduzione delle perdite e l'incremento della redditività del capitale investito. La tavola 1 riporta una selezione di casi reali⁵ in cui sono state utilizzate tecniche di ML per misurare il rischio di credito e migliorare i relativi processi di gestione, sintetizzando le aree di applicazione e le principali caratteristiche di ciascun progetto.

Come si vede, non tutti si riferiscono direttamente all'uso del ML in ambito Irb. Solo i primi due, infatti, riguardano direttamente il potenziamento e la validazione dei sistemi di rating regolamentari, mentre i successivi affrontano obiettivi gestionali come sistemi di allerta precoce, controlli interni e erogazione del credito. Anche questi ultimi, tuttavia, si trovano ad affrontare gli stessi problemi (ad esempio, qualità dei dati e interpretabilità) che devono essere risolti per utilizzare il ML ai fini Irb.

Possibili insidie nei modelli di ML. Gli algoritmi di ML soffrono anche di alcune potenziali carenze, che vanno affrontate con attenzione per ridurre al minimo il rischio di conseguenze indesiderate⁶. In particolare:

- i modelli possono dipendere eccessivamente dal campione utilizzato e rivelarsi instabili nel tempo (funzionando in modo insoddisfacente se applicati a nuovi dati). Questo overfitting è un noto svantaggio dei modelli ML e esistono diverse metodologie per circoscriverne gli effetti: tra le più note la k-fold cross-validation, il k-fold random subsampling, gli approcci leave-one-out e jack-knife (Berrar, 2019);
- i modelli possono diventare rapidamente obsoleti e richiedere una ricalibrazione/ristima per continuare a offrire prestazioni elevate. Una possibile soluzione a questo problema è il cosiddetto «apprendimento continuo» (Liu, 2017) in cui i nuovi dati vengono costantemente utilizzati per riqualificare il modello, aggiornarne i parametri o addirittura modificarne la struttura. Tale approccio può portare a prestazioni migliori quando i fattori di rischio cambiano rapidamente, poiché filtra i dati e conserva solo

⁵ Tali casi sono stati analizzati in dettaglio in Di Biasi et al. (2022).

⁶ Questo elenco di potenziali debolezze nei modelli di ML può anche essere visto come un catalogo dei costi aggiuntivi affrontati dalle Autorità di vigilanza durante la convalida di algoritmi innovativi. Come notato da Alonso e Carbo (2020), tali costi di supervisione dovrebbero essere confrontati con i vantaggi dei modelli di ML così come emergono dalla letteratura accademica.

Tavola 1

Cinque esempi di applicazione del MI al rischio di credito

Ambito	Benefici	Dati e algoritmi	Tecniche di interpretabilità	Sfide principali
Un nuovo modello Irb rivolto alle Pmi al dettaglio	Offre ai clienti un'esperienza digitale completa, concentrandosi su standard di digitalizzazione elevati e traendo vantaggio dalla Psd2 e dal Gdpr. Può agire in contesti di mercato inattesi, ad esempio post Covid-19	Dati transazionali, carte di credito, Pos e web sentiment. Alberi decisionali, random forest e gradient boosting	Grafici di dipendenza parziale, aspettativa condizionale individuale, Lime e Shap	Assumere nuovo personale, investire nell'It, rivedere le normative, implementare un framework di convalida <i>ad hoc</i>
Un challenger model per validare il modello Irb per le Pd al dettaglio	Eseguire la prima convalida (per ottenere l'autorizzazione di vigilanza) e la convalida continua (per evidenziare tempestivamente eventuali problematiche emerse dal modello della banca)	Dati transazionali innovativi basati sui conti correnti dei mutuatari. Alberi decisionali con extreme gradient boosting (Xgb)	Analisi dell'importanza delle caratteristiche (feature importance analysis)	Tempistica del progetto e scelta di un compromesso tra una challenge completa e la necessità di garantire la comparabilità dei risultati
Un sistema di allerta precoce basato su dati transazionali	Capacità di utilizzare i dati transazionali per identificare nuovi modelli di informazioni	Singole transazioni combinate con informazioni preesistenti. Random forest, Xgb e reti neurali	Analisi dell'importanza delle caratteristiche, Shap, Lime e OptiLime di Crif	Selezione delle caratteristiche, sviluppo del modello e interpretabilità
Controlli di secondo livello su un portafoglio di Pmi al dettaglio	Fornisce agli analisti del credito, che rivedono le pratiche di fido, criteri di campionamento migliori e (basati sul rischio) informazioni aggiuntive. Consente l'analisi a livello di portafoglio	Storia del conto corrente, rapporti finanziari, liquidità e titoli detenuti, trigger Ifrs 9 e prospettive del settore. Xgb a un periodo	Analisi dell'importanza delle caratteristiche e Shap	Selezione delle caratteristiche, usabilità (fornisce agli analisti del credito una «spiegazione» dei risultati del modello)
Ottimizzazione delle politiche creditizie	Automatizza ulteriormente le «politiche di credito» utilizzate per tradurre gli score creditizi in decisioni operative. Riduce i costi e aumenta la velocità	Score creditizi, decisioni operative e loro performance ex post, 30 indicatori utilizzati nelle politiche creditizie della banca. Random forest	Analisi dell'importanza delle caratteristiche e Lime	Ottimizzazione degli iperparametri

le informazioni più importanti; la spinta verso una maggiore automazione, tuttavia, può portare a una perdita di controllo sul modello. In alternativa è possibile basarsi sulla supervisione degli esperti, riqualificando i modelli quando mostrano un calo delle prestazioni e/o si verificano cambiamenti importanti nei dati di input;

- i controlli di qualità per dati non strutturati (ad esempio, stringhe di testo in linguaggio naturale) possono rivelarsi più complessi rispetto a quelli per dati costituiti da record e campi ordinatamente organizzati. Tuttavia, essi devono essere progettati, documentati, aggiornati ed eseguiti su base regolare, al fine di garantire che le informazioni disponibili soddisfino standard adeguati (Cai e Zhu, 2015). Poiché un

modello è valido solo quanto le informazioni utilizzate per costruirlo, è probabile che la qualità dei dati contribuisca maggiormente alla solidità di un algoritmo di MI rispetto all'uso di tecniche di ultimissima generazione;

- i modelli MI possono rivelarsi ostici da interpretare. Anche in presenza di risultati complessivamente corretti, i fattori che ne determinano gli esiti possono risultare ardui da individuare (o variare notevolmente tra gli individui, rendendo difficile identificare un pattern coerente)⁷. Sebbene questo sia certamente un problema, la ricerca sull'interpretabilità del MI ha identificato diverse tecniche per evitare che i modelli diventino «scatole nere». Esse includono i modelli surrogati (modelli più semplici e interpretabili,

⁷ Bracke et al. (2019), trattando un modello di previsione dei default dei mutui, propone di utilizzare tecniche di clustering per arrivare a «gruppi di spiegazione» per diverse aree dello spazio di input.

addestrati per approssimare le previsioni di un algoritmo più complesso e utilizzati per spiegare le relazioni tra i dati), la Lime (Local Interpretable Model-agnostic Explanation, una tecnica che identifica le caratteristiche che contribuiscono maggiormente a ogni singola classificazione attraverso un'approssimazione locale eseguita su versioni leggermente modificate dell'osservazione originale)⁸, i valori di Shapley (una misura di quanto ciascuna caratteristica contribuisce a una previsione⁹, basata su un gran numero di confronti tra coppie di set di caratteristiche alternativi) e la Shap (SHapley Additive exPlanations, un approccio relativamente recente che combina le funzionalità di Lime e Shapley)¹⁰;

- garantire l'interpretabilità di un modello, tuttavia, va oltre lo sviluppo di metriche sofisticate, che da sole non sono sufficienti per evitare situazioni in cui le previsioni individuali possono rivelarsi incoerenti o controintuitive. Piuttosto che affidarsi a soluzioni meramente tecniche, la ricerca dell'interpretabilità dovrebbe dunque basarsi, in primo luogo, su un paradigma di gestione del rischio per cui tutte le funzioni coinvolte della banca sono consapevoli dei principali passaggi attraverso i quali un modello raggiunge le sue valutazioni: quali variabili vengono utilizzate, perché sono state scelte, come vengono trasformate, quale struttura del modello è stata ritenuta preferibile e per quali ragioni. Un simile approccio può rivelarsi costoso, in quanto richiede che lo sviluppo e il test del modello coinvolgano anche le funzioni orientate al business e il middle management; tuttavia, apre la strada a una fruttuosa interazione tra le diverse strutture della banca e tra questa e le Autorità di vigilanza.

3. L'utilizzo del MI nei modelli Irb: il ruolo della regolamentazione

Le regole sui requisiti patrimoniali bancari. L'approccio basato sui rating interni ai requisiti patrimoniali delle banche è disciplinato dal regolamento (Ue) n. 575/2013 (Capital Requirements Regulation, Crr), in particolare dalla Parte Terza, Titolo II, Capo 3 (artt. 142 e ss.), che stabilisce

le condizioni richieste per la convalida di un sistema di rating interno.

Per cominciare, l'articolo 144 richiede che il sistema preveda una valutazione significativa delle caratteristiche del debitore e dell'operazione e stime del rischio accurate e coerenti, e inoltre svolga un ruolo essenziale nella gestione del rischio e nei relativi processi decisionali. Tali principi generali consentono la coesistenza di rating regolamentari e gestionali (cioè utilizzati solo a fini interni di gestione dei rischi): infatti, il Crr non richiede ai rating Irb di effettuare la migliore valutazione possibile, né esclude la presenza di altri sistemi di supporto alle decisioni. È evidente tuttavia che il divario tra modelli normativi e gestionali non può diventare troppo ampio; in tal senso depongono l'articolo 174 (che richiede ai modelli Irb di tipo statistico di utilizzare variabili «ragionevoli ed efficaci») e l'articolo 179 (secondo cui le stime interne dei parametri devono incorporare tutti i dati, le informazioni e i metodi rilevanti). Se usati a fini regolamentari, i modelli di MI devono soddisfare tutti i requisiti previsti dal Crr. Ad esempio, il citato articolo 174 (rivolto ai modelli statistici e ad altri metodi «meccanici») richiede che le variabili di input costituiscano una base ragionevole ed efficace, che i dati utilizzati siano rappresentativi del portafoglio della banca (che deve verificarne l'accuratezza), che l'output sia integrato dalla supervisione umana per individuare e risolvere eventuali errori, che il modello venga regolarmente validato, confrontando previsioni e risultati effettivi. Quest'ultima previsione (insieme all'articolo 185, che richiede «sistemi robusti» per convalidare l'accuratezza e la coerenza dei sistemi di rating) indica come i modelli di MI possano svolgere un ruolo anche come strumenti di convalida attraverso i quali il sistema di rating Irb è sottoposto a controlli incrociati al fine di evidenziare potenziali debolezze in termini di struttura e/o di variabili omesse. I modelli basati sul MI possono inoltre fornire uno strumento aggiuntivo per identificare e correggere le incongruenze nei database utilizzati.

Inoltre, l'articolo 171 impone che le valutazioni del rischio di credito siano debitamente documentate e trasparenti; ciò è ulteriormente rafforzato dall'articolo 175 che, nel caso dei

⁸ Cfr. Ribeiro et al. (2016). Una versione ottimizzata di Lime, che affronta i problemi di instabilità che ne minano l'affidabilità, è stata proposta da (Visani et al., 2020): con questo nuovo approccio (OptiLime [Crrf](#)), la stabilità è massimizzata per ogni possibile livello di «aderenza» (intesa come somiglianza al modello MI originale).

⁹ Il valore di Shapley può «dividere» una previsione individuale tra tutte le caratteristiche che hanno contribuito a determinarla, fornendo così una spiegazione completa del motivo per cui un certo ha ricevuto uno specifico score creditizio. Come notato in Molnar (2019), ciò può renderlo preferibile quando la normativa prevede che i clienti siano titolari di un «diritto a ricevere spiega-

zioni». Per una discussione approfondita dei valori di Shapley, cfr. ad esempio Giudici e Raffinetti (2021), che introduce anche un'estensione dell'approccio originale di Shapley (Shapley-Lorenz) basato sulle scomposizioni di Lorenz.

¹⁰ Lime, Shapley e Shap sono note come tecniche di interpretazione locale, poiché analizzano come cambiano le previsioni dei singoli modelli quando si alterano i dati di input. Sono utilizzati principalmente per produrre una rappresentazione visiva che rifletta il contributo di ciascuna caratteristica (variabile esplicativa) a una singola previsione, quantificando l'importanza delle varie caratteristiche che maggiormente influenzano l'output generato dal modello (ad esempio, una probabilità di default). Al contrario, le tecniche di interpretazione

modelli statistici, richiede esplicitamente uno schema dettagliato delle loro basi matematico/empiriche e un rigoroso processo quantitativo che includa test di performance. L'articolo 189, infine, prevede che tutti gli aspetti rilevanti dei processi di rating siano approvati dagli amministratori e dall'alta dirigenza, che devono acquisire una conoscenza generale del sistema. Il requisito della «interpretabilità» – inteso nella duplice accezione tecnica e organizzativa di cui si è detto in precedenza – risulta dunque espressamente imposto dalla normativa.

Sebbene le richieste del Crr valgano sia per i modelli statistici standard che per i modelli basati sul ML, è chiaro che esse possono rivelarsi particolarmente impegnative nel caso di questi ultimi. Ciò può spiegare perché alcune banche hanno scelto di sperimentare modelli ML solo per scopi gestionali, al di fuori dell'ambito Irb¹¹. Tuttavia, tali modelli non potranno essere esclusi indefinitamente dal sistema di rating regolamentare, posto che i sistemi Irb devono rimanere «significativi», «accurati» ed «essenziali» per i processi di governo del rischio, e devono utilizzare «tutti i dati, le informazioni e i metodi rilevanti».

L'impatto di alcune ulteriori normative. Alcune ulteriori normative (già emanate o *in fieri*) possono condizionare l'uso del ML nelle banche. Come notato dal discussion paper Eba, questo è il caso di alcune regole di conservazione dei dati adombrate dal Gdpr¹²: tale regolamento, infatti, detta un principio generale secondo cui i dati personali devono essere conservati solo per il tempo necessario a raggiungere lo scopo per il quale erano stati raccolti. Ciò potrebbe pregiudicare la capacità delle banche di accedere a banche dati esterne sulle persone fisiche che coprono l'intervallo di tempo minimo (cinque anni per le esposizioni al dettaglio) richiesto dall'articolo 180 del Crr per i rating Irb¹³.

Ulteriori vincoli alla capacità delle banche di implementare modelli di rating basati sul ML (sia all'interno che all'esterno dell'ambito Irb) potrebbero derivare dalla proposta di Artificial Intelligence Act (Aia) presentata nel 2021 dalla Commissione europea¹⁴. Tale normativa inquadra i modelli di scoring per le persone fisiche tra i sistemi di intelligenza che creano un «rischio elevato» per i diritti fondamentali degli individui,

e pertanto li assoggetta a numerosi obblighi¹⁵ in materia di gestione del rischio, governo dei dati¹⁶, documentazione, trasparenza, supervisione, robustezza, accuratezza e sicurezza.

Come rilevato dalla Bce (2021), vi è il rischio che anche i modelli statistici tradizionali come le regressioni o gli alberi decisionali vengano fatti rientrare nell'ambito dell'Aia, rendendo inutilmente onerosa la gestione di tecniche consolidate¹⁷; per questo la Banca Centrale Europea¹⁸ ha espressamente richiesto l'esenzione di tali modelli, segnalando altresì come l'imposizione di vincoli validi solo per i modelli di intelligenza artificiale possa risultare incoerente con la necessità di un approccio «technologically neutral» alla vigilanza bancaria (che promuova la sicurezza e la solidità degli enti indipendentemente dall'applicazione di specifiche tecnologie). La bozza di Aia proposta dalla Commissione solleva due preoccupazioni principali, per quanto riguarda l'uso dei modelli ML ai fini Irb:

- da un lato, la scelta di imporre gli stessi requisiti a tutti i produttori non tiene conto del fatto che le banche sono già attentamente vigilate e che l'utilizzo di modelli innovativi avviene sempre nell'ambito di un dialogo approfondito con le Autorità di vigilanza. I nuovi requisiti dovrebbero dunque essere rivolti principalmente a operatori non bancari (società di Big Data, operatori di e-commerce), in modo da garantire un'effettiva parità di condizioni;
- dall'altro, le nuove regole potrebbero rivelarsi assai più rigide dei criteri stabiliti nel Crr, che mediano in maniera equilibrata tra due esigenze contrastanti: garantire modelli sicuri e robusti e insieme evitare di fissare limiti così stringenti da scoraggiare lo sviluppo di sofisticati sistemi di rating regolamentari.

È auspicabile che simili preoccupazioni vengano affrontate dai co-legislatori impegnati nella messa a punto del testo finale.

4. Conclusioni

Il ML è ancora in forte espansione in termini di perfezionamenti metodologici e applicazioni innovative, e tuttavia vi è ormai un diffuso consenso circa le sue caratteristiche

globale mirano – in modo più «tradiczionale» – a comprendere la relazione tra ciascuna caratteristica (variabile esplicativa) di un modello e la sua variabile obiettivo.

¹¹ Cfr. Institute of International Finance (2019).

¹² General Data Protection Regulation (regolamento Ue 2016/679 del Parlamento europeo e del Consiglio).

¹³ Invero, tale requisito deve valere «almeno per una fonte di dati» utilizzata dal modello, il che lascia margini per un'interpretazione flessibile da parte delle Autorità.

¹⁴ Proposta di regolamento del Parlamento europeo e del Consiglio recante norme armonizzate sull'intelligenza artificiale, Com/2021/206.

¹⁵ Nel caso degli enti creditizi, si ritiene che molti di questi requisiti siano soddisfatti dal rispetto delle regole di governance interna stabilite dalla Direttiva sui requisiti patrimoniali (direttiva 2013/36/Ue del Parlamento europeo e del Consiglio).

¹⁶ Ad esempio, secondo l'articolo 10.3 del progetto di regolamento, i set di dati di formazione, convalida e test devono essere «rilevanti, rappresentativi, privi di errori e completi».

chiave, i suoi lati deboli e i suoi punti di forza. Da tale consenso dovrebbe derivare la messa a punto di standard ottimali per lo sviluppo, la convalida e la supervisione dei modelli basati di MI nelle banche.

Il MI è un'etichetta ad ampio raggio utilizzata per racchiudere un insieme di tecniche estremamente diverse, che non dovrebbero essere trattate allo stesso modo. Ciò è particolarmente vero quando si tratta di interpretabilità e del rischio di creare «scatole nere»: sebbene gli approcci di deep learning siano indubbiamente soggetti a tale rischio, tecniche come alberi decisionali, random forest e Xgb hanno raggiunto la piena maturità e consentono agli sviluppatori di modelli di garantire la trasparenza e evitare l'overfitting. Le tecniche di interpretabilità citate nel paragrafo 2 non devono tuttavia essere viste come un risultato, ma piuttosto come un mezzo per facilitare il dialogo con gli utenti del modello. Un confronto continuo con le parti interessate è fondamentale per garantire che tutte le implicazioni di un nuovo algoritmo siano pienamente comprese prima che esso diventi parte degli strumenti di gestione del rischio. Modelli-pilota, dashboard ed explainers non dovrebbero essere visti come semplici «contentini», usati per rendere accettabile un modello, ma come un passaggio fondamentale nel suo sviluppo e una fonte di arricchimento reciproco per sviluppatori e utenti.

I modelli basati sul MI possono essere utilizzati per perseguire obiettivi assai più ampi del calcolo del patrimonio di vigilanza, riducendo i costi, accorciando i tempi di risposta, rafforzando il monitoraggio, migliorando i canali distributivi e irrobustendo le relazioni con i clienti. Sviluppare tali modelli al di fuori del perimetro Irb può rivelarsi in qualche misura vantaggioso, in quanto consente una maggiore flessibilità e offre alle banche l'opportunità di innovare senza immediatamente sottoporsi ai vincoli normativi e agli standard di vigilanza.

Tuttavia, poiché il gap tra modelli regolamentari e gestionali non può allargarsi a dismisura, i risultati basati sul MI non possono essere esclusi indefinitamente dal perimetro Irb. Il dialogo di vigilanza è fondamentale per trovare un modo per integrarli nei sistemi di gestione dei rischi delle banche, garantendo al contempo che le stime rimangano – per citare il

Crr – «plausibili e intuitive». Le Autorità di vigilanza – Eba, Bce e supervisor nazionali – sembrano sempre più consapevoli di questo obiettivo e sono impegnate con le banche nella ricerca di un adeguato equilibrio tra la capacità del MI di gestire efficacemente nuove fonti di dati e la necessità di garantire trasparenza e stabilità nel tempo ai modelli.

Certamente, la graduale migrazione dei nuovi algoritmi all'interno dei modelli per il capitale regolamentare non deve essere ostacolata da un malinteso senso di «conservatività», che rischia di approfondire lo iato tra il modo in cui i rischi sono percepiti dal management di una banca e il modo in cui essi sono rappresentati dai suoi modelli Irb. Input floor sui parametri di rischio, output floor sui requisiti patrimoniali, nuovi vincoli sui portafogli low default sono solo esempi di questa ricerca di ulteriori ammortizzatori patrimoniali e margini di prudenza, che ha reso sempre più distanti tra loro i modelli gestionali e quelli regolamentari. In un simile contesto, le banche potrebbero trovare poco conveniente arricchire i modelli Irb con nuovi moduli basati sul MI e scegliere di limitare il machine learning alle applicazioni orientate al business, ampliando così il divario con i modelli normativi e portando a duplicare i costi di manutenzione.

Va tenuto presente, infine, che rallentare l'innovazione non è possibile. Sebbene i nuovi modelli comportino chiaramente rischi e debolezze, l'inazione ha i propri costi e pericoli. Scoraggiare l'uso di algoritmi e fonti di dati innovativi può portare le banche a competere con gli operatori non bancari tenendo un braccio legato dietro la schiena, finendo per ampliare il ruolo dello shadow banking system¹⁹ a tutto detrimento della reale efficacia della vigilanza.

Bibliografia

Alonso A., Carbó J.M. (2020), *Machine Learning in Credit Risk: Measuring the Dilemma Between Prediction and Supervisory Cost*, Documentos de trabajo No. 2032, Banco de España, Madrid.

Berrar D. (2019), «Cross-Validation», in *Encyclopedia of Bioinformatics and Computational Biology*, Elsevier, pp. 542–545.

Bracke P., Datta A., Jung C., Sen S. (2019), *Machine Learning Explainability in Finance: An Application to Default Risk Analysis*, Staff Working Paper No. 816, Bank of England, London.

¹⁷ D'altra parte, i modelli basati sul MI utilizzati per scopi di marketing o rilevamento di frodi non si qualificerebbero come sistemi di intelligenza artificiale ad alto rischio, anche se basati su tecnologie «scatola nera» e alimentati con dati personali.

¹⁸ Cfr: European Central Bank, 2021.

¹⁹ Cfr: Resti et al. (2021).

- Breeden J.** (2021), «A Survey of Machine Learning in Credit Risk», in *Journal of Credit Risk*, <https://doi.org/10.21314/JCR.2021.008>.
- Cai L., Zhu Y.** (2015), *The Challenges of Data Quality and Data Quality Assessment in the Big Data Era*, Codata 14, 2. <https://doi.org/10.5334/dsj-2015-002>.
- Chen C., Yokoyama S., Yamashita T., Kawamura H.** (2019), *Application of XGBoost to Credit Scoring*, Ipsj Sig Technical Report, Information Processing Society of Japan.
- Cover T., Hart P.** (1967), «Nearest Neighbor Pattern Classification», in *Ieee Transactions on Information Theory*, 13, pp. 21–27.
- Di Biasi P. et al.** (2022), *Machine Learning for Credit Risk Management and Irb Models: Lessons from success Case Histories*, A joint paper by Intesa Sanpaolo and **Crif**, Intesa **Sanpaolo-Crif**, Milano-Bologna.
- European Banking Authority (Eba)** (2021), *Eba Discussion Paper on Machine Learning for Irb Models*, No. Eba/Dp/2021/04, European Banking Authority, Paris.
- (2020), *Eba Report on Big Data and Advanced Analytics*, No. Eba/Rep/2020/01, European Banking Authority, Paris.
- European Central Bank** (2021), *Opinion of the European Central Bank of 29 December 2021 on a Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Con/2021/40)*, European Central Bank, Frankfurt am Main.
- Fernández-Delgado M., Cernadas E., Barro S., Amorim D.** (2014), «Do We Need Hundreds of Classifiers to Solve Real World Classification Problems?», in *The Journal of Machine Learning Research*, 15, pp. 3133-3181.
- Giudici P., Raffinetti E.** (2021), *Shapley-Lorenz eXplainable Artificial Intelligence*, *Expert Systems with Applications* 167, 114104. <https://doi.org/10.1016/j.eswa.2020.114104>.
- Institute of International Finance** (2019), *Machine Learning in Credit Risk*, Summary report, Washington Dc.
- Jones S., Johnstone D., Wilson R.** (2015), «An Empirical Evaluation of the Performance of Binary Classifiers in the Prediction of Credit Ratings Changes», in *Journal of Banking & Finance*, 56, pp. 72-85.
- Liu B.** (2017), «Lifelong Machine Learning: A Paradigm for Continuous Learning», in *Frontiers of Computer Science*, 11, pp. 359-361.
- Mitchell T.M.** (1997), *Machine Learning*, McGraw-Hill series in computer science, McGraw-Hill, New York.
- Molnar C.** (2019), *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*.
- Nagy Z.** (2018), *Artificial Intelligence and Machine Learning Fundamentals*, Packt Publishing Limited, Birmingham.
- Resti A., Onado M., Quagliariello M., Molyneux P.** (2021), *Shadow Banking: What Kind of Macroprudential Regulation Framework? From Research to Policy Actions*, Study requested by the Econ committee No. PE 662.925, European Parliament, Brussels.
- Ribeiro M.T., Singh S., Guestrin C.** (2016), «Why Should I Trust You?: Explaining the Predictions of Any Classifier».
- Visani G., Bagli E., Chesani F.** (2020), *OptiLime: Optimized Lime Explanations for Diagnostic Computer Algorithms*.