



Libera Università degli Studi Sociali - LUISS Guido Carli

Dipartimento di Impresa e Management

Corso di Laurea Magistrale in Gestione d'Impresa

Cattedra in Analisi Finanziaria

Fallimento d'Impresa: Una Prospettiva sul Ruolo del CEO

RELATORE

Chiar.ma Prof.ssa Magnanelli Barbara Sveva

CANDIDATO

Salvatore Pingitore

Matricola n° 706651

CORRELATORE

Egregio Prof. Paolone Francesco

ANNO ACCADEMICO 2019/2020

Indice dei Contenuti

Introduzione	7
Capitolo 1: Il fallimento d’Impresa	8
1.1 Il fallimento come fisiologico durante la vita dell’impresa	8
1.2 Definizione del Fallimento.....	9
1.3 Cause del fallimento	10
1.4 Conseguenze del Fallimento	12
1.5 Obiettivi della Tesi	13
Capitolo 2: I modelli Predittivi, la Corporate Governance, il CEO	15
2.1 Recap Teorico.....	15
2.1.1 Metodologie e Cronologia	15
2.1.2 Validazione ed Accuratezza.....	16
2.1.3 Indici Ricorrenti	17
2.2 Metodi Statistici Tradizionali.....	19
2.2.1 Beaver, 1966	19
2.2.2 Altman 1968, Z-Score	21
2.2.3 Altman et al. - ZETA ANALYSIS.....	24
2.2.4 Ohlson, 1980.....	25
2.2.5 Altri modelli Statistici	27
2.3 Metodi di Machine Learning ed Intelligenza Artificiale	29
2.3.1 Case-based Reasoning	29
2.3.2 Support Vector Machines	29
2.3.3 Bagging.....	30
2.3.4 Boosting.....	31
2.3.5 Random forest.....	31
2.3.6 Reti Neuronali.....	32
2.3.7 Machine Learning & Artificial Intelligence: Literature Review	33
2.4 Corporate Governance e Fallimento: il ruolo del CEO	35
2.4.1 Il problema Principal-Agent e Principal-Principal.....	35
2.4.2 Definizione di Corporate Governance.....	37

2.4.3 Il ruolo del CEO	39
2.4.4 Corporate Governance e Fallimento: Literature Review.....	41
2.4.5 Considerazioni Finali	43
Capitolo 3: Lo Studio Empirico.....	44
3.1 Formulazione delle ipotesi.....	44
3.2 Descrizione del Sample.....	45
3.3 Profilazione delle Variabili	48
3.4 Metodologia: la Regressione Logistica.....	51
3.5 Risultati della Regressione Logistica.....	54
3.6 Classificazione del modello completo	56
3.7 La potenza predittiva del Modello 1	60
Capitolo 4: Conclusione	62
4.1 Discussione dei risultati	62
4.2 Implicazioni Pratiche	64
4.3 Limitazioni	65
4.4 Spunti per ulteriori ricerche	66
Sommario	68
Bibliografia.....	80

Indice delle Tabelle

Tabella 1: Modelli con accuratezza massima (Fonte: Gissel et Al., 2007).....	17
Tabella 2: Indicatori presentati nello studio di Beaver (Fonte: Beaver,1966)	20
Tabella 3: Risultati ai test di Holdout e Lachenbruch (Fonte: Altman et al., 1977).....	25
Tabella 4: Variabili per il gruppo delle imprese non fallite al tempo “t”	46
Tabella 5: Variabili per il gruppo delle imprese fallite al tempo “t”	46
Tabella 6: Variabili per il gruppo delle imprese non fallite al tempo “t-1”	46
Tabella 7: Variabili per il gruppo delle imprese fallite al tempo “t-1”	47
Tabella 8: Variabili per il gruppo delle imprese non fallite al tempo “t-2”	47
Tabella 9: Variabili per il gruppo delle imprese fallite al tempo “t-2”	47
Tabella 10: Valori del VIF sulle variabili di regressione	49
Tabella 11: Matrice di correlazione per le variabili di regressione	50
Tabella 12: Risultato della regressione logistica sulla variabile CEO-Tenure.....	54
Tabella 13: Risultati della regressione logistica sulla variabile CEO-Duality	55
Tabella 14: Risultati della regressione logistica - Modello 1	56
Tabella 15: Risultati di classificazione della regressione logistica - Modello 1	58
Tabella 16: Risultati della regressione logistica - Modello 2.....	59
Tabella 17: Risultati della classificazione della regressione logistica – Modello 2	59
Tabella 18: Risultati ai test holdout ed anni precedenti	60

Introduzione

Una delle questioni più cruciali che ogni business ben strutturato ed organizzato deve prima o poi fronteggiare è come valutare l'affidabilità creditizia dei partner. Difatti, una parte crescente del totale delle transazioni non ha luogo mediante esborso immediato di liquidità, bensì tramite il meccanismo di crediti e debiti, che intreccia le imprese a tal punto da poter dar luogo ad un cosiddetto "rischio sistemico", o a catena, di fallimenti (pensiamo ad esempio alle banche ed al mercato interbancario dei prestiti).

Ogni qualvolta si dà luogo ad un atto transattivo – luogo giuridico a titolo oneroso - ottenendo in cambio una promessa di pagamento, ci si pone la domanda di "quanto" questa promessa sia affidabile, e "come" valutare l'affidabilità della stessa.

La ricerca delle cause del fallimento e delle sue determinanti è oggetto di numerosi studi che si sono susseguiti dal 1960 in poi. La letteratura è fiorente di articoli scientifici riguardanti i modelli di classificazione e predittivi del fallimento aziendale. Il motivo ricorrente degli articoli è sempre lo stesso: capire quali variabili osservabili e pubblicamente disponibili possano o meno essere considerate predittive della situazione di difficoltà finanziaria, e con che grado di affidabilità nel tempo. Diversi sono i filoni della ricerca, i quali si differenziano per il tipo di variabili esplicative impiegate, e per tecniche statistiche utilizzate.

Seguendo l'approccio dei lavori più accreditati e la tecnica statistica più utilizzata in letteratura, questa tesi si pone come obiettivo finale di dimostrare il nesso tra le qualità del CEO e la probabilità di fallimento, e di valutare la potenza predittiva e la classificazione operata del modello creato.

Il resto dell'elaborato è strutturato come segue: nel Capitolo 1 si darà una panoramica generale del tema, e si discuterà della definizione di fallimento, delle sue cause e conseguenze, e di quali considerazioni hanno fatto sì che si arrivasse allo stabilire l'obiettivo finale della tesi; nel Capitolo 2 verranno illustrati in dettaglio i lavori fondamentali sul tema del fallimento, operando una divisione tra gli stessi usando come criteri le tecniche statistiche utilizzate e le variabili coinvolte; nel Capitolo 3, si discuterà dell'analisi statistica effettuata in questa tesi, e finalmente nel Capitolo 4 si discuteranno i risultati, nonché le implicazioni e delle limitazioni presenti nello studio.

Capitolo 1: Il fallimento d'Impresa

1.1 Il fallimento come fisiologico durante la vita dell'impresa

Ogni impresa nasce, cresce e muore. Questo concetto è a grandi linee espresso dalla teoria del “*Business Life Cycle*”. La teoria, che vede come primo grande contributore Chandler (1962), descrive il processo di vita di un'impresa come *continuum* su quattro fasi distinte. Ad ogni fase sono connesse le relative difficoltà, sfide organizzative e di mercato, ed una certa probabilità di rimanere profittevolmente sul mercato o sopperire alle sue forze.

La prima fase è quella del lancio del prodotto o del servizio, detta di *seed*. L'impresa è appena nata ed è tipicamente di ridotte dimensioni e con una struttura di governance informale. Il problema principale in questa fase è di ottenere clienti e stabilire come fare arrivare loro il prodotto o servizio lanciato. Durante questa fase le vendite sono basse, ma lentamente e stabilmente crescenti. Il focus dell'attività sarà pertanto sul marketing ai segmenti di clientela prescelti, pubblicizzando il proprio vantaggio competitivo e la *value proposition*. A livello finanziario, data la fase iniziale, i costi risulteranno alti e i ricavi bassi, e questo porterà ad una situazione di perdita d'esercizio. In questa fase, date le perdite che si accumulano e la difficoltà a reperire finanziamenti in quanto il business risulta rischioso e immaturo agli occhi dei finanziatori, possono verificarsi problemi di cassa che possono risolversi nel fallimento d'impresa.

Nella seconda fase, detta di crescita, l'impresa, se sopravvissuta alla prima fase, sperimenterà una rapida crescita delle vendite e auspicabilmente avrà profitti positivi. L'impresa è pertanto tentata ad aumentare velocemente la dimensione del business tramite incremento degli investimenti e specularmente del capitale di debito. La decisione, se non ben ponderata o spropositata a livello quantitativo, ha il potenziale di sfociare in difficoltà finanziarie qualora la prevista tendenza di crescita delle vendite (e quindi dei profitti e della cassa funzionale al ripagamento del debito) dovesse invertirsi o non realizzarsi al livello desiderato.

Nella terza fase, ovvero della maturità, le imprese sopravvissute sono caratterizzate dall'aver vendite e profitti stabili o leggermente decrescenti. Il focus principale in questa fase sta nel non perdere quote di mercato, e possibilmente nello sfociare in mercati nuovi e/o con prodotti diversi.

L'ultima fase è quella del declino. Qui le vendite, i profitti ed i flussi di cassa sono stabilmente decrescenti e l'impresa agisce sotto difficoltà finanziarie ed organizzative. Se fallisce nel reinventarsi, l'impresa dovrà uscire dal mercato, con conseguente fallimento o liquidazione.

1.2 Definizione del Fallimento

L'esatta definizione di "*financial distress*" non è ancora ben definita in letteratura (Shi e Li, 2019). In generale, il fallimento viene definito come lo stato in cui un'impresa o un imprenditore individuale, risulta insolvente rispetto le sue obbligazioni alle loro rispettive scadenze.

L'insolvenza può avere diverse gradazioni, passando da difficoltà temporanea ad onorare gli impegni finanziari a scadenza, alla più seria e definitiva bancarotta (Sun, Li, Huang & He, 2014).

A livello operativo, un'impresa è considerata fallita al verificarsi di uno o più dei seguenti eventi: bancarotta, *bond default*, mancato rientro da scoperto bancario, o un mancato pagamento di dividendi su azioni privilegiate (Beaver, 1966).

All'interno del perimetro giuridico italiano, il fallimento si caratterizza come procedura concorsuale a carattere liquidatorio, finalizzata alla soddisfazione dei creditori mediante la liquidazione del patrimonio di pertinenza dell'imprenditore.

In questa sede, adotteremo la definizione di fallimento operativa, che è da ritenersi più calzante ai fini di una valutazione basata su indicatori di bilancio. Il campione delle imprese fallite nell'analisi empirica del capitolo 3 sarà basato sulle imprese che hanno presentato la documentazione per l'apertura delle procedure fallimentari. Non prenderemo in considerazione pertanto la temporanea difficoltà finanziaria (che sarebbe comunque di difficile individuazione mediante qualsivoglia banca dati). Nel

prosiegua i termini bancarotta, fallimento e insolvenza verranno utilizzati come sinonimi.

1.3 Cause del fallimento

Come postulato dalla teoria del *business life cycle*, il fallimento può essere un passaggio fisiologico della vita dell'impresa. Come illustrato nel primo paragrafo di questo capitolo, il passaggio da un'impresa sana ad un'impresa fallita è caratterizzato da una fase intermedia di declino. Damodaran (2009) sintetizza i tratti comuni alle imprese in declino in cinque caratteristiche, specificando che non tutte le imprese in declino possiedono tutte e cinque le caratteristiche ma condividono abbastanza di queste in molti casi, permettendo dunque la generalizzazione. Le cinque caratteristiche sono:

- 1) Vendite stagnanti o decrescenti: la mancata crescita delle vendite o la stagnazione delle stesse su più periodi d'esercizio è probabilmente l'indicatore più eloquente. Vendite stagnanti o che crescono ad un tasso minore rispetto a quello d'inflazione del paese ove l'impresa opera sono segno inequivocabile di inefficienza operativa.
- 2) Margini operativi ristretti o negativi: è conseguente logico della stagnazione delle vendite. In presenza di vendite stagnanti l'impresa perde il potere di fissare prezzi che garantiscono margini accettabili, e tende ad abbassare il prezzo dei prodotti o servizi che offre in modo da prevenire un ulteriore abbassamento delle vendite. Questa combinazione si risolve nel deterioramento o nella negativizzazione dei margini operativi, che l'impresa cercherà occasionalmente di evitare effettuando vendite di assets.
- 3) Disinvestimenti: se gli assets di una impresa in declino sovente hanno un valore specifico maggiore se utilizzati da altre entità, le quali intendono utilizzarli in modo differente e migliore, è ragionevole supporre che i disinvestimenti saranno più frequenti nelle imprese in declino rispetto alle imprese che si trovano in fase di *seed* o di *growth*. Il bisogno di disinvestire sarà maggiore all'aumentare del capitale di debito che l'impresa dovrà ripagare.

- 4) Distribuzione massiccia di dividendi e *buybacks*: le imprese in declino effettuano pochi o nessun investimento che genera valore e generalmente non possiedono assets esistenti che possano generare flussi di cassa positivi. Se il livello del debito dell'impresa è accettabile e non genera preoccupazione che possa essere causa del fallimento, è ipotizzabile che queste imprese non solo paghino ingenti dividendi (per rimanere a galla rispetto gli azionisti), ma che effettuino anche operazioni di riacquisto di azioni proprie.
- 5) Financial Leverage elevato: in presenza di profitti decrescenti e stagnanti derivanti dagli *assets-in-place* e poco potenziale per una futura crescita, molte imprese in declino fanno fronte ad ingenti obbligazioni di debito acquisite durante la fase di crescita. In aggiunta al debito pregresso che diventa insostenibile, vi possono essere problemi addizionali quali la rinegoziazione dei termini per rifinanziare il debito esistente (che generalmente a questo punto risultano più sfavorevoli rispetto ai termini con i quali il capitale è stato acquisito nella fase di crescita).

In aggiunta alle cinque caratteristiche individuate da Damodaran, le quali riguardano la prospettiva interna e prettamente finanziaria delle le imprese in fase di declino, alcuni filoni della letteratura sul tema dividono in cause di fallimento esterne all'impresa, cause interne e manageriali.

Argenti (1976) identifica le cause interne con le seguenti: assenza di una risposta adeguata ad uno *shift* tecnologico, comunicazione tra gli attori interni inadeguata, frode, inadeguata valutazione dei fattori di costo, conoscenza insufficiente a livello finanziario, e alto *leverage*.

Ooghe & Prijcker (2008) delineano un framework concettuale composto da cinque *clusters* di fattori sia interni che esterni (ambiente immediato, ambiente generale, *management*, politiche interne e caratteristiche specifiche dell'impresa) ed associano ad ogni fase di vita dell'impresa diverse determinanti della bancarotta. Secondo gli autori, i quattro casi distinti e le loro determinanti sarebbero:

- 1) Fallimento di una start-up: a) la mancanza di esperienza nell'industry, b) la debolezza del business plan, c) le inapproprietezze manageriali e d) la mancanza di un vantaggio strategico rispetto la concorrenza. La compresenza di questi fattori guida a vendite basse, spese esagerate in attivo fisso e alla sottostima dei

costi operativi. Di conseguenza si verifica un influsso di cassa non sufficiente, problemi di liquidità, aumento del debito e da ultimo la bancarotta (la mancanza di liquidità come conseguenza alle inefficienze, e la seguente bancarotta sono tratti comuni a tutti e quattro i casi).

- 2) Per un'impresa definita "ambiziosamente in crescita", i fattori che contribuiscono al fallimento risultano: a) l'expertise, la persuasività e la propensione al rischio del management, b) la sovrastima delle vendite e c) l'incapacità di reagire a cambiamenti dell'ambiente esterno. La somma di queste inadeguatezze porta a vendite insufficienti, e *overcapacity* degli impianti.
- 3) Il terzo caso individuato dagli autori è quello di un'impresa "*dazzled growth*" che potrebbe tradursi nella crescita dimensionale senza un fine strategico puntuale. Gli autori riconducono al fallimento di questo tipo d'impresa fattori quali: a) la perdita del senso di realismo da parte del management, b) ingenti esborsi di capitale e acquisizioni di altre imprese e c) struttura manageriale inadeguata.
- 4) Il quarto ed ultimo caso è rappresentato da un'impresa "apatica e stagnante". Le determinanti del fallimento di queste imprese sono secondo gli autori: a) una mancanza di impegno e motivazione del management derivante dalla stagnazione, b) la non consapevolezza dei cambiamenti dell'ambiente e c) la perdita del vantaggio strategico.

1.4 Conseguenze del Fallimento

In generale, ogni impresa cercherà a tutti i costi di evitare la situazione fallimentare. Ad ogni modo, il rischio sopportato per chi concede liquidità, è essenzialmente il fallimento del debitore, e dunque la (parziale o totale a seconda della effettiva consistenza patrimoniale del fallito) inesigibilità del credito. Il premio per il rischio, dal suo lato, è di norma rappresentato dalla remunerazione del capitale sotto forma di interessi per rapporti di tipo finanziario, e sconti e abbuoni nei rapporti di tipo operativo.

Il rischio in capo al creditore è tanto più significativo e concreto quanto più quest'ultimo risulta esposto in termini di capitale al rischio di credito delle controparti, e varia a seconda delle modalità con le quali è esposto.

Non tutti i creditori sono infatti uguali: abbiamo creditori assistiti da privilegio e creditori chirografari. I creditori privilegiati sono coloro i quali concedono credito all'impresa ed ottengono in cambio, oltre alla promessa di restituzione del capitale a scadenza congiuntamente ai dovuti interessi, anche una garanzia reale sul credito concesso. In altre parole, i creditori privilegiati sono coloro il quale credito viene soddisfatto per primo nell'atto di liquidazione dei beni aziendali in caso di fallimento, liquidando i beni oggetto di garanzia. I creditori chirografari dall'altro lato, non sono assistiti da garanzie reali, e saranno gli ultimi soddisfatti qualora, dopo la soddisfazione *totale* dei creditori privilegiati, la consistenza dell'attivo lo consenta in tutto o in parte.

Il problema del fallimento è di rilievo non soltanto per gli istituti creditizi¹ (che tipicamente sono creditori privilegiati), i quali se non adottano politiche atte a minimizzare il rischio di credito rischiano di incorrere in gravi problemi quali ad esempio la ricapitalizzazione forzata o il fallimento stesso a loro volta in caso di grossa esposizione verso un singolo o più clienti fortemente intrecciati tra di loro, ma anche per l'impresa stessa e i suoi azionisti o soci, per i lavoratori dipendenti, e più in generale per tutti gli *stakeholders* che hanno come centro di gravità l'impresa in questione.

Da queste considerazioni nasce la necessità di disporre di modelli che riescano a prevedere lo stato di insolvenza per una qualsiasi impresa, possibilmente mediante l'utilizzo e la manipolazione di informazioni di pubblico dominio.

Eppure, non esiste una maniera univoca per esprimere il concetto di rischio di credito, né un set predeterminato di regole che siano strumentali a valutare il rischio di fallimento delle controparti, anche se, come vedremo, la letteratura in merito nonché i modelli predittivi sono abbondanti.

1.5 Obiettivi della Tesi

Diversi sono i modelli predittivi che la letteratura offre al giorno d'oggi. Essi si differenziano principalmente per metodologia di analisi e per le variabili utilizzate.

¹ I quali generalmente per erogare ingenti somme di denaro non richiedono solo garanzie reali, ma anche il rispetto di determinati parametri di bilancio atti a garantire che possano fare fronte agli impegni assunti (covenants).

Abbiamo infatti i primi filoni di letteratura risalenti agli anni '60 con i due articoli di Beaver, che trovano nella analisi degli indici di bilancio una prima approssimazione di ciò che verrà ampiamente studiato e discusso in seguito. Beaver parte dall'analisi di indicatori finanziari costruiti a partire dai bilanci pubblicamente disponibili, contrapponendo le medie dei gruppi tra indicatori di imprese fallite e non fallite, con un sample costruito con la tecnica del *pair-matched sample*.

In seguito, ci si approccerà a modelli statisticamente più complessi quali lo Z-score di Altman (1968), dove viene impiegata la metodologia statistica nota come analisi dei discriminanti multipli (*Multi Discriminant Analysis* o MDA), e nel quale vengono affiancate alle classiche informazioni contabili anche informazioni derivate dal mercato borsistico.

Inoltre, passeremo in rassegna le tecniche più moderne di predizione del fallimento, definite di "Machine Learning & Artificial Intelligence" (ML & AI di seguito), tra i quali si annoverano le *Artificial Neural Network* (ANN), i *Support Vector Machine* (SVM), i *Decision Trees*, il *Boosting* ed il *Bagging*. Questi ultimi metodi si rivelano in media più precisi rispetto ai metodi statistici tradizionali ma hanno la pecca di richiedere una grande mole di dati e grande potenza computazionale.

Da ultimo, vedremo come un ulteriore filone sviluppatosi in seguito, considera le informazioni di *corporate governance* affiancandole a quelle finanziarie-contabili per ottenere una maggiore accuratezza nella classificazione dicotomica impresa fallita – impresa non fallita. Ed è esattamente in questo filone di ricerca lo studio in questa tesi vuole collocarsi.

Scopo di questa tesi di laurea è di indagare con uno studio sperimentale il nesso tra variabili di corporate governance ed economico-finanziarie ed il rischio di fallimento. Nello specifico, delle ipotesi verranno esplicitate e testate per valutare l'incidenza delle variabili descrittive le qualità del CEO sul rischio di bancarotta. Da ultimo si discuterà della potenza classificatoria del modello costruito e della sua forza predittiva con un test di *holdout*.

Capitolo 2: I modelli Predittivi, la Corporate Governance, il CEO

2.1 Recap Teorico

Questa sezione offre un riepilogo a grandi linee dei principali *Research Papers* che si sono susseguiti durante gli anni, le loro metodologie, la loro accuratezza e gli indici utilizzati.

2.1.1 Metodologie e Cronologia

Come accennato in precedenza, esiste una grande varietà di modelli predittivi del fallimento i quali si differenziano per metodologia, per fattori esplicativi impiegati, aree geografiche e di *industries* considerate.

Alcuni modelli sono più focalizzati e meno generalisti di altri, ad esempio: Altman (1968) sviluppa il suo modello per le imprese manifatturiere; Edminster (1972) per le piccole-medio imprese; Sinkey (1975) indaga le variabili cruciali coinvolte nei fallimenti bancari; Wang (2004), sviluppa un modello per le imprese digitali.

Su un piano prettamente cronologico, durante gli anni '60 due soli studi sono stati pubblicati: Beaver (1966) e Altman (1968). L'articolo di Altman avrà grande seguito ed innescherà un incremento degli studi sul tema. Infatti, 28 studi sono stati pubblicati negli anni '70; 53 negli anni '80; 70 studi negli anni '90, 11 dal 2000 al 2004, fino ad arrivare ad $n=496$ lavori pubblicati al 2017 (Gissel et al., 2007). In particolare, gli studi riguardanti il fenomeno della bancarotta hanno visto un incremento dell'83,5% dal 2008 (anno in cui si fa risalire l'inizio della crisi mondiale con il fallimento di Lehman Brothers) al 2017 (Shi & Li, 2019).

A livello metodologico, l'MDA² è stata largamente impiegata dal 1968 in poi. Le regressioni Logit e Probit appaiono dagli anni '70 in poi, ma non sostituiscono l'MDA in popolarità fino alla fine degli anni '80. La principale differenza tra Logit e Probit è che la quest'ultima richiede stime non-lineari (Dimitras et al., 1996). A fine anni '80, appaiono per la prima volta i metodi di Reti Neuronali, i quali sono

² Il focus sulle principali differenze dei metodi statistici verrà effettuato nell'esplicazione dei singoli modelli più importanti.

“disegnati al fine di emulare il comportamento umano nella funzione di ricognizione di *patterns*” (Anandrarajan et al., 2004).

Come indicato nella *Systematic Literature Review* effettuata da Shi e Li (2019), i metodi statistici tradizionali sono largamente più impiegati rispetto ai metodi di Machine Learning & Artificial Intelligence. Abbiamo infatti che il metodo più popolare (impiegato in 123 articoli su 321 analizzati) è la regressione logistica con il 38,3% del totale, al secondo posto troviamo il metodo delle Reti Neurali (17,5%) e da terzo l'analisi dei discriminanti (16,1%).

2.1.2 Validazione ed Accuratezza

Una ulteriore parentesi è necessaria per discutere la validazione dei modelli nonché la loro accuratezza.

I modelli predittivi del fallimento vengono stimati su un campione³ di imprese fallite, alle quali vengono accoppiate le imprese (una o più) non fallite al fine di evidenziare differenze strutturali tra i due gruppi. L'obiettivo è poter utilizzare il modello stimato sul campione originale per fare previsioni su imprese che *non* appartengono a quest'ultimo. A tal fine, si parla di tecniche di validazione, dalle quali dipende e discende l'accuratezza stimata dei modelli.

Come tecnica di validazione, molti studi utilizzano il “*Lachenbruch*” o metodo *Jackknife*, che consiste nel ricalcolare più volte la stessa grandezza statistica stimata lasciando fuori dal campione una osservazione alla volta. Il ricalcolo si effettua ogni volta sul nuovo insieme (ossia il campione privato ogni volta di una diversa osservazione). Quest'ultimo metodo è chiaramente adatto e spesso richiesto se il *sample* è di dimensioni ridotte, anche se una validazione migliore si ottiene mediante la tecnica dell'*hold-out* proposta da Jones (1987). Con questa, il modello viene stimato sul *sample* di “*training*” e testato su un nuovo *sample* che è indipendente da quello utilizzato in fase di calcolo detto di “*test*” o di *holdout*. La tecnica consente quindi di testare l'abilità predittiva del modello in maniera più precisa e sistematica.

³ Sarebbe infatti utopico pensare di poter studiare l'intera popolazione di imprese fallite accostandogli tutte le imprese non fallite, anche se alcuni studi (pur circoscrivendo l'indagine ad un determinato paese) ci vanno molto vicini.

Per quanto concerne l'accuratezza dei vari modelli, la letteratura è solita distinguere tra errori del "Tipo I", ed errori del "Tipo II". Gli errori del Tipo I avvengono quando un modello classifica come non-fallite le imprese fallite, e viceversa l'errore di Tipo II. È generalmente accettato che gli errori di Tipo I sono i più "costosi" per diverse ragioni, tra le quali possiamo citare potenziali citazioni in giudizio (Koh, 1987) o perdita di reputazione dell'impresa. Gli studi che hanno raggiunto la maggiore accuratezza con le tecniche statistiche sono riassunti nella Tabella 1.

Tabella 1: Modelli con accuratezza massima (Fonte: Gissel et Al., 2007)

	<u>Lowest Accuracy</u>	<u>Highest Accuracy</u>	<u>Studies which obtained Highest Accuracy</u>
MDA	32%	100%	Edmister [1972]; Santomero and Vinso [1977]; Marais [1980]; Betts and Belhoul [1982]; El Hennawy and Morris [1983]; Izan [1984]; Takahashi et al. [1984]; Frydman et al. [1985]; Patterson [2001]
Logit analysis	20%	98%	Dambolena and Shulman [1988]
Probit analysis	20%	84%	Skogsvik [1990]
Neural networks	71%	100%	Messier and Hansen [1988]; Guan [1993]; Tsukuda and Baba [1994]; El-Temtamy [1995]

2.1.3 Indici Ricorrenti

Il numero di fattori (o indici indistintamente di seguito) considerati nei modelli varia da 1 a 57. Parlando della frequenza di utilizzo degli indici impiegati nei vari studi, il più popolare (che appare in 54 studi al 2004) è "Utile Netto / Totale Attivo" (*Return on Assets*); il secondo "Attivo Corrente / Passivo Corrente" (*Current Ratio*), che appare in 51 studi. Ad ogni modo, la media degli indicatori considerati in ogni studio varia da 8 a 10 (Gissel et al., 2007).

Nella sezione seguente, passeremo in rassegna i *milestones*, ovvero quegli studi che hanno gettato le fondamenta per test e studi successivi: verranno divisi in a) Metodi Statistici Tradizionali e b) Metodi di Machine Learning & Artificial Intelligence. Un focus verrà infine effettuato sui principali studi che affiancano le tradizionali variabili finanziarie a quelle di corporate governance.

2.2 Metodi Statistici Tradizionali

In questa sezione saranno illustrati in dettaglio i principali lavori sul tema della predizione del fallimento, i quali usano come tecnica di modellazione la statistica classica ed hanno costituito le fondamenta per esplorazioni successive.

2.2.1 Beaver, 1966

Il primo studio di grande rilevanza sul tema della predizione del fallimento si fa risalire a William H. Beaver con il suo articolo “*Financial Ratios As Predictors of Failure (1966)*”. Beaver considera sei macro-gruppi di indicatori costruibili con informazioni di pubblico dominio, descrittivi: a) Flusso Finanziario, b) Utile Netto, c) Debito su Totale Attivo, d) Liquidità su Totale Attivo, e) Liquidità su Passività Correnti ed f) Turnover, per un totale di 30 indicatori. I dati sono estratti dal database COMPUSTAT® andando indietro di cinque anni a partire dalla data del fallimento.

I 30 indicatori sono stati scelti in base a tre criteri:

- 1) Popolarità: da intendersi come frequenza di utilizzo nella letteratura precedente sul tema;
- 2) Performance: gli indicatori avevano già performato bene in studi precedenti, portando ad una discreta accuratezza dei risultati;
- 3) Flusso Finanziario: in precedenza gli indicatori riguardanti il flusso finanziario non erano stati utilizzati nei modelli predittivi.

Nella pagina seguente la tabella riepilogativa dei 30 indicatori selezionati.

Tabella 2: Indicatori presentati nello studio di Beaver (Fonte: Beaver, 1966)

<p>GROUP I (CASH-FLOW RATIOS)</p> <ol style="list-style-type: none"> 1. Cash flow to sales 2. Cash flow to total assets 3. Cash flow to net worth 4. Cash flow to total debt <p>GROUP II (NET-INCOME RATIOS)</p> <ol style="list-style-type: none"> 1. Net income to sales 2. Net income to total assets 3. Net income to net worth 4. Net income to total debt <p>GROUP III (DEBT TO TOTAL-ASSET RATIOS)</p> <ol style="list-style-type: none"> 1. Current liabilities to total assets 2. Long-term liabilities to total assets 3. Current plus long-term liabilities to total assets 4. Current plus long-term plus preferred stock to total assets <p>GROUP IV (LIQUID-ASSET TO TOTAL-ASSET RATIOS)</p> <ol style="list-style-type: none"> 1. Cash to total assets 2. Quick assets to total assets 3. Current assets to total assets 4. Working capital to total assets 	<p>GROUP V (LIQUID-ASSET TO CURRENT DEBT RATIOS)</p> <ol style="list-style-type: none"> 1. Cash to current liabilities 2. Quick assets to current liabilities 3. Current ratio (current assets to current liabilities) <p>GROUP VI (TURNOVER RATIOS)</p> <ol style="list-style-type: none"> 1. Cash to sales 2. Accounts receivable to sales 3. Inventory to sales 4. Quick assets to sales 5. Current assets to sales 6. Working capital to sales 7. Net worth to sales 8. Total assets to sales 9. Cash interval (cash to fund expenditures for operations) 10. Defensive interval (defensive assets to fund expenditures for operations) 11. No-credit interval (defensive assets minus current liabilities to fund expenditures for operations)
--	---

Nel suo studio, il campione è composto da 79 imprese fallite, alle quali accoppia 79 imprese non fallite seguendo come criterio di accoppiamento il valore totale dell'attivo e l'appartenenza ad un determinato segmento di industria identificato dal codice SIC (*Standard Industry Classification*), definito dal dipartimento di commercio statunitense.

Come giustificazione teorica all'uso degli indicatori, Beaver propone una visione d'impresa come "riserva di liquidità", la quale può essere "drenata" o "pompata".

A parità di condizioni quindi, per una qualsiasi impresa, ci si dovrebbe aspettare che:

- 1) Maggiore è la liquidità, minore è la probabilità di fallimento;
- 2) Maggiore è il flusso finanziario, minore è la probabilità di fallimento;
- 3) Maggiore è l'ammontare del debito, maggiore è la probabilità di fallimento;
- 4) Maggiore è l'assorbimento di cassa a supporto delle *current operations*, maggiore è la probabilità di fallimento.

I risultati comparativi delle medie tra i due gruppi per ogni singolo indicatore vanno esattamente nella direzione ipotizzata da Beaver, confermando così le 4 ipotesi sopracitate.

Segue un test di classificazione dicotomica tra imprese fallite e non fallite. Al fine di fare predizioni ed inferenza, ogni indicatore viene ordinato in un *array*, tracciando una linea visuale di *cut-off* tale che la somma degli errori di classificazione viene minimizzata⁴. Ricordiamo che gli errori possibili sono del “Tipo I” se un’impresa fallita viene classificata come non fallita, e del “Tipo II” se un’impresa non fallita viene classificata come fallita. Se un’impresa si trova sotto il punto di *cut-off* verrà classificata come fallita, altrimenti come non fallita.

Beaver trova che l’indice che meglio discrimina i due gruppi risulta essere “Flusso Finanziario/ Totale Debito”, che classifica correttamente l’87% delle imprese del campione all’anno precedente il fallimento. L’indice second-best risulta essere “Utile Netto / Totale Attivo”, anch’esso con un’accuratezza dell’87% ad un anno dal fallimento.

2.2.2 Altman 1968, Z-Score

Edward Altman, con il suo articolo “*Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy*” è il primo ad utilizzare la statistica multivariata, precisamente l’analisi dei discriminanti, a fini predittivi del fallimento.

La Multi Discriminant Analysis (MDA) è una tecnica statistica multivariata che consente di classificare a priori in uno o più gruppi mutualmente esclusivi un set di osservazioni con determinate caratteristiche. È una metodologia largamente impiegata quando la variabile dipendente oggetto di studio (nel nostro caso fallimento-non fallimento) appare in forma non quantitativa ma qualitativa. Dopo che i dati per ogni gruppo sono raccolti, la MDA cerca di derivare la combinazione lineare degli indicatori considerati che meglio discrimina tra i gruppi, assegnando ad ogni indicatore un coefficiente discriminante. La MDA ha il vantaggio di considerare

⁴ Si noti che lo stesso Beaver chiarisce che l’individuazione del punto di *cut-off* segue un processo del tipo *trial – error*.

simultaneamente l'intero profilo di variabili comuni ai gruppi in oggetto, considerando anche le interazioni tra queste (Altman 1968).

Nel suo studio, Altman considera 66 imprese in totale, 33 fallite (imprese che hanno invocato il Chapter X del Bankruptcy National Act) e 33 non fallite, stratificando il campione di non fallite usando come criteri il segmento industriale e la dimensione dell'attivo.

Dopo aver compilato e testato una lista di 22 indicatori, la funzione che meglio discrimina i due gruppi, nota come Z-score, assegna un punteggio da -4 a +8 derivante dagli indicatori e dai coefficienti discriminanti, e risulta la seguente:

$$Z - Score = .012 X1 + .014 X2 + .033 X3 + .006 X4 + .999 X5,$$

Dove:

$X1 = \text{Capitale Circolante} / \text{Totale Attivo}$, viene assunto come misura della liquidità netta in rapporto al totale dell'attivo;

$X2 = \text{Utile Portato a Nuovo} / \text{Totale Attivo}$, viene assunto come misura della profittabilità cumulata, e quindi anche come proxy dell'età dell'impresa;

$X3 = \text{EBIT} / \text{Totale Attivo}$, è assunto come indicatore della produttività aziendale, astruendo così il risultato da fattori esogeni e non, quali tasse e leverage;

$X4 = \text{Valore di mercato dell'Equity} / \text{Valore contabile del debito}$ indica di quanto il valore totale dell'attivo può diminuire prima di superare il valore totale del debito facendo precipitare l'impresa in insolvenza⁵;

$X5 = \text{Vendite} / \text{Totale attivo}$ è misura dell'abilità dell'impresa di generare vendite sufficienti sfruttando l'attivo.

Per testare la potenza discriminante individuale delle cinque variabili della funzione Z, Altman esegue un *F-Test*. Il test mette in relazione a livello univariato le differenze tra i valori medi delle variabili in ogni gruppo alla variabilità dei valori tra i gruppi.

⁵ In tal caso infatti, data l'equazione contabile Attivo = Passivo + Patrimonio Netto, avremmo che il patrimonio netto avrebbe valore 0 o negativo.

Altman trova che tutte le variabili sono significative al livello dell'1%, eccetto X5 la quale non risulta particolarmente significativa.

Al fine di ottenere la profilazione finale delle variabili e considerare l'interazione tra queste, Altman crea un vettore scalato, moltiplicando ogni variabile per la radice quadrata degli elementi diagonali della matrice varianze-covarianze. Sorprendentemente, gli indici che più contribuiscono alla separazione tra i due gruppi sono X3, X4 ed X5⁶. Una probabile spiegazione della significatività su base multivariata di X5 è data dalla forte correlazione negativa (-0.78) che si osserva tra X3 e X5 nel gruppo delle imprese fallite. La logica retrostante la forte correlazione negativa tra X3 ed X5 è che “al soffrire di perdite d'esercizio e deteriorando verso il fallimento, gli *assets* non sono rimpiazzati con la frequenza dei tempi sani e le perdite cumulate riducono ulteriormente l'*asset-size*” (Altman 1968).

Il modello viene testato due volte su due sample diversi: il primo sample è composto da 33 imprese fallite e da 33 non fallite. Il modello raggiunge un'accuratezza del 95% su entrambi i gruppi ad un anno dal fallimento. Il secondo sample è composto da 25 imprese fallite e 66 non fallite. Validando tramite la procedura dell'hold-out, il modello classifica correttamente il 96% delle imprese fallite (24 su 25), ed il 79% delle imprese non fallite.

A fini di inferenza, un'impresa rientrerà tra le fallite se avrà un punteggio inferiore a 1,81, mentre rientrerà tra le non fallite se otterrà un punteggio maggiore di 2,99. L'intervallo intercorrente tra questi due valori è definito da Altman come “*Zone of Ignorance*”, ovvero l'intervallo entro il quale ricadono gli errori di classificazione.

L'articolo ha avuto largo seguito e la funzione Z è tutt'oggi utilizzata come *proxy* del fallimento, anche affiancata anche da altri modelli ben più complessi.

⁶ Ricordiamo che X5 risultava non significativo a livello univariato.

2.2.3 Altman et al. - ZETA ANALYSIS

Nel 1977, Altman pubblica un nuovo Research paper dal titolo “ZETA ANALYSIS: A new model to identify Bankruptcy risk of corporations”, con lo scopo di “costruire, analizzare e testare un nuovo modello di predizione del fallimento che consideri esplicitamente nuovi sviluppi relativi ai fallimenti aziendali (Altman 1977)”.

Le ragioni che spingono Altman a rivisitare l’originale Z-score, sono le seguenti:

- 1) Dimensione aziendale delle imprese fallite in aumento rispetto all’articolo del 1968; nell’originale Z-score l’*asset size* medio del campione è pari a USD 20 milioni, contro gli USD 100 milioni del campione nell’articolo ZETA Analysis;
- 2) Adattare il modello agli anni più recenti alla data: l’articolo ZETA Analysis copre gli anni 1969-1975;
- 3) I modelli precedenti si erano concentrati o su una larga definizione di industria intesa come “manifattura generica”, o su industrie specifiche, da qui la necessità di creare un modello adatto anche alle *private companies*.
- 4) Considerare l’adattamento dei *financial statements* ai nuovi standard contabili.

Dopo un processo iterativo su 27 variabili, gli autori optano per un modello a 7 variabili, le quali non solo si sono dimostrate di grande affidabilità nel discriminare i due gruppi, ma sono state anche le più affidabili nei processi di validazione del modello.

Le sette variabili incluse nel modello sono le seguenti:

- 1) $X1 = \text{EBIT} / \text{Totale Attivo}$, la quale si è già dimostrata di grande rilievo in diversi studi precedenti (Altman 1968, Beaver 1967);
- 2) $X2 = \text{Stabilità degli utili}$: misurata come l’errore standard normalizzato attorno alla tendenza media di 10 anni in $X1$;
- 3) $X3 = \text{Indice di copertura del debito}$ (i.e. $\text{EBIT} / \text{Interessi Passivi}$);
- 4) $X4 = \text{Profittabilità cumulata}$, misurata come $\text{Utili a Nuovo} / \text{Totale Attivo}$
- 5) $X5 = \text{Attivo Corrente} / \text{Passivo Corrente}$

- 6) X6 = Azioni ordinarie / Valore di mercato dell'Equity: misurate entrambe come media dei prezzi di scambio delle azioni ordinarie e totali nei 5 anni precedenti;
- 7) X7 = Dimensione aziendale, misurata come Totale Attivo;

Gli autori propongono due tipi di regressioni, la quadratica e la lineare, e testano il modello tramite *hold-out* e tramite Lachenbruch. I risultati sono esposti nella Tabella 3.

Tabella 3: Risultati ai test di Holdout e Lachenbruch (Fonte: Altman et al., 1977)

Overall classification accuracy (in percent).

Years prior to bankruptcy	Bankrupt firms		Non-bankrupt firms		Total	
	Linear	Quadratic	Linear	Quadratic	Linear	Quadratic
1 Original sample	96.2	94.3	89.7	91.4	92.8	92.8
1 (Lachenbruch validation test)	(92.5)	(85.0)	(89.7)	(87.9)	(91.0)	(86.5)
2 Holdout	84.9	77.4	93.1	91.9	89.0	84.7
3 Holdout	74.5	62.7	91.4	92.1	83.5	78.9
4 Holdout	68.1	57.4	89.5	87.8	79.8	74.0
5 Holdout	69.8	46.5	82.1	87.5	76.8	69.7

Il modello è caratterizzato da un'accuratezza sostanzialmente paragonabile allo Z-score originale nell'anno 1, ma maggiore accuratezza degli anni 2, 3, 4 e 5 dall'evento bancarotta.

2.2.4 Ohlson, 1980

Il lavoro di James A. Ohlson si contrappone per metodologia statistica a quello di Altman (1968). Esso infatti muove tre specifiche critiche alla MDA, ovvero: 1) vi sono alcuni requisiti statistici che devono sussistere ai fini di una buona analisi⁷; 2) l'output risultante da un'analisi MDA è un punteggio che ha poca applicabilità pratica in quanto distribuisce in modo ordinale le osservazioni considerate; 3) potrebbero presentarsi problemi di arbitrarietà dei criteri in fase di accoppiamento delle imprese non fallite.

⁷ Per esempio, le variabili indipendenti dovrebbero seguire la distribuzione normale.

Di contro, Ohlson propone l'analisi Logit come alternativa ai problemi sopracitati, riducendo l'analisi essenzialmente alla seguente domanda: dato un set predeterminato di imprese, qual è la probabilità che ognuna di queste fallisca in un dato periodo di tempo?

Così facendo, tramite l'analisi Logit si stima una probabilità che l'impresa fallisca, rendendo il risultato facilmente interpretabile.

Ohlson identifica quattro dimensioni statisticamente significative nell'influenzare la probabilità di fallimento, ovvero:

- 1) La dimensione dell'impresa;
- 2) Una o più misure della struttura finanziaria;
- 3) Una o più misure di performance;
- 4) Una o più misure di liquidità.

Gli indici utilizzati da Ohlson, 9 in totale, sono i seguenti:

- 1) *SIZE*: definita come logaritmo di Totale Attivo / Prodotto Nazionale Lordo;
- 2) *TLTA*: Debito totale / Totale Attivo;
- 3) *WCTA*: Capitale Circolante / Totale Attivo;
- 4) *CLCA*: Passività correnti / Attivo Corrente;
- 5) *OENEG*: 1 se il debito totale eccede il totale dell'attivo, 0 altrimenti;
- 6) *NITA*: Reddito Netto / Totale Attivo;
- 7) *FUTL*: Flusso finanziario / Totale Debito;
- 8) *INTWO*: 1 se l'utile netto è negativo negli ultimi due anni di osservazione, 0 altrimenti;
- 9) *CHIN*: $(UN_t - UN_{t-1}) / (|UN_t| + |UN_{t-1}|)$, dove "UN t" è l'Utile netto dell'ultimo periodo di osservazione. La variabile è atta a misurare la crescita o decrescita nell'utile netto.

Il campione considerato da Ohlson comprende 2163 imprese nel periodo 1970-1976, di cui 105 fallite e 2058 in attività alla data dello studio.

La funzione che meglio si adatta allo scopo di separazione dei due gruppi, detta O-score e derivante dal modello Logit, risulta essere:

$$O - score = -1.3 - 0.4 X1 + 6.0 X2 - 1.4 X3 + 0.8 X4 - 2.4 X5 - 1.8 X6 \\ + 0.3 X7 - 1.7 X8 - 0.5 X9$$

Per costruzione, la funzione ricerca la *log-odd* della probabilità fallimento, e dunque maggiore è l'O-score, maggiore è la probabilità di fallimento.

Ohlson trova che per minimizzare la somma degli errori di Tipo I e di Tipo II, il punto ottimale di cut-off è $P = 0.038$.

Il modello discrimina correttamente il 96,12% delle osservazioni ad un anno dall'evento fallimento, ed il 95,55% a due anni dallo stesso.

2.2.5 Altri modelli Statistici

Zmijewski (1984) utilizza la regressione Probit. Il suo modello si basa sullo studio di 40 imprese fallite e 800 non fallite. Stima un'accuratezza del 99%, ma non riporta i risultati derivanti dal test di validazione; utilizza le seguenti tre variabili: 1) "Utile Netto / Totale Attivo", 2) "Totale Debito / Totale Attivo", e 3) "Attivo Corrente / Passivo Corrente".

Edminster (1972) costruisce un modello a 7 variabili pensato per prevedere il fallimento delle piccole e medie imprese tramite analisi MDA; ottiene un'accuratezza del 93% ad un anno dall'evento.

Taffler (1984) crea un modello basato su sullo Z-score (Altman, 1968) per le imprese britanniche. Il sample delle fallite consiste in 46 imprese manifatturiere quotate al London Stock Exchange, accoppiate con il criterio una impresa fallita-una impresa non fallita basandosi su *industry* e *size*. Il modello ottiene un'accuratezza del 97% ad un anno dal fallimento.

Cultrera & Brédart (2015), si focalizzano sulle PMI belghe usando il modello Logit includendo variabili di controllo quali dimensione ed età delle imprese. Essi concludono che includere gli indici di profittabilità e liquidità migliora l'affidabilità del modello.

Alaminos et al. (2016), sviluppano un modello “globale” introducendo variabili dummy descrittive la regione geografica. Gli autori concludono che un modello globale è possibile ed il loro modello arriva ad un’accuratezza oscillante tra l’89% e il 92% ad un anno dalla bancarotta a seconda della regione geografica considerata.

Gli studi fin qui esposti rappresentano solo una piccola frazione della letteratura, ed utilizzano solo tecniche statistiche. Si può comunque concludere che i *ratios* ottenibili dai bilanci combinati alle tecniche statistiche di Logit, Probit ed MDA consentono di ottenere un quadro chiaro della probabilità di fallimento di un’impresa, e con un ragionevole grado di affidabilità nel tempo.

Nella prossima sezione si passeranno in rassegna i principali lavori che utilizzano come tecniche Machine Learning e Artificial Intelligence.

2.3 Metodi di Machine Learning ed Intelligenza Artificiale

I metodi di machine learning ed intelligenza artificiale (ML & AI) sono considerati tra i più importanti avanzamenti in matematica applicata, con significanti implicazioni nella classificazione dei problemi (Tian, Shi & Liu, 2012). Questa sezione passerà in rassegna le principali e più moderne tecniche di predizione del fallimento tramite modelli basati su Machine Learning e Artificial Intelligence. Seguirà una discussione sui principali riscontri ottenuti.

2.3.1 Case-based Reasoning

Il Case-based Reasoning (CBR) è concettualmente simile ai processi cognitivi che la mente umana segue nel *problem-solving* (L. Kolodner, 1993). Esso rientra tra i metodi di intelligenza artificiale.

È un processo di risoluzione di nuovi problemi basato sull'applicazione delle soluzioni a problemi simili risolti precedentemente. Il CBR è stato formalizzato come processo diviso in quattro fasi (Aamodt e Plaza, 1994):

- 1) Recupero: dato un problema, si recupera la memoria sui casi rilevanti per risolverlo;
- 2) Riutilizzo: si mappa la soluzione del caso precedente al problema attuale;
- 3) Revisione: dopo aver mappato la soluzione precedente al caso attuale, si prova la nuova soluzione e se necessario la si adatta e rivaluta;
- 4) Conservazione: dopo che la soluzione è stata adattata con successo al problema attuale, la si memorizza come nuovo caso.

2.3.2 Support Vector Machines

Il *Support Vector Machine* (SVM), è un modello di *machine learning* associato ad algoritmi di apprendimento per la regressione e la classificazione. Dato un *training-set* contenente i dati per l'addestramento della funzione e diviso in due *sets* mutualmente esclusivi, un modello SVM assegna i nuovi esempi (*test-set*) ad una delle

due classi. Più precisamente, l'ottimizzazione del modello SVM è basata sulla trasformazione di una funzione matematica mediante un'altra funzione denominata *kernel*, tramite la quale si identificano le maggiori distanze tra le osservazioni più simili che sono classificate come opposte (Noble, 2006). Risulta quindi nella rappresentazione degli esempi come punti in uno spazio, mappati in modo tale che le due categorie sono divise da un divario più ampio possibile. I nuovi esempi saranno quindi re-mappati sullo stesso piano, e cadranno nell'una o nell'altra categoria a secondo del loro posizionamento nel piano.

2.3.3 Bagging

La tecnica del *Bagging* (da **B**ootstrap **agg**regating) è una tecnica di machine learning proposta da Breiman (1994). Il *bagging* coinvolge classificatori indipendenti che usano porzioni del *training-set* mediante modellazione delle medie, portando al risultato più efficiente di classificazione. Il *bagging* crea nuovi *subsets* mediante campionamento *random* dal *dataset* originale, generando una stima di intervallo di confidenza (Frigini, Savona & Vezzoli, 2016).

L'obiettivo del bagging è di ridurre l'*overfitting*⁸ di una classe su un modello. Invece che utilizzare il *test-set* per controllare se il modello stimato è *overfitting*, tramite il *bagging* il *training-set* è ricombinato n volte al fine di produrre una classificazione migliore (Barboza, 2017).

⁸ L'*overfitting*, in statistica, è quel fenomeno per cui la produzione di un'analisi (e quindi del correlato modello) corrisponde troppo esattamente ad un particolare set di dati, e per questo il modello stimato fallisce quando riapplicato ad osservazioni che non rientrano nel *training-set*.

2.3.4 Boosting

Il *boosting* è una tecnica di machine learning che consiste nel ripetuto utilizzo di una base di regole predittive o funzioni su diversi sottoinsiemi del *training-set*. La tecnica assegna un peso diverso ad ogni caratteristica (o classificatore) del *subset* del *training-set*.

Il *boosting* è basato sulla questione posta da Kearns e Valiant (1988, 1989), ovvero: “può un set di classificatori deboli⁹ creare un singolo classificatore forte?”. Numerosi sono gli algoritmi di *boosting* sviluppati dalla comunità scientifica fin oggi, ma il più adatto alla classificazione binaria è *AdaBoost*. L’algoritmo generale è descritto come segue:

- 1) Formare un largo set di caratteristiche (ovvero classificatori deboli, o variabili);
- 2) Assegnare come peso iniziale per ogni m -esima caratteristica $1/m$;
- 3) Per T volte: a) si addestra un classificatore alla volta e si valuta l’errore di classificazione; b) si sceglie il classificatore con il più piccolo margine d’errore; c) si ripesa il classificatore debole a seconda del margine d’errore;
- 4) Si forma il classificatore finale “forte” descrivendolo come combinazione lineare dei T classificatori deboli.

Dopo aver applicato l’algoritmo, un classificatore costruito da 200 caratteristiche può ottenere il 95% di classificazione corretta (Viola & Jones, 2001).

Il metodo non è di facile applicazione in quanto è necessario disporre di una grande mole di dati e potenza computazionale al fine di identificare e descrivere i classificatori.

2.3.5 Random forest

Il *Random Forest* è una tecnica di machine learning che rientra nella più ampia branca dei *Decision Trees*. I *decision trees* sono dei grafici ad albero rappresentanti un *set* di decisioni e le loro possibili conseguenze.

⁹ Un classificatore debole è definito come un classificatore che è solo leggermente correlato con la vera classificazione.

Nel machine learning gli alberi decisionali vengono utilizzati a fini predittivi e di classificazione. Si passa da un *set* di osservazioni su un elemento, solitamente rappresentato nei rami (e.g. indici di bilancio per un'azienda), al “*target value*” dell'elemento stesso (rappresentato nelle foglie). Se la variabile target è discreta, si parla di *classification trees*.

Con la tecnica del *random forest*, proposta per la prima volta da Tin Kam (1995), si costruiscono una moltitudine di *decision trees* (ogni *decision tree* rappresenta un diverso modello con diverse caratteristiche e variabili) e si ottiene come *output* un unico modello che ha la caratteristica di essere molto più accurato e preciso dei singoli modelli che lo compongono.

La ragione per cui il *random forest* è più preciso dei singoli *trees* è che componendo un singolo modello come combinazione di n modelli, il margine di errore è notevolmente ridotto in quanto se alcuni *trees* possono classificare male i gruppi, altri *trees* arriveranno alla soluzione corretta. Il prerequisito fondamentale affinché il *random forest* funzioni è che i modelli che lo compongono devono essere debolmente correlati o ancora meglio, non correlati.

2.3.6 Reti Neuronal

Le reti neurali artificiali (ANN) rientrano tra le tecniche di *Artificial Intelligence*. Esse rappresentano un tentativo di simulare tramite l'utilizzo di *computers* la rete neurale umana con le sue interconnessioni tra neuroni, al fine di rendere la macchina capace di imparare e fare decisioni come farebbe un cervello umano. Si compongono di algoritmi di auto-apprendimento.

Le ANN vengono addestrate dagli esempi che le si danno come *input* e dai risultati che a questi corrispondono, e formano un modello probabilistico che connette *input* e *output*. L'addestramento della rete è tipicamente condotto determinando la differenza tra l'output processato dalla rete stessa e il target output desiderato, valutando il margine di errore. Tramite una serie di successivi aggiustamenti l'*output* prodotto dalla

rete per un dato *input* sarà sempre più vicino al valore desiderato. Una volta raggiunto un buon livello di accuratezza, la rete è in grado di identificare i nuovi esempi restituendo il *target output*.

Tra i principali problemi di applicabilità pratica delle ANN, tre in particolare vengono segnalati da Zhao et al. (2014): 1) la performance della rete risente di dati non bilanciati¹⁰, 2) l'accuratezza aumenta all'aumentare delle osservazioni che si danno in pasto alla rete¹¹ e 3) all'aumentare dei livelli aumenta la precisione e contestualmente il tempo computazionale richiesto per l'output.

2.3.7 Machine Learning & Artificial Intelligence: Literature Review

Diversi sono gli studi che si sono susseguiti utilizzando tecniche di ML & AI. In questo paragrafo passeremo in rassegna alcuni di questi, evidenziando i *findings* e le tecniche utilizzate.

Kumar & Ravi (2007), effettuano una *systematic literature review* sui *Research Papers* pubblicati dal 1968 al 2005, passando in rassegna tecniche statistiche e di ML & AI. Essi concludono che le tecniche c.d. “*ensambled*¹²” sovraperformano consistentemente i classificatori singoli.

Tsai et al. (2014) considerano tre diversi *datasets* per tre differenti paesi: Australia, Germania e Giappone. I tre *datasets* si compongono rispettivamente di 690, 1000 e 690 imprese, 307/383, 700/300 e 307/383 (fallite/non fallite). Come tecnica di validazione utilizzano la “*10-fold cross-validation*”, che consiste nel dividere ogni dataset in 10 parti uguali e addestrare e testare il modello su ognuna di queste 10 parti. Gli autori propongono diverse tecniche di modellazione, tra cui *SVM-bagging*, *SVM-boosting*, e *decision tree*. Concludono che il miglior metodo è il *decision tree*, che porta il loro modello ad un'accuratezza dell'86%.

¹⁰ Per esempio, nel nostro caso l'output dovrebbe consistere nel classificare le imprese come fallite o non fallite. Un importante sbilanciamento della numerosità delle osservazioni a favore di uno dei due gruppi nel *training-set* riduce sensibilmente l'affidabilità del modello nel *test-set*.

¹¹ È esattamente per questo motivo che le ANN richiedono un'enorme mole di dati al fine di un corretto funzionamento.

¹² Ovvero quelle tecniche derivanti dalla combinazione di singoli classificatori (boosting, bagging, random forest etc.)

Heo & Yang (2014), utilizzano un campione di 2762 imprese coreane (metà fallite e metà non fallite), costruiscono un modello sfruttando l'algoritmo di *boosting* AdaBoost, e lo paragonano ad ANN, SVM, *Decision trees*, e lo Z-Score di Altman. Gli autori concludono che il metodo più preciso in assoluto risulta AdaBoost, che classifica correttamente il 93,8% del campione totale.

Danenas & Garsva (2015), trovano che SVM sovraperforma i metodi Logit e ANN, con un'accuratezza media del 93% su un campione di 21847 imprese statunitensi comprendente in 9 anni di osservazione.

Barboza, Kimura & Altman (2017), effettuano uno studio paragonando modelli utilizzanti SVM (con due diversi *kernel*), Bagging, Boosting, Random Forest, ANN ed i classici MDA e Logit. Raccolgono dati per aziende statunitensi e canadesi dal 1985 al 2013 utilizzando il database COMPUSTAT®. Compongono il *training-set* con 449 imprese fallite e lo stesso numero di non fallite. Per il *test-set*, includono 133 imprese fallite e 13300 imprese attive. Gli autori evidenziano che le tecniche di machine learning sovraperformano consistentemente le tecniche statistiche (MDA e Logit) nel *test-set*, con Boosting e Random Forest rispettivamente al primo e secondo posto con un'accuratezza dell'86,31% e 86,04%.

Come si può notare dai risultati esposti sopra, le tecniche di ML & AI vengono largamente impiegate negli studi recenti, portando ad un alto livello di affidabilità, anche se non c'è unanimità nella comunità scientifica su quale sia la migliore tecnica da impiegare in ogni contesto. Altrettanto desumibile è che queste ultime richiedono una grande mole di dati per funzionare a dovere, e possono soffrire di diversi *bias* in fase di costruzione del *sample* (come d'altronde le tecniche statistiche). Quando però il *sample* è "ben costruito" e sufficientemente ampio, queste tecniche sovraperformano e sono mediamente più accurate dei loro parenti statistici.

2.4 Corporate Governance e Fallimento: il ruolo del CEO

In questa sezione, si darà un quadro generale della Corporate Governance, e verranno passati in rassegna i principali lavori che considerano le variabili di Corporate Governance come migliorative dei modelli predittivi del fallimento classici, i quali, come abbiamo visto, impiegano prevalentemente informazioni economico-finanziarie di pubblico dominio. Un focus verrà effettuato sul ruolo dello *Chief Executive Officer* (CEO) all'interno dell'impresa.

2.4.1 Il problema Principal-Agent e Principal-Principal

Non tutte le imprese sono uguali, non tutte hanno la stessa dimensione. Nelle piccole e medie imprese, vi è la tendenza alla presenza della figura dell'imprenditore, inteso come proprietario ed amministratore dell'impresa. Nelle grandi imprese (ancor di più nel caso delle imprese quotate) vi è solitamente la separazione tra proprietà e *management*. Questa separazione può far emergere dei conflitti di interesse tra azionisti ed amministratori (conflitto *principal-agent*), e tra azionisti di maggioranza ed azionisti di minoranza (conflitto *principal-principal*).

La teoria che per prima studia il conflitto di interesse all'interno delle organizzazioni è nota come teoria dell'agenzia (Jensen & Meckling, 1976). Essa si occupa di studiare i conflitti d'interesse che si verificano tra parti contraenti che hanno interessi opposti. Per ogni contratto infatti, vi sono sempre almeno due parti opposte. Gli agenti contraenti sono supposti essere razionali, nel senso che puntano a massimizzare i loro *payoffs* individuali. Gli azionisti vorranno massimizzare la loro ricchezza, mentre i *managers* saranno più interessati ad accrescere i loro introiti e la loro reputazione.

Il conflitto tra azionisti e *managers* concerne per lo più aziende caratterizzate da una grande dimensione e ad azionariato frammentato. In queste organizzazioni, il singolo azionista possiede di norma una porzione infinitesima delle azioni totali, e dunque scarso potere di controllo sugli amministratori e sugli organi delegati, i quali potrebbero "distrarre" risorse di pertinenza del patrimonio aziendale a fini privati. Classici esempi di comportamenti adottati dal *management* a danno degli *shareholders* sono rinvenibili nella vendita di *assets* a prezzi non di mercato, nell'utilizzo di beni aziendali a fini personali, e nel non corretto utilizzo dei flussi di cassa al fine di aumentare la propria

retribuzione ed accrescere la dimensione aziendale. Quest'ultimo esempio può verificarsi nel riacquisto di azioni proprie: una parte consistente della retribuzione del *top management* è costituito dalla parte variabile, la quale è spesso legata a dei parametri di performance quali l'aumento del prezzo delle azioni; in presenza di consistenti e cospicui flussi di cassa, il *management* potrebbe essere tentato ad effettuare un riacquisto delle azioni in circolazione al fine di aumentare il prezzo delle stesse, e quindi incassare un bonus maggiore.

Il secondo conflitto d'interesse che si verifica all'interno delle imprese è tra azionisti di maggioranza ed azionisti di minoranza (conflitto *principal-principal*). Il potere degli azionisti di maggioranza nel contesto della grande impresa è considerevole. Nella situazione di presenza contemporanea di azionariato diffuso e di un singolo azionista che detiene il pacchetto di controllo, il potere del singolo azionista appartenente all'insieme dell'azionariato diffuso è quasi nullo. Nel contesto assembleare quindi, il *management* sarà nominato essenzialmente dall'azionista di controllo, e sarà a questo che risponderà. Questa circostanza crea una situazione sfavorevole in capo agli azionisti di minoranza, i quali potranno vedersi indirettamente "espropriati" del valore aziendale grazie a decisioni che favoriscono l'interesse dell'azionista di controllo a danno dell'interesse collettivo societario.

I conflitti di agenzia, se non correttamente gestiti, comportano quindi la incorretta allocazione delle risorse di cui l'impresa dispone e pongono la base per la riduzione o la stagnazione del valore aziendale nel lungo periodo. Al fine di mitigare i problemi di agenzia all'interno delle organizzazioni, un contributo fondamentale è dato dallo schema e dalla struttura della *Corporate Governance*, e dalla vigilanza sull'osservanza del mantenimento dei principi che da essa discendono.

2.4.2 Definizione di Corporate Governance

La *Corporate Governance* è definita come quell'insieme di regole, processi ed istituzioni che governano e dirigono la vita delle imprese. La struttura della *governance* identifica e distribuisce obblighi, doveri e responsabilità tra i partecipanti dell'impresa, ed include un corpo di regole codificate funzionali al prendere decisioni di *business*. La *corporate governance* include i processi tramite i quali le imprese stabiliscono e perseguono i loro obiettivi nel contesto sociale, regolatorio e competitivo.

In genere, la struttura della *corporate governance* implica la presenza di tre organi distinti, i quali in ragione dell'osservanza del sistema giuridico entro il quale agisce l'impresa assumono nomi e ruoli differenti. Questi tre organi sono:

- 1) **Assemblea dei soci:** è l'organo in cui si riuniscono i proprietari od azionisti dell'impresa ed ha generalmente il compito di approvare gli atti più importanti (ad esempio fusioni, acquisizioni, approvazione del bilancio), e di nominare il Consiglio di Amministrazione e il Collegio Sindacale.
- 2) **Consiglio di Amministrazione (o *Board of Directors*):** è l'organo amministrativo dell'impresa. All'interno di esso si riuniscono gli amministratori nominati dall'Assemblea. Essi sono deputati della gestione esclusiva dell'impresa e sono incaricati di stabilire e perseguire l'oggetto sociale. Il Consiglio di Amministrazione può delegare una o più delle sue attribuzioni ai così detti organi delegati (CEO, CFO, COO e così via). Gli organi delegati agiscono in nome e per conto del consiglio di amministrazione, e a questo fanno rapporto. Nelle grandi imprese, di fatto il potere decisionale e gestionale è per lo più in mano agli organi delegati, ed il ruolo del consiglio di amministrazione è spesso ridotto ad un mero potere di sorveglianza sugli organi delegati.
- 3) **Collegio sindacale od Organo di Sorveglianza:** è l'organo deputato alla sorveglianza sui consiglieri di amministrazione. Il compito del consiglio è di assicurarsi che gli atti compiuti dal Consiglio di Amministrazione siano conformi alla legge ed allo statuto societario. Valutano inoltre l'adeguatezza dell'assetto organizzativo, amministrativo e contabile adottato dalla società ed al suo concreto funzionamento in funzione del raggiungimento dell'oggetto

sociale. I componenti del consiglio rispondono direttamente all'assemblea dei soci.

La compresenza dei tre organi è tesa a garantire la minimizzazione dei conflitti d'interesse ed alla massimizzazione della ricchezza degli azionisti.

Assieme alla struttura della corporate governance illustrata sopra, diverse sono le linee guida sui principi che devono trovare applicazione da parte degli agenti coinvolti nella governance. L'OCSE, in un documento redatto nel 2004, ne esplicita sei (comunque di carattere non obbligatorio). Secondo questi: l'assetto del governo societario dovrebbe promuovere mercati trasparenti ed efficienti, essere coerente con la legge ed articolare chiaramente la suddivisione delle responsabilità fra le diverse autorità preposte alla supervisione, alla regolamentazione ed alla garanzia dell'applicazione delle norme; l'assetto del governo societario dovrebbe proteggere i diritti degli azionisti e facilitarne l'esercizio; l'assetto societario dovrebbe assicurare l'equo trattamento di tutti gli azionisti, compresi quelli di minoranza e quelli stranieri. Ad ogni azionista dovrebbe essere riconosciuta la possibilità di disporre di efficaci rimedi giuridici per la violazione dei propri diritti; l'assetto del governo societario dovrebbe riconoscere i diritti degli stakeholder previsti dalla legge o da mutui accordi e incoraggiare un'attiva cooperazione fra le imprese e tali soggetti al fine di creare ricchezza e posti di lavoro, e di assicurare lo sviluppo durevole di imprese finanziariamente solide; l'assetto di governo societario dovrebbe assicurare una tempestiva e accurata informazione su tutte le questioni rilevanti riguardanti la società, comprese la situazione finanziaria, le performance, gli assetti proprietari, e la governance della società; ed infine, l'assetto del governo societario deve assicurare la guida strategica della società, l'efficace monitoraggio della gestione da parte del consiglio d'amministrazione e che il consiglio di amministrazione risponda del suo operato alla società e agli azionisti.

L'interesse sulla regolamentazione della *Corporate Governance* è stato mantenuto vivo da scandali quali fallimenti del calibro di Enron e MCI Inc. (WorldCom) negli Stati Uniti, e Parmalat in Italia.

2.4.3 Il ruolo del CEO

Il CEO (*Chief Executive Officer*, o Amministratore Delegato) è, all'interno dell'azienda e della struttura di *Corporate Governance*, colui che è incaricato del ruolo più apicale. Il CEO è responsabile per il successo dell'impresa ed in generale per ogni decisione ai livelli più alti dell'organizzazione; essi possono infatti chiedere ed ottenere consigli ma sono in ultima istanza i responsabili per le decisioni prese.

Il CEO è un organo delegato nominato dal *Board of Directors*. I consiglieri decidono le sue attribuzioni e deleghe, ed il CEO risponderà direttamente a questi ultimi. Il CEO può essere anche membro del board e in alcuni casi anche il presidente del consiglio stesso. Si parla in quest'ultimo caso di *CEO-duality*.

Le responsabilità del CEO possono variare da entità ad entità ed a seconda delle attribuzioni che ad esso sono state delegate. Esso è in genere incaricato di sviluppare ed eseguire le strategie a lungo termine dell'impresa, con l'obiettivo di accrescere il valore aziendale e di conseguenza il valore per gli azionisti. Nelle piccole-medio imprese il CEO, oltre a definire la visione strategica, generalmente assume un ruolo più operativo occupandosi delle operazioni *day-to-day*. Nelle grandi imprese di contro, il CEO si occupa solitamente di decisioni di alto livello, delegando le altre operazioni a *managers* e dipartimenti appositi.

Quando si discute del ruolo del CEO, lo si contrappone spesso al ruolo del Presidente del Board. Quest'ultimo è in posizione gerarchica superiore rispetto al CEO, considerato che il CEO è in ultima istanza dipendente e nominato dal *Board* e dal Presidente e che non può prendere decisioni rilevanti senza l'approvazione di questi. Nella maggioranza dei casi ed al crescere della dimensione aziendale, il Presidente non è direttamente coinvolto nelle operazioni di *business* e lascia grande flessibilità di agire al CEO. In alcuni casi, come accennato in precedenza, i ruoli di CEO e presidente del Board possono coincidere. Questo è per la maggior parte vero al diminuire della dimensione aziendale. Il coincidere dei due ruoli può far emergere dei conflitti di interessi, per esempio: a) nelle votazioni concernenti l'*Executive Compensation* e b) nella valutazione della performance del CEO e degli organi delegati ad opera del consiglio d'amministrazione. In questi due casi, qualora ci fosse dualità tra il CEO ed

il Presidente, vi sarebbe un interesse da parte del CEO-Presidente rispettivamente ad alzare l'*Executive Compensation* anche in caso di cattiva *performance* ed a mitigare od offuscare le *performance* del CEO qualora queste si rivelassero inadeguate, portando ad un'effettiva diminuzione del valore aziendale. Per questa ragione solitamente una buona *Corporate Governance* prescrive una separazione tra i due ruoli.

Da queste considerazioni nasce la necessità di indagare il nesso tra qualità del CEO e la struttura di Corporate Governance con la performance dell'impresa. Come vedremo nel prossimo paragrafo infatti, molti studi si sono concentrati sul nesso tra qualità del CEO e struttura della *Governance*, e bancarotta (che rappresenta l'estremo peggiore della cattiva *performance* d'impresa).

2.4.4 Corporate Governance e Fallimento: Literature Review

Uno dei primi studi ad indagare la connessione tra Corporate Governance e fallimento d'impresa, risale a Chaganti et al., 1985. Gli autori studiano la correlazione tra fallimento, numerosità del CdA (*board size* di seguito), proporzione degli amministratori esterni¹³ (%OUTSIDE di seguito) all'interno del CdA e CEO-*duality*. Il sample in oggetto di studio è composto da 42 imprese nel settore del commercio al dettaglio, di cui 21 fallite e 21 non fallite, nel periodo 1970-1976. Al fine di discriminare i due gruppi di imprese, gli autori trovano che delle tre variabili prese in considerazione, solo la variabile *board size* risulta significativa. Le imprese non fallite tendono infatti ad avere un CdA più numeroso.

Hambrick & D'Aveni (1992), analizzano un *sample* di 114 imprese (57 in bancarotta e 57 attive). Dai risultati emerge che un CEO con uno stile di *leadership* dominante porta ad una maggiore probabilità di fallimento.

Daily & Danton (1994a), impiegano la regressione Logit su un campione di 100 imprese (50 fallite, 50 non fallite) durante il 1990. Includono tre variabili di *corporate governance*: 1) CEO-*duality* (definita come CEO = Presidente del CdA), 2) % OUTSIDE, e 3) numero assoluto di amministratori indipendenti. Inoltre, impiegano cinque variabili finanziarie di controllo esprimenti la dimensione dell'impresa, la liquidità, la profittabilità, il *leverage* ed il capitale circolante espresso in percentuale sulle vendite. Gli autori trovano che impiegare gli indicatori di corporate governance migliora consistentemente l'affidabilità e la precisione dei modelli (60,1% contro 70,13% di accuratezza ad un anno dal fallimento, e sottolineano l'importanza di variabili come CEO-*duality* (correlata positivamente al fallimento d'impresa) e %OUTSIDE.

In un altro studio di Daily & Danton (1994b), applicando la regressione Logit su un campione di 114 imprese tra il 1972 e il 1982, gli autori trovano che CEO-*duality* e la proporzione degli amministratori indipendenti sono significative al fine di discriminare in maniera predittiva le imprese fallite e non.

¹³ Per amministratori interni intendiamo direttori che hanno qualche tipo di rapporto in essere con la società stessa: pensiamo ad esempio ad azionisti-amministratori, CEO, rappresentati di altre parti di interesse (e.g. istituti di credito). Per amministratori esterni od indipendenti intendiamo direttori che non hanno alcun tipo di rapporto in essere con la società, né direttamente né indirettamente.

Elloumi & Gueyié (2001), studiano un campione composto da 92 imprese canadesi (46 fallite e 46 non fallite). Le variabili di corporate governance prese in considerazione sono: 1) %OUTSIDE, 2) CEO-*duality* (intesa come CEO = Presidente del *Board*), 3) AUDITCOM (variabile *dummy* con valore 1 se tutti i membri dell'*audit committee* sono indipendenti, con valore 0 in caso contrario), e 4) BLOCKHLD (che esprime la percentuale di azioni cumulata detenuta dagli azionisti che hanno più del 20% e che non sono affiliati con il CdA). Includono anche variabili di controllo esprimenti la liquidità ed il *leverage*. Gli autori concludono che combinando indicatori di Corporate Governance ad indicatori economico-finanziari migliora la potenza discriminante del modello (70,6% contro 72,1%).

Ciampi (2015) sviluppa un modello su misura per la piccole e medie imprese, studiando un campione di 3210 PMI italiane (1605 fallite e 1605 non fallite) utilizzando la regressione Logit. L'autore utilizza un *mix* di indicatori di corporate governance ed economico-finanziari; nello specifico: 1) CEO-*duality*, 2) $0 < \% \text{ OUTSIDE} < 0.5$, 3) $\% \text{ OUTSIDE} > 0.5$, 4) *Board Size*, 5) $\text{GUO} > 50\%$ ¹⁴, 6) ROI, 7) *EBITDA* / Vendite, 8) Interessi Passivi / Vendite e 9) Debito Finanziario / Patrimonio Netto. Ciampi trova una correlazione negativa tra CEO-*duality* e fallimento delle PMI, una correlazione negativa tra la variabile n. 2 e la probabilità di fallimento ed una correlazione positiva tra la variabile n. 3 e il fallimento. In altre parole, all'aumentare degli amministratori indipendenti sopra il 50% del totale del *board*, di pari passo cresce la probabilità di bancarotta. Ciampi trova inoltre, consistentemente con altri studi, che combinare le variabili di corporate governance a quelle economico-finanziarie aumenta l'affidabilità del modello.

Süsi & Lukason (2018), propongono uno studio effettuato su 67058 piccole e medie imprese estoni con dati del 2015, utilizzando la regressione Logit. Le variabili considerate sono: 1) *Board Size*, 2) *Board Gender*, 3) *Board Tenure*, 4) Età del *Top Management*, 5) *Multiple Directorship* (il numero di altri incarichi che i membri del *board* hanno in altre imprese), 6) *Ownership Concentration* (variabile *dummy* con valore 1 se l'azionista di maggioranza ha più del 50% delle azioni, 0 altrimenti), 7) Azioni detenute dai membri del board. Dall'analisi dei risultati, emerge che: a)

¹⁴ Le variabili 2), 3), e 5) sono espresse come *dummy*, con valore 1 se la condizione espressa è vera, e valore 0 altrimenti,

all'aumentare del *board size* cresce il rischio di fallimento, b) la variabile *Board Gender* e *Board Tenure* ed *Ownership Concentration* risultano irrilevanti, c) il fallimento è negativamente correlato con l'età del top management, d) il fallimento è positivamente correlato con la variabile *Multiple Directorship* ed e) la probabilità di fallire decresce al crescere del capitale azionario posseduto dai membri del board.

2.4.5 Considerazioni Finali

Come possiamo notare dalle evidenze fornite dagli studi esposti nel precedente paragrafo, non vi è unanimità dei risultati per quanto riguarda la connessione tra variabili di *corporate governance* e probabilità di fallire. Al contrario, a seconda del tipo di imprese analizzate e delle variabili considerate, gli studi possono portare a risultati discordanti. I risultati esposti da Süsi & Lukason (2018) per quanto riguarda la variabile *Board Size* sono in aperta contraddizione con quanto ottenuto da Chaganti et al. (1985). Secondo i primi all'aumentare della numerosità del board cresce la probabilità di fallimento, mentre per i secondi decresce. Altro esempio può essere esposto per la variabile CEO-duality, che per Chaganti et al. (1985) risulta irrilevante, per Daily e Danton (1994a) risulta discriminante e correlata positivamente al fallimento, e per Ciampi (2015) risulta negativamente correlata alla probabilità di fallimento.

Altra considerazione doverosa è che gli studi sopraesposti sono solo una piccola parte della letteratura sul tema, e che la significatività delle variabili potrebbe variare nel corso degli anni, così come a seconda del tipo di imprese (settore e dimensione) e area geografica considerata. Inoltre, tutti gli studi sono concordi e unanimi nell'affermare che l'aggiunta delle variabili di corporate governance migliora l'affidabilità dei modelli predittivi del fallimento.

Scopo del prossimo capitolo sarà di costruire diversi modelli unendo variabili economico-finanziarie a quelle di *corporate governance*, e paragonarli tra loro al fine di arrivare al modello finale che auspicabilmente avrà una buona forza predittiva nei test di validazione.

Capitolo 3: Lo Studio Empirico

In questo capitolo verrà delineato e discusso lo studio empirico effettuato in questa tesi. Si discuterà della formulazione delle ipotesi, delle variabili considerate, della metodologia adottata ed infine dei riscontri ottenuti.

3.1 Formulazione delle ipotesi

Come esposto nel capitolo precedente in seguito alla literature review, la letteratura risulta spesso discorde riguardo le variabili che incidono sulla probabilità di fallimento: questo è più vero sulle variabili di *corporate governance*, e sul ruolo del CEO.

Abbiamo infatti risultati discordanti al variare della dimensione aziendale e della struttura proprietaria delle imprese considerate, ma abbiamo risultati concordanti sul fatto che l'aggiunta delle variabili di corporate governance migliori l'accuratezza del modello. In tal proposito, visto e considerato il suo ruolo apicale nell'organizzazione aziendale, siamo interessati a capire che incidenza abbiano le variabili descrittive delle qualità del CEO sulla probabilità di fallimento. La dimensione dei fallimenti aziendali in Europa riguarda per lo più le piccole e medie imprese, dove spesso l'azionariato è concentrato e si ha dualità tra azionista di maggioranza e CEO, e di conseguenza dualità tra presidente del *board* e CEO. Appare quindi ragionevole ipotizzare che nelle piccole e medie imprese, il CEO-proprietario sia interessato alla massimizzazione del valore aziendale per sé stesso, e che i conflitti di interesse in questa situazione diventino per lo più trascurabili, in quanto non si verifica la classica separazione tipica delle grandi imprese tra azionariato e management.

Per studiare l'incidenza delle qualità del CEO sulla probabilità di fallimento, due ipotesi distinte saranno testate:

Ipotesi 1: All'aumentare della tenure del CEO in azienda diminuisce la probabilità di fallimento e,

Ipotesi 2: La CEO-duality (quindi CEO anche presidente del board of directors) è negativamente correlata alla probabilità di fallimento.

3.2 Descrizione del Sample

Il processo di costruzione del sample oggetto di studio ha seguito diverse fasi. Da primo, utilizzando il database ORBIS si è composta una lista di imprese Europee fallite dal 2015 al 2018 ed appartenenti ai settori manifatturiero e dei servizi, utilizzando come criteri di selezione la forma giuridica standardizzata a S.p.A. Italiana non quotata, e la dimensione degli *assets* superiore ad un milione di euro . La numerosità del campione in seguito a questo passaggio è risultata di 458 imprese.

I dati finanziari sono stati estratti per tutte le 458 imprese. Il sample di imprese fallite ha subito una notevole riduzione man mano che ai dati finanziari si sono accostati i dati di *corporate governance* oggetto delle ipotesi sopra menzionate, non disponibili per una parte notevole delle imprese selezionate.

Similmente alla maggior parte degli studi precedenti (e.g. Beaver 1966, Altman 1968), si è proceduto ad accoppiare ad ogni impresa fallita una impresa attiva seguendo come criterio di accoppiamento la dimensione degli *assets* ed il segmento del business principale, identificato dal codice NACE. Seguendo questo procedimento ogni impresa fallita è accostata ad una “gemella” non fallita. Il sample finale si compone quindi di 316 imprese europee, di cui 158 dichiarate fallite e 158 ancora attive alla data della raccolta dati. Lo spazio temporale considerato varia dal 2013 al 2018. Al fine di eseguire la procedura di validazione *holdout*, il *sample* è stato diviso in *training-set* e *test-set*. I due *sets* sono stati ottenuti mediante campionamento *random* sul *set* completo. Il primo contiene 266 imprese ed il secondo altre 50 imprese, equamente divise in fallite e non fallite in ognuno dei due *set*.

Per ogni impresa, sono stati raccolti i dati di bilancio per l’anno del fallimento (anno “t”), un anno prima all’evento (“t-1”) e due anni prima (“t-2”). I dati finanziari sono stati estratti automaticamente dal database ORBIS, mentre per i dati di *Corporate Governance* si è proceduto al calcolo delle *tenures* manualmente utilizzando la data di nomina come inizio *tenure*.

Di seguito delle tabelle che sintetizzano l'andamento statistico delle variabili che saranno utilizzate nelle regressioni, per ogni gruppo d'impresa al tempo "t", "t-1" e "t-2" (*assets* in migliaia di euro).

Tabella 4: Variabili per il gruppo delle imprese non fallite al tempo "t"

. summarize status assets ceo_duo ceo_ten liquidity ebitass assturn tdta rigidity

Variable	Obs	Mean	Std. Dev.	Min	Max
status	158	0	0	0	0
assets	158	30667.36	63134.89	326.3618	582109
ceo_duo	158	.8291139	.3776062	0	1
ceo_ten	158	10.26251	5.052592	3.859726	29.28
liquidity	158	.3747613	.6682085	.0000818	3.541439
ebitass	158	.076083	.1015414	-.2450552	.6767974
assturn	158	1.818086	1.517854	.0092627	12.33738
tdta	158	.6006145	.2838756	.0243798	2.375695
rigidity	158	.282736	.2345372	0	.9847177

Tabella 5: Variabili per il gruppo delle imprese fallite al tempo "t"

. summarize status assets ceo_duo ceo_ten liquidity ebitass assturn tdta rigidity

Variable	Obs	Mean	Std. Dev.	Min	Max
status	158	1	0	1	1
assets	158	28357.9	59730.19	310.2109	570507.2
ceo_duo	158	.5126582	.5014291	0	1
ceo_ten	158	7.034549	4.578131	.5424657	40.45
liquidity	158	.0529066	.0984673	.0000308	.6548287
ebitass	158	-.2742609	1.042034	-10.9484	.7814027
assturn	158	1.45753	2.042519	.0091079	22.48886
tdta	158	1.207088	1.132391	.0766015	12.16741
rigidity	158	.3650596	.2552891	0	.9814569

Tabella 6: Variabili per il gruppo delle imprese non fallite al tempo "t-1"

. summarize status assets ceo_duo ceo_ten liquidity ebitass assturn tdta rigidity

Variable	Obs	Mean	Std. Dev.	Min	Max
status	158	0	0	0	0
assets	158	31248.32	66716.64	314.5781	636367
ceo_duo	158	.8291139	.3776062	0	1
ceo_ten	158	9.262513	5.052592	2.859726	28.28
liquidity	158	.3833449	.8169258	.0000212	6.761621
ebitass	158	.0579188	.0989491	-.5548449	.4129767
assturn	158	1.843107	1.728827	.0780779	15.36315
tdta	158	.6084633	.2751049	.0189905	1.915254
rigidity	158	.2857275	.239434	0	.9686562

Tabella 7: Variabili per il gruppo delle imprese fallite al tempo "t-1"

. summarize status assets ceo_duo ceo_ten liquidity ebitass assturn tdta rigidity

Variable	Obs	Mean	Std. Dev.	Min	Max
status	158	1	0	1	1
assets	158	31361.93	61050.51	295.2705	585198.8
ceo_duo	158	.5126582	.5014291	0	1
ceo_ten	158	6.080067	4.528212	.2315	39.45
liquidity	158	.0932093	.2334379	8.55e-08	2.170088
ebitass	158	-.0706036	.1923072	-.9687414	.1619083
assturn	158	1.281727	1.073757	.0084785	7.403985
tdta	158	.895593	.3765627	.056586	3.010852
rigidity	158	.3604779	.2382632	.0015079	.9817156

Tabella 8: Variabili per il gruppo delle imprese non fallite al tempo "t-2"

. summarize status assets ceo_duo ceo_ten liquidity ebitass assturn tdta rigidity

Variable	Obs	Mean	Std. Dev.	Min	Max
status	158	0	0	0	0
assets	158	28965.77	59367.13	227.8611	570824
ceo_duo	158	.8291139	.3776062	0	1
ceo_ten	158	8.262513	5.052592	1.859726	27.28
liquidity	158	.3738192	.7519665	.0001045	5.016528
ebitass	158	.0513715	.0901821	-.4146675	.4087321
assturn	158	1.779876	1.607022	.0596549	14.86876
tdta	158	.6041305	.2613468	.0291437	1.732835
rigidity	158	.297713	.2419962	.0019021	.9519197

Tabella 9: Variabili per il gruppo delle imprese fallite al tempo "t-2"

. summarize status assets ceo_duo ceo_ten liquidity ebitass assturn tdta rigidity

Variable	Obs	Mean	Std. Dev.	Min	Max
status	158	1	0	1	1
assets	158	32689.46	64365.16	232.9608	596663.5
ceo_duo	158	.5126582	.5014291	0	1
ceo_ten	158	5.097392	4.507683	.0219178	38.45
liquidity	158	.0771536	.166417	.0000106	1.544642
ebitass	158	-.055219	.2755788	-2.558829	.1561514
assturn	158	1.298858	1.0886	.0082003	7.213766
tdta	158	.8284885	.2842229	.0907664	2.474707
rigidity	158	.3719078	.240434	.0008719	.9828144

3.3 Profilazione delle Variabili

La variabile dipendente sulla quale verranno effettuate le regressioni, è **Status**, codificata con 1 se l'impresa è fallita, e con 0 altrimenti.

Al fine di confermare o rigettare le ipotesi esplicitate nel paragrafo 3.1, le variabili indipendenti oggetto di studio sono CEO Tenure e CEO-*duality*.

CEO Tenure misura il tempo intercorrente tra la nomina del CEO e la data di raccolta dati.

CEO-Duality è una variabile *dummy* che assume valore 1 se il CEO è anche Presidente del *Board of Directors*, e valore 0 altrimenti.

Per arrivare alla profilazione delle variabili di controllo, si sono seguiti due criteri: il primo, la popolarità nella letteratura precedente. In particolare, si è optato similmente a Ohlson (1980) e Beaver (1966), di selezionare un indice per ogni macroarea (liquidità struttura, profittabilità e dimensione - la quale è già stata tenuta in considerazione in fase di *matching*). Il secondo criterio è stato l'eliminazione della possibile multicollinearità tra gli indici in questione.

La multicollinearità è il fenomeno per cui le variabili indipendenti di un modello di regressione risultano correlate. Se il grado di correlazione tra le variabili è sufficientemente alto, può causare problemi quando si stima il modello e quando si interpretano i risultati. I coefficienti di regressione infatti indicano in media di quanto si muove la variabile dipendente al muoversi di una variabile indipendente *mantenendo tutte le altre costanti*. Se le variabili indipendenti sono altamente correlate tra di loro, un movimento in una di queste causerà un opposto o uguale movimento in quelle correlate; questo rende difficile per il modello effettuare una stima corretta della relazione tra la variabile dipendente e le variabili indipendenti, perché le variabili correlate tendono a muoversi all'unisono. Dunque, in presenza di multicollinearità tra le variabili indipendenti del modello, due sono i principali effetti indesiderati:

- 1) I coefficienti stimati possono muoversi sensibilmente risultando poco precisi a seconda di quali variabili vengano incluse o escluse nella regressione;

- 2) La multicollinearità, riducendo la precisione dei coefficienti, indebolisce il potere statistico del modello. I *p-values* infatti potrebbero essere inaffidabili nell'identificare le variabili statisticamente significative.

Al fine di eliminare la multicollinearità, si è proceduto a delle regressioni lineari sulla variabile dipendente via *software* STATA®, e ci si è avvalsi dello strumento VIF (*Variance Inflator Factor*). Il VIF regredisce ogni singola variabile esplicativa con tutte le altre presenti nel modello. Tramite questo procedimento si ottengono i valori dell' R-quadro, che saranno presenti al denominatore della formula del VIF, dove

$$VIF = \frac{1}{1 - R_i^2}$$

Il VIF può assumere tutti i valori superiori ad 1, e la sua interpretazione è la seguente: un VIF di 1.2 ci dice che la varianza di un particolare coefficiente è del 20% più grande che nel caso in cui non vi fosse alcuna multicollinearità tra i regressori considerati. Non vi è una regola generale per l'interpretazione dei valori risultanti dal VIF, ma in genere si fa corrispondere al valore 1 la non correlazione tra i regressori, a valori compresi tra 1 e 5 moderata correlazione, e per valori maggiori di 5 si parla di alta correlazione. In questa tesi, si è optato di escludere tutte le variabili che risultassero in un valore del VIF superiore a 5. Di seguito la tabella con i valori del VIF e la matrice delle correlazioni per le variabili selezionate.

Tabella 10: Valori del VIF sulle variabili di regressione

. vif

Variable	VIF	1/VIF
ebitass	5.00	0.199902
tdta	4.88	0.205121
assturn	1.75	0.570533
rigidity	1.14	0.880071
liquidity	1.12	0.890608
ceo_ten	1.08	0.925256
ceo_duo	1.06	0.942778
Mean VIF	2.29	

Tabella 11: Matrice di correlazione per le variabili di regressione

```
. correlate status ceo_ten ceo_duo liquidity ebitass tdta assturn rigidity
(obs=266)
```

	status	ceo_ten	ceo_duo	liquid~y	ebitass	tdta	assturn	rigidity
status	1.0000							
ceo_ten	-0.2884	1.0000						
ceo_duo	-0.3554	0.1275	1.0000					
liquidity	-0.3091	0.1317	0.1404	1.0000				
ebitass	-0.2291	0.0748	0.0650	0.0826	1.0000			
tdta	0.3492	-0.1705	-0.1262	-0.2092	-0.8749	1.0000		
assturn	-0.0986	-0.0419	0.0572	-0.0276	-0.6023	0.5428	1.0000	
rigidity	0.2074	-0.1053	-0.0406	-0.1114	0.1032	-0.0720	-0.2972	1.0000

Il profilo delle variabili di controllo risulta quindi contenere un indice di liquidità, uno di efficienza operativa o di profittabilità in senso lato, uno di competitività, uno di struttura finanziaria ed uno di struttura dell'attivo.

Di seguito una descrizione delle cinque variabili di controllo:

- 1) **Liquidity:** È una misura di liquidità risultante dal quoziente “Cassa e mezzi equivalenti” / “Passività Correnti”. Dà indicazioni di quanto le passività correnti siano liquidabili immediatamente con mezzi contanti o equivalenti. Ricordiamo che le passività correnti sono quelle passività destinate ad essere estinte nell'arco dei 12 mesi. Più alto è Liquidity, più alta è la solvibilità a breve termine dell'impresa.
- 2) **EBIT / Assets:** È una misura di efficienza operativa che indica quanto l'impresa è efficace a “trasformare” gli assets in profitti. L'EBIT viene utilizzato al posto dell'utile netto in quanto il primo indica il profitto operativo dell'impresa, e tralascia quindi componenti esogene quali tasse e struttura finanziaria. Maggiore è l'EBIT in confronto agli assets, più efficace è l'impresa a trasformare gli assets in profitti. Anche EBIT/Assets similmente a Current e Liquidity si paragona per l'impresa “n” con la media del settore e con gli anni precedenti.
- 3) **Asset Turnover:** È misura di competitività. L'indice è calcolato come “Vendite” / “Totale attivo”, e dà un'idea di quanto l'impresa sia efficace nel generare vendite mediante gli assets. Maggiore è l'indice, meglio l'impresa sta performando, in quanto l'indice esprime quante vendite si effettuano per ogni euro di assets. L'asset turnover è paragonato a livello settoriale e tra gli anni.
- 4) **Totale Debito / Totale Attivo:** È una misura della struttura finanziaria o di *leverage* dell'impresa. L'indice descrive la proporzione di mezzi impiegati nell'impresa

finanziati tramite debito. Maggiore è l'indice, maggiore è il rischio finanziario associato all'attività dell'impresa. Inoltre, un alto valore dell'indice è indicativo della bassa potenzialità di contrarre ulteriore debito dell'impresa, il che risulta in una bassa flessibilità finanziaria.

- 5) **Rigidity Index:** L'indice di rigidità è una misura della struttura dell'attivo di un'impresa. Esso è calcolato come "Attivo Non Corrente" / "Totale Attivo", la porzione dell'attivo immobilizzato rispetto al totale dello stesso. L'indice varia da settore a settore e viene comparato tramite i valori medi degli stessi.

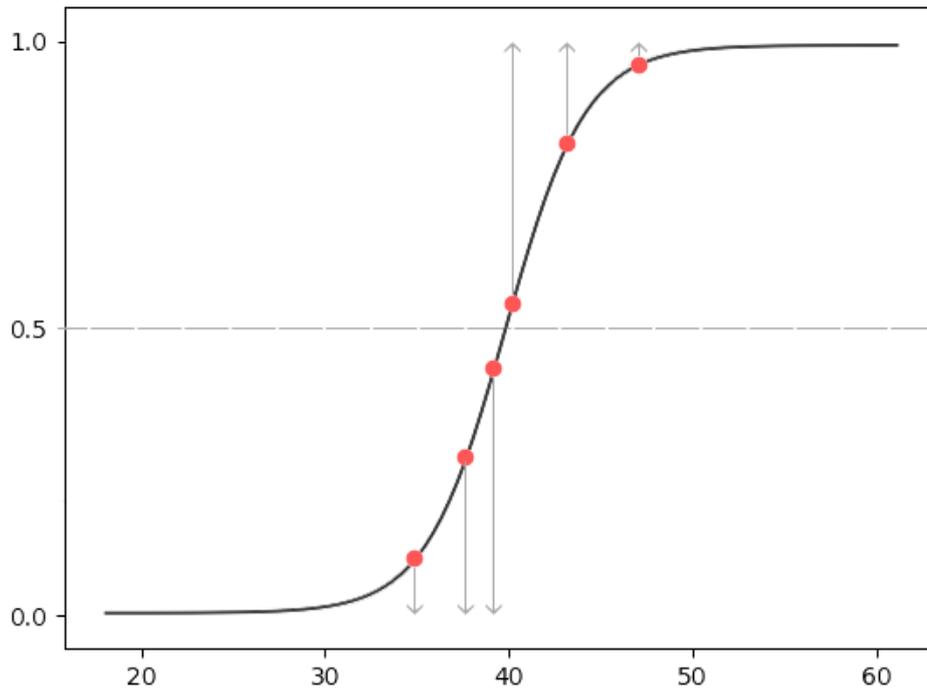
3.4 Metodologia: la Regressione Logistica

Similmente alla maggior parte dei lavori precedenti sul tema, si è scelto di effettuare le regressioni sulla variabile dipendente mediante la regressione logistica. La scelta è dettata dalla semplicità applicativa del metodo ed alla sua facile interpretazione.

Come accennato in precedenza, la regressione logistica è un tipo di regressione particolarmente adatta quando la variabile dipendente è espressa in forma dicotomica, 0 ed 1. L'obiettivo della regressione è stabilire la probabilità con cui le variabili indipendenti e di controllo possano generare 0 od 1 sulla variabile dipendente. Essa restituisce un valore compreso tra 0 ed 1 che descrive la probabilità dell'elemento X di appartenere ad una delle due classi. Nel caso specifico degli studi sul fallimento, per ogni impresa la regressione Logistica ritorna la probabilità che queste rientrino nell'universo delle fallite o delle attive.

Graficamente, come si può vedere dalla Figura 1, il modello Logit si concreta in una sigmoide con estremi pari a 0 ed 1 (entro la linea ricadono tutte i valori delle probabilità).

Figura 1) Rappresentazione grafica del modello Logit



Nello specifico, il modello Logistico generalizzato per l'intera popolazione risulta essere dato da:

$$E[Y|\mathbf{X}] = Pr(Y = 1 | X_1, \dots, X_k) = \lambda(\mathbf{X}^T \boldsymbol{\beta}) =$$

$$= \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}} = p \quad (1)$$

E tramite manipolazione algebrica si ottiene:

$$p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)}} \quad (2)$$

Dove:

$Pr = p$ è la probabilità di ottenere come risultato 1 nella variabile indipendente;

Y è la variabile indipendente dicotomica;

\mathbf{X} è il vettore di variabili indipendenti o regressori X_1, \dots, X_k ;

$\boldsymbol{\beta}$ è il vettore dei parametri o dei coefficienti dei regressori β_0, \dots, β_k ;

λ è la funzione di ripartizione della distribuzione logistica standard ed e è il numero di Eulero, pari a circa 2,71828.

Per ogni osservazione del campione $i = 1, \dots, n$ si dispone di una determinazione Y e di k determinazioni delle variabili indipendenti X_1, \dots, X_k (nel nostro caso le X dalle variabili indipendenti sono illustrate nel paragrafo precedente). La regressione logistica cerca quindi una relazione non lineare tra le variabili indipendenti X_1, \dots, X_k e la variabile dipendente Y , sfruttando la funzione di ripartizione λ della distribuzione logistica standard, stimando il valore dei coefficienti β_0, \dots, β_k mediante il metodo della massima vero-somiglianza. Tramite il metodo della massima vero-somiglianza, il modello massimizza la funzione di log-vero-somiglianza, trovando i coefficienti che danno la più alta probabilità di avere stime di classificazione corrette sulla variabile dipendente.

I coefficienti della regressione logistica sono il logaritmo naturale dell'*odds ratio*, dove l'*odds ratio* è definito come:

$$\text{Odds Ratio} = \frac{p}{1 - p}$$

L'*odds ratio* misura quanto è larga la probabilità che $y=1$ rispetto alla probabilità che $y=0$.

L'interpretazione dei coefficienti della regressione logistica è che un incremento della variabile esplicativa X_k , avrà effetto di rendere la probabilità dell'evento $Y=1$ più o meno probabile (a seconda del segno del coefficiente). Per questo motivo, l'interpretazione dell'output della regressione ci dice soltanto se la probabilità dell'evento aumenterà o diminuirà, non fornendo informazioni circa la forza dell'effetto.

Una volta calcolato il vettore dei coefficienti β_0, \dots, β_k , tramite la funzione (1) o (2) è possibile procedere alla stima della probabilità p che l'osservazione Y rientri nella classe 1.

3.5 Risultati della Regressione Logistica

Una volta specificato il modello di regressione da usare e le variabili coinvolte, è possibile procedere allo studio sperimentale. Per tutte le regressioni è stato utilizzato il software statistico STATA® versione 15.2.

In prima analisi si sono effettuate due regressioni singole sulle variabili indipendenti unitamente alle cinque variabili di controllo illustrate in precedenza, utilizzando i dati del primo anno di osservazione del *training-set*; per la variabile **CEO-Tenure** il risultato della regressione è illustrato nella Tabella 12.

Tabella 12: Risultato della regressione logistica sulla variabile CEO-Tenure

```

. logit status ceo_ten liquidity ebitass tdtas assturn rigidity

Iteration 0:   log likelihood = -184.37715
Iteration 1:   log likelihood = -121.46984
Iteration 2:   log likelihood = -102.03997
Iteration 3:   log likelihood = -97.039506
Iteration 4:   log likelihood = -96.937652
Iteration 5:   log likelihood = -96.937297
Iteration 6:   log likelihood = -96.937297

Logistic regression               Number of obs   =           266
                                  LR chi2(6)         =           174.88
                                  Prob > chi2         =           0.0000
                                  Pseudo R2           =           0.4742

Log likelihood = -96.937297
    
```

status	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
ceo_ten	-.128806	.0449746	-2.86	0.004	-.2169546 -.0406575
liquidity	-2.416669	1.302547	-1.86	0.064	-4.969614 .1362753
ebitass	-7.947679	1.929447	-4.12	0.000	-11.72933 -4.166033
tdtas	3.421777	.8306221	4.12	0.000	1.793787 5.049766
assturn	-.4949135	.1768958	-2.80	0.005	-.8416228 -.1482042
rigidity	1.831071	.8266593	2.22	0.027	.2108488 3.451294
_cons	-1.03595	.8700602	-1.19	0.234	-2.741237 .6693366

Ciò che si evince dal risultato della regressione è la correlazione negativa tra la variabile indipendente “ceo_ten” e la variabile dipendente “status”, con un *p-value* significativo all’1%. Considerando che di *default* il modello logistico ricerca la probabilità che si verifichi l’evento codificato con “1” (quindi che si verifichi l’evento fallimento) e dato il coefficiente di correlazione negativo tra la variabile “ceo_ten” e la variabile dicotomica “status”, è possibile concludere che l’ipotesi n.1 risulta confermata, ovvero:

All’aumentare della CEO-Tenure, diminuisce la probabilità di fallimento.

La seconda regressione è stata effettuata sulla variabile **CEO-Duality**, mantenendo invariate le variabili di controllo. I risultati della regressione sono esposti nella tabella 13.

Tabella 13: Risultati della regressione logistica sulla variabile CEO-Duality

```
. logit status ceo_duo liquidity ebitass tdtas assturn rigidity
```

```
Iteration 0: log likelihood = -184.37715
Iteration 1: log likelihood = -116.6876
Iteration 2: log likelihood = -97.870307
Iteration 3: log likelihood = -92.296326
Iteration 4: log likelihood = -92.093463
Iteration 5: log likelihood = -92.092603
Iteration 6: log likelihood = -92.092603
```

```
Logistic regression                Number of obs    =      266
                                LR chi2(6)         =      184.57
                                Prob > chi2          =      0.0000
Log likelihood = -92.092603        Pseudo R2        =      0.5005
```

status	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
ceo_duo	-1.982115	.4749255	-4.17	0.000	-2.912952	-1.051278
liquidity	-2.427233	1.42899	-1.70	0.089	-5.228002	.3735349
ebitass	-7.878125	1.755139	-4.49	0.000	-11.31813	-4.438117
tdtas	3.968272	.8441924	4.70	0.000	2.313685	5.622859
assturn	-.5191622	.1812373	-2.86	0.004	-.8743808	-.1639436
rigidity	1.909135	.8441728	2.26	0.024	.2545864	3.563683
_cons	-.9962561	.8017797	-1.24	0.214	-2.567715	.5752032

Anche nel caso della variabile dipendente “**ceo_duo**”, il coefficiente di regressione risulta negativo e la variabile significativa con un *p-value* inferiore allo 0.1%. Questo risultato consente di confermare l’ipotesi n.2, e quindi:

La CEO-duality è correlata negativamente con la probabilità di fallimento.

3.6 Classificazione del modello completo

Utilizzando congiuntamente le due variabili indipendenti CEO-Tenure e CEO-duality e le variabili di controllo in un unico modello, è possibile ottenere un modello statistico utilizzabile a fini di classificazione. Similmente agli studi precedenti, considererò adesso tutte le variabili utilizzate fin ora come regressori indipendenti (Modello 1), e paragonerò i risultati con un altro modello che considera solo le variabili finanziarie (Modello 2). Tutti i dati sui quali sono effettuate queste due regressioni si riferiscono all'anno "t", e con le 266 osservazioni contenute nel *test-set*. I risultati del Modello 1 sono esposti nella tabella 14.

Tabella 14: Risultati della regressione logistica - Modello 1

```
. logit status ceo_ten ceo_duo liquidity ebitass tdtta assturn rigidity
```

Iteration 0: log likelihood = -184.37715
Iteration 1: log likelihood = -111.79427
Iteration 2: log likelihood = -92.952965
Iteration 3: log likelihood = -88.075337
Iteration 4: log likelihood = -87.863379
Iteration 5: log likelihood = -87.862841
Iteration 6: log likelihood = -87.862841

Logistic regression	Number of obs	=	266
	LR chi2(7)	=	193.03
	Prob > chi2	=	0.0000
Log likelihood = -87.862841	Pseudo R2	=	0.5235

status	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
ceo_ten	-.1219514	.048012	-2.54	0.011	-.2160532 -.0278496
ceo_duo	-1.926331	.4853181	-3.97	0.000	-2.877537 -.9751251
liquidity	-2.347732	1.381103	-1.70	0.089	-5.054644 .3591804
ebitass	-8.413086	1.902914	-4.42	0.000	-12.14273 -4.683442
tdtta	3.649933	.8620647	4.23	0.000	1.960318 5.339549
assturn	-.5695864	.1875338	-3.04	0.002	-.937146 -.2020269
rigidity	1.870735	.8686314	2.15	0.031	.1682493 3.573222
_cons	.3443629	.9377086	0.37	0.713	-1.493512 2.182238

I risultati della regressione logistica in questo caso ci indicano un profilo di variabili altamente significative al fine di discriminare tra i due gruppi. Le variabili, ceo_duo, ebitass e tdtta sono significative al livello dell 0,1%; la variabile assturn è significativa al livello dell'1%; le variabili rigidity e ceo_ten, risultano significative al livello del 5%; infine, la variabile liquidity risulta significativa al livello del 10%. Inoltre, la voce Prob > chi2 = 0.0000 ci indica che il modello è statisticamente altamente significativo.

Dai risultati della regressione notiamo anzitutto che i segni dei coefficienti stimati per tutte le variabili considerate sono in linea quanto evidenziato nella letteratura precedente sul tema (fatta eccezione per le variabili descrittive il CEO che, come già puntualizzato in precedenza, sono oggetto di risultati discordanti in letteratura). Abbiamo infatti che:

- 1) all'aumentare della variabile Liquidity, e quindi con l'aumentare della cassa in proporzione alle passività correnti, diminuisce la probabilità di fallimento: all'aumentare della liquidity infatti l'impresa diventa più solvibile;
- 2) all'aumentare della remunerazione operativa degli assets (ebitass) diminuisce la probabilità di fallire: all'aumentare di questo indice l'impresa è in grado di "trasformare" più velocemente gli assets in profitti operativi, garantendosi margini per la remunerazione dell'equity e per gli investimenti;
- 3) all'aumentare dell'asset turnover (assturn), e quindi delle vendite come percentuale degli assets diminuisce la probabilità di fallimento: anche qui, aumentando le vendite in percentuale sugli assets, si presume che aumentino di pari passo i profitti e quindi la remunerazione di tutti gli *stakeholders*;
- 4) all'aumentare dell'incidenza percentuale del debito sul totale dei mezzi impiegati (tdta) aumenta la probabilità di fallimento;
- 5) all'aumentare della Rigidity, e quindi dell'incidenza dell'attivo non corrente rispetto all'attivo totale, aumenta la probabilità di fallimento. L'impresa è infatti più "rigida" nello smobilizzare capitali per effettuare un eventuale ripagamento del debito.

Una volta ottenuti i coefficienti di regressione, è possibile procedere alla classificazione delle osservazioni in uno dei due gruppi. Utilizzando come punto di *cut-off* 0.5, otteniamo classificate come fallite le imprese che presentano una $p > 0.5$, e come non fallite le imprese che ottengono $p \leq 0.5$; utilizzando il comando "*estat classification*" dopo aver regredito il modello in STATA®, otteniamo la percentuale di classificazione corretta (Tabella 15).

Tabella 15: Risultati di classificazione della regressione logistica - Modello 1

`. estat classification`

Logistic model for status

Classified	True		Total
	D	~D	
+	110	13	123
-	23	120	143
Total	133	133	266

Classified + if predicted $\Pr(D) \geq .5$

True D defined as status != 0

Sensitivity	$\Pr(+ D)$	82.71%
Specificity	$\Pr(- \sim D)$	90.23%
Positive predictive value	$\Pr(D +)$	89.43%
Negative predictive value	$\Pr(\sim D -)$	83.92%
False + rate for true ~D	$\Pr(+ \sim D)$	9.77%
False - rate for true D	$\Pr(- D)$	17.29%
False + rate for classified +	$\Pr(\sim D +)$	10.57%
False - rate for classified -	$\Pr(D -)$	16.08%
Correctly classified		86.47%

Il Modello 1 raggiunge un'accuratezza totale dell'86,47% nel primo anno di osservazione. La voce "*Sensitivity*" della Tabella 15 si riferisce agli errori del Tipo I (ovvero quando un'impresa nella realtà è fallita ed il modello la classifica come non fallita), mentre la voce "*Specificity*" si riferisce agli errori del Tipo II (impresa non fallita classificata come fallita).

Paragonando il Modello 1 con il Modello 2, è interessante notare, che il modello con le sole variabili finanziarie raggiunga un'accuratezza minore (tabelle 16 e 17). Abbiamo infatti un'accuratezza totale stimata del Modello 1 all'86,47% contro l'82,71% del Modello 2.

Tabella 16: Risultati della regressione logistica - Modello 2

```
. logit status liquidity ebitass tdta assturn rigidity

Iteration 0:  log likelihood = -184.37715
Iteration 1:  log likelihood = -126.45181
Iteration 2:  log likelihood = -107.43987
Iteration 3:  log likelihood = -102.35344
Iteration 4:  log likelihood = -102.24328
Iteration 5:  log likelihood = -102.24289
Iteration 6:  log likelihood = -102.24289

Logistic regression      Number of obs      =      266
                        LR chi2(5)                   =      164.27
                        Prob > chi2                     =      0.0000
Log likelihood = -102.24289      Pseudo R2          =      0.4455
```

status	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
liquidity	-2.542895	1.357028	-1.87	0.061	-5.202621	.1168315
ebitass	-7.122455	1.687511	-4.22	0.000	-10.42992	-3.814993
tdta	3.773325	.8043015	4.69	0.000	2.196923	5.349727
assturn	-.4488789	.1702067	-2.64	0.008	-.782478	-.1152799
rigidity	1.922173	.8024887	2.40	0.017	.3493242	3.495022
_cons	-2.510457	.7356157	-3.41	0.001	-3.952237	-1.068677

Tabella 17: Risultati della classificazione della regressione logistica – Modello 2

```
. estat classification
```

```
Logistic model for status
```

Classified	True		Total
	D	~D	
+	108	21	129
-	25	112	137
Total	133	133	266

```
Classified + if predicted Pr(D) >= .5
```

```
True D defined as status != 0
```

Sensitivity	Pr(+ D)	81.20%
Specificity	Pr(- ~D)	84.21%
Positive predictive value	Pr(D +)	83.72%
Negative predictive value	Pr(~D -)	81.75%

False + rate for true ~D	Pr(+ ~D)	15.79%
False - rate for true D	Pr(- D)	18.80%
False + rate for classified +	Pr(~D +)	16.28%
False - rate for classified -	Pr(D -)	18.25%

Correctly classified	82.71%
----------------------	--------

3.7 La potenza predittiva del Modello 1

Una volta profilate le variabili ed eseguito la regressione con i dati del primo anno del *training-set*, otteniamo i coefficienti di regressione (Tabella 16, colonna Coef.). Mediante questi ultimi, è possibile estendere l'analisi sia agli anni precedenti al primo, sia al test di validazione *holdout*. Ricordando che i coefficienti risultanti dalle regressioni altro non sono che i *log-odds* della probabilità di fallimento, è possibile invertire la formula e dai coefficienti estrapolare la probabilità di fallimento compresa tra 0 ed 1.

$$p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)}} \quad (2)$$

Il processo di generazione della probabilità p utilizzando la formula scritta sopra è stato eseguito cinque volte: sui tre anni del *test-set*, e sugli anni “t-1” e “t-2” del *training-set*. Il punto di *cut-off* è stato impostato coerentemente con il primo anno di analisi ad un valore $p > 0,5$ per le imprese da classificare come fallite e $p < 0,5$ per le imprese da classificare come attive. I risultati per l'accuratezza e la quantità di errori per le loro due possibili specificazioni sono esposti nella Tabella 18.

Tabella 18: Risultati ai test holdout ed anni precedenti

Training Set				Test Set			
	Accuracy	Type I	Type II		Accuracy	Type I	Type II
<i>Anno t</i>	86.4%	17.3%	9.8%	<i>Anno t</i>	82.0%	8.0%	28.0%
<i>Anno t-1</i>	85.0%	16.5%	13.5%	<i>Anno t-1</i>	74.0%	16.0%	36.0%
<i>Anno t-2</i>	81.6%	21.1%	15.8%	<i>Anno t-2</i>	70.0%	28.0%	32.0%
<i>Sample Size</i>	266			<i>Sample Size</i>	50		

I risultati nella tabella danno un quadro generale della potenza predittiva del modello. Sul *training-set*, sul quale i coefficienti sono stati stimati, l'accuratezza passa dall'86,4% nel primo anno all'81,6% al terzo anno. Sul *test-set* i risultati del primo anno di osservazioni si rilevano di poco inferiori rispetto al *training-set* (82,0% del *test-set* contro gli 86,4% del *training-set* di accuratezza totale). Negli anni successivi al primo, il *test-set* performa decisamente peggio rispetto al training, con uno scarto percentuale che arriva all'11,6% di classificazioni corrette (anno “t-2”).

Nel complesso, si può affermare che il Modello 1 abbia una buona potenza predittiva entro un anno: la percentuale di classificazione corretta è infatti in linea con le percentuali ottenute da altri modelli in letteratura. La regressione logistica conferma dunque la sua validità ad oggi come strumento di discriminazione per eventi discreti.

Capitolo 4: Conclusione

In questo capitolo si discuterà dei risultati raggiunti nella tesi, delle loro implicazioni pratiche e delle limitazioni dello studio. Da ultimo uno spazio sarà ritagliato per la discussione di nuovi spunti di ricerca.

4.1 Discussione dei risultati

In questa tesi si è condotto uno studio sperimentale sulle determinanti del fallimento d'impresa. Si è partiti con un inquadramento puramente teorico del fenomeno del fallimento e si è discusso delle sue conseguenze e della sua rilevanza. Una *review* sistematica della letteratura rilevante al tema, e dei principali *findings* è stata portata avanti, e da questa sono seguite le ipotesi oggetto di ricerca.

Su un *sample* composto da 316 imprese di cui 158 fallite e 158 attive, si è dimostrato che le qualità del CEO sono determinanti del fallimento d'impresa. Nello specifico, due ipotesi sono state formulate e dimostrate, e le stesse sono in linea con i riscontri ottenuti da Ciampi (2015) e in contraddizione con quelli ottenuti da Daily & Danton (1994a) e Chaganti (1985). La variabile CEO-Duality infatti, non solo risulta significativa a livello del 0,1%, ma è negativamente correlata con la probabilità di fallimento (ipotesi 1). Inoltre, all'aumentare degli anni di servizio dello stesso nell'impresa, e quindi dell'esperienza maturata all'interno, diminuisce la probabilità di fallimento (variabile CEO-Tenure, ipotesi 2).

Le altre variabili incluse nel modello 1, si rivelano altamente significative e dall'effetto in linea con gli studi precedenti. Gli indici di debito e di efficienza operativa (rispettivamente *tdta* ed *ebitass*) sono quelli che hanno la maggiore potenza discriminatoria, con *z-value* rispettivamente di 4.23 e -4.42. Ciò implica che in generale che, a parità di altri fattori, le imprese più indebitate hanno più probabilità di fallire e che all'aumentare della remunerazione operativa dell'attivo diminuisce la probabilità di incorrere nel fallimento. Ulteriore conferma in linea con la letteratura è data dalla variabile di liquidità e dall'*asset turnover*, correlate negativamente alla probabilità di fallimento. Da ultimo, l'indice di rigidità ci indica che imprese con un alto valore di attivo immobilizzato hanno più probabilità di incorrere nel fallimento, poiché gli *assets*

liquidabili in tempi brevi al fine di fare fronte alle obbligazioni di debito potrebbero non avere la consistenza sufficiente.

Discutendo della potenza classificatoria, il modello 1 (che unisce le variabili economico-finanziarie a quelle descrittive le qualità del CEO) restituisce risultati molto accurati sul primo anno del *training-set* (86,4% di classificazione corretta). Ciò è in linea con altri risultati che utilizzano la stessa metodologia di regressione (Barboza et al. 2017, Ciampi 2015). La potenza predittiva del modello è stata testata da primo nei due anni precedenti la prima regressione sul *training-set*, ed in seguito sui tre anni del campione di *holdout*. Anche su questi test i risultati sono incoraggianti, indicando una discreta potenza predittiva che raggiunge l'82,0% nel primo anno del *test-set*. È da sottolinearsi inoltre, come l'unione delle variabili descrittive le qualità del CEO a quelle economico finanziarie migliori notevolmente la potenza classificatoria del modello (86,4% contro 82,71%).

Nell'andare indietro negli anni, i *test* su entrambi i *set* indicano una perdita di parte dell'accuratezza del modello. Il motivo principale per questo fenomeno potrebbe essere che nell'andare indietro nel tempo dall'evento fallimento, i valori dei parametri di bilancio per le imprese fallite dovrebbero essere più "vicini" ai valori dei parametri di bilancio delle imprese non fallite, e quindi potrebbero perdere la loro significatività statistica portando a risultati meno accurati (ricordiamo che i *test* sugli anni precedenti al primo sono stati effettuati utilizzando gli stessi coefficienti ottenuti nella regressione del primo anno sul *training-set*).

4.2 Implicazioni Pratiche

Come accennato nel capitolo 1, il fallimento d'impresa è un tema di grande rilevanza nella vita dell'impresa. Tutti gli *stakeholders* che gravitano attorno all'impresa avranno conseguenze diverse ma severe derivanti dall'eventuale fallimento. I creditori perderanno il loro credito, i dipendenti il loro lavoro, gli azionisti la loro ricchezza, ed in generale la società perderà parte del suo benessere derivante dall'avere un'impresa attiva e profittevole.

Poter quindi disporre di un modello caratterizzato da *inputs* costruibili a partire da informazioni di pubblico dominio può sensibilmente mitigare per gli *stakeholders* i rischi derivanti del fallimento d'impresa, fornendo loro una leva per agire in tempo e chiarendo loro quali parametri dovrebbero osservare e tenere sotto controllo.

In tal proposito, lo studio effettuato in questa tesi sottolinea l'importanza della figura del CEO nelle piccole e medie imprese, e suggerisce che per queste avere un CEO anche presidente del consiglio di amministrazione e con una lunga esperienza all'interno dell'azienda alle spalle, mitighi sensibilmente il rischio di fallire. Esso sarà infatti orientato verso strategie di crescita a lungo termine e sostenibili.

Inoltre, il modello 1 ribadisce l'importanza delle variabili finanziarie ai fini della determinazione del fallimento. Le imprese pertanto dovrebbero prestare attenzione al livello del debito accumulato, cercare di massimizzare le vendite minimizzando gli *assets* immobilizzati, e mantenere un livello di cassa o mezzi equivalenti sufficiente al ripagamento del debito a scadenza.

Infine, per i concessionari di credito alle imprese questa tesi suggerisce che includere le variabili di *corporate governance* nei modelli di valutazione del rischio di credito aumenti la precisione della valutazione effettuata.

4.3 Limitazioni

Nonostante questa tesi aggiunga un contributo alla ricerca sul tema del fallimento non è esente da limitazioni. Tre sono le principali limitazioni dello studio qui effettuato.

Da primo è conveniente discutere delle qualità del sample che ha formato la base per lo studio statistico. Il campionamento ha visto coinvolte per lo più piccole e medie imprese. L'*asset-size* medio infatti è attorno ai 30 milioni di euro. Come discusso in precedenza, quando le aziende non sono di grandi dimensioni e non quotate in borsa, vi è una tendenza ad avere la figura dell'imprenditore-azionista, il quale avrà tutto l'interesse al che l'impresa sia ben performante sia nel breve che nel lungo periodo. I conflitti d'interesse tipicamente descritti nella teoria dell'agenzia non sono così dominanti come nel caso delle grandi aziende dall'azionariato frazionato e dove quindi il singolo azionista ha potere decisionale nullo. Questo potrebbe essere motivo per il quale al cambiare della dimensione aziendale del *sample*, portandolo con una incidenza maggiore di aziende di maggiore dimensione, i *findings* sulla figura del CEO evidenziati in questa tesi potrebbero non essere validi. È pur vero però che la dimensione aziendale delle imprese coinvolte nei fallimenti è per lo più piccola e media, e questo resta un dato di fatto.

In secondo luogo, c'è da segnalare il campione è di dimensione limitata. Date le variabili di interesse CEO-duality e CEO-tenure non sempre disponibili sul *database* consultato, ed avendo seguito la procedura di accoppiamento una impresa fallita-una impresa non fallita (pur considerando come criteri l'*asset size* e il settore di appartenenza), il *sample* è risultato di sole 316 imprese in totale. Il modello costruito pertanto, non è generalizzabile a tutto l'universo delle imprese fallite. *Samples* di dimensioni maggiori potrebbero evidenziare altre correlazioni e con altre intensità.

Infine, le imprese considerate in questo studio non comprendono quelle facenti parte del settore finanziario-assicurativo, ma solo manifatturiere e di servizi. Il modello pertanto non risulta applicabile ed estendibile a questo tipo di imprese.

4.4 Spunti per ulteriori ricerche

Questo studio ha tentato di dimostrare la correlazione tra CEO-duality e CEO-tenure, e la probabilità di fallimento. Durante ogni fase di costruzione dell'elaborato sono emerse delle ulteriori domande di ricerca che saranno di seguito discusse.

Anzitutto il campione è composto da imprese caratterizzate da una dimensione media dell'attivo di circa 30 milioni di euro. Sarebbe interessante verificare se le correlazioni evidenziate in questa tesi sono ancora vere in presenza di imprese di maggiore dimensione. Inoltre, le imprese considerate in questo studio sono solo 316. Un ampliamento dell'indagine sul CEO e la probabilità di fallimento potrebbe considerare pertanto un campione di dimensioni maggiori.

Un ulteriore spunto per nuove ricerche può considerarsi l'inserimento nel modello, oltre alle variabili del CEO, di altre variabili di *corporate governance* descrittive le qualità *board*, il tipo di schema adottato, gli organi di controllo e l'azionariato, fornendo quindi un quadro complessivo più chiaro della distribuzione delle responsabilità di gestione e di controllo all'interno dell'impresa e delle implicazioni di queste nel fallimento d'impresa.

Inoltre, il modello 1, del quale i coefficienti sono stati utilizzati per operare la classificazione nei test di *holdout*, non tiene in considerazione informazioni provenienti dal rendiconto finanziario. Includere una o più misure del flusso di cassa potrebbe migliorare l'accuratezza del modello e tenere in conto la dimensione dinamica di ingressi ed uscite di liquidità

Infine, la metodologia statistica utilizzata in questa tesi è la regressione logistica, la quale ha dimostrato un buon livello di accuratezza. Altre tecniche di classificazione, quali regressione Probit, *random forest*, *decision trees*, SVM ed ANN potrebbero portare ad un miglioramento della potenza classificatoria del modello qui costruito.

Sommario

Introduzione

Una delle questioni più cruciali che ogni business ben strutturato ed organizzato deve prima o poi fronteggiare è come valutare l'affidabilità creditizia dei partner.

Ogni qualvolta si dà luogo ad un atto transattivo – luogo giuridico a titolo oneroso - ottenendo in cambio una promessa di pagamento, ci si pone la domanda di “quanto” questa promessa sia affidabile, e “come” valutare l'affidabilità della stessa

La ricerca delle cause del fallimento e delle sue determinanti è oggetto di numerosi studi che si sono susseguiti dal 1960 in poi. Seguendo l'approccio dei lavori più accreditati e la tecnica statistica più utilizzata in letteratura, questa tesi si pone come obiettivo finale di dimostrare il nesso tra le qualità del CEO e la probabilità di fallimento, e di valutare la potenza della classificazione operata del modello creato.

Capitolo 1: Il fallimento

Ogni impresa nasce, cresce e muore. Questo concetto è a grandi linee espresso dalla teoria del “*Business Life Cycle*”. La teoria, che vede come primo grande contributore Chandler (1962), descrive il processo di vita di un'impresa come *continuum* su quattro fasi distinte.

La prima fase è quella del lancio del prodotto o del servizio. L'impresa è appena nata ed è tipicamente di ridotte dimensioni e con una struttura di governance informale. Il problema principale in questa fase è di ottenere clienti e stabilire come fare arrivare loro il prodotto o servizio lanciato. In questa fase, gli ingenti investimenti di start-up e la scarsa base di utenza, le perdite di esercizio si accumulano e risulta difficoltoso reperire finanziamenti in quanto il business risulta rischioso e immaturo agli occhi dei finanziatori. Questa situazione può degenerare fino al verificarsi di problemi di cassa che possono risolversi nel fallimento d'impresa.

Nella seconda fase, detta di crescita, l'impresa, se sopravvissuta alla prima fase, sperimenterà una rapida crescita delle vendite e auspicabilmente avrà profitti positivi. L'impresa è pertanto tentata ad aumentare velocemente la dimensione del business tramite incremento degli investimenti e specularmente del capitale di debito. La decisione ha il potenziale di sfociare in difficoltà finanziarie qualora la prevista tendenza di crescita delle vendite dovesse invertirsi o non realizzarsi al livello desiderato.

Nella terza fase, ovvero della maturità, le imprese sopravvissute sono caratterizzate dall'aver vendite e profitti stabili o leggermente decrescenti. Il focus principale in questa fase sta nel non perdere quote di mercato, e possibilmente nello sfociare in mercati nuovi e/o con prodotti diversi.

L'ultima fase è quella del declino. Qui le vendite, i profitti ed i flussi di cassa sono stabilmente decrescenti e l'impresa agisce sotto difficoltà finanziarie ed organizzative. Se fallisce nel reinventarsi, l'impresa dovrà uscire dal mercato, con conseguente fallimento o liquidazione.

L'esatta definizione di "*financial distress*" non è ancora ben definita in letteratura (Shi e Li, 2019). In generale, il fallimento viene definito come lo stato in cui un'impresa o un imprenditore individuale, risulta insolvente rispetto le sue obbligazioni alle loro rispettive scadenze.

A livello operativo, un'impresa è considerata fallita al verificarsi di uno o più dei seguenti eventi: bancarotta, bond default, mancato rientro da scoperto bancario, o un mancato pagamento di dividendi su azioni privilegiate (Beaver, 1966). Pertanto, in questa sede adotteremo la definizione di fallimento operativa, che riteniamo più calzante ai fini di una valutazione basata su indicatori di bilancio.

Come postulato dalla teoria del *business life cycle*, il fallimento può essere un passaggio fisiologico della vita dell'impresa. Come illustrato nel primo paragrafo di questo capitolo, il passaggio da un'impresa sana ad un'impresa fallita è caratterizzato da una fase intermedia di declino. Damodaran (2009) sintetizza i tratti comuni alle

imprese in declino in cinque caratteristiche, specificando che non tutte le imprese in declino possiedono tutte e cinque le caratteristiche ma condividono abbastanza di queste ed in molti casi per permettere la generalizzazione. Le cinque caratteristiche sono: a) Vendite stagnanti o decrescenti, b) Margini operativi ristretti o negativi, c) Disinvestimenti, d) Distribuzione massiccia di dividendi e *buybacks* ed e) Financial Leverage elevato.

In aggiunta alle cinque caratteristiche individuate da Damodaran, le quali riguardano la prospettiva interna e prettamente finanziaria delle le imprese in fase di declino, alcuni filoni della letteratura sul tema dividono in cause di fallimento esterne all'impresa, cause interne e manageriali.

Argenti (1976) identifica le cause interne con le seguenti: assenza di una risposta adeguata ad uno *shift* tecnologico, comunicazione tra gli attori interni inadeguata, frode, inadeguata valutazione dei fattori di costo, conoscenza insufficiente a livello finanziario, e alto *leverage*.

Ooghe & Prijcker (2008) delineano un framework concettuale composto da cinque *clusters* di fattori sia interni che esterni (ambiente immediato, ambiente generale, management, politiche interne e caratteristiche specifiche dell'impresa) ed associano ad ogni fase di vita dell'impresa diverse determinanti della bancarotta.

In generale, ogni impresa cercherà a tutti i costi di evitare la situazione fallimentare. Ad ogni modo, il rischio sopportato per chi concede liquidità, è essenzialmente il fallimento del debitore, e dunque la (parziale o totale a seconda della effettiva consistenza patrimoniale del fallito) inesigibilità del credito. Il rischio in capo al creditore è tanto più significativo e concreto quanto più esso risulta esposto in termini di capitale al rischio di credito delle controparti e varia a seconda delle modalità con le quali è esposto.

Il problema del fallimento è di rilievo non soltanto per gli istituti creditizi, ma anche per l'impresa stessa e i suoi azionisti o soci, per i lavoratori dipendenti, e più in generale per tutti gli *stakeholders* che hanno come centro di gravità l'impresa in questione.

Scopo di questa tesi di laurea è di indagare con uno studio sperimentale il nesso tra variabili di *corporate governance* ed economico-finanziarie ed il rischio di fallimento.

Nello specifico, delle ipotesi verranno esplicitate e testate sull'incidenza delle variabili descrittive le qualità del CEO sul rischio di bancarotta. Da ultimo si discuterà della potenza classificatoria del modello costruito e della sua forza predittiva con un test *holdout*.

Capitolo 2: I Modelli Predittivi, la Corporate Governance, il CEO

Esiste una grande varietà di modelli predittivi del fallimento i quali si differenziano per metodologia statistica utilizzata, per fattori esplicativi impiegati, aree geografiche e di industries considerate. La divisione operata in questa tesi vede: a) Modelli statistici Tradizionali, b) Metodi di Machine Learning e Artificial Intelligence e c) Modelli Ibridi di *Corporate Governance*

Come indicato nella *Systematic Literature Review* effettuata da Shi e Li (2019), i metodi statistici tradizionali sono largamente più impiegati rispetto ai metodi di ML & AI. Abbiamo infatti che il metodo più popolare (impiegato in 123 *papers* su 321 analizzati) è la regressione logistica (Logit) con il 38,3% del totale, al secondo posto troviamo il metodo delle Reti Neurali (17,5%) e da terzo l'analisi dei discriminanti (16,1%).

I modelli predittivi del fallimento vengono stimati su un campione di imprese fallite, alle quali vengono accoppiate le imprese (una o più) non fallite al fine di riscontrare differenze strutturali tra i due gruppi. L'obiettivo è poter utilizzare il modello stimato sul campione originale per fare previsioni su imprese che *non* appartengono a quest'ultimo. Le due tecniche di validazione principali sono il metodo Jackknife e il metodo dell'*holdout*. Tramite il primo, la stessa grandezza statistica stimata viene ricalcolata n volte, lasciando fuori dall'insieme di campionamento una osservazione alla volta. Tramite il secondo invece, i coefficienti stimati nelle regressioni sul *training-set* si applicano al *test-set* (che è indipendente e contiene osservazioni diverse dal *training-set*) e si opera la classificazione.

Nei metodi statistici tradizionali rientrano gli studi di Beaver (1966), Altman (1968 e 1977) ed Ohlson (1980).

Beaver (1966) scrive il primo articolo di grande rilevanza sul tema; stila una lista di 30 indicatori finanziari (si veda la Tabella 2 a pagina 20), descrittivi ogni area dell'attività aziendale. Tramite l'analisi delle medie e dell'analisi statistica univariata, Beaver trova che l'indice che meglio discrimina tra i due gruppi di imprese risulta essere Flusso Finanziario / Totale debito, mentre il secondo Utile Netto / Totale Attivo.

Altman (1968) è il primo ad utilizzare l'analisi statistica multivariata tramite l'analisi dei discriminanti multipli (MDA). L'autore, dopo aver compilato e testato la potenza discriminatoria di una lista di 22 variabili, arriva ad una funzione discriminatoria a cinque variabili (si veda pagina 20 per la descrizione di queste ultime) detta Z-Score. La funzione restituisce un numero reale che va da -4 a +8: se un'impresa ottiene uno score inferiore a 1,81 verrà classificata come fallita; per uno score superiore a 2,99 l'impresa è classificata come non fallita; tra 1,81 e 2,99 ricadono la maggior parte degli errori di classificazione e pertanto l'intervallo in questione viene definito come “*zone of ignorance*”. Il modello raggiunge un'accuratezza del 95% sul *training-set* e del 96% sul *test-set* (riferito alla classificazione delle imprese fallite).

Il lavoro di Ohlson (1980), si contrappone per metodologia statistica a quello di Altman (1968). Esso infatti utilizza la regressione logistica, contrapponendola all'MDA. La regressione logistica stima la probabilità che un'impresa fallisca, ed è di più facile interpretazione rispetto allo Z-score di Altman. L'autore opta per un modello a 9 variabili (si veda pagina 26). Il suo modello raggiunge un'accuratezza del 96,12% al primo anno, e del 95,55% al secondo anno di osservazione.

Tra i metodi di Machine Learning ed Intelligenza Artificiale (ML & AI), si annoverano: a) *Case-based Reasoning*, b) *Support Vector Machines (SVM)*, c) *Bagging*, d) *Boosting*, e) *Random Forest*, ed f) Reti Neurali (per una descrizione di ogni metodo si rimanda alla digressione da pagina 29 a pagina 33). Kumar & Ravi (2007) tramite una *systematic literature review* trovano che i metodi di ML & AI sovraperformano i metodi statistici. Tsai et al. (2014), sperimentano sullo stesso *set* di osservazioni i metodi *SVM-bagging*, *SVM-boosting* e *Decision Trees*: concludono che il miglior metodo è il *decision tree* con un'accuratezza dell'86%. Heo & Yang (2014) paragonano *boosting*, ANN, SVM, *decision trees* e lo Z-score

di Altman: concludono che il metodo più preciso è il *boosting* con un'accuratezza del 93,8%. Danenas & Garsva (2015) stimano un'accuratezza del 93% utilizzando il metodo SVM. Barboza, Kimura & Altman (2017) ottengono un'accuratezza dell'86,31% e 86,04% con *boosting* e *random forest*. In ultima battuta, i metodi di ML & AI si rivelano molto precisi, ma richiedono *data-sets* molto estesi e grande potenza computazionale.

Un filone di ricerca che affianca variabili finanziarie e variabili non finanziarie è quello riguardante la *corporate governance*. Chaganti et al. (1985), mettono in relazione *board size*, il numero di amministratori indipendenti e la CEO-duality con la probabilità di fallimento: trovano che tutte le variabili sono insignificanti eccetto *board size*, la quale risulta negativamente correlata con il fallimento. Hambrick & D'Aveni (1992), sottolineano come lo stile di *leadership* dominante sia correlato positivamente con la probabilità di bancarotta. Daily & Danton (1994a) trovano che la CEO-duality è positivamente correlata al fallimento. Ciampi (2015) trova una correlazione negativa tra CEO-duality e fallimento, ed una correlazione positiva tra gli amministratori in dipendenti presenti nel board in misura superiore al 50% e fallimento. Elloumi & Gueyié (2001) concludono che includere indicatori di corporate governance affiancandoli a quelli finanziari, aumenta la precisione del loro modello (72,1% contro 70,6%). Non vi è dunque unanimità di risultati di correlazione come per i modelli che utilizzano le sole variabili finanziarie (su tutte CEO-duality); il fenomeno può essere dovuto alla diversa dimensione delle imprese considerate nei singoli modelli.

Capitolo 3: Lo Studio Empirico

Notando la discordanza di risultati in letteratura sulle variabili di *corporate governance*, ma avendo l'unità di risultato che indica che l'aggiunta di queste variabili migliora l'affidabilità del modello, siamo ora interessati a capire che incidenza abbiano le variabili descrittive delle qualità del CEO sulla probabilità di fallimento. Due ipotesi sono introdotte:

Ipotesi 1: All'aumentare della tenure del CEO in azienda diminuisce la probabilità di fallimento e,

Ipotesi 2: La CEO-duality (quindi CEO anche presidente del board of directors) è negativamente correlata alla probabilità di fallimento.

Il *sample* oggetto di studio si compone di 316 imprese europee (anni 2013-2018), equamente divise tra fallite e non fallite. Per ogni impresa si sono raccolti tre anni di dati: per l'anno del fallimento e per i due anni precedenti a quest'ultimo. La dimensione media dell'attivo delle imprese del campione risulta intorno ai 30 milioni di euro. Il *sample* è stato in seguito diviso in *training-set*, contenente 133 imprese fallite e 133 imprese non fallite sulle quali stimare i parametri del modello, e *test-set* contenente 25 imprese fallite e 25 non fallite per poter applicare i coefficienti ottenuti nel *training* al fine di testare la potenza predittiva del modello.

La variabile indipendente è Status, codificata con 1 se l'impresa è fallita, e con 0 se è attiva. Le variabili indipendenti sono CEO-Tenure e CEO-duality. La prima descrive il tempo intercorrente tra la nomina del CEO e il tempo di raccolta dati; la seconda assume valore 1 se il CEO è anche presidente del Board e 0 altrimenti.

Le variabili di controllo (si vedano le pagine 48, 49 e 50 per il procedimento di profilazione di queste), risultano:

- 1) Liquidity: Misura di liquidità risultante dal quoziente "Cassa e mezzi equivalenti / Passività Correnti";
- 2) EBIT / Assets: Misura di efficienza operativa che indica quanto l'impresa è efficiente nel "trasformare" gli assets in risultato operativo;

- 3) Asset Turnover: Misura di competitività risultante dal quoziente “Vendite / Totale Attivo”. L’indice esprime quante vendite vengono effettuate per ogni euro in assets:
- 4) Totale Debito / Totale Attivo: Misura di struttura finanziaria;
- 5) Rigidity Index: risulta da “Attivo Non Corrente / Totale Attivo” ed è misura di struttura dell’attivo

Similmente alla maggior parte degli studi precedenti, si è scelto di effettuare le regressioni mediante la regressione logistica. La scelta è dettata anche dalla semplicità applicativa del metodo ed alla facile interpretazione dei risultati.

Da primo, sono state effettuate due regressioni con entrambe le variabili indipendenti (si vedano le tabelle 12 e 13). I risultati confermano le due ipotesi. In particolare, per entrambe le regressioni operate prima sulla variabile CEO-Tenure e sulla variabile CEO-Duality poi, risultano coefficienti con segno negativo, ed una significatività rispettivamente all’1% ed allo 0,1%. Questo conferma che:

- 1) La CEO-duality è correlata negativamente con la probabilità di fallimento e
- 2) All’aumentare della CEO-Tenure diminuisce la probabilità di fallimento.

In seconda battuta, sono stati creati due modelli: il primo (Modello 1) unendo le variabili del CEO a quelle finanziarie, ed il secondo utilizzando le sole variabili finanziarie (Modello 2). Tutte le variabili inserite nel Modello 1 e nel Modello 2 sono considerate come regressori indipendenti.

I risultati di classificazione dei due modelli, confermano quanto trovato in studi precedenti, ovvero che aggiungere le variabili di *Corporate Governance* aumenta la precisione della classificazione operata dalla regressione (86,47% di classificazione corretta del Modello 1 contro l’82,71% del Modello 2, si vedano le tabelle 14, 15, 16 e 17).

Da ultimo, al fine di validare il Modello 1, è stato eseguito un test di *holdout*. Il test di *holdout* consiste nel riapplicare i coefficienti ottenuti nella regressione del primo anno del *training-set* (Modello 1) sia agli anni precedenti dello stesso *set*, sia al campione del *test-set* per tutti e tre gli anni (tabella 18). I risultati del test consentono di affermare nel complesso che il Modello 1 è dotato di una buona forza

predittiva all'anno del fallimento (82,0% di classificazione corretta nel primo anno del test *holdout*).

Capitolo 4: Conclusione

In questa tesi si è dimostrato che le qualità del CEO sono determinanti del fallimento d'impresa. Due ipotesi sono state ricercate e dimostrate, e le stesse sono in linea con i riscontri ottenuti da Ciampi (2015) e in contraddizione con quelli ottenuti da Daily & Danton (1994a) e Chaganti (1985). La variabile CEO-Duality infatti, non solo risulta significativa a livello del 0,1%, ma è negativamente correlata con la probabilità di fallimento (ipotesi 1). Inoltre, all'aumentare degli anni di servizio dello stesso nell'impresa, e quindi dell'esperienza maturata all'interno, diminuisce la probabilità di fallimento (variabile CEO-Tenure, ipotesi 2).

Per quanto concerne la classificazione operata e la potenza predittiva del Modello 1, questo è molto accurato sul primo anno del *training-set* (86,4% di classificazione corretta). La potenza predittiva del modello è stata testata da primo nei due anni precedenti la prima regressione sul *training-set*, ed in seguito sui tre anni del campione di *holdout* (tabella 18). Anche su questi test i risultati sono incoraggianti, indicando una discreta potenza predittiva che raggiunge l'82,0% nel primo anno sul *test-set*.

Le variabili finanziarie incluse nel modello 1 si rivelano significative e dall'effetto in linea con gli studi precedenti. Gli indici di debito e di efficienza operativa sono quelli che offrono la maggiore potenza discriminatoria, indicando che in generale le imprese dovrebbero costantemente tenere sotto controllo il livello del debito ed avere particolare riguardo nell'aver una remunerazione operativa dell'attivo sufficiente. Ulteriore conferma in linea con la letteratura è data dalla variabile di liquidità e dall'asset turnover, correlate negativamente alla probabilità di fallimento. Da ultimo, l'indice di rigidità ci indica che imprese con un alto valore di attivo immobilizzato hanno più probabilità di incorrere nel fallimento.

Come accennato nel capitolo 1, il fallimento d'impresa è un tema di grande rilevanza nella vita dell'impresa. Tutti gli stakeholders che gravitano attorno all'impresa avranno conseguenze diverse ma severe derivanti dall'eventuale fallimento. Poter disporre di un modello caratterizzato da *inputs* costruibili a partire da informazioni di pubblico dominio può sensibilmente mitigare i rischi derivanti del fallimento d'impresa, fornendo agli *stakeholders* una leva per agire in tempo e chiarendo loro quali parametri dovrebbero osservare e tenere sotto controllo.

A tal proposito, lo studio effettuato in questa tesi sottolinea l'importanza della figura del CEO nelle piccole e medie imprese, e suggerisce che per queste avere un CEO anche presidente (e azionista di maggioranza) e con una lunga esperienza all'interno dell'azienda alle spalle, mitighi sensibilmente il rischio di fallire. Inoltre, il modello 1, ribadendo l'importanza delle variabili finanziarie ai fini della determinazione del fallimento, suggerisce che le imprese dovrebbero prestare attenzione al livello del debito accumulato, cercare di massimizzare le vendite minimizzando gli assets immobilizzati, e mantenere un livello di cassa o mezzi equivalenti sufficiente al ripagamento del debito a scadenza. Da ultimo, per i concessionari di credito alle imprese questa tesi indica che includere le variabili di corporate governance nei modelli di valutazione del rischio di credito aumenti la precisione della valutazione effettuata.

Tre sono le principali limitazioni dello studio: da primo, la dimensione qualitativa del campione. Esso infatti è composto da imprese con un *asset-size* medio di 30 milioni di euro, che le qualifica come imprese medio-piccole. I riscontri sulla figura del CEO qui ottenuti potrebbero non essere più validi su un campione caratterizzato da imprese per lo più grandi. In secondo luogo, il *sample* è di dimensione limitata. *Samples* con numerosità maggiori potrebbero portare ad evidenziare altre correlazioni e con diverse intensità. Infine, le imprese considerate nel campione non appartengono al settore finanziario-assicurativo e pertanto il modello non è estendibile a questa tipologia di imprese.

In fase di costruzione dell'elaborato sono emerse delle ulteriori domande di ricerca che saranno di seguito discusse. Anzitutto dato il campione di indagine caratterizzato dalla prevalenza di imprese medie e piccole, sarebbe interessante verificare se le correlazioni evidenziate in questa tesi sono ancora vere in presenza

di imprese di maggiore dimensione. Secondo, un ampliamento dell'indagine sul ruolo del CEO e la probabilità di fallimento potrebbe considerare un campione di dimensioni numericamente maggiori. Un ulteriore spunto per nuove ricerche può considerarsi l'inserimento nel modello, oltre alle variabili del CEO, di altre variabili di corporate governance descrittive le qualità board, il tipo di schema adottato, gli organi di controllo e l'azionariato, fornendo quindi un quadro complessivo più chiaro della distribuzione delle responsabilità di gestione e di controllo all'interno dell'impresa e delle implicazioni di queste nel fallimento d'impresa. Inoltre, includere una o più misure del flusso di cassa nel modello potrebbe migliorare l'accuratezza del modello e tenere in conto la dimensione dinamica di ingressi ed uscite di liquidità. Infine, utilizzare altre tecniche di classificazione, quali regressione Probit, *random forest*, *decision trees*, SVM ed ANN potrebbero portare ad un miglioramento della potenza classificatoria del modello qui costruito.

Bibliografía

Aamodt, E. Plaza (1994); Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. AI Communications. IOS Press, Vol. 7: 1, pp. 39-59.

Alaminos D, Del Castillo A, Fernández MÁ. A Global Model for Bankruptcy Prediction [published correction appears in PLoS One. 2018 Nov 28;13(11):e0208476]. PLoS One. 2016;11(11):e0166693. Published 2016 Nov 23. doi:10.1371/journal.pone.0166693

Altman, E. 1968: "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy". The Journal of Finance 23(4): 589-609.

Altman, Edward I. & Haldeman, Robert G. & Narayanan, P., 1977. "ZETA analysis A new model to identify bankruptcy risk of corporations," Journal of Banking & Finance, Elsevier, vol. 1(1), pages 29-54, June.

Argenti J. (1976): "Corporate Collapse: The Causes and Symptoms". London, UK: McGraw-Hill Book Company (UK) Ltd.

Beaver, W. (1966): "Financial Ratios As Predictors of Failure". Journal of Accounting Research, 4, 71-111. doi:10.2307/2490171

Chandler, A.D. (1962): "Strategy and Structure: Chapters in the History of American Enterprise". MIT Press, Boston.

Ciampi, Francesco, 2015. "Corporate governance characteristics and default prediction modeling for small enterprises. An empirical analysis of Italian firms," Journal of Business Research, Elsevier, vol. 68(5), pages 1012-1025.

Cultrera L., Brédart X., 2015: "Bankruptcy Prediction: The case of Belgian SMEs". Review of Accounting and Finance 15(1): 101-119, DOI: 10.1108/RAF-06-2014-0059

Daily, C.M., & Dalton, D.R. (1994a). Bankruptcy and corporate governance: The impact of board composition and structure. *Academy of Management Journal*, 37(6), 1603–1617.

Daily, C.M., & Dalton, D.R. (1994b). Corporate governance and the bankrupt firm: An empirical assessment. *Strategic Management Journal*, 15(8), 643–654.

Damodaran, Aswath: “Valuing Declining and Distressed Companies” (June 23, 2009). Available at
SSRN: <https://ssrn.com/abstract=1428022> or <http://dx.doi.org/10.2139/ssrn.1428022>

Danenas, P. and Garsva, G. (2015) Selection of Support Vector Machines Based Classifiers for Credit Risk Domain. *Expert Systems with Applications*, 42, 3194-3204.

<https://doi.org/10.1016/j.eswa.2014.12.001>

Edmister, R. 1972: “An empirical test of financial ratio analysis for small business failure prediction”. *Journal of Financial and Quantitative Analysis* 7(2): 1477-1493.

Elloumi, F., & Gueyié, J.P. (2001). Financial distress and corporate governance: An empirical analysis. *Corporate Governance: The International Journal of Business in Society*, 1(1), 15–23.

Figini S., Savona R., Vezzoli M. (2016): “Corporate Default Prediction Model Averaging: A Normative Linear Pooling Approach”. Volume23, Issue1-2

Flavio Barboza, Herbert Kimura, Edward Altman, 2017. “Machine learning models and bankruptcy prediction”. *Expert Syst. Appl.* 83, C (October 2017), 405-417. Doi: <https://doi.org/10.1016/j.eswa.2017.04.006>

Gissel, Jodi L.; Giacomino, Don; and Akers, Michael D., "A Review of Bankruptcy Prediction Studies: 1930-Present" (2007). *Accounting Faculty Research and Publications*. 25.

Hambrick, Donald C., and Richard A. D'Aveni. "Top Team Deterioration As Part of the Downward Spiral of Large Corporate Bankruptcies." *Management Science*, vol. 38, no. 10, 1992, pp. 1445–1466.

Heo, J. , & Yang, J. Y. (2014). AdaBoost based bankruptcy forecasting of Korean construction companies. *Applied Soft Computing*, 24 , 4 94–4 99 .

Jensen, Michael C. and Meckling, William H., *Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure* (July 1, 1976). Harvard University Press, December 2000, *Journal of Financial Economics (JFE)*, Vol. 3, No. 4, 1976.

Jones, F. 1987. "Current techniques in bankruptcy prediction". *Journal of Accounting Literature* 6: 131-164.

Koh, H. 1987. "Prediction of going-concern status: A probit model for the auditors". Ph.D. dissertation, Virginia Polytechnic Institute and State University.

Kolodner L., 1993: "Case-based Learning", Morgan Kaufmann Publishers

Micheal Kearns (1988); "Thoughts on Hypothesis Boosting", Unpublished manuscript (Machine Learning class project, December 1988).

Michael Kearns; Leslie Valiant (1989). Cryptographic [sic] limitations on learning Boolean formulae and finite automata. *Symposium on Theory of Computing*. 21. ACM. pp. 433–444. doi:10.1145/73007.73049. ISBN 978-0897913072.

Noble, W. What is a support vector machine?. *Nat Biotechnol* 24, 1565–1567 (2006). <https://doi.org/10.1038/nbt1206-1565>

Ohlson, J.A. (1980) "Financial Ratios and the Probabilistic Prediction of Bankruptcy". *Journal of Accounting Research*, 18, 109-131. <http://dx.doi.org/10.2307/2490395>

Ooghe, H. and De Prijcker, S., 2008: "Failure processes and causes of company bankruptcy: a typology". *Management decision*.

P. Viola, M.Jones, “Robust Real-time Object Detection”, 2001

Ravi Kumar P, Ravi V. 2007. Bankruptcy prediction in banks and firms via statistical and intelligent techniques. *Eur J Oper Res* 180: 1-28.

Shi, Y., & Li, X. (2019): “An overview of bankruptcy prediction models for corporate firms: A systematic literature review”. *Intangible Capital*, 15(2), 114-127.

<https://doi.org/10.3926/ic.1354>

Sinkey, Jr., J. 1975: “A multivariate statistical analysis of the characteristics of problem banks”. *Journal of Finance* 30(1): 21-36.

Sun, J., Li, H., Huang, Q., & He, K. (2014): “Predicting financial distress and corporate failure: A review from the state-of-the-art definitions, modeling, sampling, and featuring approaches” *knowledge-Based Systems*, 57, 41-56.

<https://doi.org/10.1016/j.knosys.2013.12.006>

Süsi, V. and Lukason, O. (2019), "Corporate governance and failure risk: evidence from Estonian SME population", *Management Research Review*, Vol. 42 No. 6, pp. 703-720

Taffler, Richard J., 1984. "Empirical models for the monitoring of UK corporations," *Journal of Banking & Finance*, Elsevier, vol. 8(2), pages 199-227, June.

Tian, Y. , Shi, Y. , & Liu, X. (2012). Recent advances on support vector machines research. *Technological and Economic Development of Economy*, 18 (1), 5–33 .

Tsai, C.-F. , Hsu, Y.-F. , & Yen, D. C. (2014). A comparative study of classifier ensembles for bankruptcy prediction. *Applied Soft Computing*, 24 , 977–984 .

Wang, B. 2004: “Strategy changes and internet firm survival”. Ph.D. dissertation, University of Minnesota.

Zhao, Z. , Xu, S. , Kang, B. H. , Kabir, M. M. J. , Liu, Y. , & Wasinger, R. (2014). Investi- gation and improvement of multi-layer perception neural networks for credit scoring. *Expert Systems with Applications*, in press.

Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*, 22, 59–82.