LUISS

Corso di laurea magistrale in

Amministrazione, Finanza e Controllo

Cattedra di Revisione Aziendale, Tecnica e Deontologia Professionale

I modelli predittivi della crisi d'impresa: profili metodologici e riscontri empirici

Relatore:
Prof. Alessandro Mechelli

Correlatore:
Prof.ssa Francesca Di Donato

Candidato:
Antonio Infante
Matricola: 788191

Anno Accademico 2024/2025

Sommario

Introduzione

Il tema della crisi d'impresa rappresenta oggi una delle questioni più rilevanti e complesse nel panorama economico e giuridico contemporaneo. La frequenza con cui le imprese si trovano ad affrontare difficoltà economico-finanziarie, l'instabilità crescente dei mercati, l'aumento della concorrenza e le sfide derivanti dalla trasformazione digitale impongono una riflessione profonda sulla capacità delle organizzazioni di anticipare, prevenire e gestire situazioni di squilibrio. In tale contesto, diventa sempre più strategico il ricorso a strumenti predittivi in grado di segnalare tempestivamente le condizioni che possono condurre un'impresa verso uno stato di crisi.

Negli ultimi decenni, l'approccio alla crisi si è progressivamente trasformato. Da una concezione tradizionale, fondata sulla gestione ex post dell'insolvenza e su logiche sanzionatorie, si è passati a un modello proattivo, orientato alla continuità aziendale e alla prevenzione. Questo cambiamento è stato favorito sia dall'evoluzione della dottrina economico-aziendale, sia dalle più recenti innovazioni normative, in particolare con l'introduzione del Codice della Crisi d'Impresa e dell'Insolvenza (CCII). Il nuovo impianto normativo ha reso obbligatoria per gli organi amministrativi l'adozione di assetti organizzativi adeguati alla rilevazione tempestiva della crisi e ha istituito sistemi di allerta precoce fondati su indicatori quantitativi.

Parallelamente, anche sul piano metodologico si è assistito a una profonda trasformazione: dai primi modelli statistici degli anni '60-'80 si è arrivati a sistemi predittivi sempre più complessi, in grado di integrare approcci probabilistici, regressioni logistiche e, più recentemente, algoritmi di machine learning. Questi ultimi, sfruttando la potenza computazionale e la disponibilità di big data, offrono la possibilità di individuare pattern nascosti e prevedere con maggiore accuratezza l'insorgere di situazioni di crisi.

Alla luce di questo scenario, la presente tesi ha l'obiettivo di analizzare e confrontare i principali modelli predittivi della crisi d'impresa, mettendone in luce i profili metodologici e i riscontri empirici. Il lavoro è articolato in tre capitoli, ciascuno dei quali affronta una dimensione specifica del fenomeno.

Nel primo capitolo, si delinea il quadro teorico, normativo e concettuale della crisi d'impresa. Dopo aver ricostruito l'evoluzione del concetto di crisi in ambito aziendale ed economico, si analizzano le principali tipologie di crisi (economica, finanziaria, patrimoniale, sistemica) e le cause sottostanti, distinguendo tra fattori endogeni ed

esogeni. Viene poi esaminato il ruolo degli indicatori di squilibrio come strumenti di allerta precoce, con riferimento sia alla letteratura classica, sia alla normativa italiana recente, con particolare attenzione alle disposizioni del CCII, al sistema di indicatori elaborato dal Consiglio Nazionale dei Dottori Commercialisti e al funzionamento dell'OCRI. Vengono inoltre esaminati nel dettaglio il modello Z-Score di Altman, il modello logit di Ohlson, il modello probit di Zmijewski e l'hazard model di Shumway. Di ciascun modello vengono presentate le ipotesi teoriche di fondo, le variabili utilizzate, le formule di calcolo, i risultati empirici e i limiti metodologici. L'obiettivo di questo capitolo è mostrare come la crisi d'impresa possa essere oggetto di previsione quantitativa attraverso strumenti basati sull'analisi dei dati di bilancio, ma anche evidenziare le criticità legate alla rigidità dei modelli, alla dipendenza da dati storici e all'inadeguatezza in contesti complessi e dinamici.

Il secondo capitolo è incentrato sui modelli predittivi moderni basati sul machine learning. Dopo una introduzione teorica alle logiche e ai fondamenti del machine learning supervisionato, vengono analizzati i principali algoritmi utilizzati nella previsione della crisi d'impresa: Random Forest, Support Vector Machine (SVM), reti neurali artificiali e boosting. Per ciascuna tecnica vengono illustrati il funzionamento, le caratteristiche distintive, le metriche di valutazione e i principali risultati emersi dalla letteratura. Vengono inoltre presentati alcuni studi empirici significativi che confrontano le performance dei modelli di machine learning con quelle dei modelli tradizionali, evidenziandone la maggiore capacità predittiva, la flessibilità e l'adattabilità a dataset complessi.

Il terzo capitolo è dedicato alla verifica empirica delle capacità predittive di diversi modelli applicati alla previsione della crisi d'impresa: i modelli confrontati sono lo Z-Score di Atlman, come benchmark tradizionale, il Random Forest. basato su logiche di apprendimento supervisionato e l'XGBoost, uno dei modelli più avanzati e performanti del machine learning.

Il dataset utilizzato è denominato *Polish Companies Bankruptcy*, fornito da EMIS, nel dataset sono presenti dati finanziari di aziende polacche attive tra il 2000 e il 2013, e contiene 64 variabili economico-patrimoniali e una variabile target binaria che indica lo stato dell'impresa (fallita o attiva). L'analisi si è focalizzata in particolare sui file riferiti

a uno e due anni prima del fallimento, in linea con l'orizzonte previsionale del modello di Altman.

In sintesi, la tesi intende offrire un'analisi completa ed equilibrata dell'evoluzione dei modelli predittivi della crisi d'impresa, combinando rigore teorico, approfondimento metodologico e lettura empirica. Lo scopo ultimo è fornire un contributo alla costruzione di strumenti sempre più efficaci e tempestivi per la rilevazione degli squilibri aziendali, a beneficio della stabilità dell'impresa e del sistema economico nel suo complesso.

1 – LA CRISI D'IMPRESA: NATURA, CONCETTO E INTERPRETAZIONI GIURIDICHE ED ECONOMICHE

1.1 – Premessa

Il concetto di crisi d'impresa ha progressivamente assunto, nel corso degli ultimi decenni, un ruolo centrale all'interno della riflessione economico-aziendale, giuridica e manageriale. In passato spesso percepita come un evento improvviso, ineluttabile e irrimediabile, oggi viene letta come un processo degenerativo, che si manifesta attraverso una serie di segnali premonitori, più o meno visibili, e che può essere anticipato, contenuto e in taluni casi anche superato, grazie all'adozione di strumenti di diagnosi precoce e a una governance aziendale orientata al monitoraggio del rischio.

La crisi non è un fatto isolato, ma piuttosto l'esito finale di un percorso di progressivo squilibrio che coinvolge le dimensioni economica, finanziaria, patrimoniale e strategica dell'impresa. È la rottura degli equilibri aziendali che segna il passaggio da una gestione ordinaria a una gestione critica, talvolta emergenziale, in cui la continuità operativa può risultare compromessa. La natura progressiva e stratificata del fenomeno ha indotto studiosi e operatori a rivedere radicalmente l'approccio con cui la crisi viene analizzata, affrontata e – idealmente – prevenuta.

Dal punto di vista giuridico, l'approccio dei legislatori ha subìto un profondo mutamento. La normativa italiana, a lungo ancorata alla legge fallimentare del 1942[1], ha recepito solo tardivamente l'esigenza di una gestione anticipata della crisi, focalizzandosi per decenni su logiche sanzionatorie e liquidatorie. È solo con il Codice della Crisi d'Impresa e dell'Insolvenza (D.lgs. 14/2019), entrato in vigore nel 2022, che si assiste a una vera svolta concettuale: la crisi viene definita non più come insolvenza attuale, ma come probabilità futura di insolvenza, e gli amministratori vengono investiti di obblighi positivi di prevenzione, diagnosi e attivazione tempestiva di strumenti di allerta. L'enfasi si sposta quindi sul concetto di continuità aziendale, sulla adeguatezza degli assetti organizzativi (art. 2086 c.c.) e sulla costruzione di un contesto in cui il rischio d'impresa venga gestito in modo proattivo.

[1] Regio Decreto n.267 del 16 marzo 1942.

Sotto il profilo economico-aziendale, si è passati da una visione meramente contabile, centrata su indici statici e dati storici, a un approccio più dinamico e sistemico, che interpreta la crisi come un fallimento della strategia, della governance e della capacità dell'impresa di adattarsi all'ambiente competitivo. Già a partire dagli anni '60, con i primi studi di Beaver e poi con Altman, si afferma l'idea che il fallimento non sia un evento imprevedibile, ma che esistano pattern ricorrenti e identificabili nei dati aziendali che precedono l'insolvenza. La letteratura successiva, da Ohlson (1980) a Shumway (2001), ha perfezionato questa impostazione, introducendo modelli sempre più sofisticati, basati su logiche probabilistiche, dati di mercato, approcci hazard e, più recentemente, metodologie di machine learning.

In questo scenario evolutivo, gli strumenti predittivi della crisi assumono un ruolo cruciale. Essi non sono soltanto strumenti di analisi retrospettiva, ma diventano veri e propri sistemi di allerta precoce, capaci non solo di guidare le scelte strategiche del management, ma anche di orientare la politica creditizia delle banche e supportare l'intervento tempestivo delle autorità. La possibilità di prevedere l'insorgere della crisi, monitorando con continuità indicatori contabili, finanziari e comportamentali, costituisce oggi una leva fondamentale per garantire non solo la sopravvivenza dell'impresa, ma anche la stabilità complessiva del sistema economico.

In definitiva, il tema della crisi d'impresa non può più essere trattato esclusivamente in termini reattivi o liquidatori: esso deve essere affrontato con un approccio integrato, intersettoriale e basato sull'anticipazione. La capacità di individuare tempestivamente i segnali di squilibrio, analizzarli con strumenti quantitativi adeguati e agire con prontezza rappresenta la nuova frontiera della gestione aziendale e della regolazione giuridica dell'attività economica.

1.2 – Il quadro normativo italiano: il Codice della Crisi d'impresa e dell'Insolvenza (CCII): principali innovazioni e finalità

Il Codice della Crisi d'Impresa e dell'Insolvenza (CCII) rappresenta il punto di arrivo di un processo evolutivo che ha interessato per decenni il diritto concorsuale italiano, segnato dalla necessità di superare l'impianto tradizionale della legge fallimentare del

1942, ormai considerata inadeguata a fronteggiare le complesse dinamiche del sistema economico contemporaneo. L'obiettivo del legislatore non è stato solo quello di riformare le procedure esecutive e concorsuali, ma soprattutto di ridefinire la cultura giuridico-economica della crisi, ponendo l'accento sulla prevenzione e sulla gestione anticipata del rischio di insolvenza.

In questa prospettiva, il Codice ha introdotto un'architettura normativa profondamente innovativa, incentrata sul concetto di crisi come "probabilità di futura insolvenza", e non più come evento già avvenuto. Si afferma così un modello ispirato alla gestione del rischio aziendale e alla responsabilizzazione degli organi gestori, che diventano i primi presìdi interni di monitoraggio e intervento. Il riferimento al rinnovato art. 2086 c.c., che impone agli imprenditori collettivi l'adozione di assetti organizzativi adeguati anche alla rilevazione tempestiva della crisi, segna l'introduzione di un vero e proprio dovere di prevenzione della crisi. Tale obbligo assume rilievo non solo giuridico, ma anche organizzativo e strategico, in quanto presuppone un sistema informativo evoluto, una cultura del controllo e una capacità decisionale proattiva.

La filosofia del Codice è coerente con le linee guida europee, in particolare con la Direttiva UE 2019/1023, la quale promuove meccanismi di allerta precoce e strumenti di ristrutturazione preventiva. Secondo le indicazioni della Commissione Europea[2] e dell'OCSE[3], i sistemi di gestione anticipata delle crisi sono fondamentali per ridurre l'impatto sistemico delle insolvenze e preservare la competitività del tessuto produttivo. L'intervento normativo italiano si pone quindi nel solco delle migliori pratiche internazionali, avvicinandosi a modelli più moderni e orientati alla sopravvivenza dell'impresa e non solo alla tutela dei creditori.

In dottrina, il nuovo impianto normativo è stato letto come un passaggio dalla logica del "fallimento come sanzione" a quella della "crisi come opportunità di risanamento"[4].

Questo cambio di paradigma implica una valorizzazione delle competenze manageriali e una maggiore integrazione tra diritto e strumenti aziendalistici di analisi, monitoraggio e previsione.

---

[2] Commissione Europea (2016). *Comunicazione della Commissione: Quadro europeo per la ristrutturazione preventiva*. Bruxelles.

[3] OECD (2017), *OECD Corporate Governance Factbook 2017*, OECD Publishing.

[4] Panizza, R. (2021). *La gestione della crisi d'impresa tra prevenzione e ristrutturazione*. Giappichelli.

Uno degli elementi di maggiore rilievo introdotti dal Codice della Crisi d'Impresa e dell'Insolvenza è rappresentato dalla previsione di un *sistema di indicatori significativi di crisi*, elaborato dal Consiglio Nazionale dei Dottori Commercialisti e degli Esperti Contabili (CNDCEC), con l'obiettivo di fornire uno strumento tecnico-operativo atto a supportare l'attività degli organi di amministrazione e controllo nell'individuazione tempestiva di segnali di squilibrio economico-finanziario. La logica alla base di questo impianto è fortemente anticipatoria: non si tratta di misurare l'insolvenza in senso classico, ma di individuare *disequilibri potenziali*, che, se trascurati, possono evolvere in crisi conclamata.

Gli indicatori definiti dal CNDCEC si articolano in due livelli principali: da un lato, un indicatore principale di sintesi, rappresentato dal **DSCR – Debt Service Coverage Ratio**, e dall'altro, indicatori secondari legati a voci fondamentali di bilancio (patrimonio netto, indebitamento, sostenibilità degli oneri finanziari, ecc.). Il DSCR, calcolato come rapporto tra i flussi finanziari attesi nei sei mesi successivi e gli impegni finanziari previsti nel medesimo orizzonte temporale, assume un valore soglia di **1**: se tale valore risulta inferiore, il rischio che l'impresa non riesca a far fronte ai propri obblighi nel breve periodo è considerato elevato, rappresentando un campanello d'allarme.[5]

In caso di indisponibilità o non attendibilità del DSCR, la normativa prevede il ricorso a cinque ulteriori indicatori alternativi, tra cui:

- **Patrimonio netto negativo**, indice di una progressiva erosione della base patrimoniale;
- **Indice di sostenibilità degli oneri finanziari**, calcolato come rapporto tra oneri finanziari e ricavi, utile a valutare l'incidenza del debito sul volume d'affari;
- **Indice di adeguatezza patrimoniale**, volto a misurare la coerenza tra struttura patrimoniale e investimenti;
- **Indice di ritorno liquido dell'attivo**, legato alla redditività netta in termini di flussi finanziari;
- **Indice di indebitamento previdenziale ed erariale**, che segnala eventuali esposizioni verso l'erario o gli enti previdenziali.

[5] Consiglio Nazionale dei Dottori Commercialisti e degli Esperti Contabili (CNDCEC). (2019). *Gli indici dell'allerta ex art. 13, co. 2 Codice della Crisi e dell'Insolvenza*.

Per ognuno di questi indicatori il CNDCEC ha individuato soglie quantitative predefinite, differenziate in base al settore economico di appartenenza dell'impresa (secondo codici ATECO), alla dimensione aziendale e al ciclo di vita. In presenza del superamento delle soglie per almeno tre indicatori su cinque, si ritiene sussistere una probabile situazione di crisi, che impone all'organo amministrativo una valutazione approfondita e, se del caso, l'attivazione tempestiva degli strumenti di composizione previsti dal Codice. L'impostazione degli indicatori, oltre a fornire un sistema di alert precoce, ha anche il merito di promuovere una maggiore consapevolezza gestionale e responsabilizzazione del management, favorendo l'adozione di assetti organizzativi più solidi e il potenziamento del sistema di controllo interno.[6] In questo contesto assume un ruolo centrale la **creazione dell'OCRI** (Organismo di Composizione della Crisi d'Impresa), istituito presso ciascuna Camera di Commercio. L'OCRI rappresenta uno strumento operativo e istituzionale volto a supportare le imprese in fase preventiva, prima che la crisi si trasformi in un'insolvenza conclamata. Si tratta di un elemento cardine del nuovo modello di governance anticipatoria della crisi, coerente con i principi espressi dalla Direttiva (UE) 2019/1023 in tema di ristrutturazioni preventive.

L'OCRI ha funzioni di ricezione, valutazione e gestione delle segnalazioni relative agli indicatori di crisi, nonché di attivazione della procedura di **composizione negoziata della crisi**, che si pone come alternativa non giudiziale alle classiche procedure concorsuali. In pratica, in caso di superamento delle soglie d'allerta definite dagli indicatori CNDCEC, l'organo di controllo societario (sindaci o revisori) o i creditori pubblici qualificati (Agenzia delle Entrate, INPS, Agenzia delle Entrate-Riscossione) possono trasmettere un'apposita segnalazione all'OCRI, che a sua volta convoca l'imprenditore per avviare un confronto riservato e guidato.[7]

Il procedimento non è contenzioso, bensì ispirato alla logica della mediazione e della riservatezza, nella prospettiva di individuare soluzioni negoziali o **piani di risanamento** sostenibili. L'OCRI nomina un esperto indipendente, selezionato da un apposito elenco gestito dalla Camera di Commercio, con il compito di facilitare il dialogo tra l'impresa e i creditori e di verificare la possibilità di recupero della continuità aziendale. Questo

---

[6] Il Sole 24 Ore. (2019). *Gli indici di allerta dei commercialisti*. (https://i2.res.24o.it/pdf2010/Editrice/ILSOLE24ORE/QUOTIDIANO_FISCO/Online/_Oggetti_Correlati/Documenti/2019/10/25/ebook.pdf)

[7] Camera di Commercio Treviso-Belluno. *Come opera l'OCRI? Da chi è costituito? (https://www.tb.camcom.gov.it/content/14805/Dachicostituito/).*

esperto svolge un ruolo tecnico e terzo: non propone soluzioni vincolanti, ma assiste il debitore nel valutare scenari di riequilibrio, anche attraverso strumenti di regolazione concordata come gli accordi di ristrutturazione, i piani attestati o le convenzioni di moratoria.[8]

L'istituzione dell'OCRI risponde quindi alla necessità di costruire un'infrastruttura pubblica territoriale, accessibile e competente, in grado di accompagnare le imprese nel percorso di emersione e gestione precoce della crisi, superando lo stigma dell'insolvenza e della procedura fallimentare. Inoltre, essa rafforza la collaborazione tra istituzioni, professionisti e sistema camerale, configurandosi come punto di riferimento tecnico-operativo per l'intero processo di allerta e prevenzione. Si può quindi sicuramente affermare che il Codice della Crisi d'Impresa e dell'Insolvenza ha segnato una profonda cesura rispetto all'impianto tradizionale della precedente legge fallimentare, sostituendo una logica punitiva e liquidatoria con un approccio improntato alla prevenzione, alla continuità aziendale e alla valorizzazione del tessuto produttivo. Il Codice ha introdotto una logica anticipatoria e gestionale, che si manifesta non solo con il fatto che l'emersione tempestiva dei segnali di difficoltà diventa una responsabilità condivisa tra organo amministrativo, organo di controllo e sistema pubblico di vigilanza, ma anche con la possibilità di attivare, per gli imprenditori in difficoltà, percorsi riservati di gestione e risanamento mediante il coinvolgimento di un esperto indipendente.

Il cambio di paradigma si riflette anche sul piano culturale: la crisi non è più considerata un'anomalia da reprimere, ma un evento fisiologico della vita aziendale, che richiede strumenti tecnici, managerialità consapevole e risposte rapide. In tale contesto, il CCII si configura come una riforma di sistema, ispirata ai principi europei di efficienza, responsabilizzazione e tutela della continuità, contribuendo a modernizzare in profondità la disciplina italiana dell'insolvenza.

[8] Allianz Trade. *L'OCRI e la procedura di composizione assistita della crisi*. (https://www.allianz-trade.com/it_IT/news-e-approfondimenti/community/codice-crisi-d-impresa/locri-e-la-procedura-di-composizione-assistita-della-crisi.html).

1.3 – La crisi d'impresa sotto il profilo economico-finanziario

La crisi d'impresa rappresenta un momento critico e complesso nel ciclo di vita aziendale, ed è oggetto di studio trasversale in numerosi ambiti disciplinari; essa può essere definita come una situazione di persistente squilibrio economico-finanziario che compromette, in misura variabile, la capacità dell'impresa di garantire la continuità aziendale nel medio-lungo periodo.

Secondo la letteratura classica di economia aziendale, la crisi è una rottura degli equilibri fondamentali della gestione[9], e in quanto tale costituisce un sintomo di inefficienza strutturale o congiunturale. La crisi non si manifesta quasi mai in modo improvviso, ma è spesso il risultato di un processo graduale e cumulativo, nel quale diversi segnali di difficoltà emergono progressivamente fino a condurre a una situazione di instabilità aziendale.

Il fenomeno può assumere diverse connotazioni e svilupparsi attraverso stadi evolutivi successivi, ognuno caratterizzato da specifiche problematiche gestionali e finanziarie. Generalmente, in un primo momento si osserva una perdita progressiva della redditività operativa, che può derivare da fattori interni, come inefficienze produttive, errori strategici o incapacità di innovare, oppure da elementi esterni, quali mutamenti del contesto competitivo, fluttuazioni della domanda o shock macroeconomici. Questa fase iniziale spesso si traduce in una contrazione della marginalità, che compromette la capacità dell'impresa di autofinanziare la propria crescita e di mantenere livelli adeguati di investimenti.[10]

Se la riduzione della redditività non viene prontamente affrontata con interventi correttivi, si entra in una fase più avanzata del processo di crisi, caratterizzata dall'erosione della liquidità aziendale. La minore generazione di cassa comporta una difficoltà crescente nel far fronte agli impegni finanziari correnti, portando a un aumento dell'esposizione verso fornitori e istituti di credito. Tale situazione può generare un circolo vizioso, in cui il peggioramento del merito creditizio dell'impresa riduce ulteriormente la sua capacità di accesso al finanziamento, aumentando il rischio di tensioni finanziarie.

Nel caso in cui questi squilibri non vengano risolti, la crisi evolve in una situazione di instabilità patrimoniale e finanziaria che compromette definitivamente la continuità

[9] Airoldi, G., Brunetti, G., Coda, V. (1994). *Economia aziendale*. Il Mulino.
[10] Argenti, J. (1976). *Corporate Collapse: The Causes and Symptoms*. McGraw-Hill

aziendale. La perdita di fiducia da parte degli stakeholder, il deterioramento del valore del capitale proprio e l'aggravarsi dell'indebitamento possono sfociare in uno stato di insolvenza, in cui l'impresa non è più in grado di adempiere alle proprie obbligazioni. In questa fase, diventa determinante la scelta del management e degli organi di governance riguardo all'attivazione di strumenti di risanamento, ristrutturazione del debito o, nei casi più gravi, di procedure concorsuali volte a gestire l'uscita controllata dal mercato.[11]

Nel paradigma sistemico dell'impresa, la crisi è interpretata come il risultato del progressivo deterioramento della capacità dell'azienda di interagire efficacemente con il proprio ambiente competitivo ed operativo. Essendo l'impresa concepita come un sistema aperto caratterizzato da una continua interdipendenza con i mercati, i clienti, i fornitori, le istituzioni finanziarie e gli altri stakeholder, nel momento in cui i meccanismi di adattamento a queste variabili esterne si indeboliscono o vengono trascurati, si innesca un processo degenerativo che può condurre alla perdita della competitività e, successivamente, a una crisi strutturale: si parla in tal senso di *"fallimento della gestione strategica"*[12], facendo riferimento a tutte quelle situazioni in cui il management si mostra incapace di formulare e implementare scelte adeguate rispetto alle sfide poste dall'ambiente esterno. Questo tipo di fallimento può derivare da fattori interni, quali un'errata allocazione delle risorse, una strategia incoerente o obsoleta, una governance inefficace o una gestione inadeguata delle innovazioni tecnologiche. Tuttavia, può anche essere il risultato di fattori esterni non adeguatamente previsti o fronteggiati, quali cambiamenti normativi, crisi macroeconomiche, rivoluzioni tecnologiche o mutamenti nei comportamenti dei consumatori.

Un aspetto cruciale della crisi in ottica sistemica riguarda il concetto di resilienza aziendale, ovvero la capacità dell'impresa di assorbire gli shock esterni e adattarsi alle nuove condizioni di mercato. Secondo il framework della *Dynamic Capabilities Theory*[13] la sostenibilità a lungo termine di un'impresa dipende dalla sua capacità di sviluppare competenze dinamiche, ovvero la possibilità di riconfigurare rapidamente le proprie risorse e strategie per rispondere ai cambiamenti ambientali. Quando questo meccanismo

[11] Tale prospettiva evidenzia come la crisi d'impresa non debba essere intesa come un evento improvviso, ma piuttosto come un fenomeno prevedibile e gestibile attraverso adeguati strumenti di monitoraggio, controllo e intervento tempestivo.

[12] Invernizzi, G. (2000). *Strategie e politiche per il governo dell'impresa in crisi*. Egea.

[13] Teece, D. J., Pisano, G., & Shuen, A. (1997). *Dynamic Capabilities and Strategic Management*. Strategic Management Journal, 18(7), 509-533.

di adattamento fallisce – ad esempio a causa di rigidità organizzative, resistenze al cambiamento o miopia strategica – l'impresa diventa vulnerabile a shock interni ed esterni, aumentando il rischio di crisi.[14]

Dunque, nel paradigma sistemico, la crisi aziendale non è semplicemente il risultato di difficoltà economico-finanziarie, ma rappresenta il sintomo di un'incapacità più profonda del management di governare la complessità e il cambiamento. La prevenzione della crisi passa, quindi, attraverso un approccio strategico proattivo, basato sull'analisi costante dei segnali deboli di mercato, sull'adozione di una cultura organizzativa flessibile e orientata all'innovazione e sull'applicazione di strumenti di governance efficaci.

Nel dibattito internazionale, autori come Altman[15] e Wruck[16] pongono l'accento sul concetto di *distress finanziario*, ossia su quella condizione in cui l'impresa non è più in grado di generare sufficiente cash flow operativo per far fronte agli obblighi contrattuali: questo stato di difficoltà non implica necessariamente l'insolvenza conclamata, ma rappresenta un livello intermedio di rischio, in cui la solvibilità futura è compromessa, pur rimanendo formalmente sostenuta nel breve termine. Il *distress finanziario* può dunque essere definito come *un'area grigia* in cui l'impresa è esposta a tensioni di liquidità croniche, in presenza di un deterioramento strutturale della performance economica.

Tale visione, sebbene fortemente orientata alla dimensione finanziaria, è stata progressivamente arricchita da una lettura multifattoriale del concetto di crisi, nella quale il deterioramento del cash flow non è soltanto una conseguenza di dinamiche economico-contabili, ma anche il riflesso di disfunzioni strategiche, organizzative e sistemiche. In questo senso, si afferma, come anticipato, una visione della crisi come fallimento dell'equilibrio complessivo dell'impresa, in cui concorrono fattori endogeni (quali governance inefficace, leadership debole, mancanza di innovazione) ed esogeni (tensioni competitive, mutamenti normativi, volatilità macroeconomica).

---

[14] Tra i casi più eclatanti di crisi aziendali riconducibili ad un fallimento della gestione strategica si possono citare *Kodak*, *Blockbuster* e *BlackBerry*, le quali si sono mostrate incapaci di innovare nonostante disponessero delle competenze tecnologiche per farlo.

[15] Altman, E.I. (1968). *Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy*. Journal of Finance, 23(4), 589--609.

[16] Wruck, K. H. (1990). *Financial Distress, Reorganization, and Organizational Efficiency*. Journal of Financial Economics, 27(2), 419-444.

L'apporto di Wruck è particolarmente rilevante, in quanto mette in evidenza il legame tra *distress finanziario* ed efficienza organizzativa, sostenendo che il momento di crisi può anche costituire un'opportunità di riallocazione delle risorse, ridefinizione delle strutture decisionali e rinnovamento della mission aziendale. Tuttavia, perché ciò accada, è necessario che il management sia in grado di affrontare tale fattispecie con strumenti analitici avanzati, un'adeguata cultura del rischio e un approccio orientato alla creazione di valore sostenibile nel tempo.

In linea generale, la classificazione tipologica della crisi consente di distinguere tra:

- **Crisi economica:** si manifesta quando l'impresa non è più in grado di generare un flusso reddituale sufficiente a coprire i costi di produzione e remunerare adeguatamente i fattori produttivi impiegati;[17]
- **Crisi finanziaria:** stando alla definizione di Altman e Hotchkiss[18], è lo stato in cui si entra quando l'impresa, pur presentando talvolta risultati economici positivi, non è in grado di far fronte tempestivamente agli impegni di pagamento verso terzi per carenza di liquidità o flussi di cassa;
- **Crisi patrimoniale:** situazione in cui l'impresa subisce una riduzione rilevante del proprio patrimonio netto, che compromette l'equilibrio finanziario e la capacità di attrarre capitale di rischio o di credito[19];
- **Crisi globale o sistemica:** quando si combinano squilibri di natura economica, finanziaria e patrimoniale, configurando una situazione di instabilità generale e persistente che minaccia la continuità aziendale.[20]

Alla luce delle considerazioni esposte, emerge chiaramente come la crisi d'impresa non possa essere interpretata unicamente come una fase di difficoltà finanziaria transitoria, bensì come un fenomeno complesso, multidimensionale e sistemico. Essa rappresenta, in ultima analisi, la manifestazione di una rottura progressiva dell'equilibrio gestionale e della capacità dell'impresa di perseguire in modo efficace e sostenibile le proprie finalità economiche, strategiche e sociali. Il concetto di equilibrio, da sempre centrale nella teoria economico-aziendale, viene messo in discussione quando l'impresa perde la propria

[17] Airoldi, G., Brunetti, G., Coda, V. (1994). *Economia aziendale*. Il Mulino.
[18] Altman, E.I., Hotchkiss, E. (2006). *Corporate Financial Distress and Bankruptcy*. Wiley.
[19] Bastia, P., Campedelli, B. (2020). *Bilancio e controllo della crisi d'impresa*. FrancoAngeli.
[20] Invernizzi, G. (2000). *Strategie e politiche per il governo dell'impresa in crisi*. Egea.

capacità di adattamento, di generazione di valore e di interazione virtuosa con il contesto ambientale.

In questo senso, la crisi non è un evento improvviso, bensì il risultato di un processo degenerativo cumulativo, che si sviluppa nel tempo e attraversa differenti livelli dell'organizzazione: dalla perdita della redditività alla compromissione della liquidità, fino all'indebolimento patrimoniale e all'insolvenza.

A fronte di questa complessità, risulta indispensabile per il management adottare un approccio integrato alla diagnosi e alla prevenzione della crisi, che combini strumenti quantitativi e qualitativi, visione strategica e capacità di governo. In tale ottica, il ruolo della corporate governance, dei sistemi informativi aziendali, dei presidi di controllo interno e degli assetti organizzativi assume una rilevanza decisiva per la rilevazione precoce dei segnali di deterioramento. La capacità di cogliere tempestivamente tali segnali – spesso deboli e diffusi – rappresenta il presupposto fondamentale per impostare efficaci strategie di risposta e piani di risanamento, pienamente in linea con l'approccio introdotto dal CCII.

Alla luce di ciò, il passaggio successivo di questa analisi consiste nell'esaminare criticamente gli indicatori di crisi e i principali segnali di allerta precoce, con l'obiettivo di comprendere come sia possibile, nella pratica, misurare e anticipare il rischio di crisi prima che esso evolva in uno stato conclamato di insolvenza, evidenziando i punti di forza e di debolezza dei singoli modelli.

1.4 – Indicatori di crisi: segnali precoci e misurazione quantitativa

Il riconoscimento tempestivo dello stato di crisi rappresenta una delle competenze gestionali più strategiche nell'attuale contesto imprenditoriale. La possibilità di intervenire in una fase precoce del processo degenerativo, prima che esso si trasformi in una condizione irreversibile di insolvenza, consente all'impresa di preservare il proprio valore economico, proteggere le relazioni con gli stakeholder e salvaguardare la continuità aziendale. È proprio la natura non improvvisa della crisi a rendere cruciale la capacità di intercettarne i segnali iniziali: come ampiamente riconosciuto dalla letteratura economico-aziendale, la crisi raramente si manifesta come un evento repentino e isolato, ma è spesso anticipato da sintomi riconoscibili sotto forma di anomalie contabili, squilibri gestionali, inefficienze operative e tensioni relazionali con il contesto esterno.

In questa prospettiva, assume un'importanza fondamentale l'adozione di sistemi di allerta precoce (*early warning systems*), strumenti diagnostici che consentono di monitorare l'evoluzione dell'equilibrio aziendale e di rilevare tempestivamente situazioni di rischio. I primi modelli si basavano esclusivamente su dati di bilancio storici e su indicatori economico-finanziari sintetici, come i classici indici di redditività, liquidità e indebitamento. Tuttavia, il progresso della dottrina e l'evoluzione della pratica hanno portato allo sviluppo di modelli più sofisticati, che integrano variabili qualitative, informazioni forward-looking, dati settoriali e indicatori comportamentali. Tali sistemi risultano oggi ancora più potenti grazie all'introduzione delle tecnologie di analisi predittiva, del machine learning e della business intelligence, che consentono di individuare pattern anomali nei flussi informativi aziendali prima che i sintomi diventino evidenti nei bilanci.

La possibilità di intercettare per tempo il deterioramento gestionale offre al management un fondamentale margine di manovra per mettere in atto strategie di contenimento, recupero e turnaround. Le azioni correttive possono spaziare dalla ridefinizione del modello di business alla ristrutturazione dell'indebitamento, dalla modifica degli assetti organizzativi al ricambio della leadership, e così via. In tal senso, l'emersione anticipata della crisi non deve essere vista come una minaccia, ma piuttosto come una *finestra di opportunità* per il riposizionamento strategico dell'impresa e per il rafforzamento della sua resilienza.

La prevenzione efficace della crisi, dunque, non si basa soltanto sulla disponibilità di strumenti analitici, ma soprattutto sulla capacità del management di interpretare correttamente i segnali deboli, di assumere decisioni tempestive e di adottare una cultura organizzativa orientata al monitoraggio continuo dei rischi. È proprio questa capacità proattiva di lettura e intervento che può fare la differenza tra un'impresa destinata al dissesto e una che riesce a invertire la traiettoria negativa, trasformando una fase critica in un'occasione di rilancio strutturale.

Beaver, in uno studio pionieristico che ha segnato l'inizio della letteratura quantitativa sulla previsione dell'insolvenza[21], dimostrò come alcune grandezze contabili – in particolare il rapporto tra cash flow operativo e debiti totali – possedessero un'elevata

[21] Beaver, W.H. (1966). *Financial Ratios as Predictors of Failure*. Journal of Accounting Research, 4, 71-111.

capacità discriminante tra imprese in equilibrio economico-finanziario e imprese prossime al fallimento. Il suo approccio, basato sull'analisi univariata[22] di indicatori contabili, ha evidenziato che, già diversi anni prima dell'insolvenza conclamata, era possibile osservare una differenziazione statisticamente significativa nel comportamento dei parametri tra imprese "fallite" e "non fallite". In tal modo, Beaver fu tra i primi a concepire il fallimento non come un evento improvviso e imprevedibile, ma come un fenomeno anticipabile sulla base di segnali misurabili e sistematici.

Il contributo innovativo di Beaver risiede anche nell'aver introdotto una logica predittiva basata sulla probabilità condizionata: in altri termini, egli analizzò le probabilità di sopravvivenza aziendale nel tempo, costruendo modelli in grado di stimare, sulla base di un singolo indicatore, il rischio di default nei successivi esercizi. Questa impostazione ha rappresentato una svolta metodologica per l'epoca, ponendo le basi per lo sviluppo successivo di modelli multivariati (come quelli di Altman e Ohlson) e dell'odierna analisi di scoring creditizio.

Va evidenziato che, sebbene l'approccio di Beaver fosse limitato a un set ristretto di variabili contabili e non tenesse conto delle interazioni tra di esse, il suo lavoro ha aperto un filone di ricerca fondamentale, ancora oggi centrale nei sistemi di rating, nelle valutazioni di merito creditizio e nei modelli di allerta precoce. In un contesto moderno, i principi introdotti da Beaver si ritrovano nei modelli di machine learning supervisionati (oggetto del presente elaborato nei capitoli successivi), che estendono la logica discriminante attraverso l'analisi di dataset molto più complessi e multidimensionali.

In definitiva, il merito scientifico di Beaver non risiede solo nell'individuazione di specifici rapporti contabili significativi, ma nell'aver dimostrato che l'insolvenza aziendale è, con gli strumenti giusti, statisticamente prevedibile e dunque, almeno in parte, gestibile attraverso un'adeguata azione preventiva.

L'economista e accademico statunitense Edward I. Altman introdusse nel 1968 il celebre modello **Z-Score**[23], destinato a diventare uno dei più noti e diffusi strumenti di previsione del rischio di insolvenza a livello internazionale. Fondando il proprio lavoro sull'assunto che esistessero differenze strutturali identificabili tra imprese in equilibrio e imprese a

[22] Analisi statistica che si focalizza soltanto sulla variabile oggetto d'indagine, senza analizzare relazioni con altre variabili

[23] Altman, E. I. (1968). *Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy*. Journal of Finance, 23(4), 589–609.

rischio di insolvenza, e che tali divergenze potessero essere rilevate attraverso un'attenta analisi dei dati contabili e finanziari ricavabili dai bilanci aziendali, il professore newyorkese selezionò una lista iniziale di 22 indicatori, ritenuti potenzialmente efficaci nel rappresentare il rischio di default. Dopo un'attenta valutazione empirica, giunse alla conclusione che fosse possibile sintetizzare tali variabili in **cinque rapporti chiave**, i quali si dimostrarono in grado di spiegare con elevata efficacia la probabilità di fallimento delle imprese.

Statisticamente basato sull'applicazione dell'**analisi discriminante lineare**[24], il modello utilizza una funzione discriminante capace di massimizzare la distanza tra due o più gruppi predefiniti (imprese fallite vs. imprese sane), combinando linearmente le 5 variabili contabili con pesi determinati empiricamente.

Le variabili oggetto d'esame sono:

1. Capitale circolante netto / Attivo totale (**misura della liquidità a breve termine**);
2. Utili trattenuti / Attivo totale (**capacità di autofinanziamento**);
3. Utile operativo / Attivo totale (**redditività operativa**);
4. Valore di mercato del patrimonio netto / Debiti totali (**leva finanziaria corretta per il rischio di mercato)**;
5. Fatturato / Attivo totale (**efficienza nell'utilizzo degli asset**);

combinati nella seguente funzione:

$$Z = 1{,}2X_1 + 1{,}4X_2 + 3{,}3X_3 + 0{,}6X_4 + 1{,}0X_5$$

In base al valore ottenuto, l'impresa viene classificata in una delle seguenti tre categorie:

- Zona sicura ($Z > 2.99$): bassa probabilità di default;
- Zona grigia ($1.81 < Z < 2.99$): situazione intermedia, rischio non trascurabile;
- Zona a rischio ($Z < 1.81$): elevata probabilità di insolvenza entro due anni.

L'innovazione metodologica del modello di Altman non risiede soltanto nella scelta delle variabili, ma nell'approccio integrato e comparativo che consente di sintetizzare in un solo indice diverse dimensioni della solidità aziendale. Inoltre, l'utilizzo del valore di mercato del capitale proprio rappresentò, all'epoca, un'evoluzione rispetto ai modelli

[24] Anche nota come LDA, è una tecnica statistica usata per classificare osservazioni in due o più gruppi, sulla base di variabili predittive continue con l'obiettivo di trovare combinazioni lineari delle variabili che massimizzano la separazione tra i gruppi

puramente contabili, poiché introdusse un elemento di valutazione "di mercato" all'interno della diagnostica predittiva.

Nonostante il modello originale sia stato costruito su un campione di aziende manifatturiere statunitensi quotate, esso ha dato origine a numerose versioni adattate, come lo **Z'-Score** per aziende non quotate[25], particolarmente utilizzata nel contesto europeo nei modelli di analisi del rischio delle PMI[26] e lo **Z''-Score**, con l'obiettivo di rendere il modello più stabile, generalizzabile e valido in ottica internazionale, indipendentemente dalla quotazione in borsa e dal settore di appartenenza dell'impresa, superando così i limiti delle due versioni precedenti, rimuovendo il rapporto tra fatturato e attivo totale al fine di rendere l'indice applicabile anche ad imprese operanti in settori con cicli di fatturato irregolari o soggetti a crescita non lineare[27].

Tuttavia, per quanto affinati, i modelli originari (Z, Z', Z'') essendo sviluppati su campioni di aziende statunitensi – e quindi operanti in specifici contesti macroeconomici, normativi e contabili particolarmente stabili – vedono ridotta la loro efficacia quando questi vengono applicati ad aziende operanti in mercati emergenti a causa delle fisiologiche differenze strutturali tra economie sviluppate e non, quali l'incompletezza informativa, una minore trasparenza nei bilanci, mercati poco sviluppati ed elevata volatilità politica e valutaria.

Per superare tali limiti, lo stesso Altman formulò una ulteriore versione dello Z-Score che prende il nome di **EMS Score (Emerging Market Score)**[28] al fine di rispecchiare meglio le dinamiche aziendali e i rischi sistemici nei Paesi in via di sviluppo.

Il modello si mostra particolarmente flessibile in quanto non vi è una formula universalmente accettata dell'EMS Score siccome il modello viene adattato caso per caso a seconda del Paese, del campione di imprese e della disponibilità dei dati. Tuttavia, le variabili sono simili a quelle dello Z''-Score, ma con pesi e coefficienti rivisti in base al contesto locale: negli anni Altaman ha collaborato con istituzioni locali, banche ed

---

[25] Altman, E. I. (1983). *Corporate Financial Distress: A Complete Guide to Predicting, Avoiding, and Dealing with Bankruptcy*. Wiley.

[26] Il modello principale viene adattato al fine di renderlo utilizzabile anche per le imprese non quotate; questa versione prevede l'eliminazione del valore di mercato del patrimonio netto (difficilmente reperibile nelle aziende private) e una ricalibrazione dei coefficienti della funzione discriminante su un campione di aziende non quotate.

[27] Altman, E. I. (2000). *Predicting financial distress of companies: Revisiting the Z-score and ZETA models.* Working paper, NYU Stern School of Business.

[28] Altman, E. I. (2002). *Revisiting the Z-Score and ZETA Models – EMS Adaptations*. Working paper, NYU Stern School.

università per creare versioni nazionali o addirittura regionali del modello.[29]
I vantaggi principali del modello Z-Score e delle sue varianti risiedono nella semplicità di calcolo, nella chiara interpretabilità dei risultati, nella solidità empirica dimostrata da numerose applicazioni nel tempo e nella possibilità di calibrare il modello su campioni locali. Tuttavia, il modello presenta anche alcune criticità, tra cui:

- la dipendenza dai dati di bilancio, che sono informazioni storiche e soggette a manipolazione;
- la staticità del modello, che non tiene conto della variabilità dinamica dell'impresa nel tempo;
- la linearità della funzione discriminante, che non coglie interazioni non lineari tra variabili;
- la minore capacità predittiva in settori atipici.

Nonostante ciò, lo Z-Score continua a essere il modello di riferimento per analisti finanziari, revisori contabili, banche, società di consulenza e tribunali fallimentari, anche in affiancamento ad altri indicatori e modelli quantitativi. In Italia, ad esempio, il modello è talvolta richiamato nelle prassi di valutazione del rischio di crisi ai fini dell'adeguatezza degli assetti organizzativi (art. 2086 c.c.) e nei sistemi interni di scoring bancario.
Nel presente elaborato, il modello Z-Score e le sue principali varianti vengono riprese e analizzate nel dettaglio nel successivo capitolo, con particolare attenzione all'efficacia predittiva in contesti reali. In tale sede, vengono illustrate le formule originali di ciascuna versione del modello con un confronto tra i risultati ottenuti attraverso il calcolo dei punteggi con l'obiettivo di valutare coerenze, divergenze e capacità diagnostica le successive evoluzioni metodologiche come i modelli logit, probit e i metodi non lineari basati sul machine learning.
Nonostante lo Z-Score – con le sue varianti – rappresenti il principale modello predittivo, negli anni vi si sono affiancati numerosi modelli alternativi.

[29] Esempi significativi sono lo *Z-Score Brazil* sviluppato in collaborazione con la Fundação Getulio Vargas nel quale vengono utilizzati i dati contabili secondo i BR-GAAP incorporando un peso maggiore alla leva finanziaria e alla posizione di cassa, ancora oggi molto utilizzato dalle istituzioni di credito brasiliane per valutare il credit scoring delle PMI, oppure lo *Z-Score for Indian SMEs* supportato dalla Bank of India e varie università locali, nel quale viene dato particolare peso a dati qualitativi come la compliance fiscale e i ritardi nei pagamenti, data la scarsa trasparenza contabile di molte PMI.

Nel 1980, James A. Ohlson pubblicò un importante contributo nella letteratura sulla previsione del fallimento d'impresa, introducendo per la prima volta in questo campo l'applicazione della regressione logistica (*logit model*) per stimare la probabilità di default. Questo modello si pone in netta discontinuità rispetto ai precedenti modelli discriminanti tra cui la prima versione dello Z-Score: la regressione logistica, a differenza dell'analisi discriminante lineare (utilizzata nello Z-Score), non richiede l'assunzione di distribuzioni normali multivariate, né di omoschedasticità[30] tra i gruppi: questo rende il modello più robusto e adatto a trattare la natura discreta dell'evento oggetto di previsione, in questo caso il fallimento: 1 se l'impresa è a rischio di fallimento, 0 altrimenti.

Inoltre, Ohlson abbandona la logistica classificatoria e introduce una valutazione probabilistica: anziché stabilire una soglia arbitraria, il modello assegna a ciascuna impresa una probabilità di default compresa tra 0 e 1.

Nel suo studio, Ohlson utilizzò un campione di 2.163 imprese statunitensi (tra fallite e non fallite) nel periodo 1970–1976. Sulla base di un'analisi statistica rigorosa, selezionò nove variabili esplicative che si dimostrarono significativamente associate al rischio di default:

1. **SIZE** – logaritmo del totale attivo: misura della dimensione aziendale
2. **TLTA** – total liabilities / total assets: leva finanziaria complessiva
3. **WCTA** – working capital / total assets: liquidità a breve termine
4. **CLCA** – current liabilities / current assets: capacità di copertura del passivo corrente
5. **OENEG** – dummy (1 se equity negativa, 0 altrimenti): solvibilità patrimoniale
6. **NITA** – net income / total assets: redditività netta
7. **FUTL** – funds from operations / total liabilities: capacità di autofinanziamento
8. **INTWO** – dummy (1 se utile netto negativo in due anni consecutivi)
9. **CHIN** – variazione percentuale del reddito netto: tendenza reddituale

[30] Per omoschedasticità si intende che la varianza delle variabili indipendenti è uguale tra i gruppi che si stanno confrontando.

L'O-Score viene calcolato come segue:

$$O = -1.32 - 0.407 \cdot SIZE + 6.03 \cdot TLTA - 1.43 \cdot WCTA + 0.0757 \cdot CLCA - 2.37 \cdot OENEG - 1.83 \cdot NITA + 0.285 \cdot FUTL - 1.72 \cdot INTWO - 0.521 \cdot CHIN$$

mentre la probabilità di fallimento è pari a:

$$P = \frac{1}{1 + e^{-O}}$$

Una delle innovazioni più importanti del modello di Ohlson è la sua interpretabilità statistica: ciascun coefficiente misura l'effetto marginale di una determinata variabile sul log-odds di fallimento.[31]

Questo approccio permette di effettuare valutazioni più flessibili e adattabili a differenti contesti economici, rispetto ai modelli basati su soglie rigide.

Il modello logit di Ohlson presenta numerosi punti di forza che ne hanno determinato l'ampia diffusione sia in ambito accademico che professionale. Innanzitutto, una delle sue caratteristiche distintive è l'approccio probabilistico alla previsione del fallimento: diversamente dai modelli discriminanti come lo Z-Score di Altman, che classificano rigidamente le imprese in categorie predefinite sulla base di soglie arbitrarie, il modello di Ohlson stima una probabilità continua di default. Questo significa che ciascuna impresa riceve un valore compreso tra 0 e 1, che riflette il grado di rischio di insolvenza in termini quantitativi e non meramente qualitativi. Una simile impostazione consente una lettura più sfumata e graduale del rischio, riducendo il rischio di errore nelle soglie e migliorando la sensibilità alle specificità aziendali.

Un ulteriore vantaggio risiede nella maggiore robustezza statistica della regressione logistica rispetto all'analisi discriminante. Il modello logit, infatti, non richiede che i dati siano normalmente distribuiti o che le varianze siano omogenee tra gruppi. Ciò lo rende più adatto all'analisi di campioni reali e disomogenei, come quelli che si trovano nella

[31] Il modello restituisce una probabilità di fallimento compresa tra 0 e 1 grazie all'uso della regressione logistica, che stima l'impatto di ogni variabile (come l'indebitamento, la redditività, ecc.) ha sulla probabilità che l'impresa vada in default (a differenza dello Z-Score, che prevede delle soglie fisse come risultato). In particolare, ogni coefficiente che appare nella formula del modello indica quanto varia la probabilità di fallimento al variare della singola variabile corrispondente (i.e. se il coefficiente legato alla redditività è negativo, significa che all'aumentare della redditività la probabilità di fallire si riduce).

pratica quotidiana, dove le imprese presentano caratteristiche molto diverse per dimensione, settore, struttura patrimoniale e performance.

Ohlson offre inoltre un sistema di analisi più interpretabile dal punto di vista economico: ogni coefficiente stimato nel modello indica l'effetto marginale di una determinata variabile finanziaria sul rischio di fallimento, permettendo di comprendere in che misura, ad esempio, un aumento dell'indebitamento o una riduzione della redditività contribuiscano all'aumento della probabilità di default. Questo aspetto conferisce al modello una duplice valenza, rendendolo utile non solo come strumento di previsione, ma anche come supporto decisionale, poiché consente al management di individuare con precisione le leve gestionali più critiche da monitorare o correggere.

Nel 1984, Mark E. Zmijewski propose un approccio alternativo "parsimonioso" (basato soltanto su 3 variabili) ai modelli discriminanti e logistici presentati finora, criticandone alcuni limiti metodologici e introducendo una nuova metodologia basata su una regressione probit.[32]

Zmijewski selezionò soltanto 3 variabili indipendenti particolarmente significative per la previsione del rischio di crisi:

1. **ROA** → misura della redditività complessiva dell'impresa;
2. **Leverage** → grado di indebitamento;
3. **Current Ratio** → misura della liquidità a breve termine

La funzione probit stimata ha la seguente forma generale:

$$P(Y = 1) = \Phi(\beta_0 + \beta_1 \cdot ROA + \beta_2 \cdot Leverage + \beta_3 \cdot Current\ Ratio)$$

Dove:

- **P(Y=1)** è la probabilità che l'impresa sia in stato di distress;
- **Φ** rappresenta la funzione cumulativa normale standard;
- **β_0...β_3** sono i coefficienti stimati;

Uno dei principali meriti del modello di Zmijewski risiede, come anticipato, nella sua parsimonia: utilizza poche variabili altamente rappresentative, rendendolo facile da calcolare e da interpretare, pur mantenendo una buona capacità predittiva. Inoltre, il ricorso alla regressione probit consente di evitare alcuni problemi di distorsione statistica

[32] A differenza del modello logit, il probit impiega la funzione di distribuzione normale cumulativa standard; anche in questo caso la variabile dipendente è dicotomica.

che possono emergere con la regressione logistica, specialmente nei campioni sbilanciati (cioè con un numero molto diverso di imprese fallite e non fallite).

In merito al modello Z-Score, una delle critiche più forti riguarda il fatto che Altman costruì il proprio campione scegliendo a posteriori un numero uguale di imprese fallite e non fallite e che tale scelta non rispecchiasse la reale distribuzione della popolazione aziendale, in cui le imprese fallite rappresentavano, ai tempi, una minoranza assoluta: tale scelta, secondo Zmijewski portava inevitabilmente il modello a sottoperformare quando applicato ai dati reali.

Inoltre, secondo Zmijewski gli assunti di base del modello Z-Score[33] sono difficilmente riscontrabili in dati reali di bilancio, rendendo l'analisi discriminante statisticamente inappropriata e/o poco robusta.

Ulteriore oggetto di attenzione riguarda il fatto che lo Z-Score non restituisce una vera probabilità di default ma soltanto un'indicazione classificatoria, rendendo il modello rigido: Zmijewski ha proposto così un approccio probabilistico con modelli probit in cui è possibile quantificare il rischio in termini continui, adattandolo meglio a contesti diversi e a differenti soglie di tolleranza.

In merito al modello O-Score di Ohlson, invece, riteneva che questo non affrontasse adeguatamente il problema della selezione delle variabili, siccome molte delle nove selezionate (i.e. logaritmo dell'attivo, variazione del reddito netto, indicatori dummy) non sono ben giustificati sotto il profilo teorico, ma sono state incluse soltanto per via della loro significatività statistica nel campione, rischiando sia di compromettere la robustezza esterna del modello, sia di creare overfitting.[34]

Zmijewski nel suo studio mostra perplessità sull'uso di alcune variabili qualitative binarie (es. utile netto negativo per due anni, equity negativa) presenti nell'O-Score, perché potrebbero non essere universalmente rappresentative del rischio di default, soprattutto se applicate in contesti internazionali o in periodi economici anomali.[35]

---

[33] Come anticipato, il modello Z-Score basandosi sull'analisi discriminante lineare assume sia una normalità multivariata delle variabili indipendenti, sia un'uguaglianza delle varianze e covarianze tra i gruppi (in questo caso fallite o non fallite).

[34] Per *robustezza esterna* si intende la capacità di un modello di mantenere la propria affidabilità predittiva anche quando viene applicato a dati o contesti diversi da quelli utilizzati per costruirlo; tale concetto è strettamente connesso all'overfitting, il quale si verifica quando un modello statistico si adatta troppo bene ai dati di addestramento (cogliendo anche rumore, anomalie o varianze casuali) presentando quindi elevate prestazioni su dati noti ma scarse capacità predittive su nuovi dati.

[35] Zmijewski, M. E. (1984). *Methodological Issues Related to the Estimation of Financial Distress Prediction Models*. Journal of Accounting Research, 59–82.

Nel 2001 T. Shumway pubblicò uno degli articoli più influenti degli ultimi decenni nel campo della previsione di default[36], in cui propose un cambio di paradigma metodologico, introducendo l'uso di modelli di tipo *hazard*[37], già diffusi in campo medico e attuariale, ma ancora poco utilizzati nella letteratura economico-finanziaria. Questo approccio, noto come *time-to-event analysis* è particolarmente adatto per modellare fenomeni che si sviluppano nel tempo, coerente con la crisi d'impresa, che non è un evento istantaneo ma un processo graduale.

Shumway mostrò che i modelli hazard superano i limiti dei modelli "tradizionali" presentati finora, i quali trattano dati come "cross-section" statici, ignorando la componente temporale; questo modello utilizza dati panel (più anni per ogni azienda) e considera l'evoluzione nel tempo della probabilità di fallimento.

Il modello adotta una specificazione log-lineare della funzione di rischio (hazard rate), simile alla regressione logistica dell'O-Score, ma applicata su dati longitudinali.

Ulteriore novità è rappresentata dal fatto che per la prima volta le variabili indipendenti includono, oltre agli indicatori contabili, variabili di mercato che spesso anticipano le difficoltà finanziarie e offrono segnali forward-looking più tempestivi rispetto al bilancio.

Le principali sono:

1. NITA → redditività netta / totale attivo;
2. TLTA → debiti totali / totale attivo
3. Cash flow / Totale attivo;
4. Excess return → rendimento azionario oltre il rendimento di mercato;
5. Standard deviation of return → deviazione standard del rendimento, ovvero volatilità del titolo

Rispetto ai modelli di Altman, Ohlson l'hazard model riduce il rischio di overfitting grazie ad una struttura più parsimoniosa e ad una calibrazione basata su eventi che si verificano nel tempo, anziché su confronti ex post tra imprese fallite e non.[38] Tuttavia, il modello richiede una maggiore disponibilità e qualità dei dati, soprattutto di dati di mercato aggiornati, difficilmente reperibili o affidabili in caso di imprese non quotate.

---

[36] Shumway, T. (2001). *Forecasting Bankruptcy More Accurately: A Simple Hazard Model*. Journal of Business, 74(1), 101–124.

[37] Gli hazard models stimano la probabilità che un evento – il default, in questo caso – si verifichi in un certo istante temporale, condizionatamente al fatto che non si sia ancora verificato fino a quel momento.

[38] Bauer, J., Agarwal, V. (2014). *Are hazard models superior to traditional bankruptcy prediction approach? A comprehensive test.* Journal of Banking and Finance, (40), 432-442.

In termini di accuratezza predittiva, Shumway dimostra empiricamente che il suo modello è in grado di prevedere i default aziendali in modo significativamente più preciso rispetto a Ohlson e Altman, soprattutto nel breve-medio termine.[39] Questa superiorità si manifesta soprattutto nella capacità di catturare variazioni annuali nel rischio, laddove i modelli statici offrono una visione più "fotografica" e meno dinamica della solvibilità aziendale.

In conclusione, l'analisi dei principali modelli di previsione della crisi d'impresa ha messo in luce un'evoluzione significativa nelle metodologie adottate per individuare, con sufficiente anticipo, le condizioni di deterioramento economico-finanziario delle imprese. A partire dai modelli discriminanti di prima generazione, come lo Z-Score di Altman (1968), fondati su combinazioni lineari di indici di bilancio, si è passati a sistemi più flessibili e probabilistici, come il modello logit di Ohlson (1980), che ha introdotto una lettura continua e probabilistica del rischio di fallimento. La riflessione metodologica di Zmijewski (1984) ha ulteriormente arricchito il dibattito, portando l'attenzione su questioni di robustezza statistica, parsimonia e corretto disegno campionario, mentre il contributo innovativo di Shumway (2001) ha segnato una svolta verso approcci dinamici e temporali, attraverso l'adozione del modello hazard e l'integrazione di variabili di mercato.

Nel complesso, emerge come non esista un modello "perfetto" o universalmente valido: ciascuna metodologia presenta vantaggi e limiti, che variano in funzione del contesto applicativo, della qualità dei dati disponibili, del tipo di impresa e del momento storico considerato. Tuttavia, il progresso metodologico ha reso sempre più sofisticati gli strumenti a disposizione del management, della revisione e degli organi di controllo, migliorando la capacità di intercettare per tempo le dinamiche di crisi. In questa prospettiva, la previsione della crisi non è più un esercizio accademico astratto, ma uno strumento operativo fondamentale per supportare decisioni strategiche, valutazioni creditizie e politiche pubbliche di prevenzione.

[39] Shumway, T. (2001). *Forecasting Bankruptcy More Accurately: A Simple Hazard Model*. Journal of Business, 74(1), 101–124

2 – MODELLI PREDITTIVI MODERNI: INTRODUZIONE AL MACHINE LEARNING

2.1 – Premessa

Negli ultimi due decenni, l'evoluzione delle tecnologie computazionali e la disponibilità crescente di grandi volumi di dati contabili, finanziari e di mercato hanno reso possibile lo sviluppo di metodi predittivi più sofisticati, capaci di superare alcune delle rigidità tipiche dei modelli tradizionali.

Alcune delle criticità tipiche dei modelli tradizionali evidenziate in precedenza potrebbero essere efficacemente mitigate mediante il ricorso a tecnologie di ultima generazione fondate su algoritmi di intelligenza artificiale, primo tra tutti il *machine learning* (ML)[40]. Quest'ultimo si configura quale tecnica di apprendimento automatico, in grado di emulare processi decisionali complessi e di anticipare l'insorgere di situazioni di criticità aziendale, attraverso l'elaborazione e l'analisi dei dati storici disponibili. L'algoritmo è concepito per rilevare pattern ricorrenti e strutture latenti all'interno dei dati, cogliendo analogie nei percorsi evolutivi dello stato di salute economico-finanziaria di insiemi di imprese accomunate da specifici fattori, tanto qualitativi quanto quantitativi. L'impiego di tali modelli predittivi, basati sull'intelligenza artificiale, può dunque costituire un prezioso ausilio per gli organi amministrativi, coadiuvandoli nell'attività interpretativa dei dati e nell'individuazione precoce di segnali di squilibrio, sin dalle fasi prodromiche del declino aziendale.[41]

La differenza sostanziale rispetto ai modelli classici come lo Z-Score o l'O-Score risiede nella capacità del machine learning di modellare relazioni complesse e non lineari tra variabili, evitando l'imposizione di vincoli rigidi sulle distribuzioni statistiche o sull'omoschedasticità dei gruppi. I modelli ML non richiedono assunzioni ex ante sulla forma funzionale che lega input e output, ma si basano su processi iterativi di apprendimento che consentono di adattarsi dinamicamente ai dati, migliorando la performance predittiva anche in presenza di rumore informativo o variabilità elevata.

---

[40] Sebbene machine learning e intelligenza artificiale siano usati spesso come sinonimi, in realtà il primo è un'applicazione pratica della seconda. Un algoritmo di intelligenza artificiale basato sul ML è in grado di imparare e auto modificarsi, consentendo di ricevere risposte per le quali non è stato programmato direttamente.

[41] Di Martino, M. (2024). *Intelligenza artificiale e sistemi di segnalazione della crisi d'impresa.* De Iustizia ISSN 2974-7562

Inoltre, l'applicazione di tecniche di machine learning consente di integrare fonti eterogenee di dati – dati contabili, indicatori di mercato, informazioni qualitative, news testuali o segnali sociali – offrendo una visione più completa e tempestiva del profilo di rischio di un'impresa.[42]

In particolare, nell'ambito della previsione delle crisi aziendali, l'adozione di algoritmi di apprendimento supervisionato ha mostrato risultati promettenti in termini di accuratezza, robustezza e capacità di intercettare situazioni borderline che sfuggirebbero agli strumenti tradizionali. Diversi studi empirici riportati in seguito, tra cui quello di Ha, Dang e Tran[43], dimostrano come algoritmi basati su reti neurali, support vector machine (SVM), random forest e boosting superino i modelli lineari convenzionali, soprattutto in scenari ad alta dimensionalità e forte eterogeneità tra imprese. Tali approcci permettono non solo di classificare correttamente le imprese a rischio, ma anche di stimare la probabilità continua di default e di aggiornare i modelli nel tempo, apprendendo dalle nuove osservazioni.

In questo scenario, la previsione della crisi d'impresa non è più un esercizio statico e unidimensionale, bensì un processo adattivo e multilivello, in grado di riflettere la complessità crescente dell'ambiente economico-finanziario. Il ricorso al machine learning si pone quindi come una naturale evoluzione metodologica, capace di dialogare con i modelli tradizionali, integrarli e superarli laddove necessario, aprendo nuovi orizzonti per la gestione anticipata del rischio e la protezione della continuità aziendale.[44]

2.2 – Logiche e fondamenti del machine learning applicati alla finanza

Il machine learning è quindi un sottoinsieme dell'intelligenza artificiale fondato sul presupposto che i sistemi informatici possano apprendere dai dati, migliorando progressivamente le proprie prestazioni predittive senza necessitare di una programmazione esplicita e rigida a priori. Tale caratteristica rende il ML uno strumento particolarmente adatto per affrontare problemi complessi e multivariati, come la

[42] Tipaldi, A. (2022). *Crisi d'impresa e modelli predittivi: il ruolo dell'I.A. e della tecnologia blockchain nella prevenzione e nella composizione della crisi.* Rivista di diritto dell'impresa, Edizioni Scientifiche Italiane.

[43] Ha, H. H., Dang, N. H., & Tran, M. D. (2023). *Financial distress forecasting with a machine learning approach*. Corporate Governance and Organizational Behavior Review, 7(3), 90–104.

[44] Agnoli, N., Zamboni, M. (2021). *Intelligenza artificiale e previsione delle crisi aziendali.* Amministrazione e finanza, n. 12.

previsione della crisi d'impresa, che coinvolge molteplici dimensioni aziendali (economica, finanziaria, patrimoniale, strategica) e relazioni non lineari tra le variabili.
I principali metodi tradizionali presentati nel capitolo precedente, basandosi su analisi discriminante e modelli logit o probit richiedono forti assunzioni sia sulla distribuzione dei dati (normalità, omoschedasticità), sia sulla linearità delle relazioni tra le variabili. Per tale motivo questi, pur rimanendo pionieristici ed utilizzati ancora oggi in molte realtà professionali ed aziendali, tendono a semplificare quella che è la realtà aziendale, presentando limiti in termini di flessibilità e adattabilità ai mutamenti economici. In tale ottica il ML sembra essere lo strumento più adatto per superare tali limiti, distinguendosi per la sua capacità di apprendere patterns complessi, anche altamente non lineari, attraverso meccanismi di ottimizzazione iterativa.[45]
Nel contesto della finanza aziendale, e in particolare nella previsione del rischio di fallimento, il machine learning si applica prevalentemente in modalità **supervisionata**, dove un algoritmo viene addestrato su un dataset etichettato in cui ogni osservazione è associata a un output noto (es. "fallita" vs "non fallita"). Il modello viene quindi validato su dati non visti (c.d. *test set*) per verificarne la capacità predittiva. Il vantaggio chiave è che tali algoritmi, una volta addestrati, possono generalizzare su nuove osservazioni senza necessità di riformulare il modello teorico.
Un importante contributo viene offerto dallo studio di Bzdol, Altman e Krzywinski[46], il quale si rivolge non solo a ricercatori in ambito computazionale, ma anche a studiosi e professionisti operanti in settori applicativi – come la medicina, la finanza o le scienze sociali – che intendano comprendere i principi e le potenzialità di questi strumenti predittivi, fornendo una panoramica accessibile ma rigorosa dell'apprendimento supervisionato, evidenziandone la centralità all'interno dell'odierno panorama delle tecniche di A.I.
L'apprendimento supervisionato si fonda sull'idea che, attraverso l'analisi di un insieme di dati etichettati (ossia, in cui ogni osservazione è associata a un output noto), un algoritmo possa apprendere la relazione tra input e output, e generalizzarla per prevedere risultati su dati non ancora osservati. Questo approccio è particolarmente adatto a compiti

[45] Athey, S. (2018). *The impact of machine learning on economics.* The Economics of Artificial Intelligence: An Agenda. University of Chicago Press.
[46] Bzdok, D., Altman, N., & Krzywinski, M. (2018). Supervised machine learning: a brief primer. Nature Methods, 15(12), 969–970

di classificazione (es. “in crisi” vs “non in crisi”) e di regressione (es. previsione dei flussi di cassa, del punteggio di rating, ecc.).

Una volta che l’algoritmo è stato addestrato su un dataset rappresentativo che prende il nome di *training set,* può applicare quanto imparato anche a dati nuovi sui quali non era mai stato allenato, con buoni risultati. Tale processo prende il nome di **generalizzazione**: in questo modo il modello non memorizza i dati, ma apprende quelle che vengono definite *strutture nascoste,* che può poi applicare a casi futuri: apprendendo questi patterns il modello cerca di estrarre regolarità, schemi e relazioni statistiche tra le variabili piuttosto che memorizzare i dati di addestramento.

Nel loro lavoro, gli autori offrono una chiara sistematizzazione degli algoritmi supervisionati più comunemente impiegati, che verranno approfonditi in seguito:

- **Regressione lineare e logistica**, adatti per modellare rispettivamente variabili continue e dicotomiche, ampiamente diffusi in ambito economico-finanziario per la loro semplicità e interpretabilità;
- **Alberi decisionali e foreste casuali**, che utilizzano strutture gerarchiche per prendere decisioni sequenziali, capaci di gestire variabili sia numeriche che categoriche;
- **Support Vector Machines (SVM)**, modelli che identificano l’iperpiano ottimale per separare le classi in uno spazio multidimensionale, ideali per contesti ad alta dimensionalità;
- **Reti neurali artificiali**, in grado di catturare relazioni non lineari molto complesse, ma meno interpretabili rispetto ai modelli lineari.

2.3 – Valutazione delle prestazioni predittive

Una sezione importante del paper è dedicata alla descrizione del processo di modellazione, che include la preparazione dei dati (pulizia, normalizzazione), la suddivisione del dataset in set di training, validazione e test, la selezione del modello più adatto e la sua calibrazione tramite tecniche di ottimizzazione degli iperparametri. Gli autori sottolineano l’importanza della valutazione delle prestazioni predittive attraverso metriche robuste, quali *l’accuratezza, la precisione, il recall, il punteggio F1 e l’area*

sotto la curva ROC (AUC), particolarmente utile quando le classi sono sbilanciate (ad esempio poche imprese in crisi rispetto a quelle sane).

L'AUC (Area under the curve) è un grafico che rappresenta il compromesso tra:

1) **True Positive Rate (TFR)** → la percentuale di positivi correttamente individuati dal modello;

$$\text{TPR} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}}$$

La formula esprime la percentuale di imprese effettivamente in crisi che il modello ha individuato come tali. Ad esempio se 100 imprese sono effettivamente in crisi e il modello ne rileva correttamente 80, si potrà affermare che il modello riconosce l'80% delle crisi;

2) **False Positive Rate (FPR)** → la percentuale di negativi classificati erroneamente come positivi

$$\text{FPR} = \frac{\text{False Positives (FP)}}{\text{False Positives (FP)} + \text{True Negatives (TN)}}$$

Rappresenta invece la percentuale di imprese sane che il modello sbaglia e classifica come "a rischio". Ad esempio se 900 imprese sono sane e il modello segnala erroneamente 90 di esse come "a rischio", si potrà affermare che il modello ha un 10% di "falsi allarmi".

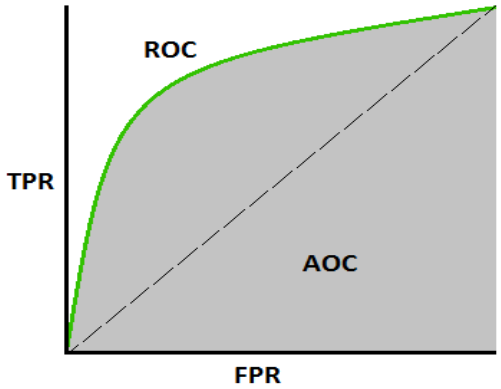

Figura 1

Fonte: https://www.domsoria.com/en/2023/07/what-are-auc-roc-curves/

L'asse delle ascisse (FPR – False Positive Rate) rappresenta la percentuale di falsi allarmi, ovvero quante imprese sane vengono erroneamente etichettate come "a rischio". L'asse delle ordinate (TPR – True Positive Rate) indica invece la sensibilità del modello, cioè la capacità di individuare correttamente le imprese effettivamente in crisi.

La curva verde mostra il comportamento del modello al variare della soglia di classificazione. Un modello efficace produce una curva che si avvicina all'angolo in alto a sinistra, dove il TPR è massimo e il FPR minimo. L'area sottesa alla curva, denominata **AUC** (Area Under the Curve), fornisce una misura sintetica della bontà del modello: un valore pari a 1 indica una classificazione perfetta, mentre un valore di 0.5 (linea tratteggiata) equivale a un modello casuale. Nel caso illustrato, l'area grigia (AUC > 0.8) suggerisce un'elevata capacità discriminante del modello.

In conclusione, il lavoro di Bzdok[47] fornisce una solida base concettuale e operativa per l'utilizzo dell'apprendimento supervisionato, presentandolo come un metodo potente e versatile, ma che richiede al contempo rigore metodologico, consapevolezza critica e attenzione all'impatto delle scelte algoritmiche sui contesti reali. Il contenuto, pur centrato sulla salute mentale, è facilmente estendibile anche ad ambiti come la finanza predittiva e la diagnosi precoce della crisi d'impresa, ove l'apprendimento dai dati storici e il riconoscimento di pattern deboli risultano centrali.

Più in linea con il lavoro del lavoro del presente elaborato risulta invece essere lo studio di Ha, Dang e Tran[48] i quali propongono un confronto sistematico tra modelli di apprendimento supervisionato e modelli tradizionali, con particolare riferimento all'analisi discriminante (Z-score) e al modello logit di Zmijewski. Lo studio, condotto su un ampio campione di imprese quotate presso la Borsa del Vietnam tra il 2009 e il 2020 (per un totale di 4.936 osservazioni), mira a valutare l'accuratezza e la solidità predittiva delle diverse metodologie nella classificazione delle imprese in stato di crisi finanziaria. Come anticipato, i modelli tradizionali presuppongono relazioni stabili e lineari tra le variabili, risultando poco adatti a cogliere le complessità e le interazioni non lineari tipiche dei fenomeni di crisi.

---

[47] du Jardin, P. (2015). *Bankruptcy prediction using terminal failure processes*. European Journal of Operational Research, 242(1), 286–303.

[48] Ha, H. H., Dang, N. H., & Tran, M. D. (2023). *Financial distress forecasting with a machine learning approach.* Corporate Governance and Organizational Behavior Review, 7(3), 90–104.

Per superare tali limiti, gli autori adottano una prospettiva basata sull'apprendimento supervisionato, testando l'efficacia di algoritmi avanzati come gli alberi decisionali, le foreste casuali e i modelli di boosting. Questi algoritmi, a differenza del modello di Altman e Zmijewski, non richiedono ipotesi stringenti sulle distribuzioni dei dati e sono in grado di apprendere pattern nascosti e interazioni complesse tra le variabili.

L'analisi empirica condotta mostra che i modelli di machine learning offrono performance predittive sensibilmente superiori, in particolare in termini di accuratezza complessiva e capacità di ridurre i falsi negativi (cioè i casi di imprese in crisi che "sfuggono" al modello). In particolare, l'algoritmo di Random Forest ha dimostrato di raggiungere un'accuratezza superiore al 98%, superando quella dei modelli di Altman e Zmijewski, i quali – pur mantenendo una buona capacità esplicativa – mostrano una maggiore rigidità nell'adattamento ai diversi contesti settoriali e dimensionali.

No.	*Method*	*Model 1 — Altman*	*Model 2 — Fich and Slezak*	*Model 3 — Zmijewski*
1	Logistic regression	0.93	0.89	0.97
2	Support vector machine	0.94	0.89	0.97
3	Decision tree	0.97	0.86	0.98
4	Random forest	**0.98**	0.90	**0.98**
5	K-nearest neighbors	0.81	0.84	0.93
6	Bayesian network	0.87	0.81	0.81

Tabella 1

Fonte: https://virtusinterpress.org/IMG/pdf/cgobrv7i3p8.pdf

Come mostrato in Tabella 1, il Random Forest ottiene un'accuratezza del 98% con il modello di Zmijewski e del 97%-98% anche con i modelli di Altman e Fich & Slezak. Al contrario, altri algoritmi come K-Nearest Neighbors e Bayesian Network presentano performance inferiori, specialmente in combinazione con i modelli Fich & Slezak.

Models		*Precision*	*Recall*	*F1-score*
Model 1 — Altman	Normal	0.98	0.97	0.97
	Financial distress	0.97	0.98	0.97
Model 2 — Fich and Slezak	Normal	0.92	0.95	0.93
	Financial distress	0.84	0.75	0.79
Model 3 — Zmijewski	Normal	0.99	0.99	0.99
	Financial distress	0.93	0.90	0.92

Tabella 2

Fonte: https://virtusinterpress.org/IMG/pdf/cgobrv7i3p8.pdf

Per valutare la qualità predittiva dei modelli, si ricorre a metriche che vadano oltre la semplice accuratezza. In particolare, la *precisione* misura la percentuale di osservazioni classificate come "in crisi" che effettivamente lo sono, mentre la *sensibilità* (recall)

esprime la capacità del modello di individuare correttamente tutte le imprese effettivamente in difficoltà.

Tuttavia, poiché queste due metriche possono entrare in tensione (un modello può avere alta precisione ma basso recall, o viceversa), si utilizza spesso un indice composito: *l'F1-score*, calcolato come media armonica tra precisione e sensibilità. Quest'ultimo offre una valutazione sintetica dell'efficacia complessiva del modello, ed è particolarmente utile in presenza di classi sbilanciate, come nel caso in cui le imprese in crisi siano in netta minoranza rispetto a quelle sane.

Il modello di Zmijewski raggiunge i valori più elevati sia in termini di precisione (0.99 per le imprese sane, 0.93 per quelle in crisi), che di recall (rispettivamente 0.99 e 0.90), con un F1-score globale di 0.99 e 0.92. Questi risultati indicano una capacità eccellente di individuare le imprese in crisi, riducendo al minimo sia i falsi positivi che i falsi negativi. Anche il modello di Altman si dimostra solido (F1 = 0.97), mentre quello di Fich & Slezak risulta più debole nella classe "distress" (F1 = 0.79), suggerendo una minore capacità di intercettare correttamente segnali precoci di deterioramento.

Lo studio evidenzia inoltre che le variabili con il maggiore potere predittivo risultano essere: il rapporto debito/capitale proprio, il turnover degli attivi e il margine di profitto. Tali indicatori, pur presenti anche nei modelli tradizionali, vengono valorizzati in maniera più efficace dagli algoritmi di apprendimento supervisionato, grazie alla loro capacità di combinare in modo dinamico ed evolutivo le informazioni derivanti dai dati.

Model 1 — Altman		***Model 2 — Fich and Slezak***		***Model 3 — Zmijewski***	
Variables	***Coeff.***	***Variables***	***Coeff.***	***Variables***	***Coeff.***
X13: Asset turnover	-3.5960	*X8: Margin*	-0.4141	*X5: Debt-to-equity ratio*	2.1389
X5: Debt-to-equity ratio	2.6649	*X9: Marginal operating income*	-4.6605	*X1: Current ratio*	-3.1924
X3: Receivable turnover	0.0261	*X11: Net return on equity*	-13.0787	*X8: Margin*	29.6188
X1: Current ratio	-1.0672	*X5: Debt-to-equity ratio*	1.0379	*X9: Marginal operating income*	-34.4099
X14: Inventory turnover period	-0.0077	*X10: Net return to equity book value*	-7.5967	*X2: Quick ratio*	-2.4527
X2: Quick ratio	0.9964	*X12: Ratio of operating income to book value of equity*	-7.4179	*X3: Receivable turnover*	-0.2399
X7: Operating cash flow to total debt	-0.1734	*X22: Earnings per share*	0.0001	*X14: Inventory turnover period*	0.0002
X15: Fixed asset turnover	-0.0010	*X1: Current ratio*	-0.1625	*X13: Asset turnover*	-0.9261

Tabella 3

Fonte: https://virtusinterpress.org/IMG/pdf/cgobrv7i3p8.pdf

Sul piano applicativo, i risultati della ricerca suggeriscono l'opportunità di affiancare – se non sostituire – i modelli tradizionali con strumenti di machine learning all'interno dei sistemi di early warning aziendali e bancari. Tali strumenti possono infatti offrire agli amministratori, ai revisori e agli enti finanziatori una visione più tempestiva e dettagliata dello stato di salute delle imprese, consentendo l'adozione di misure correttive in fase precoce.

In sintesi, lo studio di Ha, Dang e Tran[49] conferma che l'apprendimento supervisionato rappresenta un'evoluzione metodologica significativa rispetto all'approccio statistico classico, soprattutto per la capacità di adattarsi a contesti caratterizzati da elevata variabilità e incertezza, come quelli dei mercati emergenti.

2.4 – Principali algoritmi di ML per la previsione della crisi

2.4.1 – Random Forest (RF)

Tra gli algoritmi di apprendimento supervisionato più efficaci nella classificazione predittiva si annovera il Random Forest, sviluppato da Leo Breiman presso l'Università della California, Berkeley.[50]

Questo metodo si fonda sull'idea di costruire una moltitudine di alberi decisionali (da cui il termine "foresta"), ognuno dei quali contribuisce in maniera indipendente al processo decisionale. Il risultato finale viene ottenuto attraverso un meccanismo di voto maggioritario: ciascun albero "vota" per la classe di appartenenza più probabile, e la classe che ottiene più voti rappresenta la previsione del modello complessivo.

La forza del Random Forest risiede nella sua capacità di ridurre l'errore di classificazione aggregando i risultati di molti modelli deboli (i singoli alberi), ciascuno dei quali viene costruito in modo leggermente diverso dagli altri. A tal fine, durante la fase di addestramento, il modello genera vettori casuali che guidano la crescita di ogni singolo albero in maniera indipendente. Per ogni albero viene infatti prodotto un vettore casuale Vk, utilizzato per selezionare in modo casuale e ripetuto i dati e le variabili su cui costruire

[49] Ha, H. H., Dang, N. H., & Tran, M. D. (2023). *Financial distress forecasting with a machine learning approach*. Corporate Governance and Organizational Behavior Review, 7(3), 90–104.
[50] Breiman, L. (2001), *Random Forests,* Statistics Department University of Berkeley, CA.

l'albero stesso. Nonostante ciascun vettore sia generato in modo indipendente dagli altri, essi seguono la stessa distribuzione statistica, garantendo coerenza e varietà all'intero insieme.

Ogni albero, dunque, produce una propria classificazione basata su un sottoinsieme del dataset e su un percorso decisionale guidato dal relativo vettore Vk. Dopo aver costruito un numero sufficiente di alberi (tipicamente nell'ordine delle centinaia), il modello aggrega le decisioni individuali e determina la classe finale in base al criterio della maggioranza semplice.[51]

Questo approccio consente al modello di essere estremamente robusto, flessibile e preciso, soprattutto in presenza di grandi quantità di dati e di relazioni complesse tra le variabili. Inoltre, la Random Forest è meno soggetta all'overfitting rispetto ai singoli alberi decisionali, rendendola una delle tecniche più utilizzate nella previsione di eventi aziendali critici, come lo stato di crisi d'impresa.[52]

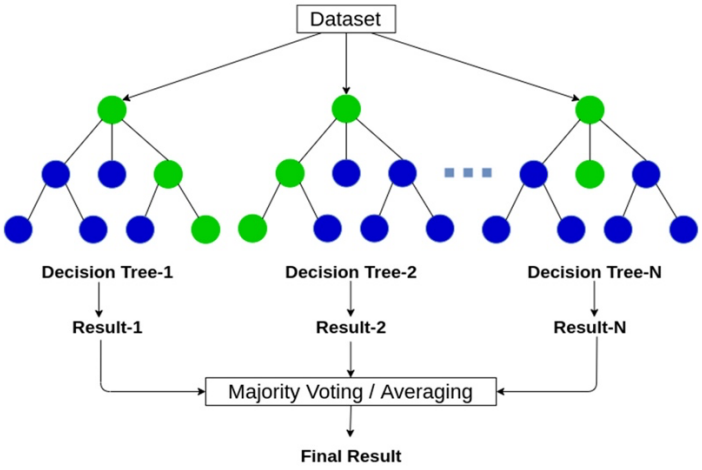

Figura 2

Fonte: https://datasciencedojo.com/blog/random-forest-algorithm/

[51] Baiju, B., & Rufsana, A. R. (2023). *A review on bankruptcy prediction using machine learning techniques*. In Proceedings of the International Conference Emerging Trends in Engineering

[52] Abikoye, B. (2021). *Machine Learning Models and AI for Predicting Financial Crises: Applications and Accuracy.*

Ipotizzando di avere un dataset iniziale contente dati storici di aziende con l'obiettivo di prevedere il rischio di crisi di un'impresa utilizzando indicatori economico-finanziari come:

- X1: indebitamento;
- X2: liquidità corrente;
- X3: ROI;
- X4: grado di copertura degli interessi;
- X5: Fatturato / attivo totale.

Il modello di RF creerà molti alberi decisionali (*Decision Tree-1, Tree-2, … Tree-N)* ognuno dei quali verrà addestrato su un sottocampione casuale del dataset usando un sottoinsieme casuale delle variabili in ogni nodo (ad esempio Tree-1 potrà usare soltanto ROI e leverage, Tree-2 solo indebitamento e liquidità corrente, e così via…).
Il modello restituirà un risultato individuale per ogni albero, ad esempio:

- Tree-1 → "crisi";
- Tree-2 → "non crisi";
- Tree-3 → "crisi";

Al termine di tale processo, se la maggioranza degli alberi prevede "crisi", il modello assegnerà rischio di fallimento; se invece si intende ottenere un risultato in termini continui (ad esempio la probabilità di crisi), si prende la media delle previsioni restituite dal modello. [53]
Come premesso, il RF si mostra particolarmente valido per la sua:

- Robustezza: combinando molti modelli, riduce l'eventuale errore dovuto al singolo albero (si ha così una riduzione dell'overfitting);
- Stabilità: l'influenza del – fisiologico – rumore presente nel dataset (errori di registrazione in fase di redazione del bilancio, un calo di fatturato dovuto a cause esterne e momentanee, bonus fiscali una tantum etc…) viene mitigata dal fatto che il modello media su molti alberi;
- Flessibilità: il modello restituisce risultati attendibili su dataset complessi con molte variabili.

[53] Alanis, E., Chava, S., & Shah, A. (2022). *Benchmarking machine learning models to predict corporate bankruptcy*

Tali considerazioni trovano riscontro nello studio condotto da Barboza, Kimura ed Altman[54], riferimento imprescindibile per i lavori successivi, il quale evidenzia come i modelli di ML – in particolare il RF – offrano performance superiori rispetto alle tecniche statistiche tradizionali. Il lavoro confronta diversi modelli di ML (SVM, bagging, boosting, RF) con metodi tradizionali quali l'analisi discriminante lineare (Z-Score di Altman) e la regressione logistica (O-Score di Ohlson) per la previsione della bancarotta aziendale, usando dati di società nordamericane dal 1985 al 2013, con l'obiettivo di verificare se i metodi avanzati offrano maggiore accuratezza predittiva, specialmente integrando nuove variabili finanziarie rispetto al classico modello Z-Score, quali:

1) Margine operativo;
2) Crescita attivi;
3) Crescita vendite;
4) Crescita dipendenti;
5) Variazione ROE;
6) Variazione price-to-book.

Il dataset utilizzato contiene oltre 10.000 osservazioni aziendali/anno, integrate da Compustat e Salomon Center Database, suddivise come segue:

- Training sample → 1985-2005 (449 imprese fallite + 449 imprese non fallite);[55]
- Testing sample → 2006-2013 (133 imprese fallite + 13.167 imprese sane).[56]

[54] Barboza, F., Kimura, H., Altman, E.I. (2017), *Machine learning models and bankruptcy prediction.* Expert System with Applications*,* 83, 405-417.

[55] Utilizzato per addestrare l'algoritmo a riconoscere le relazioni tra le variabili indipendenti e l'output atteso.

[56] Impiegato per testare le capacità predittive del modello su dati mai visti prima e per ottimizzarne la configurazione, al fine di prevenire fenomeni di overfitting e garantire una buona generalizzazione.

Training sample Model	TP	TN	FP	FN	Type I Error (%)	Type II Error (%)	AUC (%)	ACC (%)
SVM-Linear	419	306	143	30	6.68	31.85	NA	80.73
SVM-RBF	421	376	73	28	6.24	16.26	NA	88.75
Boosting	434	430	19	15	3.34	4.23	NA	96.21
Bagging	448	447	2	1	0.22	0.45	NA	99.67
Random forest	449	449	0	0	0.00	0.00	NA	100.00
Neural networks	431	331	118	18	4.01	26.28	NA	84.86
Logit	414	329	120	35	7.80	26.73	NA	82.74
MDA	361	221	228	88	19.60	50.78	NA	64.81
Testing sample Model	TP	TN	FP	FN	Type I Error (%)	Type II Error (%)	AUC (%)	ACC (%)
SVM-Linear	123	9,389	3,778	10	7.52	28.69	67.2	71.52
SVM-RBF	105	10,505	2,662	28	21.05	20.22	85.17	79.77
Boosting	108	11,417	1,750	25	18.80	13.29	92.97	86.65
Bagging	110	11,284	1,883	23	17.29	14.30	92.48	85.67
Random forest	111	11,468	1,699	22	16.54	12.90	92.92	87.06
Neural networks	124	9,582	3,585	9	6.77	27.23	90.08	72.98
Logit	118	10,028	3,139	15	11.28	23.84	90.10	76.29
MDA	86	6,854	6,313	47	35.34	47.95	63.68	52.18

Tabella 4

Fonte: https://pdf.sciencedirectassets.com/

Come si evince dalla tabella 4, RF e Bagging classificano quasi perfettamente i dati su cui sono stati addestrati: tuttavia, se è vero che questo è indice di un ottimo apprendimento, è altrettanto vero che non garantisce generalizzazione e vi è un forte rischio di overfitting.

Sul Validation Set a performare meglio sono gli algoritmi più moderni: spicca tra questi il RF che presenta la maggior accuratezza (87,06%) e la migliore AUC (92,92%) dimostrandosi il modello più bilanciato ed affidabile, ma anche Bagging e Boosting presentano degli ottimi risultati. L'analisi discriminante lineare (Z-Score) presenta invece prestazioni molto basse, con accuratezza appena del 52,18%.

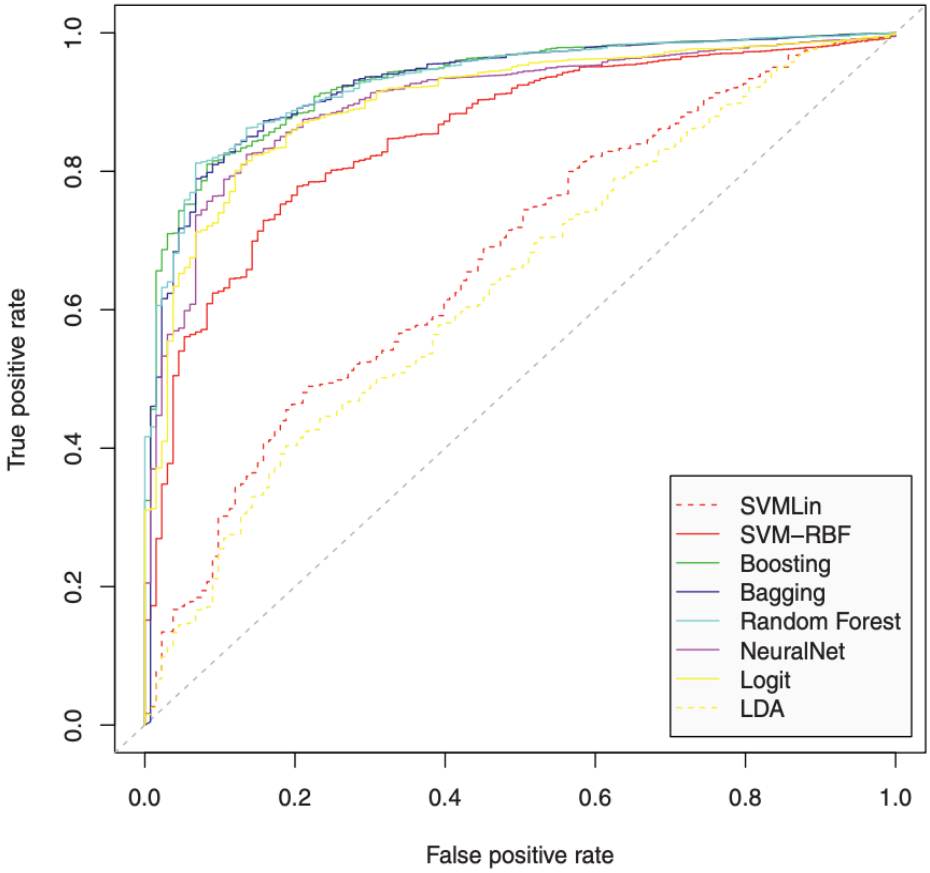

Figura 3

Fonte: Barboza, F., Kimura, H., Altman, E.I. (2017), *Machine learning models and bankruptcy prediction.* Expert System with Applications, 83, 405-417.

La figura 3 mostra le curve ROC ed offre una rappresentazione visiva della capacità discriminante dei modelli predittivi testati. Tra questi, il modello Random Forest si distingue per l'elevata qualità della curva, che si avvicina all'angolo in alto a sinistra, indicando un'eccellente capacità di identificare correttamente le imprese in crisi con un tasso contenuto di falsi positivi. Anche Boosting e Bagging mostrano performance molto competitive. Al contrario, modelli tradizionali come la regressione logistica e soprattutto l'analisi discriminante lineare (LDA) presentano curve significativamente inferiori, riflettendo una minore capacità di discriminazione, quasi vicine alla casualità. Questi risultati rafforzano l'evidenza che i metodi di machine learning, in particolare quelli basati su ensemble learning, siano più efficaci nella previsione del rischio di fallimento rispetto agli approcci statistici classici.

2.4.2 – Support Vector Machine (SVM)

Il modello Support Vector Machines (SVM) identifica l'iperpiano che separa al meglio le classi di rischio. SVM è particolarmente efficace quando i dati non sono linearmente separabili, grazie alla possibilità di mappare i dati in spazi dimensionali più alti tramite kernel.[57]

Mediante questo algoritmo è possibile prevedere se un'azienda è in crisi (classe 1) o non in crisi (classe 2) utilizzando variabili economico-finanziarie. Ipotizzando di avere due classi:

- Classe 1 (triangoli blu in figura 4) → aziende in crisi;
- Classe 2 (cerchi verdi in figura 4) → aziende sane.

L'SVM cerca l'iperpiano ottimale (riga nera centrale) che massimizza il margine tra i due gruppi. Questo margine è definito dalle linee rosse tratteggiate e dai *support vector*: essendo i punti più vicini al confine, rappresentano gli esempi più "al limite": spostando anche solo uno di questi punti, cambia l'intera linea di separazione, mentre gli altri, più lontani, non influenzano direttamente l'iperpiano.

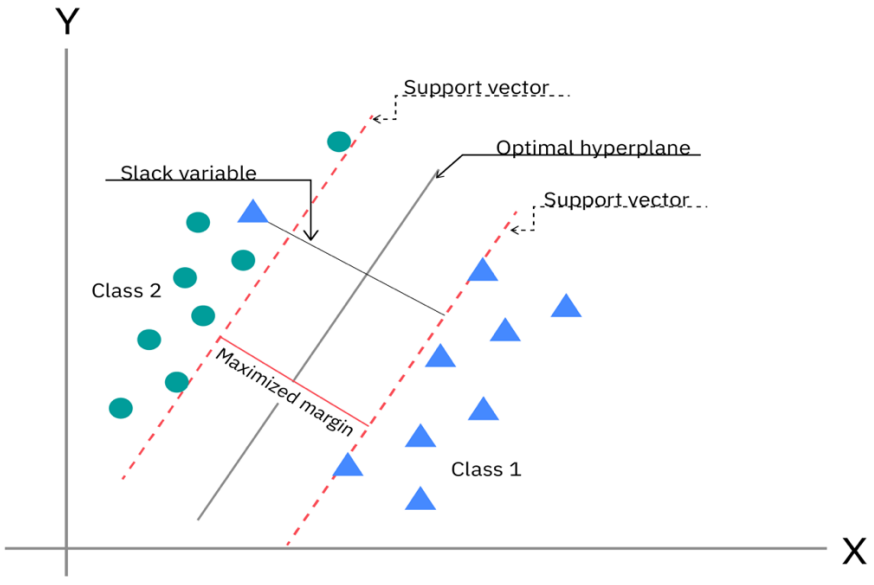

Figura 4

Fonte: https://www.ibm.com/think/topics/support-vector-machine

[57] Ruotolo, M. (2013). *Analisi di SVM per la classificazione automatica.* Università di Torino.

Applicato ad un contesto predittivo della crisi d'impresa, i support vectors rappresentano le aziende "di confine", cioè quelle più vicine al punto di crisi: il modello impara proprio da questi casi critici, che fungono da base per costruire una previsione efficace.
Componente fondamentale del modello è rappresentata dalla *funzione kernel*, uno strumento matematico che consente di trasformare i dati in uno spazio a dimensionalità superiore, dove diventa più facile separare le classi anche quando queste non sono separabili in modo lineare nello spazio originale. Mediante il c.d. *kernel trick*, la funzione calcola direttamente le distanze tra i punti come se fossero già nello spazio trasformato, risparmiando così tempo e potenza di calcolo.

Tipo di kernel	Formula (semplificata)	Quando si usa
Lineare	$K(x, x') = x \cdot x'$	Quando i dati sono già separabili
Polinomiale	$K(x, x') = (x \cdot x' + c)^d$	Per relazioni più complesse
Radiale (RBF/Gauss.)	$K(x, x') = e^{-\gamma\|x-x'\|^2}$	Quando la separazione è molto non lineare
Sigmoid	$K(x, x') = \tanh(\alpha x \cdot x' + c)$	Meno usato, simile a reti neurali

L'SVM si mostra particolarmente adatto alla previsione della crisi in quanto nella realtà le aziende non sono sempre chiaramente separabili siccome i dati finanziari non sempre seguono schemi lineari. Un kernel, ad esempio radiale, può modellare confini complessi tra aziende sane e in crisi, migliorando la precisione del modello.
Il vantaggio dell'SVM risiede proprio nel fatto che, lavorando in spazi trasformati grazie alle funzioni kernel, riesce a separare le classi anche in presenza di relazioni non lineari tra le variabili non predittive.[58]

[58] Horak, J., Vrbka, J., & Suler, P. (2020). *Support vector machine methods and artificial neural networks used for the development of bankruptcy prediction models and their comparison.* Journal of Risk and Financial Management, 13(3)

In linea con questa visione, Dellepiane, Marcantonio, Laghi e Renzi[59] affrontano il tema della previsione della crisi d'impresa attraverso l'uso dei SVM con particolare attenzione al contesto macroeconomico della crisi finanziaria del 2008.

Lo studio si distingue per aver integrato – attraverso tecniche di feature automatizzata – nei modelli predittivi non solo ulteriori indicatori contabili (come EBIT/Total Asset, Net Income/Total Asset), ma anche variabili macroeconomiche specifiche per paese, quali tassi sovrani a 10 anni e spread di rating, al fine di migliorare l'accuratezza predittiva e adattarsi alle dinamiche sistemiche del contesto post crisi.

Dopo un confronto iniziale tra i metodi classici (Z-Score, regressione logistica) gli autori utilizzando le SVM come metodo di apprendimento supervisionato per la classificazione binaria di imprese insolventi e non. Il dataset è composto da dati di 6.929 aziende internazionali, con due sottocampioni bilanciati (Sample 1 e Sample 2) per testare la capacità predittiva a uno e a due anni di distanza dal default.

Per ogni sottocampione vengono riportati:

- cc/all: accuratezza complessiva del modello (% di osservazioni correttamente classificate);
- NS/aNS: accuratezza nella classificazione di imprese fallite, ovvero i recall della classe critica.

[59] Dellepiane, U., Di Marcantonio, M., Laghi, E., & Renzi, S. (2015). *Bankruptcy Prediction Using Support Vector Machines and Feature Selection During the Recent Financial Crisis*. INTERNATIONAL JOURNAL OF ECONOMICS AND FINANCE, 182-196.

Sample		LDA		Logit		SVM	
		NS	S	NS	S	NS	S
	NS	20	13	26	7	28	5
Sample 1A	S	10	28	7	31	8	30
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		67.61%	60.61%	80.28%	78.79%	81.69%	84.85%
		NS	S	NS	S	NS	S
	NS	15	18	22	11	28	5
Sample 1B	S	4	34	4	34	7	31
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		69.01%	45.45%	78.87%	66.67%	83.10%	84.85%
		NS	S	NS	S	NS	S
	NS	14	12	19	7	25	1
Sample 1C	S	6	35	6	35	8	33
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		76.06%	63.33%	83.10%	80.00%	83.10%	86.67%
		NS	S	NS	S	NS	S
	NS	14	13	21	6	23	4
Sample 1D	S	4	41	8	37	19	26
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		77.46%	53.85%	78.87%	73.08%	71.83%	96.15%
		NS	S	NS	S	NS	S
	NS	21	13	25	9	30	4
Sample 1E	S	2	35	2	35	8	29
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		78.87%	61.76%	84.51%	73.53%	83.10%	88.24%
Mean		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		73.80%	57.00%	81.13%	74.41%	80.56%	88.15%

Tabella 5

Fonte: https://doi.org/10.5539/ijef.v7n8p

La Tabella 5 presenta un confronto tra tre modelli predittivi – LDA, regressione logistica (Logit) e SVM – applicati a cinque sottoinsiemi del campione (Sample 1A–1E), con l'obiettivo di prevedere lo stato di crisi aziendale a un anno di distanza, utilizzando un set completo di variabili contabili e macroeconomiche.

1) SVM:

È il modello più performante su entrambe le metriche:

- cc/all → 80,56%;
- NS/aNS → 88,15%

In ogni sotto-campione analizzato, l'SVM mantiene una performance superiore o pari all'83% di accuratezza complessiva, con un massimo del 96,15% di recall nel Sample 1D.

2) ***Regressione Logistica (Logit):***

Si presenta come secondo modello più efficace con i seguenti risultati:

- cc/all → 81,13% (leggermente superiore a SVM)
- NS/aNS → 74,41%

Ciò significa che il modello tende a classificare meglio le imprese sane (elevata accuratezza totale), ma mostra una minore sensibilità nel riconoscere i fallimenti, con valori di recall inferiori a SVM in tutti i sotto-campioni.

Rispetto all'SVM, è un modello meno "sensibile", quindi più conservativo ma rischia di non rilevare molte crisi reali.

3) ***LDA (Linear Discriminant Analysis):***

Risulta, ancora una volta, essere il modello meno performante dei tre:

- cc/all → 73,80%
- NS/aNS → 57,00%

Le prestazioni del modello sono estremamente altalenanti, con recall molto basso in alcuni sotto-campioni (es. solo 45,45% in Sample 1B).

L'LDA fatica chiaramente a identificare le imprese in difficoltà, confermando i limiti dei modelli lineari in contesti con elevata variabilità e relazioni complesse tra le variabili.

I risultati della tabella 2.5 confermano che il modello Support Vector Machine, specialmente quando arricchito con variabili macroeconomiche e sottoposto a una selezione ottimale delle feature, è il più efficace nella previsione del rischio di fallimento a breve termine. Esso offre un buon compromesso tra accuratezza generale e capacità di intercettare correttamente le situazioni di crisi (recall). Al contrario, modelli statistici più tradizionali come LDA mostrano performance sensibilmente inferiori, specialmente nella gestione dei falsi negativi, e risultano quindi meno adatti a contesti dinamici e complessi.

Sample		LDA		Logit		SVM	
		NS	S	NS	S	NS	S
	NS	16	13	17	12	24	5
Sample 2A	S	4	35	11	28	10	29
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		75.00%	55.17%	66.18%	58.62%	77.94%	82.76%
		NS	S	NS	S	NS	S
	NS	17	11	18	10	26	2
Sample 2B	S	7	33	12	28	8	32
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		73.53%	60.71%	67.65%	64.29%	85.29%	92.86%
		NS	S	NS	S	NS	S
	NS	18	9	21	6	24	3
Sample 2C	S	5	36	4	37	6	35
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		79.41%	66.67%	85.29%	77.78%	86.76%	88.89%
		NS	S	NS	S	NS	S
	NS	14	13	21	6	23	4
Sample 2D	S	0	41	7	34	11	30
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		80.88%	51.85%	80.88%	77.78%	77.94%	85.19%
		NS	S	NS	S	NS	S
	NS	18	12	21	9	26	4
Sample 2E	S	3	35	7	31	8	30
		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		77.94%	60.00%	76.47%	70.00%	82.35%	86.67%
Mean		cc/all	NS/aNS	cc/all	NS/aNS	cc/all	NS/aNS
		77.35%	58.88%	75.29%	69.69%	82.06%	87.27%

Tabella 6

Fonte: https://doi.org/10.5539/ijef.v7n8p

La tabella 6 illustra la capacità predittiva dei tre modelli nell'identificare le imprese a rischio di default con due anni di anticipo, su cinque sotto-campioni. Rispetto alla previsione a un anno della tabella 5, questo test rappresenta sicuramente un esercizio di previsione più sfidante data la maggiore incertezza nel lungo periodo.

Lo studio ha restituito i seguenti risultati:

1) ***SVM***

Il modello conferma la sua superiorità rispetto agli altri anche con un orizzonte di previsione più esteso presentando:

- cc/all → 82,06%;
- NS/aNS → 87,27%.

Raggiungendo picchi particolarmente elevati in Sample 2B (92,86% di NS/aNS) e Sample 2C (88,89%): questo indica che, nonostante il maggior orizzonte temporale, l'SVM è in grado di intercettare correttamente quasi 9 imprese su 10.

2) ***Logit***

Meno performante dell'SVM ma mantiene risultati solidi:

- cc/all → 75,29%;
- NS/aNS → 69,69%.

Il modello mostra una certa stabilità tra i campioni, ma tende a classificare con minor efficacia le imprese in crisi, rispetto all'SVM.

3) ***LDA***

Si riconferma il modello con le prestazioni peggiori:

- cc/all → 77,35% (molto vicino al modello Logit);
- NS/aNS → 58,88%

Pur avendo in alcuni casi una buona accuratezza generale (80,88% nel Sample 2D), l'incapacità di rilevare tempestivamente i fallimenti aziendali (51,85% di NS/aNS in 2D) lo rende poco adatto all'analisi predittiva in ottica gestionale e di risk management.

La previsione del fallimento a due anni rappresenta un obiettivo ambizioso, ma i risultati dimostrano che le SVM sono in grado di mantenere elevata precisione e sensibilità, anche in un orizzonte temporale più esteso. Con una media dell'87,27% di recall sulla classe "fallita", l'SVM si conferma come il modello più affidabile nella previsione anticipata della crisi d'impresa, rendendolo uno strumento particolarmente utile per l'implementazione di sistemi di early warning e per il supporto alle decisioni manageriali. Al contrario, l'uso di modelli lineari come LDA risulta meno efficace in contesti ad alta complessità e variabilità, come quello economico-finanziario post-crisi.

2.4.3 – Gradient Boosting Machines (GBM)

Il Gradient Boosting è una metodologia particolarmente indicata per la previsione della crisi d'impresa in quanto consente di apprendere in modo progressivo e raffinato gli errori dei modelli precedenti. A differenza di tecniche lineari come la regressione logistica, il Gradient Boosting può cogliere pattern non lineari e interazioni complesse tra variabili, ad esempio tra l'indebitamento, la variazione del fatturato e la presenza di ritardi nei pagamenti. L'uso di questa tecnica consente di ottenere algoritmi con un elevato grado di accuratezza, in grado di classificare correttamente non solo le imprese manifestamente in crisi, ma anche quelle che presentano segnali più deboli o ambigui. Questo aspetto è di fondamentale importanza nei modelli di *early warning*, dove l'obiettivo non è tanto rilevare una crisi già conclamata, quanto anticipare un deterioramento della posizione finanziaria dell'impresa.

In particolare, il Gradient Boosting si adatta efficacemente a dataset sbilanciati, tipici degli studi sulla crisi d'impresa, in cui le aziende fallite costituiscono una minoranza rispetto a quelle in continuità. Mentre modelli lineari tendono a sovrastimare la classe dominante, il boosting consente di focalizzare l'apprendimento proprio sulle osservazioni più difficili da classificare, ovvero quelle "vicine al confine" tra stabilità e insolvenza.

Un ulteriore vantaggio risiede nella capacità di gestione automatica dei dati mancanti *(missing values),* che rappresentano una sfida ricorrente nell'analisi dei bilanci aziendali. Grazie all'uso di alberi decisionali come base dell'ensemble, il Gradient Boosting è in grado di effettuare splitting intelligenti anche in presenza di variabili incomplete, mantenendo l'efficienza predittiva.

Dal punto di vista operativo, il Gradient Boosting permette inoltre di attribuire un peso diverso alle osservazioni, offrendo la possibilità di rafforzare il segnale predittivo per imprese particolarmente rilevanti o settori ad alto rischio. Tale flessibilità è preziosa in un'ottica di risk management, in quanto consente di personalizzare il modello in base al contesto d'analisi, sia esso settoriale, regionale o dimensionale.

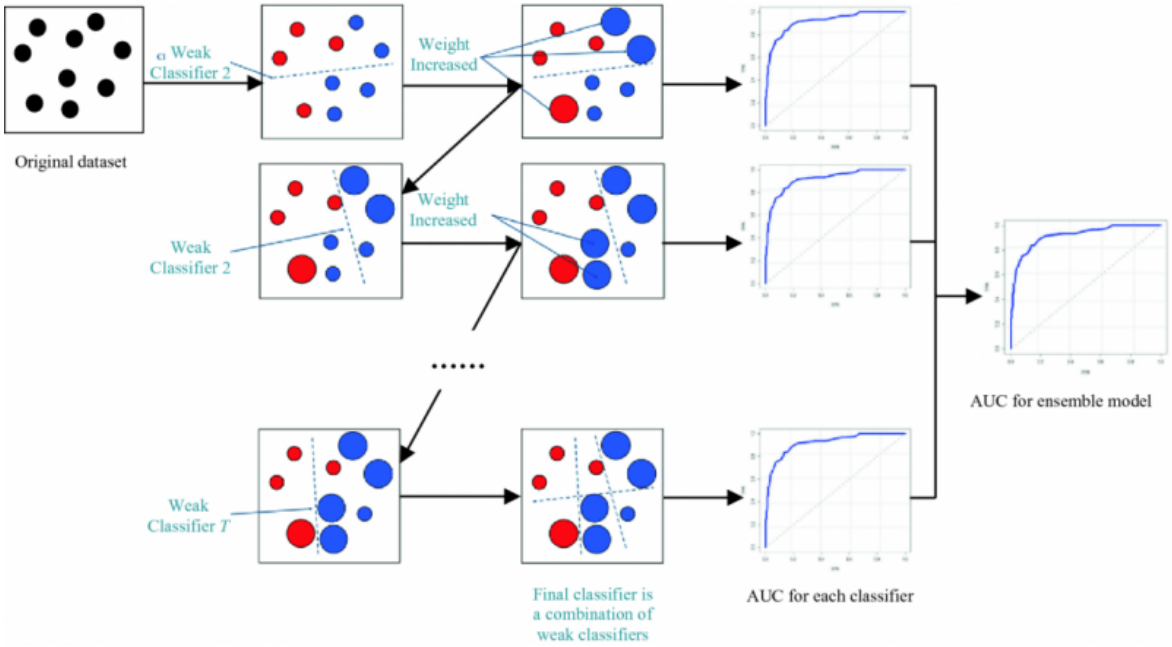

Figura 5

Fonte: https://datascience.eu/it/apprendimento-automatico/gradient-boosting-cosa-bisogna-sapere/

La figura 5 descrive efficacemente il funzionamento del Gradient Boosting come processo iterativo di apprendimento, in cui ciascun classificatore successivo corregge gli errori del precedente. Applicato alla previsione della crisi d'impresa, ciò significa che il modello impara progressivamente a riconoscere segnali anche deboli o non lineari di deterioramento finanziario. Imprese che non vengono identificate come a rischio al primo passaggio, magari perché presentano indicatori contabili in apparenza solidi ma celano squilibri più sottili – come un forte indebitamento a breve termine coperto da una temporanea liquidità, o una crescita anomala dei ricavi non accompagnata da utili – vengono via via intercettate nei passaggi successivi del boosting. Ogni nuovo classificatore apprende da questi "errori" e contribuisce a raffinare la sensibilità del modello globale, arrivando a distinguere con sempre maggiore precisione tra imprese sane e imprese in crisi.[60]

Questo approccio si dimostra particolarmente efficace nel contesto della previsione d'impresa, poiché la crisi raramente è determinata da un unico indicatore evidente: al

[60] Alanis, E., Chava, S., & Shah, A. (2022). *Benchmarking machine learning models to predict corporate bankruptcy.* arXiv

contrario, è spesso il risultato di combinazioni complesse e latenti di segnali deboli e discontinui. Il Gradient Boosting, grazie alla sua architettura basata su alberi decisionali e correzione iterativa dell'errore, è in grado di **catturare queste non linearità e interazioni tra variabili**, offrendo un modello predittivo che va oltre le semplificazioni delle tecniche statistiche classiche.

Inoltre, la fase di aggregazione finale – rappresentata nel diagramma con la curva ROC composita – dimostra che il modello finale ha una capacità discriminante (AUC) superiore rispetto ai singoli classificatori. Questo aspetto è cruciale per i sistemi di *early warning* aziendale, in quanto consente di ridurre sia i falsi negativi (imprese in crisi non identificate) che i falsi positivi (aziende sane erroneamente classificate come a rischio), con evidenti benefici per banche, revisori, creditori e autorità di vigilanza.

Infine, la logica di correzione progressiva su cui si fonda il Gradient Boosting rispecchia un principio operativo importante anche per i decisori aziendali: l'analisi predittiva non si basa su un'unica fotografia, ma su un apprendimento continuo, capace di adattarsi a nuove informazioni e a contesti dinamici. In quest'ottica, il Gradient Boosting si configura non solo come uno strumento tecnico di classificazione, ma anche come un vero e proprio alleato strategico per la gestione del rischio di crisi.

Numerosi studi empirici hanno dimostrato l'efficacia di questa tecnica in ambito finanziario e creditizio. Tra questi, nello studio *"Machine learning models for bankruptcy prediction in Italy: Do industrial variables count?"* [61] vengono applicati modelli di boosting alle imprese italiane, mostrando come l'accuratezza del modello superasse ampiamente quella dei benchmark tradizionali. Lo studio si propone di costruire modelli predittivi affidabili del fallimento aziendale in Italia, verificando se l'inclusione di variabili industriali (relative al settore di appartenenza) migliori la performance predittiva rispetto ai soli indicatori di bilancio, con l'obiettivo di sviluppare un modello predittivo, specificamente progettato per l'economia italiana, in grado di classificare le imprese come solventi o insolventi con un anno di anticipo, utilizzando il database AIDA di Bureau van Dijk relativo al periodo 2007–2015. A fianco dei consueti indicatori finanziari impiegati in letteratura (ROA, leverage, cashflow), il modello incorpora un insieme di

[61] Bragoli, D., Ferretti, C., Ganugi, P., Marseguerra, G., Mezzogori, D., & Zammori, F. (2019). *Machine learning models for bankruptcy prediction in Italy: Do industrial variables count?* (Working Paper No. 19/3). Università Cattolica del Sacro Cuore, Dipartimento di Matematica per le Scienze Economiche, Finanziarie ed Attuariali.

variabili industriali e regionali, tra cui: *densità di imprese nel settore, quota di mercato, presenza in distretti industriali e mark-up settoriale medio*.

Le metriche di valutazione sono le seguenti:

1) T1 (Type 1 error) → imprese fallite classificate erroneamente come sane;
2) T2 (Type 2 error) → imprese sane classificate erroneamente come fallite;
3) Recall → capacità di rilevare le crisi reali;
4) Precision → accuratezza tra le predette in crisi;
5) F1 e F2 Score → medie armoniche tra precisione e recall.

I risultati evidenziano che XGBoost è l'algoritmo con le migliori performance e che le variabili industriali/regionali svolgono un ruolo rilevante. In particolare, l'appartenenza a un distretto industriale, un elevato mark-up e una maggiore quota di mercato risultano associati a una minore probabilità di fallimento.

	financial ratios					
	T1	T2	F1	F2	Recall	Precision
Logistic Regression	15.26	19.12	37.39	56.24	84.74	23.99
Random Forest	16.22	10.36	50.91	66.57	83.78	36.57
XGBoost	12.31	17.42	40.57	59.87	87.69	26.39
Neural Network	17.02	11.02	49.53	65.15	82.98	35.54
	financial ratios + industrial variables					
	T1	T2	F1	F2	Recall	Precision
Logistic Regression	**15.06**	19.71	36.79	55.75	**84.94**	23.48
Random Forest	**15.55**	11.26	49.32	65.71	**84.45**	34.84
XGBoost	**10.72**	19.22	38.89	58.80	**89.28**	24.86
Neural Network	17.26	**10.89**	49.72	65.19	82.74	**35.79**

Tabella 7

Fonte: https://dipartimenti.unicatt.it/dimsefa-dime19_03.pdf

La tabella 7 mostra i risultati ottenuti da quattro modelli predittivi – Logistic Regression, Random Forest, XGBoost e Neural Network – valutati secondo diverse metriche, sia utilizzando solo indicatori finanziari (prima sezione), sia combinando indicatori finanziari e variabili industriali (seconda sezione).

- Sezione 1

XG Boost mostra la migliore capacità di recall (87.69%) tra tutti i modelli, ossia è il più efficace nel riconoscere le imprese realmente fallite; ha inoltre l'errore di tipo 1 (T1) più basso: solo il 12,31% delle imprese fallite vengono erroneamente classificate come sane. Tuttavia la precisione (26,39%) è piuttosto bassa: cioè tra le imprese previste come fallite, solo una piccola parte lo è davvero.

- Sezione 2

I risultati migliorano con l'aggiunta delle variabili industriali, e l'XG Boost si riconferma il modello con i migliori risultati, superando anche gli ottimi risultati nella sezione 1: l'errore di Tipo 1 scende dal 12,31% al 10,72%, aumentando ulteriormente il recall fino a 89,28%, confermandosi il più alto tra i modelli.

Random Forest e Neural Network migliorano, ma non superano l'XGBoost; la Neural Network ha la precisione più alta (35,79%) e ottimi valori F1 ed F2, ma un recall più basso (82,74%) rispetto a XGBoost.

I risultati confermano che l'algoritmo è altamente efficace per il rilevamento precoce della crisi d'impresa, e che le informazioni settoriali e regionali hanno un peso significativo che può agevolare la diagnosi.

Un ulteriore contributo rilevante nell'ambito della previsione della crisi d'impresa è rappresentato dallo studio di Aliaj, Anagnostopoulos e Piersanti[62], che applica tecniche di machine learning, in particolare modelli ensemble come Random Forest e Gradient Boosting, per anticipare il rischio di default delle imprese italiane. Gli autori impiegano per la prima volta un ampio set (300.000 imprese) di *dati creditizi granulari* provenienti dal Registro Centrale dei Rischi della Banca d'Italia, arricchendoli con dati pubblici di bilancio, dati inerenti ai ritardi nei pagamenti, classificazioni bancarie (L1-L12). L'obiettivo dello studio non è tanto quello di prevedere il fallimento in senso tecnico, quanto piuttosto di identificare tempestivamente lo stato di default bancario, considerato un importante indicatore precoce di deterioramento finanziario, focalizzandosi sulla condizione di default "aggiustato" come definito dalla stessa Banca d'Italia, ovvero l'incapacità di onorare gli obblighi finanziari verso l'intero sistema creditizio.

[62] Aliaj, T., Anagnostopoulos, A., Piersanti, S. (2020) *Firms Default Predictions with Machine Learning,* Lecture Notes in Computer Science, Springer.

	Pr	Re	F1	Type-I	Type-II	BACC
NAIVE	0.25	0.77	0.38	0.23	0.49	0.64
MNB	0.44	0.09	0.15	0.91	0.03	0.53
LOG	0.36	0.22	0.28	0.78	0.03	0.60
GB	0.19	0.78	0.30	0.22	0.15	0.81
RF	0.18	0.80	0.30	0.20	0.16	0.82
DT	0.10	0.71	0.18	0.29	0.26	0.72
BAG	0.17	0.75	0.27	0.25	0.17	0.79
ADA	0.18	0.76	0.29	0.24	0.16	0.80
COMB	0.19	0.84	0.31	0.16	0.16	0.84

Tabella 8

Fonte: https://arvix.org/pdf/2002.11705

La tabella 8 indica come a raggiungere i risultati migliori siano ancora una volta i modelli di Grandient Boosting, Random Forest e COMB, registrando i valori più alti in termini di precisione (Pr), recall (Re) e F1 score e i valori più bassi in termine di errore, sia di primo che di secondo tipo.

I risultati ottenuti mostrano che, soprattutto in contesti sbilanciati e complessi, gli algoritmi di boosting si mostrano particolarmente affidabili in termini di accuratezza e capacità di generalizzazione, fornendo un'evidenza empirica del valore strategico delle tecniche di apprendimento automatico per la costruzione degli *early warning system*, applicabili in ambito bancario, gestionale e di vigilanza.

Sempre in ottica di un approccio proattivo e alla necessità, oltre che strettamente operativa, voluta dal legislatore di anticipare i segnali di crisi e favorire un intervento prima che si giunga ad una fase irreversibile, è di grande uno studio sulle c.d. *zombie firms*.[63] Gli autori comparano i risultati di vari modelli, tra cui lo Z-Score, il Distance-to-Default e il più moderno XGBoost, oltre a vari modelli standard, servendosi di un dataset di 300.000 imprese italiane osservate nel periodo 2008-2017, focalizzandosi su bilanci aziendali e i *missing data* (indicatori di incompletezza dei dati), al fine di identificare le c.d. "imprese zombie": aziende formalmente attive ma in stato di crisi latente o cronica, (per almeno 3 anni consecutivi).

[63] Bargagli-Stoffi, F., Gallo, L., Rungi, A., & Rungi, M. (2023). *Machine learning for zombie hunting: Predicting distress from firms' accounts and missing values*. arXiv

Indicator (in logs)	Coeff.	Std. Error	N. obs.	Adj. R squared
Total Factor Productivity	-.213***	(.049)	568,288	.257
Labor Productivity	-.444***	(.038)	577,485	.139
Sales	-1.198***	(.055)	595,609	.189
Employees	-.760***	(.034)	578,532	.112

Tabella 9

Fonte: DBI

Dallo studio emerge che le zombie rispetto alle altre imprese:

- Hanno il 21,3% di produttività in meno se misurata come Total Factory Productivity;
- Hanno una produttività del lavoro inferiore del 44%;
- Vendono in media il 120% in meno;
- Impiegano il 76% in meno di personale.

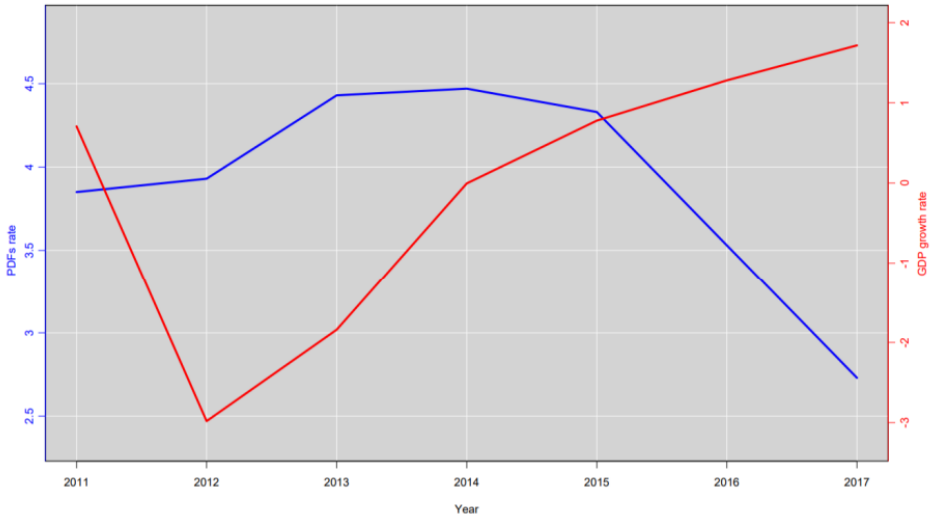

Figura 6

Fonte: Banca d'Italia

La figura 6 mostra l'andamento temporale della quota di imprese zombie (linea blu) e del tasso di crescita del PIL (linea rossa). Come prevedibile, le imprese zombie proliferano nei periodi di crisi o stagnazione e diminuiscono durante la ripresa o per l'uscita di queste dal mercato (*zombie cleansing*) o perché vengono meno i presupposti temporali e quantitativi per classificare un'impresa come tale.[64]

Prima di passare ad un confronto diretto con i nuovi modelli statistici, gli autori effettuano un confronto tra Distance-to-Default (DtD) e Z-Score in termini di precisione e FDR (False Discovery Rate) pari a 1-Precision.

Il DtD è un indicatore quantitativo utilizzato per stimare la probabilità di insolvenza di un'impresa. È basato sul modello strutturale di Merton e si fonda sull'idea che un'impresa può essere considerata insolvente quando il valore del suo attivo scende al di sotto del valore del debito, calcolato come:

$$\mathrm{DtD} = \frac{\ln\left(\frac{V_A}{D}\right) + \left(\mu - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}$$

Dove:

- V_A → valore di mercato dell'attivo dell'impresa;
- D → valore del debito, solitamente a un anno;
- μ → tasso di crescita atteso dell'attivo;
- σ^2 → varianza dell'attivo;
- T → orizzonte temporale.

Maggiore è il valore ottenuto, più lontana è l'impresa dal rischio di default, e viceversa.

[64] Andrews, D., Petroulakis, F., (2019). *Breaking the shackles: Zombie firms, weak banks and depressed restructuring in Europe*. Working Paper Series 2240, European Central Bank.

Decile	Precision DtD	FDR DtD	Precision Z-score	FDR Z-score
1st	0.2680	0.7320	0.1613	0.8387
2nd	0.2680	0.7320	0.1505	0.8495
3rd	0.2680	0.7320	0.1505	0.8495
4th	0.2258	0.7742	0.1371	0.8629
5th	0.2022	0.7978	0.1269	0.8731
6th	0.1759	0.8241	0.1185	0.8815
7th	0.1569	0.8431	0.1108	0.8892
8th	0.1467	0.8533	0.1036	0.8964
9th	0.1411	0.8589	0.0981	0.9019
10th	0.1313	0.8687	0.0969	0.9031

Tabella 10

Fonte: https://arxiv.org/pdf/2306.08165

In termini di precisione, nei primi tre decili il DtD ha una precisione costante di circa il 27%, mentre lo Z-Score si mostra molto meno preciso, con valori decrescenti dal 16% nel primo decile fino ad un 9% nel decimo. Coerentemente, il DtD ha un FDR minore pari a circa il 73% mentre quello dello Z-Score oscilla tra l'84% e il 90%, indicando una più elevata probabilità di etichettare imprese sane come a rischio.

Uno degli elementi più innovativi dello studio è stato proprio l'inserimento dei valori mancanti come predittori, sfruttando il fatto che la mancanza di dati in alcune voci di bilancio può essere essa stessa un segnale di fragilità aziendale.

L'XGBoost, come anticipato, a differenza sia del DtD che dello Z-Score è l'unico in grado di identificare e gestire tale segnale attribuendo così ai missing values un valore predittivo considerevole: ciò trova conferma nei valori dell'AUC, che sintetizza la capacità discriminante complessiva del modello. L'algoritmo XGBoost raggiunge un'AUC pari a 0,97, evidenziando un'elevata accuratezza nella classificazione delle imprese a rischio, superiore ai comunque ottimi risultati ottenuti dal DtD e dallo Z-Score, rispettivamente pari a 0,85 e 0,75.

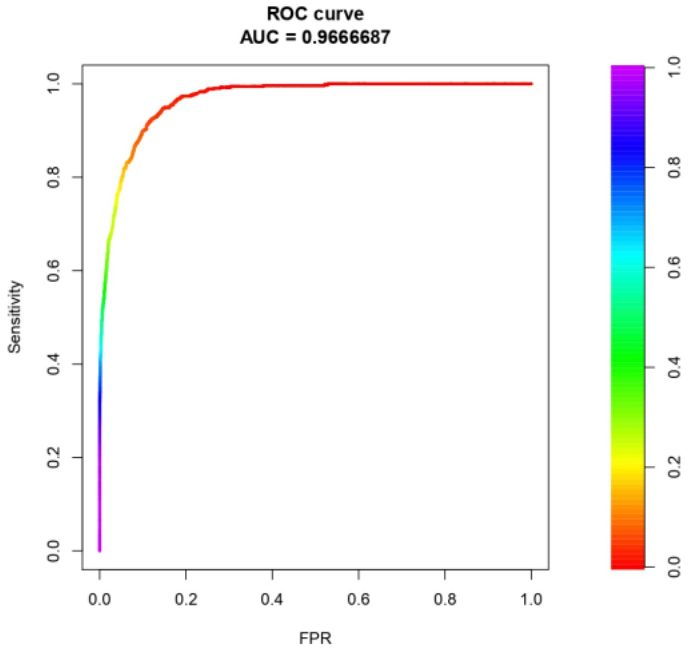

Figura 7

Fonte: https://arvix.org/pdf/2306.08165

Questo perché il Gradient Boosting è meno sensibile alla multicollinearità, richiede meno tuning parametrico rispetto modelli, al contempo, gode di un'elevata adattabilità ai diversi contesti settoriali e dimensionali, confermando la sua capacità di gestire dati mancanti e pattern complessi.

Nel complesso, l'impiego di algoritmi di boosting – in particolare XGBoost – consente di costruire modelli predittivi affidabili, scalabili e interpretabili, idonei all'integrazione all'interno di sistemi di controllo di gestione, auditing predittivo e scoring bancario. Alla luce di questi risultati, si può affermare che tali tecniche rappresentino oggi una delle soluzioni più avanzate e promettenti per il monitoraggio del rischio di crisi d'impresa, specialmente in contesti normativi che richiedono strumenti di allerta precoce, come previsto dal Codice della Crisi d'Impresa e dell'Insolvenza.

2.4.4 – Artificial Neural Network

Una Artificial Neural Network (ANN), ovvero rete neurale artificiale, è un modello computazionale ispirato al funzionamento del cervello umano. Viene utilizzato nel campo del machine learning per riconoscere schemi complessi e relazioni nei dati, e trova ampio impiego in ambiti come la previsione della crisi d'impresa, l'analisi del rischio di credito, la diagnosi medica, il riconoscimento vocale, e molto altro.

Una rete neurale è formata da nodi (neuroni artificiali) organizzati in strati (layers):

1. *Strato di input (input layer)*: riceve i dati iniziali (es. indicatori finanziari di un'azienda).
2. *Strati nascosti (hidden layers)*: elaborano le informazioni tramite connessioni pesate, trasformandole in rappresentazioni più astratte.
3. *Strato di output (output layer*): fornisce il risultato del modello, come ad esempio la probabilità che un'impresa entri in crisi.

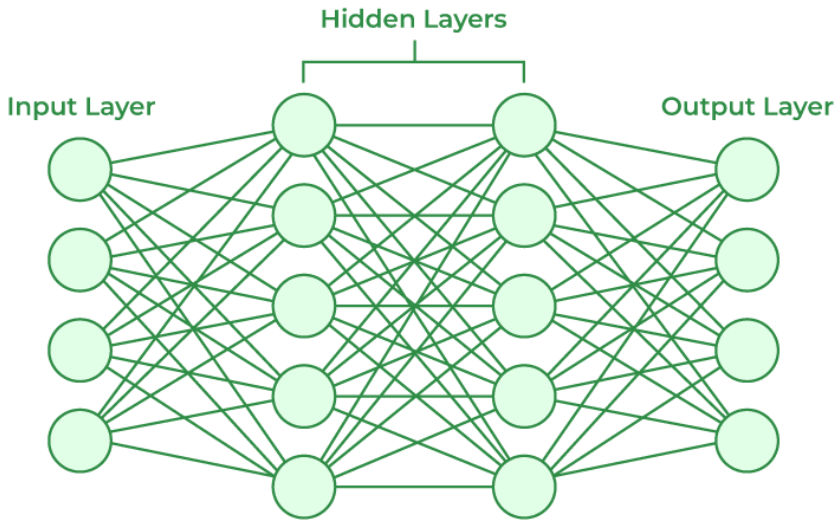

Figura 8

Fonte: https://www.geeksforgeeks.org/artificial-neural-networks-and-its-applications/

Ogni nodo riceve input da altri nodi, li pesa con dei coefficienti (weights), li somma, e applica una funzione di attivazione, come la funzione sigmoide o ReLU, per generare un output:

$$y = \sigma\left(\sum_{i=1}^{n} w_i x_i + b\right)$$

Dove:

- X_i → input;
- W_i → peso;
- b → bias (termine correttivo)
- σ → funzione di attivazione

Durante il **training**, l'ANN aggiusta questi pesi attraverso un processo chiamato **backpropagation**, minimizzando una **funzione di errore (loss function)**.

Le reti neurali artificiali (ANN) condividono i principali punti di forza dei modelli di machine learning analizzati in precedenza, come la capacità di cogliere relazioni non lineari complesse e di apprendere automaticamente dai dati, senza la necessità di definire regole esplicite. A questi vantaggi si aggiunge una notevole versatilità: le ANN si adattano efficacemente a diverse tipologie di input, non solo numerici, ma anche testuali, visivi o categoriali, ampliandone le potenzialità applicative in molteplici contesti.

L'impiego delle reti neurali artificiali presenta anche alcune limitazioni che ne riducono l'immediata applicabilità in contesti operativi. In primo luogo, tali modelli sono spesso definiti come "black box", poiché non consentono un'interpretazione trasparente delle logiche decisionali interne: le relazioni tra input e output risultano opache, rendendo difficile la comprensione delle motivazioni alla base delle previsioni. Questo rappresenta un limite significativo, soprattutto in ambiti – come quello economico-finanziario – in cui la tracciabilità e la giustificazione delle scelte analitiche è un requisito essenziale.

In secondo luogo, le reti neurali richiedono una grande quantità di dati e risorse computazionali per poter essere addestrate in modo efficace. La loro implementazione risulta pertanto meno agevole in contesti con dati limitati o infrastrutture informatiche poco avanzate.

Infine, le ANN sono soggette a un elevato rischio di overfitting, ovvero la tendenza ad adattarsi troppo ai dati di training, perdendo capacità di generalizzazione sui dati futuri.

Per contrastare questo fenomeno è necessario applicare opportuni metodi di regolarizzazione e tecniche di validazione incrociata, che richiedono competenze specialistiche.[65]

Uno degli studi più rappresentativi sull'applicazione delle reti neurali artificiali (ANN) alla previsione della crisi d'impresa è quello condotto da Zhang, Hu, Patuwo e Indro[66]. Questo lavoro propone un confronto sistematico tra le ANN e i modelli statistici tradizionali, in particolare la regressione logistica, con l'obiettivo di valutarne l'efficacia nella classificazione delle imprese fallite rispetto a quelle sane.

Gli autori utilizzano un campione bilanciato composto da 220 aziende (110 fallite e 110 non fallite) e impiegano una tecnica di cross-validation a cinque ripetizioni per garantire una valutazione robusta e affidabile delle performance predittive. Questo approccio consente di ridurre il rischio di overfitting e di testare la capacità del modello di generalizzare su dati non osservati.

I risultati mostrano chiaramente la superiorità delle reti neurali rispetto ai modelli statistici convenzionali. In particolare, la percentuale di classificazione corretta complessiva ottenuta dalle ANN è stata del 86,64%, mentre quella della regressione logistica si è attestata al 78,55%, evidenziando un miglioramento di quasi 10 punti percentuali. Analizzando le due categorie separatamente, le ANN hanno raggiunto un'accuratezza dell'85,5% nella classificazione delle aziende fallite e del 90,9% per quelle non fallite, superando rispettivamente le performance della regressione logistica che si sono fermate al 72,3% e all'84,5%. Inoltre, gli autori sottolineano che gli output generati dal modello neurale possono essere interpretati in chiave probabilistica, come stime delle probabilità a posteriori bayesiane, conferendo così al modello una solida base teorica e operativa per supportare le decisioni manageriali e finanziarie.

In sintesi, lo studio evidenzia come le ANN rappresentino una tecnologia promettente per la previsione del rischio d'insolvenza, fornendo un'alternativa più performante ai tradizionali strumenti di scoring, a patto che il loro utilizzo sia accompagnato da una corretta implementazione e da un attento processo di validazione.

---

[65] Kendirli, S., & Citak, L. (2022). *Estimating Financial Failure in Businesses Using Artificial Neural Networks.* Journal of Corporate Governance and Information Risk Management, 9(2), 1–15.

[66] Zhang, G., Hu, M. Y., Patuwo, B. E., & Indro, D. C. (1999). *Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis*. European Journal of Operational Research, 116(1), 16–32.

	Overall	
Statistics	ANN	Logistic
Mean	86.64	78.55
t-statistic	10.3807	
p-value	0.0005	

Tabella 11

Fonte: https://www.bobm.net.au/teaching/SimSS/zhang.pdf

Alla luce dell'analisi svolta, è possibile affermare che il ricorso alle tecniche di machine learning rappresenta un'evoluzione metodologica rilevante rispetto ai modelli tradizionali utilizzati per la previsione della crisi d'impresa. Algoritmi come il Random Forest, il Support Vector Machine, il Gradient Boosting e le Reti Neurali Artificiali hanno mostrato, nei numerosi studi esaminati, performance predittive superiori rispetto a modelli lineari come lo Z-Score di Altman o la regressione logistica, soprattutto in termini di accuratezza, capacità di intercettare segnali deboli e riduzione degli errori di classificazione. Tuttavia, l'adozione di queste tecniche, pur promettente, non è priva di criticità. In particolare, alcuni algoritmi – come le reti neurali – richiedono una grande quantità di dati, elevate capacità computazionali e competenze specialistiche per la loro corretta implementazione e manutenzione. Inoltre, il carattere "black box" di molti modelli avanzati rende complessa l'interpretazione dei risultati, un aspetto rilevante in contesti in cui la trasparenza del processo decisionale è imprescindibile (ad esempio nella revisione contabile o nella compliance normativa).

A fronte di ciò, emerge una necessità di equilibrio tra potenza predittiva e fruibilità operativa: modelli come il Random Forest o il Gradient Boosting, grazie alla loro combinazione di accuratezza, interpretabilità parziale e minore esigenza di tuning rispetto ad altri approcci, sembrano offrire il miglior compromesso per l'integrazione nei sistemi di early warning aziendale. Inoltre, l'evidenza empirica mostra che l'integrazione di informazioni settoriali, geografiche o comportamentali migliora ulteriormente la capacità discriminante dei modelli, suggerendo l'importanza di una progettazione ad hoc dei predittori in funzione del contesto.

In conclusione, il machine learning non deve essere inteso come un sostituto totale dei modelli tradizionali, bensì come uno strumento complementare, da integrare nei sistemi decisionali in modo graduale e consapevole. La scelta dell'algoritmo più adatto dovrà

tener conto del bilanciamento tra rigore metodologico, accessibilità tecnica e coerenza con gli obiettivi informativi, affinché l'innovazione tecnologica si traduca in un effettivo miglioramento della capacità delle imprese di anticipare e fronteggiare tempestivamente i segnali di crisi.

3 – ANALISI EMPIRICA

3.1 – Premessa

Dopo aver introdotto, nel primo capitolo, le fondamenta teoriche del concetto di crisi d'impresa, analizzando la nozione di default in chiave economico-aziendale e normativa, e aver delineato le principali manifestazioni sintomatiche della crisi nelle sue diverse fasi evolutive, si è proseguito, nel secondo capitolo, con un'approfondita disamina comparativa tra i modelli tradizionali di previsione del fallimento e le più recenti tecniche di Machine Learning, esaminate sia sotto il profilo teorico-metodologico sia attraverso la letteratura scientifica di riferimento. Tale seconda sezione si è basata su numerosi studi accademici e contributi di ricerca che, in maniera pressoché unanime, evidenziano i limiti strutturali dei modelli statici tradizionali e sottolineano, al contempo, le superiori potenzialità dei modelli predittivi di tipo algoritmico, capaci di apprendere dai dati e di adattarsi a contesti complessi e dinamici, come quelli caratterizzati da forte squilibrio informativo e da classi fortemente sbilanciate, tipici dell'ambito del fallimento aziendale.
Alla luce delle premesse teoriche e degli elementi emersi dalla rassegna critica della letteratura, questo terzo e ultimo capitolo si propone di fornire un contributo originale e sperimentale, attraverso la realizzazione di un'analisi empirica basata su codice Python eseguito all'interno dell'ambiente Google Colab. L'obiettivo principale è quello di testare operativamente i modelli di previsione trattati in precedenza e valutarne la performance su un dataset reale, costituito da informazioni finanziarie di imprese polacche, con lo scopo di verificare se le evidenze teoriche e i risultati empirici discussi dagli autori citati nel secondo capitolo trovino effettivo riscontro anche in un'applicazione concreta e replicabile.
In particolare, verrà implementato un percorso analitico rigoroso che parte dalla preparazione e pre-processing del dataset, prosegue con la costruzione e la calibrazione dei modelli predittivi (sia tradizionali sia di machine learning), e si conclude con un confronto diretto e sistematico tra le diverse metodologie, valutando ciascuna in base a specifiche metriche di performance (accuratezza, precisione, recall, F1-Score, AUC-ROC, false positive/negative rate, ecc.). I modelli saranno implementati attraverso strumenti open source di ampia diffusione mentre l'intero workflow sarà eseguito e

documentato nel contesto di Google Colab, ambiente che consente la tracciabilità, la riproducibilità e la trasparenza dell'intero processo analitico.
Attraverso questa indagine si rende possibile trarre conclusioni autonome rispetto alla capacità delle tecniche di machine learning di anticipare con maggiore efficacia le situazioni di insolvenza aziendale, e quindi confermare (o eventualmente discutere criticamente) le tesi emerse nella letteratura di settore. In definitiva, questo capitolo rappresenta l'anello di congiunzione tra teoria e pratica, tra riflessione accademica e sperimentazione operativa, dando concretezza e profondità al percorso di ricerca intrapreso nei capitoli precedenti.

3.2 – Scelta, descrizione e pulizia dei dataset

Il dataset utilizzato per l'analisi è intitolato **"Polish Companies Bankruptcy"** ed è una raccolta di dati finanziari utilizzata per la previsione del fallimento aziendale. I dati sono stati raccolti da EMIS (Emerging Markets Information Service) e comprendono informazioni su aziende polacche, sia fallite che operative, nel periodo compreso tra il 2000 e il 2013. Questo dataset è stato reso disponibile attraverso il UCI Machine Learning Repository[67] e Kaggle[68], ed è ampiamente utilizzato per studi di classificazione binaria (e non) nel campo del rischio.
Il dataset è suddiviso in cinque file distinti, ciascuno denominato in base all'orizzonte temporale di previsione del fallimento:

- 1year.arff: prevede il fallimento a 5 anni;
- 2year.arff: prevede il fallimento a 4 anni;
- 3year.arff: prevede il fallimento a 3 anni;
- 4year.arff: prevede il fallimento a 2 anni;
- 5year.arff: prevede il fallimento a 1 anno.

Ciascun file contiene 64 variabili finanziarie (denominate X1, X2, … , X64) e una variabile target binaria che indica lo stato dell'azienda:

[67] https://archive.ics.uci.edu/dataset/365/polish+companies+bankruptcy+data
[68] https://www.kaggle.com/c/companies-bankruptcy-forecast/data

- 0 se l'azienda è operativa (non fallita);
- 1 se l'azienda è fallita.

Per l'analisi empirica si è scelto di concentrare l'attenzione sui dataset **4year.arff** e **5year.arff**, corrispondenti rispettivamente alla previsione del fallimento aziendale **a due anni** e **a un anno** di distanza. Tale decisione metodologica non è casuale, bensì motivata da una coerenza logica e temporale con l'orizzonte operativo del modello di Altman, che storicamente si propone come strumento di previsione del rischio di default nel breve termine.

Come anticipato, il modello Z-Score, infatti, è concepito per individuare tempestivamente segnali di deterioramento della salute finanziaria di un'impresa, con l'obiettivo di anticipare il rischio di insolvenza entro un arco temporale di massimo due anni. Nella letteratura economico-aziendale e finanziaria, è ampiamente riconosciuto che l'efficacia del modello tende a decrescere al crescere della distanza temporale dall'evento di default, risultando quindi meno performante su orizzonti di medio-lungo periodo. Alla luce di ciò, l'utilizzo dei dataset a 1 e 2 anni consente non solo di mantenere un allineamento con la logica predittiva del modello tradizionale, ma anche di effettuare un confronto più equo e metodologicamente fondato. Questo approccio garantisce una base comune di riferimento per la valutazione delle performance, evitando il rischio di penalizzare artificiosamente il modello Altman confrontandolo su orizzonti temporali per i quali non è stato originariamente progettato. Inoltre, la previsione a breve termine riveste un ruolo centrale nella prassi gestionale e nella diagnosi precoce della crisi, poiché consente di attuare tempestivamente interventi correttivi da parte dei soggetti economici coinvolti.

Per tali ragioni, la scelta dei file **4year** e **5year** rappresenta una soluzione metodologica coerente con le finalità dell'analisi e consente di garantire omogeneità e validità comparativa nel confronto tra approcci predittivi di natura diversa.

Il primo passaggio è stato quello di installare la libreria *liac-arff,* necessaria per leggere i *file .arff,* un formato utilizzato comunemente per dataset in ambito di machine learning, come quelli usati in questa analisi. In questa cella vengono importate tutte le librerie necessarie:

- *pandas* per la gestione dei dati,
- *scipy.io.arff* per caricare i dataset,
- strumenti di *scikit-learn* per valutare e costruire i modelli,

- *matplotlib e seaborn* per i grafici,
- infine, vengono importati i due modelli di machine learning che verranno usati: Random Forest e XGBoost.

```
import pandas as pd
from scipy.io import arff

from sklearn.model_selection import train_test_split,  GridSearchCV
from sklearn.metrics import classification_report, confusion_matrix, precision_score, recall_score, f1_score, roc_auc_score, roc_curve


import matplotlib.pyplot as plt
import seaborn as sns

from sklearn.ensemble import RandomForestClassifier
import xgboost as xgb
```

Le due seguenti celle con le relative tabelle di output indicano i due dataset utilizzati per l'analisi: il primo, *5year.arff,* con un numero di osservazioni pari a 5910 aziende, di cui 5500 operative e 410 fallite.

```
data, meta = arff.loadarff('5year.arff')
df_5year = pd.DataFrame(data)
df_5year
```

	Attr1	Attr2	Attr3	Attr4	Attr5	Attr6	Attr7	Attr8	Attr9	Attr10
0	0.088238	0.55472	0.011340	1.02050	-66.5200	0.342040	0.109490	0.57752	1.08810	0.320360
1	-0.006202	0.48465	0.232980	1.59980	6.1825	0.000000	-0.006202	1.06340	1.27570	0.515350
2	0.130240	0.22142	0.577510	3.60820	120.0400	0.187640	0.162120	3.05900	1.14150	0.677310
3	-0.089951	0.88700	0.269270	1.52220	-55.9920	-0.073957	-0.089951	0.12740	1.27540	0.113000
4	0.048179	0.55041	0.107650	1.24370	-22.9590	0.000000	0.059280	0.81682	1.51500	0.449590
...	...	...	...	...	...	...	...	...	...	...
5905	0.012898	0.70621	0.038857	1.17220	-18.9070	0.000000	0.013981	0.41600	1.67680	0.293790
5906	-0.578050	0.96702	-0.800850	0.16576	-67.3650	-0.578050	-0.578050	-0.40334	0.93979	-0.390040
5907	-0.179050	1.25530	-0.275990	0.74554	-120.4400	-0.179050	-0.154930	-0.26018	1.17490	-0.326590
5908	-0.108860	0.74394	0.015449	1.08780	-17.0030	-0.108860	-0.109180	0.12531	0.84516	0.093224
5909	-0.105370	0.53629	-0.045578	0.91478	-56.0680	-0.105370	-0.109940	0.86460	0.95040	0.463670

5910 rows × 65 columns

Tabella 12

Il secondo, *4year.arff,* contenente dati di 9792 aziende, di cui 9277 operative e 515 fallite.

```
data, meta = arff.loadarff('4year.arff')
df_4year = pd.DataFrame(data)
df_4year
```

	Attr1	Attr2	Attr3	Attr4	Attr5	Attr6	Attr7	Attr8	Attr9	Attr10	...
0	0.159290	0.46240	0.077730	1.16830	-44.8530	0.467020	0.189480	0.82895	1.12230	0.38330	...
1	-0.127430	0.46243	0.269170	1.75170	7.5970	0.000925	-0.127430	1.16250	1.29440	0.53757	...
2	0.070488	0.23570	0.527810	3.23930	125.6800	0.163670	0.086895	2.87180	1.05740	0.67689	...
3	0.136760	0.40538	0.315430	1.87050	19.1150	0.504970	0.136760	1.45390	1.11440	0.58938	...
4	-0.110080	0.69793	0.188780	1.27130	-15.3440	0.000000	-0.110080	0.43282	1.73500	0.30207	...
...	...	...	...	...	...	...	...	...	...	...	...
9787	0.004676	0.54949	0.192810	1.38990	-39.0640	0.004676	0.013002	0.78627	0.97093	0.43205	...
9788	-0.027610	0.60748	-0.029762	0.90591	-20.9230	-0.027610	-0.027610	0.55161	1.00730	0.33509	...
9789	-0.238290	0.62708	0.090374	1.61250	-1.0692	-0.238290	-0.240360	0.28322	0.80307	0.17760	...
9790	0.097188	0.75300	-0.327680	0.43850	-214.2400	-0.331300	0.104280	0.32803	0.98145	0.24700	...
9791	0.021416	0.48678	0.148940	1.30670	-24.2820	0.021416	0.027253	1.05320	1.00140	0.51267	...

9792 rows × 65 columns

Tabella 13

Successivamente viene creato un dizionario con la descrizione di tutte le variabili presenti nel dataset, convertito in una tabella (*DataFrame*) in cui ogni riga rappresenta una variabile e la sua descrizione economico-finanziaria, al fine di rendere più leggibile il significato delle colonne nel dataset.

```
# Create a DataFrame
df_variables  = pd.DataFrame(list(features.items()), columns=["Attr", "Description"])

df_variables["Attr"] = df_variables["Attr"].str.replace("X", "Attr")
```

Si passa così alla *pulizia dei dati:* la riga che segue converte la colonna target class (che indica se un'azienda è fallita o no) da formato byte a numero intero; in più calcola la distribuzione percentuale tra aziende sane e fallite nel dataset a cinque anni.[69]

```
# Convert byte strings to regular strings, then to integers
df_5year['class'] = df_5year['class'].apply(lambda x: int(x.decode('utf-8')) if isinstance(x, bytes) else int(x))
df_5year['class'].value_counts() / len(df_5year)
```

[69] Il codice viene ripetuto anche per il dataset *4year*.

Si osserva, coerentemente con il reale contesto economico, un forte sbilanciamento dovuto ad una netta minoranza di imprese fallite rispetto a quelle non fallite.

- Nel dataset *5year,* poco meno del 7% delle aziende sono fallite:

class	count
0	0.930626
1	0.069374

- Nel dataset *4year,* poco più del 5%:

class	count
0	0.947476
1	0.052524

La sua natura rigida e priva di capacità di apprendimento adattivo rende lo Z-Score poco sensibile alle variazioni di distribuzione dei dati, e di conseguenza molto vulnerabile alla predominanza della classe maggioritaria. Questo si traduce in una tendenza sistematica – confermata più avanti – a classificare la maggior parte delle osservazioni come aziende non fallite, con conseguente perdita di accuratezza proprio nei casi più rilevanti da individuare: i fallimenti.

Al contrario, i modelli di machine learning offrono una flessibilità strutturale e algoritmica che consente loro di gestire efficacemente lo sbilanciamento. Questi modelli permettono infatti di introdurre strategie di pesatura delle classi, come ad esempio l'iperparametro ***class_weight='balanced'*** in Random Forest o ***scale_pos_weight*** in XGBoost, che riequilibrano l'influenza delle classi durante l'addestramento. In questo modo, il modello viene "incentivato" a prestare maggiore attenzione ai casi di fallimento, migliorando sensibilmente la capacità predittiva nei confronti della classe minoritaria, senza sacrificare la precisione sulla classe dominante.

Nel seguente blocco di codice[70] viene effettuato il *train-split test*, ovvero la suddivisione del dataset in due sottoinsiemi:

- uno destinato all'addestramento del modello (*x_train; y_train*), pari al 70% delle osservazioni;
- uno per la valutazione finale delle prestazioni (*x_test; y_test*), pari al 30% delle osservazioni.

```
[9]  X = df_5year.drop(columns=['class'])
     y = df_5year['class']

     X_train, X_test, y_train, y_test = train_test_split(
         X, y, test_size=0.3, stratify=y, random_state=42
     )
```

Nel processo di suddivisione del dataset in training set e test set, è **fondamentale mantenere la stessa proporzione tra le classi** (fallite e non fallite) in entrambi i sottoinsiemi. Questa procedura, nota come stratificazione, garantisce che la distribuzione delle classi nel set di addestramento sia rappresentativa di quella presente nell'intero dataset, e che il modello venga valutato su dati realistici.

Se questa proporzione non viene rispettata, si possono generare diversi problemi:

1. *Addestramento non rappresentativo*: se il training set contiene una distribuzione sbilanciata diversa da quella reale (ad esempio, troppe poche aziende fallite), il modello imparerà uno schema distorto, sottostimando la classe minoritaria e imparando a riconoscere soprattutto la classe più rappresentata. Questo porta a modelli poco sensibili ai casi critici (nel nostro caso, le aziende a rischio fallimento).
2. *Valutazione non affidabile*: se il test set ha una distribuzione diversa da quella del training set, le metriche di valutazione (accuratezza, precisione, recall, ecc.) possono risultare **fuorvianti**. Il modello potrebbe sembrare più o meno performante solo perché sta operando su un campione non coerente con quello su cui è stato addestrato.
3. *Confrontabilità ridotta*: la mancanza di coerenza tra train e test set rende più difficile confrontare modelli diversi o interpretare correttamente le differenze nei

[70] Anche questo ripetuto per il dataset *4year*.

risultati, poiché ogni modello potrebbe essere valutato in condizioni non comparabili.

Per questi motivi, nell'analisi viene utilizzato il parametro *stratify=y* durante la divisione dei dati, che impone al train-test split di mantenere la stessa proporzione di aziende sane e fallite in entrambi i set. Ciò consente di costruire un modello più equilibrato e di valutarne le prestazioni in modo realistico e attendibile, garantendo maggiore validità e generalizzabilità ai risultati ottenuti.

3.3 – Modello di Altman

Una volta effettuati data-frame, pulizia del dataset e train-split test, vengono selezionate solo le cinque variabili usate per il calcolo dello Z-Score, creando un sottoinsieme di dati contente solo i seguenti attributi, sia per il training set che per il test set.

```
X_train_Altman = X_train[['Attr3', 'Attr6', 'Attr7', 'Attr8', 'Attr9']]
X_test_Altman = X_test[['Attr3', 'Attr6', 'Attr7', 'Attr8', 'Attr9']]
df_variables[df_variables['Attr'].isin(list(X_test_Altman.columns))].reset_index(drop=True)
```

	Attr	Description
0	Attr3	working capital / total assets
1	Attr6	retained earnings / total assets
2	Attr7	EBIT / total assets
3	Attr8	book value of equity / total liabilities
4	Attr9	sales / total assets

Tabella 14

In modo da poter moltiplicare il singolo attributo per il relativo peso attribuitogli dalla formula e sommarlo agli altri:

```
z_score= (1.2*X_test_Altman['Attr3']
          + 1.4*X_test_Altman['Attr6']
          + 3.3*X_test_Altman['Attr7']
          + 0.6 *X_test_Altman['Attr8']
          + 1.0*X_test_Altman['Attr9'] )
```

Fissando una soglia di 1.81 (termine della *distress zone):* in questo modo il modello classifica le aziende con uno Z-Score inferiore come a rischio di fallimento (1), le altre come sane (0).

```
pred_Altman= [1 if score < 1.81 else 0 for score in z_score]
```

In merito agli output sono stati utilizzati i seguenti codici:

1) Per il calcolo delle metriche fondamentali, quali False Positive Rate, False Negative Rate, Preision, Recall ed F1-Score:

```
# Compute confusion matrix
tn, fp, fn, tp = confusion_matrix(y_test, pred_Altman).ravel()

# Compute metrics
fp_rate_altman = fp / (fp + tn) if (fp + tn) > 0 else 0
fn_rate_altman = fn / (fn + tp) if (fn + tp) > 0 else 0
precision_altman = precision_score(y_test, pred_Altman)
recall_altman = recall_score(y_test, pred_Altman)
f1_altman= f1_score(y_test, pred_Altman)

print(f"False Positive Rate: {fp_rate_altman:.3f}")
print(f"False Negative Rate: {fn_rate_altman:.3f}")
print(f"Precision: {precision_altman:.3f}")
print(f"Recall: {recall_altman:.3f}")
print(f"F1 Score: {f1_altman:.3f}")
```

2) Per la creazione della matrice di confusione;

```
classification_report_altman= classification_report(y_test, pred_Altman, output_dict= True)
print(classification_report(y_test, pred_Altman))
cm = confusion_matrix(y_test, pred_Altman)
fig, ax = plt.subplots(figsize=(8, 6))
sns.heatmap(cm, annot=True, ax=ax, cmap='Blues', fmt='g')
ax.set_title('Confusion Matrix')
ax.set_xlabel('Predicted Labels')
ax.set_ylabel('True Labels')
plt.show()
```

3) Per l' AUC-ROC:

```
# Compute AUC
auc_altman = roc_auc_score(y_test, pred_Altman)
print(f"AUC: {auc_altman:.3f}")

# Compute ROC curve
fpr, tpr, thresholds = roc_curve(y_test, pred_Altman)

# Plot ROC curve
plt.figure(figsize=(8, 6))
plt.plot(fpr, tpr, label=f"ROC Curve (AUC = {auc_altman:.3f})")
plt.plot([0, 1], [0, 1], linestyle='--', color='gray')  # Diagonal line
plt.xlabel("False Positive Rate")
plt.ylabel("True Positive Rate (Recall)")
plt.title("ROC Curve")
plt.legend(loc="lower right")
plt.grid(True)
plt.show()
```

Si ottengono così i seguenti risultati sui due dataset:

5YEAR ALTMAN

```
False Positive Rate: 0.208
False Negative Rate: 0.439
Precision: 0.167
Recall: 0.561
F1 Score: 0.257
```

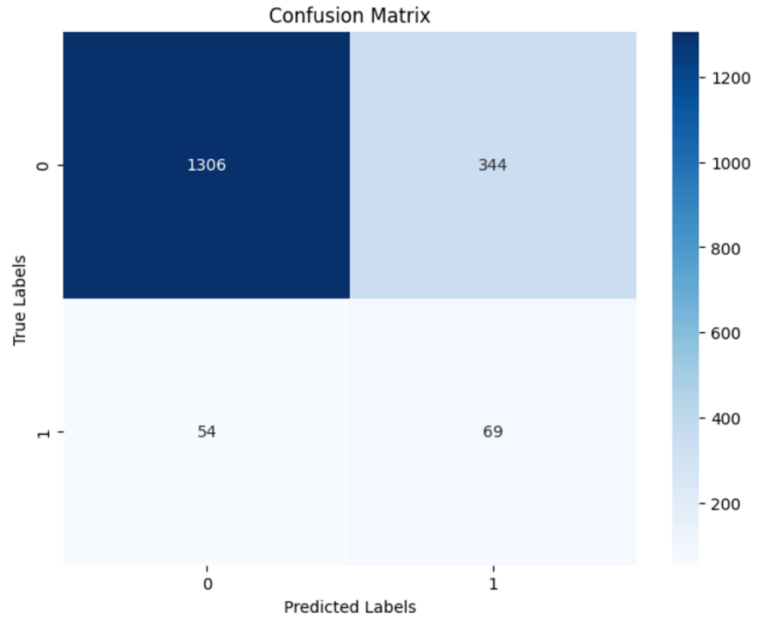

Figura 9

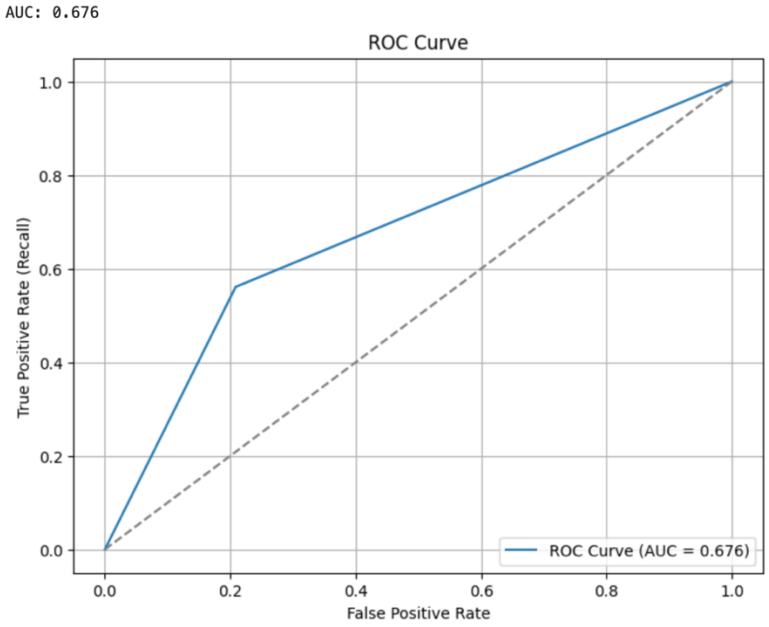

Figura 11

4YEAR ALTMAN

```
False Positive Rate: 0.238
False Negative Rate: 0.484
Precision: 0.108
Recall: 0.516
F1 Score: 0.179
```

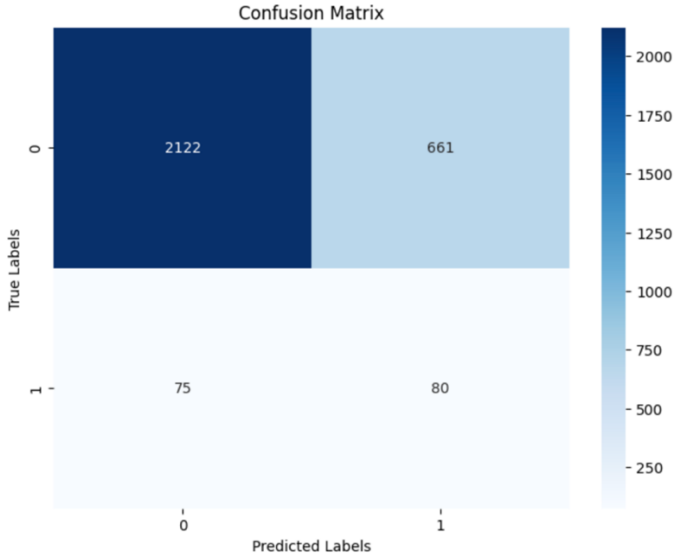

Figura 11

```
AUC: 0.639
```

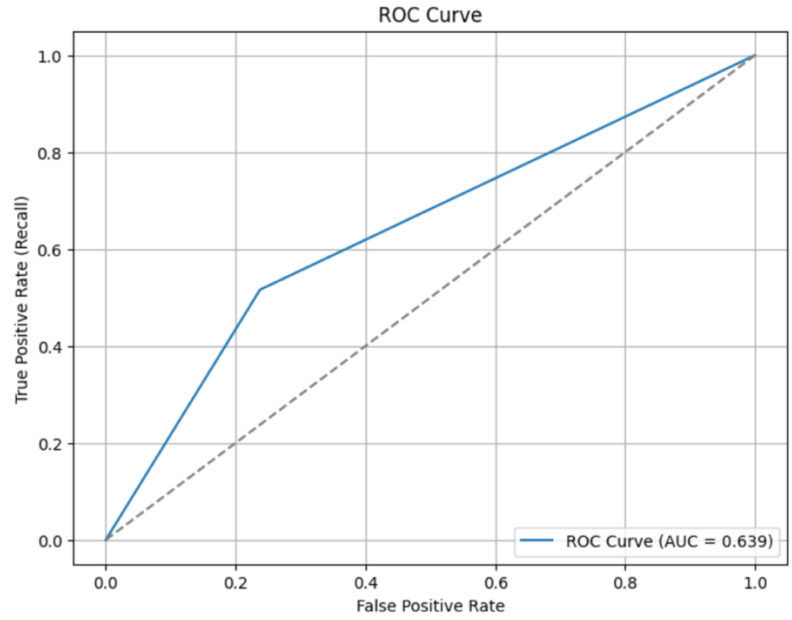

Figura 12

Il modello presenta un alto tasso di falsi negativi (43,9% per il dataset *5year* e 23,8% per il dataset *4year*), cioè molte aziende effettivamente fallite non vengono individuate, e una precisione molto bassa (16,7% e 10,8%), segno che tra le aziende previste come fallite solo poche lo sono davvero. Il valore medio di recall (56,1% e 51,6%) indica che il modello riesce ad inviduare solo poco più della metà delle aziende fallite. L'F1-Score complessivo è basso (25,7% e 17,9%), segnalando una performance debole nell'individuare la classe più critica.

La matrice di confusione evidenziano le criticità del modello nell'affrontare un problema caratterizzato da un marcato sbilanciamento tra classi:

- Vero negativo (1306 nel dataset *5year* – 2122 nel dataset *4year*) → aziende sane classificate come tali;
- Falso positivo (344 *5year* – 661 *4year*) → aziende sane erroneamente classificate come fallite;
- Falso negativo (54 *5year* – 75 *4year*) → aziende fallite che il modello non ha riconosciuto;
- Vero positivo (69 *5year* – 80 *4year*) → aziende fallite correttamente identificate.

Le evidenze indicano che il modello è molto più efficace nel riconoscere le aziende sane rispetto a quelle fallite. Infatti la maggior parte delle aziende sane viene correttamente classificata, mentre solo una piccola parte viene scambiata per fallita; al contrario **circa la metà delle aziende fallite non viene identificata dal modello,** evidente segnale di difficoltà nel cogliere i segnali di fallimento. Questo comportamento riflette la natura rigida e statica del modello, che, non essendo in grado di adattarsi alla distribuzione dei dati, tende a privilegiare la classe dominante. In un contesto in cui la capacità di intercettare tempestivamente i soggetti a rischio è cruciale, tale limitazione rappresenta un ostacolo rilevante. La matrice conferma dunque la necessità di ricorrere a modelli più flessibili, come quelli basati sul machine learning, in grado di gestire in maniera più efficace lo sbilanciamento e di offrire una maggiore sensibilità verso le classi meno rappresentate.

Le curve ROC confermano ulteriormente quanto detto finora: discostandosi solo parzialmente dalla linea di non discriminazione, creano valori AUC molto bassi, quasi prossimi alla casualità (67,6% *5year* – 63,9% *4year)*, confermando che il modello non è in grado di fornire una separazione netta tra le classi.

3.4 – Random Forest

Valutata la bontà complessiva di Altman sui dataset selezionati, lo studio procede con l'applicazione del modello Random Forest sugli stessi dati al fine di effettuare un confronto tra i risultati ottenuti.

Viene effettuata una *variable selection*, ovvero il processo mediante il quale si individuano e si utilizzano solo le variabili più rilevanti (dette anche "feature") tra tutte quelle disponibili, al fine di migliorare le prestazioni del modello e ridurne la complessità

In questo modello Random Forest sono stati selezionati degli iperparametri specifici per ottimizzare l'equilibrio tra accuratezza, generalizzazione e gestione dello sbilanciamento del dataset.

- n_estimators=200: indica il numero di alberi nella foresta. Un valore relativamente alto aumenta la stabilità delle predizioni.
- max_depth=20: limita la profondità di ciascun albero per evitare l'overfitting, cioè un eccessivo adattamento ai dati di training.
- max_features='sqrt': seleziona casualmente solo la radice quadrata del numero totale di variabili a ogni split, aumentando la diversità tra gli alberi.
- min_samples_split=5 e min_samples_leaf=2: impongono un numero minimo di osservazioni per suddividere i nodi o crearne di nuovi, utile per evitare alberi troppo frammentati.
- bootstrap=True: attiva il campionamento casuale con rimpiazzo, principio base del metodo Random Forest.
- class_weight='balanced': è un parametro fondamentale per gestire lo sbilanciamento delle classi, assegnando un peso maggiore alla classe minoritaria (aziende fallite).
- random_state=80: garantisce la ripetibilità dei risultati, fissando la casualità interna al modello.

```
[21] best_rf = RandomForestClassifier(
         n_estimators=200,
         max_depth=20,
         max_features='sqrt',
         min_samples_split=5,
         min_samples_leaf=2,
         bootstrap=True,
         class_weight='balanced',
         random_state=80
     )
```

Viene poi applicata una tecnica di selezione delle variabili basata sull'importanza delle caratteristiche (feature importance) calcolata dal modello Random Forest.

La feature importance è una misura che indica quanto ciascuna variabile contribuisce alla capacità predittiva complessiva del modello. Difatti, quando una variabile viene scelta per effettuare una divisione (split) in un albero, il modello valuta quanto migliora la "purezza" dei gruppi risultanti: più i gruppi contengono esempi omogenei (cioè della stessa classe), maggiore è il guadagno in termini di riduzione dell'impurità.

La feature importance di una variabile è quindi calcolata come la somma dei guadagni di impurità ottenuti ogni volta che quella variabile viene usata in uno split, aggregata su tutti gli alberi della foresta.

```
# Fit your model
best_rf.fit(X_train, y_train)

# Get feature importances
importances = best_rf.feature_importances_
features = X_train.columns
feat_importances = pd.Series(importances, index=features)

# Plot
feat_importances.sort_values(ascending=False).plot(kind='bar', figsize=(10,6), title='Feature Importances')
plt.tight_layout()
plt.show()

# Select top k features (e.g., top 10)
top_features = feat_importances.sort_values(ascending=False).head(10).index.tolist()
X_train_selected_rf = X_train[top_features]
X_test_selected_rf = X_test[top_features]
```

Feature importances *5year*

	Attr	Description
0	Attr16	(gross profit + depreciation) / total liabilities
1	Attr21	sales (n) / sales (n-1)
2	Attr22	profit on operating activities / total assets
3	Attr26	(net profit + depreciation) / total liabilities
4	Attr27	profit on operating activities / financial exp...
5	Attr35	profit on sales / total assets
6	Attr39	profit on sales / sales
7	Attr41	total liabilities / ((profit on operating acti...
8	Attr42	profit on operating activities / sales
9	Attr45	net profit / inventory

Tabella 15

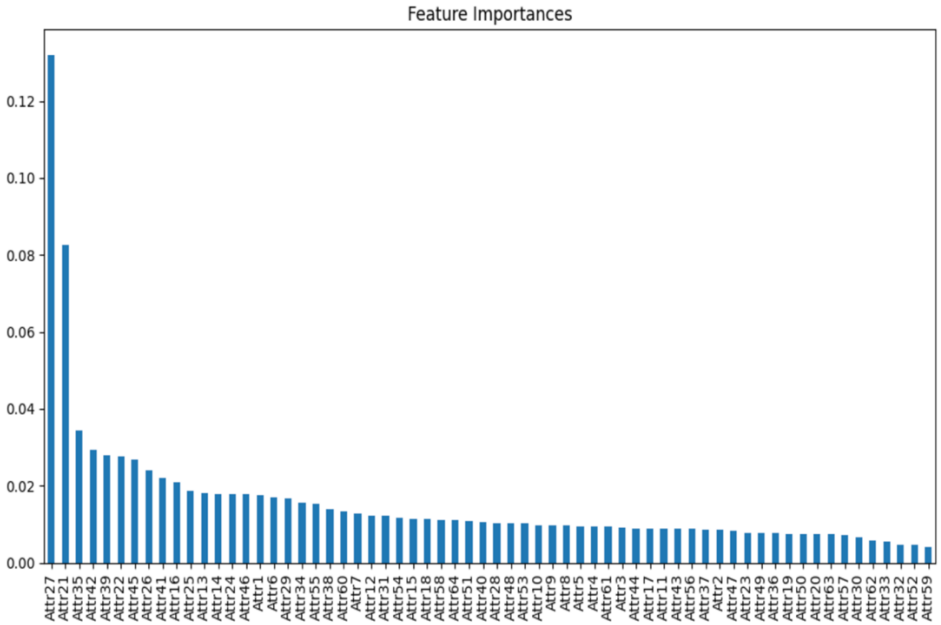

Figura 13

Feature importance *4year*

	Attr	Description
0	Attr13	(gross profit + depreciation) / sales
1	Attr16	(gross profit + depreciation) / total liabilities
2	Attr21	sales (n) / sales (n-1)
3	Attr24	gross profit (in 3 years) / total assets
4	Attr25	(equity - share capital) / total assets
5	Attr26	(net profit + depreciation) / total liabilities
6	Attr27	profit on operating activities / financial exp...
7	Attr34	operating expenses / total liabilities
8	Attr39	profit on sales / sales
9	Attr46	(current assets - inventory) / short-term liab...

Tabella 16

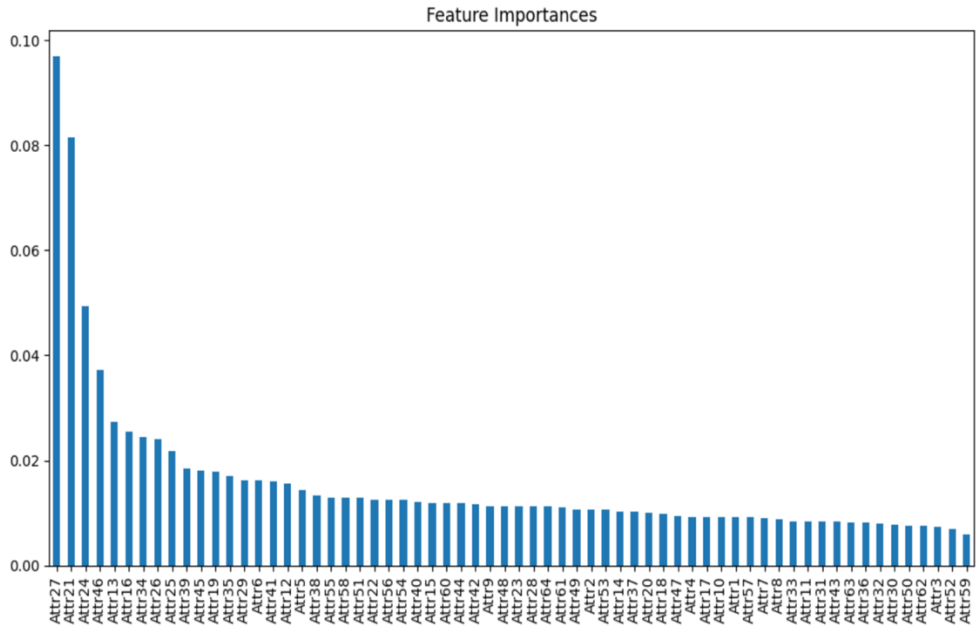

Figura 14

Il modello viene addestrato usando solo le 10 variabili più importanti; vengono poi effettuate le previsioni e visualizzate le metriche di classificazione e la matrice di confusione.

```
# Fit the model on the training data
best_rf.fit(X_train_selected_rf, y_train)

# Make predictions
y_pred_rf_variable_selection = best_rf.predict(X_test_selected_rf)

# Classification report and confusion matrix
classification_report_rf_selection = classification_report(y_test, y_pred_rf_variable_selection, output_dict=True)
print(classification_report(y_test, y_pred_rf_variable_selection))

cm = confusion_matrix(y_test, y_pred_rf_variable_selection, labels=best_rf.classes_)


# Confusion matrix heatmap
fig, ax = plt.subplots(figsize=(8, 6))
sns.heatmap(cm, annot=True, ax=ax, cmap='Blues', fmt='g')
ax.set_title('Confusion Matrix')
ax.set_xlabel('Predicted Labels')
ax.set_ylabel('True Labels')
ax.xaxis.set_ticklabels(best_rf.classes_)
ax.yaxis.set_ticklabels(best_rf.classes_)
plt.show()
```

Output *5year*

	precision	recall	f1-score	support
0	0.97	0.98	0.97	1650
1	0.69	0.55	0.62	123
accuracy			0.95	1773
macro avg	0.83	0.77	0.79	1773
weighted avg	0.95	0.95	0.95	1773

Tabella 17

Output *4year*

	precision	recall	f1-score	support
0	0.97	0.99	0.98	2783
1	0.81	0.44	0.57	155
accuracy			0.96	2938
macro avg	0.89	0.72	0.78	2938
weighted avg	0.96	0.96	0.96	2938

Tabella 18

Questo output conferma come il modello, su entrambi i dataset, pur eccellendo ulteriormente nel riconoscimento delle aziende sane (classe 0), ottiene al contempo prestazioni molto più bilanciate rispetto allo Z-Score anche sulla classe meno

rappresentata (aziende fallite, classe 1). Questo miglioramento è possibile grazie all'uso dell'iperparametro **class_weight='balanced'**, che compensa lo sbilanciamento del dataset assegnando un peso maggiore alla classe meno rappresentata. In questo modo, il modello non ignora le aziende fallite e viene spinto a imparare anche da questi casi rari ma cruciali. Questo esempio dimostra chiaramente l'importanza della scelta accurata degli iperparametri, che può fare la differenza tra un modello sbilanciato e poco utile, e uno realmente efficace nell'affrontare problemi asimmetrici come la previsione del fallimento aziendale.

Si passa così, come per lo Z-Score, al calcolo delle metriche fondamentali, della matrice di confusione e dell'AUC-ROC.

5YEAR RANDOM FOREST

```
False Positive Rate: 0.018
False Negative Rate: 0.447
Precision: 0.694
Recall: 0.553
F1 Score: 0.615
```

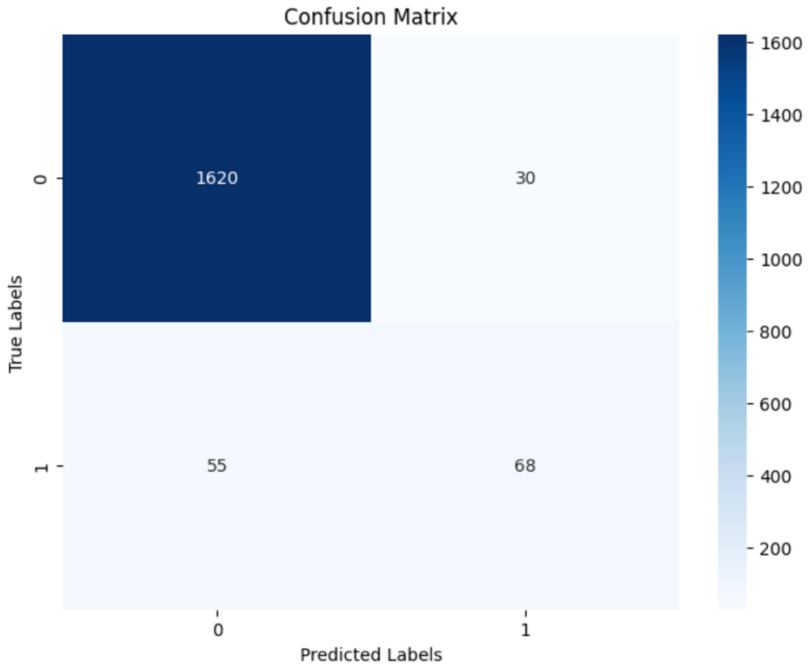

Figura 15

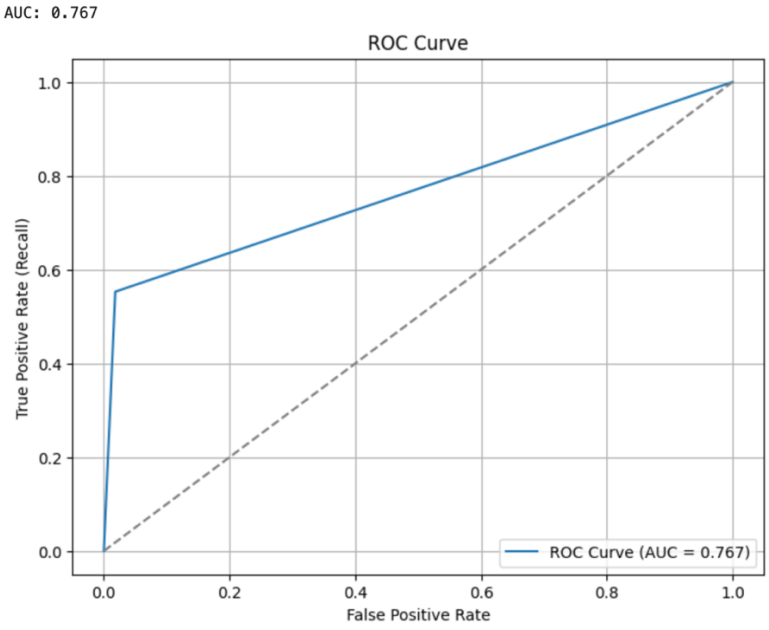

Figura 16

4YEAR RANDOM FOREST

```
False Positive Rate: 0.006
False Negative Rate: 0.561
Precision: 0.810
Recall: 0.439
F1 Score: 0.569
```

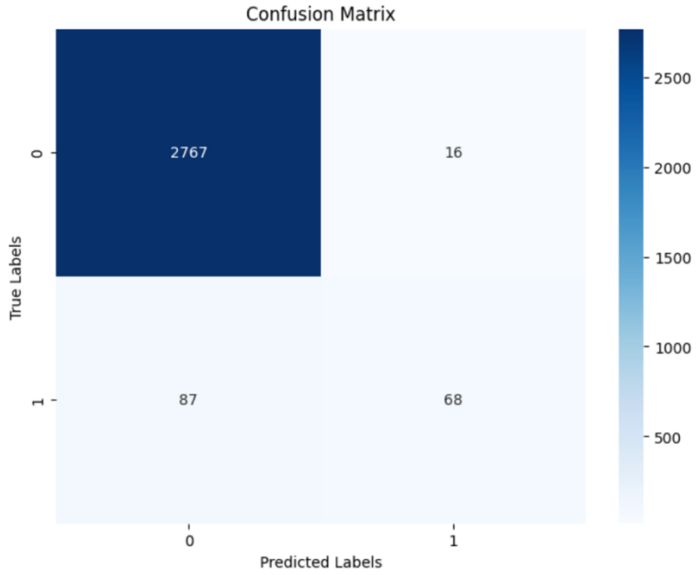

Figura 17

```
AUC: 0.716
```

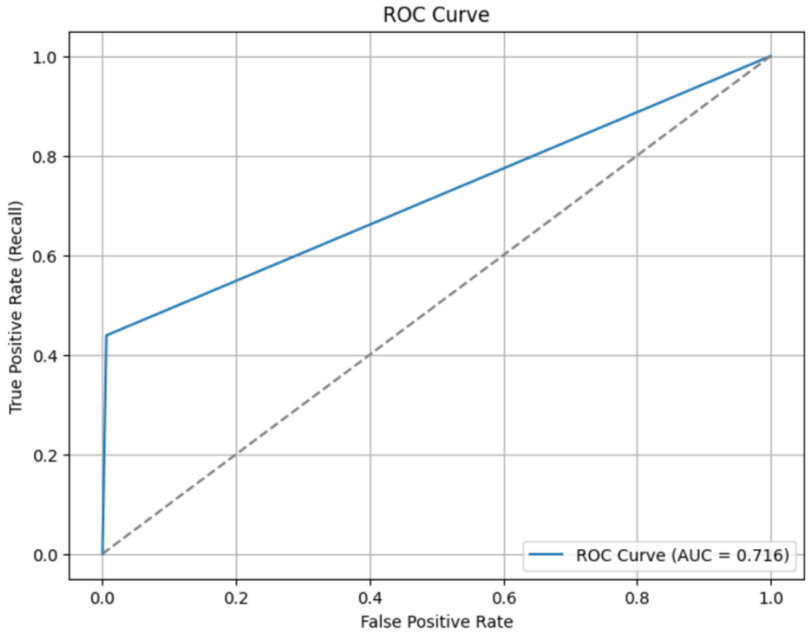

Figura 18

I risultati del RF superano di gran lunga quelli dello Z-Score:

5year

	FP Rate	FN Rate	Precision	Recall	F1 Score	AUC
Altman	0.208	0.439	0.167	0.561	0.257	0.676
Random Forest	0.018	0.447	0.694	0.553	0.615	0.767

Tabella 19

4year

	FP Rate	FN Rate	Precision	Recall	F1 Score	AUC
Altman	0.238	0.484	0.108	0.516	0.179	0.639
Random Forest	0.006	0.561	0.810	0.439	0.569	0.716

Tabella 20

- *FP RATE*: il modello RF commette molti meno falsi positivi, cioè classifica con maggiore precisione le aziende sane come tali;
- *FN RATE*: entrambi i modelli hanno una quota simile di aziende fallite non riconosciute, con un leggero svantaggio per RF;
- *Precision*: RF si mostra nettamente più preciso nel classificare le aziende fallite: quando i modello predice un fallimento, ha molte più probabilità di avere ragione;
- *Recall*: i due modelli hanno una capacità simile di individuare le aziende fallite, ma RF lo fa con molta più affidabilità (come visto dalla precisione);
- *F1-Score*: l'equilibrio tra precisione e recall è più che doppio nel modello di RF, segno di una performance molto più stabile e bilanciata.
- *AUC*: RF mostra una migliore capacità discriminante, riuscendo a separare in modo più efficace le aziende sane da quelle fallite.

Questo confronto evidenzia in modo inequivocabile la superiorità del modello RF rispetto al tradizionale Z-Score, soprattutto in termini di precisione, F1 score e AUC. Sebbene entrambi i modelli abbiano un livello simile di recall, RF risulta più affidabile e accurato nella gestione delle classi sbilanciate, grazie alla sua struttura adattiva e alla capacità di integrare il bilanciamento delle classi tramite iperparametri come **class_weight**. Tali

evidenze sottolineano il valore aggiunto offerto dai modelli di machine learning in contesti predittivi complessi e sbilanciati.

3.5 – XG Boost

Lo studio termina con la valutazione dei risultati di un modello di Boosting sui due dataset utilizzati per Z-Score e RF, mantenendo l'obiettivo di individuare quale modello meglio performa con dei dati che presentano le suddette caratteristiche.
Viene utilizzato il seguente codice per l'impostazione del modello:

```
import numpy as np
class_counts = np.bincount(y_train)
weights = class_counts[0] / class_counts[1]


fixed_params = {
    'n_estimators': 300,
    'max_depth': 6,
    'learning_rate': 0.1,
    'subsample': 0.8,
    'colsample_bytree': 0.8,
    'gamma': 0.1,
    'min_child_weight': 1,
    'random_state': 80,
    'scale_pos_weight': weights
}

xgb_model = xgb.XGBClassifier(**fixed_params)
```

I seguenti iperparametri controllano il comportamento del modello:

- *n_estimators*: numero di alberi nella sequenza (più alberi = più complessità);
- *max_depth*: profondità massima degli alberi (controlla l'overfitting);
- *learning_rate*: tasso di apprendimento (più basso = apprendimento più lento ma più preciso);
- *subsample*: frazione di osservazioni usate per ogni albero (aggiunge casualità);
- *colsample_bytree*: frazione di variabili usate per ogni albero (riduce overfitting);
- *gamma*: penalizzazione sulla complessità degli split (più alto = modello più conservativo);
- *min_child_weight*: peso minimo richiesto per dividere un nodo (regola la robustezza);
- *random_state*: per rendere i risultati riproducibili;

- *scale_pos_weight*: corregge lo sbilanciamento tra classi, migliorando la rilevazione dei casi minoritari.

Successivamente, come per il modello di RF, viene effettuata la *Feature importance*, ovvero la scelta degli iperparametri che risultano più influenti, aventi la stessa funzione del modello di RF.

```
# Feature importances
importances = xgb_model.feature_importances_
features = X_train.columns
feat_importances = pd.Series(importances, index=features)

# Plot
feat_importances.sort_values(ascending=False).plot(kind='bar', figsize=(10, 6), title='XGBoost Feature Importances')
plt.tight_layout()
plt.show()

# Select top k features
top_features = feat_importances.sort_values(ascending=False).head(10).index
X_train_selected_xgb = X_train[top_features]
X_test_selected_xgb = X_test[top_features]
```

Feature importance 5year

	Attr	Description
0	Attr6	retained earnings / total assets
1	Attr19	gross profit / sales
2	Attr21	sales (n) / sales (n-1)
3	Attr25	(equity - share capital) / total assets
4	Attr26	(net profit + depreciation) / total liabilities
5	Attr27	profit on operating activities / financial exp...
6	Attr34	operating expenses / total liabilities
7	Attr35	profit on sales / total assets
8	Attr38	constant capital / total assets
9	Attr62	(short-term liabilities *365) / sales

Tabella 21

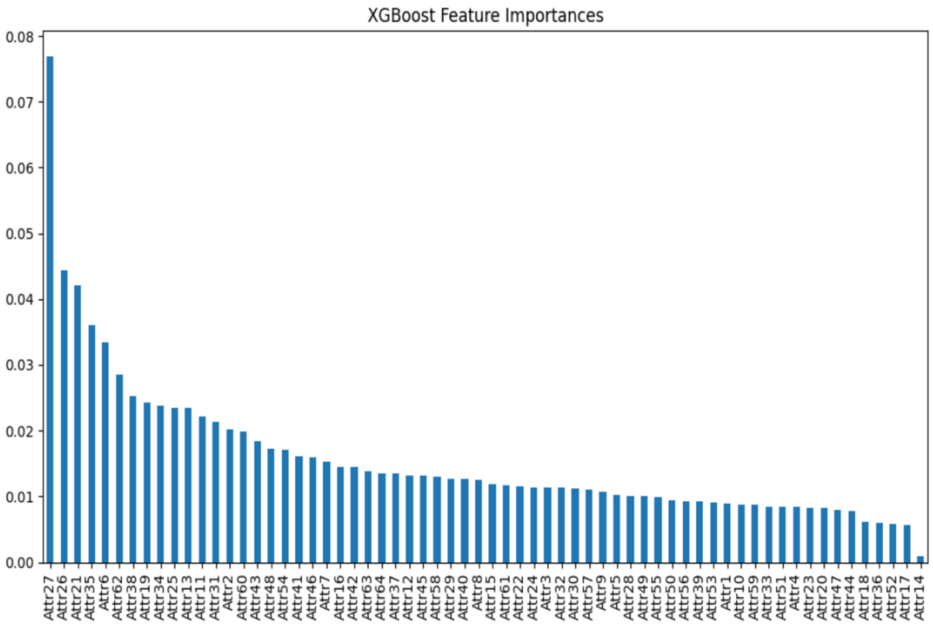

Figura 19

Feature importance 4year

	Attr	Description
0	Attr13	(gross profit + depreciation) / sales
1	Attr21	sales (n) / sales (n-1)
2	Attr22	profit on operating activities / total assets
3	Attr24	gross profit (in 3 years) / total assets
4	Attr26	(net profit + depreciation) / total liabilities
5	Attr27	profit on operating activities / financial exp...
6	Attr34	operating expenses / total liabilities
7	Attr46	(current assets - inventory) / short-term liab...
8	Attr61	sales / receivables
9	Attr62	(short-term liabilities *365) / sales

Tabella 22

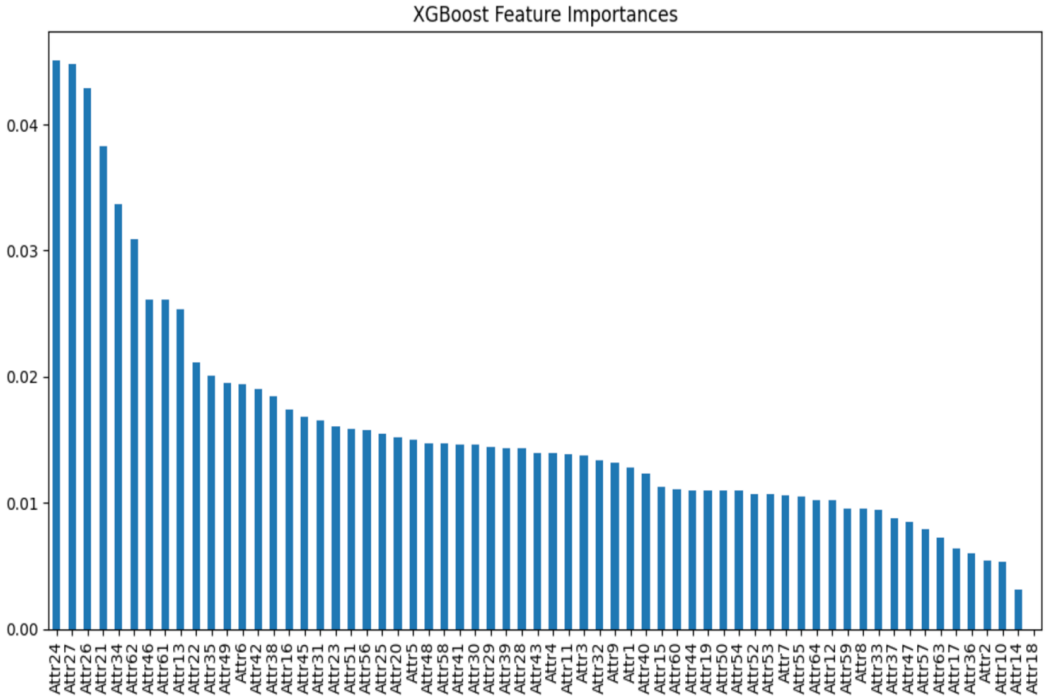

Figura 20

Come per gli altri due modelli, vengono calcolate le metriche fondamentali, che presentano i seguenti risultati:

5year

	precision	recall	f1-score	support
0	0.97	0.98	0.98	1650
1	0.71	0.62	0.66	123
accuracy			0.96	1773
macro avg	0.84	0.80	0.82	1773
weighted avg	0.95	0.96	0.95	1773

Tabella 23

4year

	precision	recall	f1-score	support
0	0.98	0.99	0.98	2783
1	0.75	0.59	0.66	155
accuracy			0.97	2938
macro avg	0.87	0.79	0.82	2938
weighted avg	0.97	0.97	0.97	2938

Tabella 24

Le prestazioni ottenute sono eccellenti: in merito alle aziende sane, il modello ottiene precisione, recall ed F1-Score superiore al 97% su entrambi i dataset, confermando un'eccellente capacità di identificare le aziende sane come tali.

In merito alle aziende fallite, il modello presenta su entrambi i dataset una precisione superiore al 70%, un recall intorno al 60% ed un F1-Score pari al 66%: quest'ultimo dato è molto significativo, in quanto il modello riesce ad individuare oltre il 60% delle aziende fallite e lo fa con un'ottima affidabilità: circa il 70% delle volte (precisione), quando predice un fallimento, ha ragione. I valori ottenuti sono simili a quelli del modello di RF ma notevolmente superiori rispetto al modello di Altman. L'accuracy complessiva risulta superiore al 96%, indicando che quasi la totalità delle previsioni sono corrette.

Una volta individuati gli iperparametri, allenato il modello e ottenute le metriche di classificazione, si passa al calcolo delle metriche fondamentali, della matrice di confusione e dell'AUC-ROC

5YEAR XGBOOST

```
False Positive Rate: 0.019
False Negative Rate: 0.382
Precision: 0.710
Recall: 0.618
F1 Score: 0.661
```

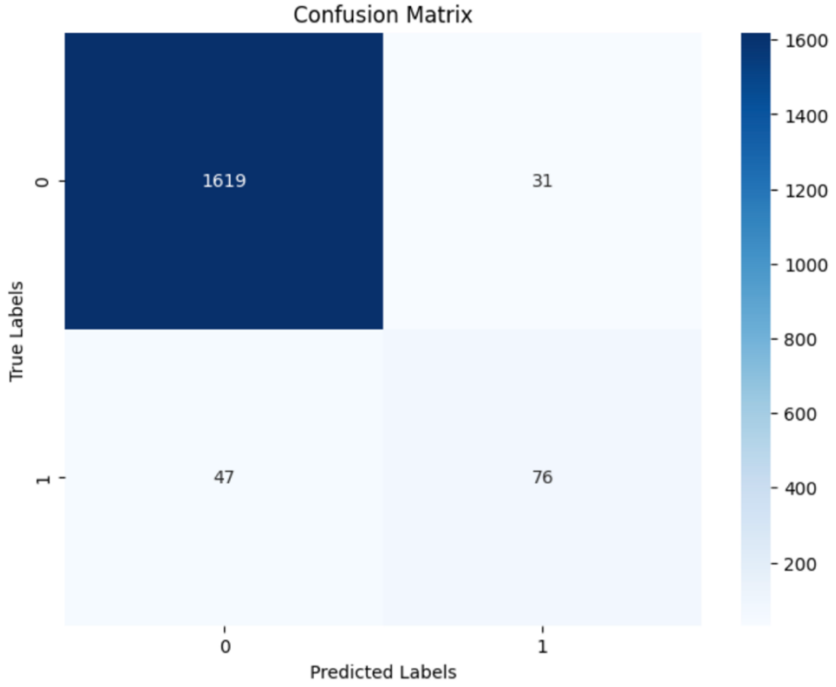

Figura 21

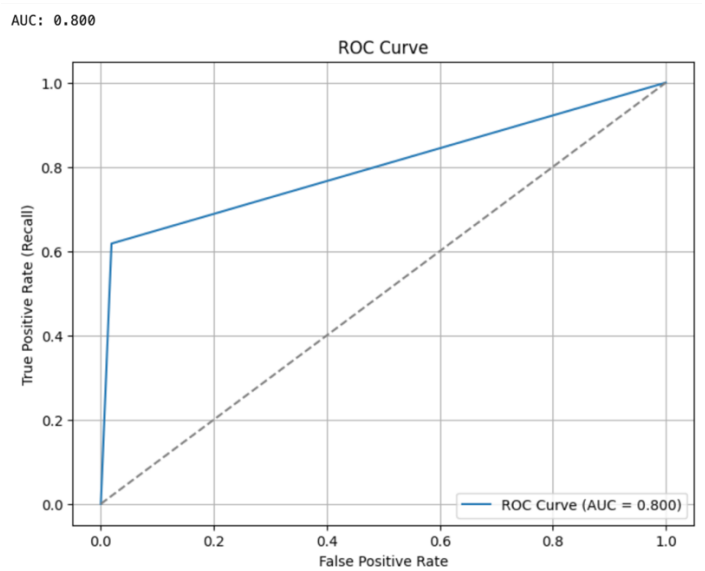

Figura 22

4YEAR XGBOOST

```
False Positive Rate: 0.011
False Negative Rate: 0.406
Precision: 0.754
Recall: 0.594
F1 Score: 0.664
```

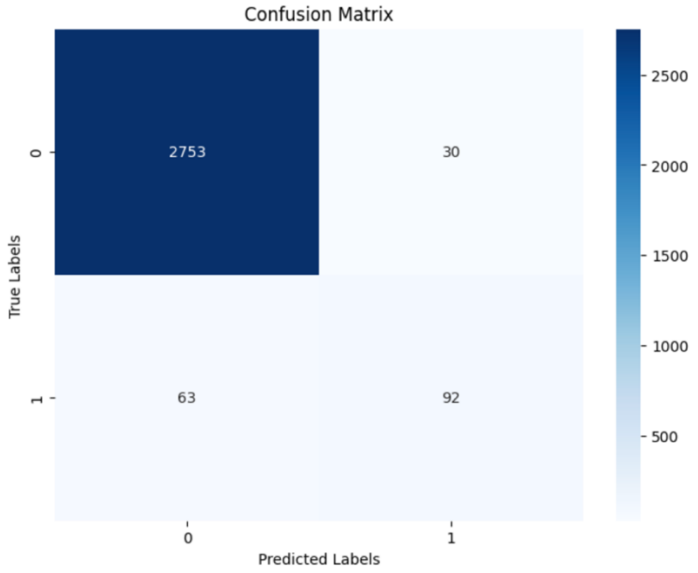

Figura 23

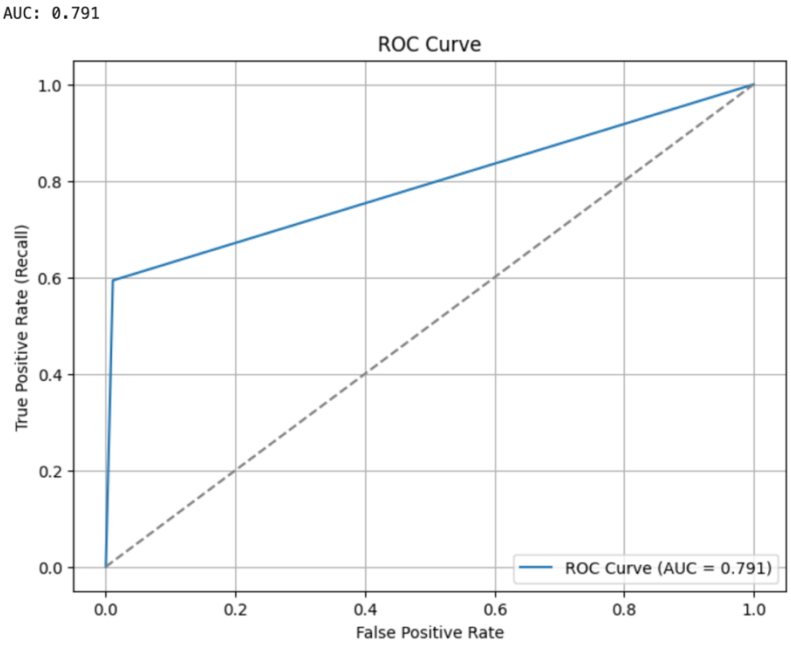

Figura 24

3.6 – Analisi comparativa delle performance predittive

L'analisi empirica condotta ha offerto un confronto esaustivo e strutturato tra tre approcci distinti alla previsione del fallimento aziendale. I due dataset utilizzati per i test sono caratterizzati da una distribuzione fortemente sbilanciata tra classi, con una netta prevalenza di imprese sane rispetto a quelle fallite: tale caratteristica ha rappresentato una sfida metodologica cruciale, rendendo necessaria l'adozione di strategie specifiche per la gestione dell'asimmetria.

5year

	FP Rate	FN Rate	Precision	Recall	F1 Score	AUC
Altman	0.208	0.439	0.167	0.561	0.257	0.676
Random Forest	0.018	0.447	0.694	0.553	0.615	0.767
Gradient Boosting	0.019	0.382	0.710	0.618	0.661	0.800

Tabella 25

4year

	FP Rate	FN Rate	Precision	Recall	F1 Score	AUC
Altman	0.238	0.484	0.108	0.516	0.179	0.639
Random Forest	0.006	0.561	0.810	0.439	0.569	0.716
Gradient Boosting	0.011	0.406	0.754	0.594	0.664	0.791

Tabella 26

Il modello Z-Score, pur vantando una lunga tradizione e una comprovata semplicità operativa, ha mostrato evidenti limiti strutturali. La sua incapacità di adattarsi alla distribuzione dei dati, unita all'impossibilità di apprendere relazioni complesse tra variabili, ha determinato performance deludenti sia in termini di precisione (spesso prevede il fallimento anche quando l'azienda è sana) sia di recall, con un F1-Score significativamente basso (indice di scarso equilibrio tra precisione e capacità di

individuare correttamente le casistiche) e con valori di AUC-ROC – i quali indicano la capacità del modello nella distinzione tra aziende sane e fallite – che si avvicinano alla casualità. Il modello tende a classificare in modo eccessivamente conservativo, privilegiando la classe maggioritaria e penalizzando fortemente la rilevazione delle aziende a rischio, come evidenziato dall'elevato False Negative Rate (non riesce ad individuare buona parte delle aziende che invece falliranno davvero).

Il classificatore Random Forest, introdotto come primo modello di machine learning, ha evidenziato un salto qualitativo rilevante. Grazie all'utilizzo dell'iperparametro *class_weight='balanced'*, il modello ha saputo mitigare lo sbilanciamento tra le classi, migliorando la sensibilità verso la classe minoritaria. La precisione risulta nettamente superiore rispetto allo Z-Score, e anche l'F1-Score evidenzia un equilibrio più robusto tra affidabilità e sensibilità. Tuttavia, si osserva una recall ancora limitata, soprattutto nei confronti delle imprese effettivamente fallite, con conseguente FNR relativamente elevato. Questo implica che, se da un lato il modello è molto affidabile nel segnalare un fallimento (quando lo fa, spesso ha ragione), dall'altro può trascurare una porzione non trascurabile di aziende a rischio reale.

Infine, il modello XGBoost – uno degli algoritmi più avanzati nel campo del gradient boosting – ha registrato le performance più solide e coerenti. L'utilizzo mirato degli iperparametri, in particolare *scale_pos_weight* per la gestione dello sbilanciamento, ha contribuito a ottenere un eccellente compromesso tra precision, recall e AUC-ROC. Il modello riesce a riconoscere oltre il 60% delle aziende fallite (recall), con una precisione superiore al 70%, raggiungendo F1-Score attorno al 66% e accuracy complessiva oltre il 96%. Anche le curve ROC, con valori AUC prossimi all'80%, indicano una forte capacità discriminante tra le classi. Ciò dimostra non solo l'accuratezza del modello, ma anche la sua robustezza in contesti reali ad alta complessità informativa.

Dal punto di vista metodologico, l'analisi conferma l'importanza della calibrazione dei modelli predittivi, specialmente in presenza di dataset fortemente sbilanciati. La capacità di bilanciare efficacemente le metriche di classificazione (minimizzando sia FPR che FNR) si è dimostrata cruciale per ottenere risultati affidabili, con particolare attenzione alle metriche più significative in un contesto applicativo sensibile come la previsione del default. In questo senso, modelli rigidi come lo Z-Score si rivelano inadeguati, mentre

approcci algoritmici basati sull'apprendimento dai dati si dimostrano altamente promettenti.

Dal punto di vista operativo, le implicazioni sono altrettanto rilevanti. L'affidabilità del modello XGBoost nel riconoscere precocemente segnali di crisi lo rende un candidato ideale per applicazioni concrete da parte di banche, investitori, revisori e autorità di vigilanza. Il fatto che tale modello riesca a individuare in modo efficace le aziende a rischio, senza compromettere la precisione sulle aziende sane, rappresenta un vantaggio strategico decisivo, permettendo interventi preventivi mirati e tempestivi.

In conclusione, i risultati ottenuti rafforzano significativamente l'ipotesi secondo cui i modelli di machine learning, se adeguatamente progettati, superano di gran lunga i modelli tradizionali nella capacità di prevedere il fallimento aziendale. In particolare, XGBoost emerge come lo strumento più efficace, grazie alla sua capacità di apprendere strutture complesse, adattarsi al contesto dei dati e gestire correttamente la classe critica del fallimento. Questo studio, integrando rigore metodologico e sperimentazione operativa, dimostra che le tecniche avanzate di AI non sono solo un'opzione teorica, ma una concreta opportunità per migliorare i sistemi di early warning finanziario, con ricadute potenzialmente rilevanti per la stabilità economica e per la prevenzione delle crisi d'impresa.

Conclusioni

La previsione della crisi d'impresa rappresenta oggi un tema cruciale per la sostenibilità economica, la continuità aziendale e la salvaguardia del tessuto produttivo. In un contesto sempre più instabile, interconnesso e competitivo, le imprese sono esposte a molteplici fattori di rischio, spesso difficili da intercettare in tempo utile. Di conseguenza, il ricorso a strumenti predittivi efficaci e tempestivi non costituisce più un'opzione accessoria, ma una necessità strategica.

Il presente lavoro ha indagato in profondità l'evoluzione dei modelli di previsione della crisi, analizzando tanto gli approcci tradizionali di matrice statistica, quanto le più recenti applicazioni dell'intelligenza artificiale. I modelli predittivi, lungi dall'essere strumenti neutri, riflettono una precisa visione del rischio d'impresa: da una parte vi sono quelli storici, basati su indici di bilancio e relazioni lineari; dall'altra, i moderni algoritmi di machine learning, capaci di apprendere da grandi moli di dati e di adattarsi dinamicamente ai cambiamenti.

Dall'analisi condotta emerge con chiarezza che i modelli più innovativi offrono prestazioni sensibilmente superiori in termini di accuratezza, precisione e capacità discriminante. Algoritmi come Random Forest e XGBoost, opportunamente calibrati, si dimostrano in grado di individuare segnali di crisi con maggiore anticipo e affidabilità rispetto ai tradizionali modelli quantitativi. La possibilità di elaborare pattern complessi, di gestire l'eterogeneità delle informazioni e di adattarsi a contesti variabili costituisce un vantaggio competitivo decisivo per chi si occupa di valutazione del rischio aziendale.

Dall'analisi emerge la necessità, per imprese e professionisti, di aggiornare gli strumenti di diagnosi precoce, integrando ai tradizionali indicatori di bilancio approcci basati su intelligenza artificiale e data science. In secondo luogo, anche il legislatore e le istituzioni finanziarie dovrebbero considerare l'utilizzo di questi strumenti nei processi di valutazione del rischio e di monitoraggio delle condizioni economico-patrimoniali delle imprese, per rafforzare l'efficacia dei sistemi di allerta e contenere l'impatto delle crisi sistemiche.

Tuttavia, questa potenza predittiva porta con sé anche nuove responsabilità. L'affidabilità dei modelli di machine learning dipende in modo critico dalla qualità dei dati, dalla trasparenza degli algoritmi, dalla capacità di interpretarne i risultati e di comprenderne i limiti. In questo senso, l'efficacia dei modelli predittivi non risiede soltanto nella loro

componente tecnica, ma anche nella loro governance e nella capacità degli operatori di integrarli nei processi decisionali in modo consapevole, etico e informato.

L'adozione di strumenti predittivi avanzati può portare benefici significativi non solo per le singole imprese, ma anche per il sistema economico nel suo complesso: un'efficace prevenzione delle crisi contribuisce infatti a ridurre i fallimenti, a tutelare i posti di lavoro, a migliorare la reputazione delle imprese e a sostenere la fiducia nei mercati. È in quest'ottica che tali modelli dovrebbero essere progressivamente integrati anche nelle prassi degli organi di controllo, delle banche, dei revisori e dei sistemi di allerta previsti dalla normativa.

Il lavoro svolto mostra dunque che la combinazione tra conoscenze economico-aziendali, capacità analitica e innovazione tecnologica è la chiave per affrontare con efficacia le sfide della prevenzione della crisi d'impresa. La sfida futura sarà quella di rendere questi strumenti accessibili, affidabili e comprensibili, costruendo un ponte tra complessità tecnica e operatività concreta.

In ultima analisi, la previsione della crisi d'impresa non è solo un esercizio quantitativo, ma un elemento essenziale di buona gestione. Sapere anticipare, comprendere e intervenire rappresenta un segno distintivo di imprese resilienti, consapevoli e orientate al lungo periodo. La qualità delle decisioni dipende oggi più che mai dalla capacità di leggere il futuro: i modelli predittivi possono e devono essere al servizio di questa lettura, con l'obiettivo di trasformare i segnali di allarme in occasioni di rilancio e rigenerazione.

Bibliografia

- Abikoye, B. (2021). *Machine Learning Models and AI for Predicting Financial Crises: Applications and Accuracy.*
- Agnoli, N., Zamboni, M. (2021). *Intelligenza artificiale e previsione delle crisi aziendali*. Amministrazione e finanza, n. 12.
- Airoldi, G., Brunetti, G., Coda, V. (1994). *Economia aziendale*. Il Mulino.
- Aliaj, T., Anagnostopoulos, A., Piersanti, S. (2020). *Firms Default Predictions with Machine Learning*. Lecture Notes in Computer Science, Springer.
- Alanis, E., Chava, S., & Shah, A. (2022). *Benchmarking machine learning models to predict corporate bankruptcy*. arXiv.
- Altman, E. I. (1968). *Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy*. Journal of Finance, 23(4), 589–609.
- Altman, E. I. (1983). *Corporate Financial Distress: A Complete Guide to Predicting, Avoiding, and Dealing with Bankruptcy*. Wiley.
- Altman, E. I. (2000). *Predicting financial distress of companies: Revisiting the Z-score and ZETA models*. Working paper, NYU Stern School of Business.
- Altman, E. I. (2002). *Revisiting the Z-Score and ZETA Models – EMS Adaptations*. Working paper, NYU Stern School.
- Altman, E. I., Hotchkiss, E. (2006). *Corporate Financial Distress and Bankruptcy*. Wiley.
- Andrews, D., Petroulakis, F. (2019). *Breaking the shackles: Zombie firms, weak banks and depressed restructuring in Europe*. ECB Working Paper Series 2240.
- Argenti, J. (1976). *Corporate Collapse: The Causes and Symptoms*. McGraw-Hill.
- Athey, S. (2018). *The impact of machine learning on economics*. The Economics of Artificial Intelligence: An Agenda. University of Chicago Press.
- Baiju, B., & Rufsana, A. R. (2023). *A review on bankruptcy prediction using machine learning techniques*. Proceedings of the International Conference Emerging Trends in Engineering.
- Barboza, F., Kimura, H., Altman, E.I. (2017). *Machine learning models and bankruptcy prediction*. Expert Systems with Applications, 83, 405-417.

- Bargagli-Stoffi, F., Gallo, L., Rungi, A., & Rungi, M. (2023). *Machine learning for zombie hunting: Predicting distress from firms' accounts and missing values*. arXiv.
- Barogoli, D., Ferretti, C., Ganugi, P., Marseguerra, G., Mezzogori, D., & Zammori, F. (2019). *Machine learning models for bankruptcy prediction in Italy: Do industrial variables count?* Working Paper No. 19/3, Università Cattolica del Sacro Cuore.
- Bastia, P., Campedelli, B. (2020). *Bilancio e controllo della crisi d'impresa*. FrancoAngeli.
- Bauer, J., Agarwal, V. (2014). *Are hazard models superior to traditional bankruptcy prediction approach?* Journal of Banking and Finance, 40, 432-442.
- Beaver, W.H. (1966). *Financial Ratios as Predictors of Failure*. Journal of Accounting Research, 4, 71-111.
- Breiman, L. (2001). *Random Forests*. Statistics Department, University of Berkeley, CA.
- Bragoli, D., Ferretti, C., Ganugi, P., Marseguerra, G., Mezzogori, D., & Zammori, F. (2019). *Machine learning models for bankruptcy prediction in Italy*. Università Cattolica.
- Camera di Commercio Treviso-Belluno. *Come opera l'OCRI? Da chi è costituito?*
- Commissione Europea (2016). *Comunicazione della Commissione: Quadro europeo per la ristrutturazione preventiva*. Bruxelles.
- Consiglio Nazionale dei Dottori Commercialisti e degli Esperti Contabili (CNDCEC) (2019). *Gli indici dell'allerta ex art. 13, co. 2 Codice della Crisi e dell'Insolvenza.*
- Dellepiane, U., Di Marcantonio, M., Laghi, E., & Renzi, S. (2015). *Bankruptcy Prediction Using Support Vector Machines and Feature Selection During the Recent Financial Crisis*. International Journal of Economics and Finance, 182-196.
- Di Martino, M. (2024). *Intelligenza artificiale e sistemi di segnalazione della crisi d'impresa*. De Iustitia, ISSN 2974-7562.
- du Jardin, P. (2015). *Bankruptcy prediction using terminal failure processes*. European Journal of Operational Research, 242(1), 286–303.

- Ha, H. H., Dang, N. H., & Tran, M. D. (2023). *Financial distress forecasting with a machine learning approach*. Corporate Governance and Organizational Behavior Review, 7(3), 90–104.
- Horak, J., Vrbka, J., & Suler, P. (2020). *Support vector machine methods and artificial neural networks used for the development of bankruptcy prediction models and their comparison*. Journal of Risk and Financial Management, 13(3).
- Il Sole 24 Ore. (2019). *Gli indici di allerta dei commercialisti.*
- Invernizzi, G. (2000). *Strategie e politiche per il governo dell'impresa in crisi*. Egea.
- Kendirli, S., & Citak, L. (2022). *Estimating Financial Failure in Businesses Using Artificial Neural Networks*. Journal of Corporate Governance and Information Risk Management, 9(2), 1–15.
- OECD (2017). *OECD Corporate Governance Factbook 2017*. OECD Publishing.
- Panizza, R. (2021). *La gestione della crisi d'impresa tra prevenzione e ristrutturazione*. Giappichelli.
- Regio Decreto n.267 del 16 marzo 1942.
- Ruotolo, M. (2013). *Analisi di SVM per la classificazione automatica*. Università di Torino.
- Shumway, T. (2001). *Forecasting Bankruptcy More Accurately: A Simple Hazard Model*. Journal of Business, 74(1), 101–124.
- Teece, D. J., Pisano, G., & Shuen, A. (1997). *Dynamic Capabilities and Strategic Management*. Strategic Management Journal, 18(7), 509-533.
- Tipaldi, A. (2022). *Crisi d'impresa e modelli predittivi: il ruolo dell'I.A. e della tecnologia blockchain nella prevenzione e nella composizione della crisi*. Rivista di diritto dell'impresa, Edizioni Scientifiche Italiane.
- Wruck, K. H. (1990). *Financial Distress, Reorganization, and Organizational Efficiency*. Journal of Financial Economics, 27(2), 419-444.
- Zhang, G., Hu, M. Y., Patuwo, B. E., & Indro, D. C. (1999). *Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis*. European Journal of Operational Research, 116(1), 16–32.
- Zmijewski, M. E. (1984). *Methodological Issues Related to the Estimation of Financial Distress Prediction Models*. Journal of Accounting Research, 59–82.