



Università
Ca' Foscari
Venezia

Corso di Laurea Magistrale
in Economia e Finanza

Tesi di Laurea

—
Ca' Foscari
Dorsoduro 3246
30123 Venezia

Data Envelopment Analysis come strumento di credit scoring anche per le PMI

Relatore

Ch. Prof. Antonella Basso

Laureando

Luca Cason

Matricola 816279

Anno Accademico

2011 / 2012

INDICE

<i>Introduzione</i>	8
Capitolo 1	
<i>Il Credit Scoring</i>	11
1.1 Credit scoring: cenni storici.....	12
1.2 Dal credito al consumo al credito alle imprese.....	15
1.3 Gli obiettivi del Credit Scoring.....	16
1.4 Le fasi del Credit Scoring.....	18
Capitolo 2	
<i>Modelli di previsione delle insolvenze</i>	22
2.1 Analisi discriminante lineare.....	22
2.2 Modello Z-Score.....	26
2.3 Modello Lineare di Probabilità.....	31
2.4 Modello Logit e Modello Probit.....	33
2.5 Campi di applicazione.....	38
2.6 Critiche e limiti dei modelli di scoring.....	40
Capitolo 3	
<i>Altri Approcci: il contributo dell'intelligenza artificiale</i>	42
3.1 Sistemi Esperti.....	42
3.2 Alberi decisionali.....	47
3.3 Reti Neurali Artificiali.....	50
3.4 Algoritmi genetici.....	55

Capitolo 4

Data Envelopment Analysis: un metodo alternativo per la valutazione del Merito

Creditizio 60

4.1 Data Envelopment Analysis.....	60
4.2 Frontiere produttive: Teorica ed Empirica.....	62
4.3 Il Modello CRS (Constant Returns to Scale).....	66
4.4 Il modello VRS (Variable Returns to Scale).....	75
4.5 Il Modello Additivo o Modello Pareto – Koopmans (PK).....	77
4.6 Altri modelli DEA.....	80
4.7 I modelli di previsione dei fallimenti aziendali.....	81
4.7.1 Il modello Z-Score di Altman.....	81
4.7.2 Modelli di regressione: Logit e Probit.....	82
4.7.3 Il metodo CUSUM.....	83
4.7.4 L’approccio CAOS.....	84
4.7.5 La metodologia DEA.....	85

Capitolo 5

Caso applicativo: DEA come strumento d’analisi della probabilità d’insolvenza delle piccole e medie imprese

88

5.1 Il campione di riferimento.....	88
5.2 Frontiera efficiente e Frontiera dei fallimenti.....	93
5.3 La selezione delle variabili.....	95
5.4 Le variabili di Input del modello per il calcolo della frontiera dei fallimenti.....	97
5.5 Le variabili di Output per la frontiera dei fallimenti.....	99
5.6 Modelli e Risultati dell’analisi.....	100
5.6.1 Modello a otto variabili.....	101
5.6.2 Modello a nove variabili: prima estensione.....	107
5.6.3 Modello a nove variabili: seconda estensione.....	112
5.6.4 Confronto tra i modelli.....	117

<i>Conclusioni</i>	119
<i>Appendice A - Variabili di input ed output impiegate nell'indagine</i>	121
<i>Appendice B – Risultati</i>	128
<i>Riferimenti bibliografici</i>	142
<i>Sitografia</i>	147

ELENCO DELLE FIGURE

FIGURA 2.1: Grafico dell'Analisi Discriminante Lineare.....	25
FIGURA 2.2: Modello Lineare e Modello Logistico.....	35
FIGURA 2.3: Confronto delle funzione di densità di probabilità del Modello Logit e del Modello Probit.....	36
FIGURA 2.4: Funzioni di ripartizione del Modello Logit e del Modello Probit.....	37
FIGURA 3.1: Schema semplificato di un sistema esperto.....	43
FIGURA 3.2: Albero Decisionale semplificato.....	47
FIGURA 3.3: Schema di un neurone artificiale j-esimo.....	51
FIGURA 3.4: Esempio di rete neurale a due strati.....	53
FIGURA 3.5: La struttura di un Algoritmo Genetico.....	58
FIGURA 4.1: Esempio di due Frontiere Produttive.....	62
FIGURA 4.2: Frontiera Teorica e Frontiera Empirica.....	63
FIGURA 4.3: Un semplice esempio di 'peer group'.....	64
FIGURA 4.4: Frontiere del Modello CRS e del Modello VRS.....	65
FIGURA 4.5: Modello CRS nel caso un input – un output.....	71
FIGURA 4.6: Modello CRS nel caso due input – un output.....	73
FIGURA 4.7: Proiezione di un'unità decisionale impropriamente invilupata.....	74
FIGURA 4.8: Modello VRS nel caso un input – un output.....	76
FIGURA 4.9: Modello Additivo nel caso un input – un output.....	78
FIGURA 5.1: Percentuali di previsioni corrette nel modello input-oriented a otto variabili.....	103
FIGURA 5.2: Percentuali di previsioni corrette nel modello output-oriented a otto variabili.....	104
FIGURA 5.3: Percentuali di previsioni corrette nel modello non oriented a otto variabili.....	105
FIGURA 5.4: Percentuali di previsioni corrette con il modello input-oriented a nove variabili.....	108

FIGURA 5.5: Percentuali di previsioni corrette con il modello output-oriented a nove variabili.....	109
FIGURA 5.6: Percentuali di previsioni corrette con il modello non oriented a nove variabili.....	110
FIGURA 5.7: Percentuali di previsioni corrette con il modello input-oriented a nove variabili (seconda estensione).....	113
FIGURA 5.8: Percentuali di previsioni corrette con il modello output-oriented a nove variabili (seconda estensione).....	115
FIGURA 5.9: Percentuali di previsioni corrette con il modello non oriented a nove variabili (seconda estensione).....	116

ELENCO DELLE TABELLE

TABELLA 2.1: Capacità diagnostica del modello di Altman al periodo t-1(dati in %)	29
TABELLA 4.1: Caso un input - un output	71
TABELLA 4.2: Caso due input - un output	72
TABELLA 5.1: Pil Regionale rispetto al Pil Totale (anno 2009, dati in mln di euro)	90
TABELLA 5.2: Totale imprese del campione suddivise per macro – attività	91
TABELLA 5.3: Variabili di Input e di Output utilizzate per il calcolo della frontiera dei fallimenti	96
TABELLA 5.4: Percentuali di previsioni corrette nel modello input-oriented a otto variabili	102
TABELLA 5.5: Percentuali di previsioni corrette nel modello output-oriented a otto variabili	104
TABELLA 5.6: Percentuali di previsioni corrette nel modello non oriented a otto variabili	105
TABELLA 5.7: Confronto tra i risultati del modello a otto variabili e quelli del modello Premachandra 2011	107
TABELLA 5.8: Percentuali di previsioni corrette con il modello input-oriented a nove variabili	108
TABELLA 5.9: Percentuali di previsioni corrette con il modello output-oriented a nove variabili	109
TABELLA 5.10: Percentuali di previsioni corrette con il modello non oriented a nove variabili	110
TABELLA 5.11: Confronto tra i risultati del modello a nove variabili (prima estensione) e quelli del modello Premachandra 2011	111
TABELLA 5.12: Percentuali di previsioni corrette con il modello input-oriented a nove variabili (seconda estensione)	113
TABELLA 5.13: Percentuali di previsioni corrette con il modello output-oriented a nove variabili (seconda estensione)	114

TABELLA 5.14: Percentuali di previsioni corrette con il modello non oriented a nove variabili (seconda estensione).....	115
TABELLA 5.15: Confronto tra i risultati del modello a nove variabili (seconda estensione) e quelli del modello Premachandra 2011.....	117
TABELLA A.1: Valori degli indici finanziari utilizzati come variabili di input e di output nei tre modelli.....	122
TABELLA B.1: Score di efficienza DEA e numero di imprese fallite per il modello output-oriented ad otto variabili, per diversi valori di λ	129
TABELLA B.2: Score di efficienza DEA e numero di imprese fallite per il modello output-oriented (prima estensione), per diversi valori di λ	133
TABELLA B.3: Score di efficienza DEA e numero di imprese fallite per il modello output-oriented (seconda estensione), per diversi valori di λ	137

INTRODUZIONE

L'obiettivo che mi sono preposto all'inizio di questo lavoro è stato quello di dimostrare come la metodologia DEA possa essere considerata dall'analista finanziario un valido strumento di credit scoring per le piccole e medie imprese (PMI).

La scelta di concentrare l'analisi sulle PMI deriva dall'osservazione della realtà che nel Nord-Est si presenta ai diversi intermediari finanziari. La clientela che si rivolge agli istituti di credito per la concessione dei finanziamenti presenta diverse conformazioni, ma nella maggior parte dei casi appartiene alla realtà delle piccole e medie aziende.

La necessità di ricorrere a strumenti quantitativi per l'analisi del rischio di credito rispecchia la nuova configurazione del mercato finanziario. In effetti, mentre in passato la concessione di finanziamenti avveniva sulla base di decisioni puramente soggettive, al giorno d'oggi ci si affida a specifiche tecniche classificatorie per poter gestire nel migliore dei modi il rischio legato alla concessione del credito.

L'adozione di un metodo di scoring porta con sé numerosi vantaggi, i quali hanno permesso ai sistemi di scoring di svilupparsi velocemente in ambito finanziario.

La capacità di classificare in maniera corretta la clientela e di attribuire ad essa un punteggio (definito score) in grado di esprimerne la rischiosità creditizia è divenuta di estrema utilità per la concessione del credito.

Infatti, in questi ultimi anni susseguitisi alla crisi finanziaria iniziata nel 2007, si è notata una maggiore attenzione nei confronti del rischio di credito, caratterizzata soprattutto da un'intensificazione dei controlli relativi alla concessione dei finanziamenti.

La crisi finanziaria degli ultimi anni non ha portato con sé solamente una netta riduzione dell'attività relativa alla concessione di crediti, ma anche un significativo deterioramento della qualità del credito.

In virtù di tali motivi, va sottolineato come i sistemi di scoring siano strumenti essenziali in quanto permettono di valutare in maniera efficiente il merito creditizio dei diversi soggetti con cui gli intermediari finanziari attuano le operazioni di finanziamento.

Al giorno d'oggi l'insieme dei modelli di credit scoring per l'analisi del rischio di credito è abbastanza esteso, però, molte volte, essi non riescono a giungere a stime soddisfacenti a

causa della complessità e della dinamicità manifestata dal fenomeno dell'insolvenza. Infatti, in aggiunta ad una pluralità di vantaggi, tali modelli presentano anche difetti più o meno marcati, che in alcuni casi possono portare a risultati distorti e a stime che non rispecchiano adeguatamente la realtà esaminata.

Proprio in questo scenario si sviluppa il seguente lavoro, il quale punta ad identificare un modello che possa essere riconosciuto dagli analisti finanziari come uno strumento facile e veloce con il quale analizzare il rischio di credito del comparto aziendale.

Il modello proposto è fondato sulla tecnica nota come Data Envelopment Analysis (DEA) introdotta da Charnes, Cooper e Rhodes nel 1978. La DEA è una tecnica di programmazione lineare che permette di calcolare l'efficienza relativa di unità decisionali o unità produttive simili (DMU - Decision Making Units) caratterizzate da un sistema di variabili di input e di output.

Rispetto ai metodi statistici che confrontano le diverse unità produttive esaminate con un'ipotetica unità produttiva media, la metodologia DEA confronta ogni singola DMU con le DMU ritenute più efficienti e ne valuta l'efficienza relativa.

In aggiunta, la tecnica DEA permette di comparare le imprese sulla base di una pluralità di input ed output, con l'intento di ottenere una misura che possa racchiudere complessivamente i diversi aspetti della realtà aziendale esaminata. Per questi e altri motivi che saranno delineati in seguito, l'intento del seguente lavoro è quello di evidenziare come la metodologia DEA può fornire un contributo essenziale nell'analizzare il rischio di credito rispetto ai classici modelli di previsione, come verrà verificato nel seguente lavoro.

La struttura attraverso cui si snoda la trattazione è la seguente.

Il Capitolo 1 introduce i principi cardine e le fasi attorno a cui si snoda il credit scoring, sottolineando l'evoluzione che esso subisce con il passare del tempo.

Nel Capitolo 2 sono presentati i modelli di previsione che adottano un approccio deduttivo all'analisi del fenomeno dell'insolvenza. Vengono descritti gli aspetti positivi e negativi di metodologie statistiche quali l'Analisi Discriminante Lineare, il Modello Lineare di Probabilità ed altri modelli di regressione.

Il Capitolo 3 illustra come il rischio relativo alla concessione dei finanziamenti venga esaminato anche dai modelli tipici dell'intelligenza artificiale. Vengono presentati e trattati brevemente i principali modelli di natura induttiva al fine di delinearne le differenze più significative rispetto ai modelli del Capitolo 2.

Nel Capitolo 4 viene introdotta la metodologia Data Envelopment Analysis, sottolineando il concetto di efficienza tramite cui opera. In aggiunta è presentato il modello che verrà poi impiegato nell'analisi, ovvero il Modello Additivo proposto da Charnes nel 1985.

Il Capitolo 5 illustra il procedimento tramite il quale è stata condotta l'analisi evidenziando le caratteristiche del campione di riferimento selezionato e delle variabili di input e di output impiegate. L'analisi che di seguito si svilupperà, ha coinvolto tre modelli differenti tra loro: il primo ispirato dagli studi di Premachandra, il secondo differisce per l'aggiunta di un indicatore, ispirato al modello Z'-Score di Altman per le imprese non quotate ed infine l'ultimo che sostituisce detto indicatore con una diversa variabile.

Concludono la trattazione due appendici: la prima riporta i valori degli indicatori economico-finanziari impiegati nell'analisi, la seconda riporta per ognuno dei tre modelli i migliori risultati ottenuti nella classificazione.

CAPITOLO 1

Il Credit Scoring

Il credit scoring costituisce lo strumento principe per la valutazione del rischio di credito, strumento di cui i diversi intermediari finanziari si avvalgono nell'analisi dei finanziamenti richiesti dalla clientela. Tale sistema automatizzato, come enunciato dalla definizione della Banca d'Italia, impiega una serie di metodi o modelli statistici per effettuare le proprie analisi:

“ Il credit scoring è un sistema automatizzato adottato dalle banche e dagli intermediari finanziari per valutare le richieste di finanziamento della clientela (in genere per la concessione del credito al consumo). Esso si basa su sistemi automatizzati che prevedono l'applicazione di metodi o modelli statistici per valutare il rischio creditizio, e i cui risultati sono espressi in forma di giudizi sintetici, indicatori numerici o punteggi, associati all'interessato, diretti a fornire una rappresentazione, in termini predittivi o probabilistici, del suo profilo di rischio, affidabilità o puntualità nei pagamenti. ”¹

Si osservi che ciò sottolinea la presenza di una linea di rottura con il passato, in cui la concessione del credito avveniva sulla base di decisioni puramente soggettive. Infatti, al giorno d'oggi la configurazione del mercato finanziario non permette più la possibilità di instaurare profondi rapporti personali tra la clientela e l'intermediario finanziario, e in aggiunta, al fine di non distorcere le valutazioni sul merito creditizio dei singoli soggetti, richiede necessariamente di disporre di tecniche quantitative adeguate ad esaminare la rischiosità delle operazioni di finanziamento.

Inoltre, l'elevato tasso di competizione all'interno del mercato spinge gli intermediari a fare pieno affidamento su queste tecniche classificatorie in quanto, grazie alla velocità nella capacità di individuare e selezionare i soggetti meritevoli di credito, vi è la possibilità di accaparrarsi una maggiore quota di mercato rispetto ai diretti concorrenti.²

Sulla base di quanto delineato finora, si può evidenziare che l'adozione di un sistema di scoring comporta una serie di vantaggi più o meno evidenti:

¹ www.bancaditalia.it

² Stanghellini E., *Introduzione ai metodi statistici per il credit scoring*, Springer-Verlag Italia 2009, Milano.

- l'oggettività del sistema, in quanto le decisioni di finanziamento non sono più sottoposte a scelte di carattere soggettivo bensì a procedimenti statistico-quantitativi fondati su set di dati. Quindi non è più il singolo intermediario a definire le variabili o i pesi da utilizzare ai fini dell'indagine, e gli strumenti basati sulle considerazioni personali lasciano spazio ai nuovi sistemi standardizzati;
- l'uniformità del sistema; con uniformità si intende la capacità di trasformare una serie di modelli statistici in punti di riferimento omogenei per la platea di utilizzatori, ovvero la capacità di tali modelli di divenire procedure largamente utilizzate ai fini del calcolo della rischiosità del credito;
- l'adeguatezza del sistema alla configurazione del mercato; l'elevato grado di automazione consente di ridurre i tempi durante la fase di accumulo delle informazioni permettendo un monitoraggio costante delle relazioni intraprese dall'intermediario con la clientela (rinegoziazione del rapporto, rottura,...).

Proprio in virtù di tali vantaggi, i sistemi di scoring hanno preso velocemente piede all'interno del comparto creditizio, infatti la loro capacità di classificare la clientela e di attribuire ad essa un punteggio (definito score) in grado di esprimerne la rischiosità creditizia è divenuta di estrema utilità per la concessione del credito. Va sottolineato che la capacità di segmentare la clientela in classi sulla base del punteggio raggiunto permette di attribuire un prezzo al finanziamento, ovvero un tasso d'interesse, coerente con la rischiosità e il merito creditizio di ciascuna classe.

Le tecniche di scoring svolgono un ruolo significativamente importante all'interno del mercato finanziario, come vedremo la loro applicabilità non si riduce solamente all'analisi del rischio nella concessione dei finanziamenti, ma costituiscono strumenti di supporto in una pluralità di ambiti finanziari.

1.1 Cenni Storici³

Il principio cardine attorno a cui si snoda il credit scoring riguarda la capacità di distinguere due gruppi appartenenti alla stessa popolazione sulla base di una serie di informazioni ad esse collegate.

³ Cenni storici tratti da: Thomas, LC (2000), *A Survey of Credit and Behavioural Scoring: Forecasting Financial Risk of Lending to Consumers*, 16(2) International Journal of Forecasting, 149-172.

Quest'approccio venne introdotto da Irving Fisher nel 1936 che lo utilizzò per distinguere due tipologie di iris basandosi solamente sulla dimensione dei fiori. I primi studi in ambito creditizio furono sviluppati da David Durand nel 1941, il quale cercò di distinguere i prestiti "buoni" da i prestiti "cattivi" avvalendosi di tecniche statistiche.

Le sue ricerche, come la maggior parte delle ricerche in materia di quel periodo, non avevano però lo scopo predittivo che viene invece richiesto attualmente ai diversi modelli di scoring. Durante il periodo della Seconda Guerra Mondiale gli intermediari finanziari riscontravano delle grandi difficoltà nella valutazione del rischio creditizio in quanto le decisioni relative alla concessione di finanziamenti erano ancora legate ai giudizi soggettivi di analisti esperti. Proprio questo fu il motivo per cui si decise di definire delle regole omogenee e condivisibili da tutti per la gestione dell'affidamento del credito.

Al termine della guerra si assistette allo sviluppo delle prime società di consulenza e nel 1958, ad opera della Fair Isaac Corporation (FICO) di San Francisco, nacque il primo modello di credit scoring per le banche e le altre società finanziarie.

Rispetto alle precedenti tecniche basate su considerazioni di carattere personale, gli strumenti di scoring permettevano quindi di differenziare in maniera migliore i clienti meritevoli di credito da quelli non meritevoli, comportando così una maggiore accuratezza nelle decisioni di concessione dei prestiti.

Come si evince dalla letteratura in merito, i modelli di scoring si svilupparono in origine nell'ambito del credito al consumo e, grazie alle loro capacità diagnostiche, consentirono di ridurre notevolmente gli errori manifestati dalle tecniche fino ad allora utilizzate.

I primi studi al di fuori dell'ambito del credito al consumo vennero realizzati da Edward I. Altman⁴, il quale formulò un modello di scoring allo scopo di prevedere il rischio di fallimento aziendale.

Si può notare che già attorno agli anni '60 le tecniche di scoring avevano trovato ampia applicazione all'interno di una pluralità di ambiti finanziari, però la loro affermazione vera e propria avvenne a metà degli anni '70 con l'approvazione degli Equal Credit Opportunity Acts. Questi atti vietavano la possibilità agli intermediari finanziari di negare il finanziamento ad un qualsiasi soggetto a meno che non fossero in grado di argomentare statisticamente la propria decisione, ed è proprio in questi atti che i sistemi di scoring trovano la loro completa affermazione quali strumenti statistici di supporto alle decisioni di finanziamento.

⁴ Noto economista statunitense, famoso per l'elaborazione della formula dello Z-Score (1968) riguardante la valutazione della probabilità di fallimento per le aziende quotate in borsa; inoltre ha continuato le proprie ricerche in merito alla probabilità di fallimento delle aziende e nel 1993, è giunto a formulare il modello Z'-Score riguardante la probabilità di fallimento delle aziende non quotate.

Gli strumenti di scoring si diffusero molto velocemente con il trascorrere degli anni e, verso la fine degli anni '90, furono applicati anche ai mutui ipotecari e ai prestiti alle piccole imprese. È ovvio che dall'accuratezza nell'analisi del rischio di credito discende la capacità di garantire la stabilità dell'intermediario finanziario, per il fatto che quest'ultimo deve stanziare un patrimonio di vigilanza il più possibile coerente con i rischi assunti.

Di fronte a quest'affermazione si può evidenziare come il primo Accordo sul Capitale, noto anche come Accordo di Basilea 1 (1988), fosse inadeguato in quanto legava il patrimonio di vigilanza a dei parametri fissi determinati in maniera esogena che non rispecchiavano affatto la rischiosità del portafoglio della singola banca. Per questo motivo nel 2006 è stato emanato il secondo Accordo sul Capitale o Basilea 2, il quale definisce i criteri in base ai quali è necessario che avvenga l'accantonamento del patrimonio di vigilanza da parte degli intermediari finanziari.

Tramite questo secondo Accordo si è cercato di creare un legame diretto tra i rischi fronteggiati dall'intermediario e il patrimonio accantonato, in maniera che quest'ultimo fosse direttamente proporzionale ai rischi assunti, incentivando una valutazione più precisa della qualità del portafoglio crediti⁵.

Questo però non è stato l'ultimo intervento in materia poiché, come sappiamo, la profonda crisi economica scoppiata negli ultimi anni ha messo a dura prova l'intero sistema finanziario internazionale. Proprio per questo è stato approvato un terzo Accordo sul Capitale o Basilea 3, il quale entrerà in vigore in maniera graduale a partire dal 2013, che mira a rafforzare la regolamentazione e la vigilanza del settore bancario al fine di evitare crisi globali future. I provvedimenti delineati in quest'accordo mirano in primo luogo a migliorare la gestione del rischio e la capacità delle banche nel fronteggiare tensioni economiche derivanti dai diversi mercati, e in secondo luogo a rafforzare le condizioni di trasparenza e di informativa dell'intero settore bancario⁶.

Inoltre, con tali riforme si cerca di intervenire sia a livello di singole banche sia a livello di sistema poiché rafforzando i singoli istituti di credito si diminuisce il rischio di incorrere in gravi shock sistemici.

Quindi, si può sottolineare come gli strumenti di scoring siano diventati sempre più significativi all'interno dell'operatività finanziaria quali strumenti di supporto decisionale in grado di valutare la rischiosità del mercato e di garantire la stabilità del sistema economico.

⁵ www.bancaditalia.it, *Questioni di Economia e Finanza*, Occasional papers, Aprile 2008.

⁶ www.bis.org

1.2 Dal credito al consumo al credito alle imprese

Come delineato nel paragrafo precedente, l'ambito in cui le tecniche di credit scoring mossero i primi passi fu quello del credito al consumo, ma a partire dagli anni '60 gli studiosi rivolsero la propria attenzione all'ambito aziendale, tentando di estendere l'applicazione di tali modelli statistici alla concessione di prestiti alle imprese⁷.

Il primo a formulare un modello per la previsione del fallimento aziendale fu Edward I. Altman; tale modello, noto come modello Z-Score⁸, utilizzava come dati di input degli indici finanziari rappresentativi della situazione in cui versava l'impresa in esame. Questo modello segnò il punto di partenza per un'estensione applicativa delle tecniche di scoring; infatti fino a quel momento l'oggetto principale dell'analisi era rappresentato dal richiedente del prestito il cui profilo differiva sensibilmente da quello dell'impresa.

Questo però non fu l'unico ambito in cui si diffusero le tecniche di scoring ma, attorno agli anni '70 alcuni studiosi ne estesero l'uso anche alla valutazione dei prestiti commerciali, ovvero quella particolare tipologia di prestiti a favore di soggetti che vogliono intraprendere un'attività commerciale o ristrutturare un'attività già esistente.

Un aspetto fondamentale, il quale porrà le basi per tutti i modelli di credit scoring, risiede nel fatto che la categoria dei clienti commerciali è estremamente eterogenea (piccole imprese, lavoratori autonomi,...) e si differenzia notevolmente dai clienti del credito al consumo.

La condizione di omogeneità della popolazione di riferimento appare per la prima volta in quest'ambito, infatti l'eterogeneità dei clienti commerciali aveva posto non pochi problemi agli studiosi nella scelta delle modalità attraverso cui applicare i modelli. Sono state proprio queste difficoltà interpretative che hanno spinto gli accademici a concentrarsi non tanto sulla valutazione della concessione o meno del credito ma su di una revisione periodica del credito concesso. Infatti uno studioso, con riferimento alla modalità attraverso cui analizzare i prestiti commerciali, ha affermato che: "Il suo vantaggio principale è nel liberare gli ufficiali di prestito e gli esaminatori bancari dalle valutazioni di routine di tutti i prestiti cosicché possano allocare il proprio tempo nella valutazione di una piccola percentuale di debitori più rischiosi"⁹.

⁷ Vitalie B. & Ashta A., *The conceptual framework of credit scoring from its origins to microfinance*, Second European Research Conference on Microfinance, Groningen, The Netherlands, June 2011.

⁸ Edward I. Altman, *Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy*, Journal of Finance, Vol. 23, No. 4, (September 1968), pp. 589-609.

⁹ Orgler, Y. E. (1970), *A Credit Scoring Model for Commercial Loans*, Journal of Money, Credit & Banking (Ohio State University Press) 2, 435-445.

Questa nuova ottica attraverso cui concepire l'uso delle tecniche di scoring, ovvero concentrandosi su un controllo periodico del credito concesso, aiuta molto gli intermediari finanziari nel monitorare costantemente la salute delle imprese. Nel momento in cui viene concesso un prestito ad una qualsiasi impresa, l'orizzonte temporale normalmente adottato è quello di lungo periodo, quindi al fine di tenere periodicamente sotto controllo il rischio di credito in cui l'impresa incorre è necessario valutare in maniera adeguata l'insieme delle informazioni che mano a mano pervengono all'intermediario finanziario (situazione dei pagamenti, comportamento del cliente,...).

La possibilità di monitorare e, proprio grazie a questo, di revisionare i prestiti concessi dai diversi istituti di credito durante tutto l'arco temporale dell'esposizione è un elemento estremamente importante per la stabilità e la vitalità dell'istituto stesso in quanto permette di ridefinire le caratteristiche della posizione (ad esempio il tasso d'interesse, l'orizzonte temporale,...) in maniera da evitare eventuali perdite.

Il passaggio delle tecniche di scoring dal credito al consumo al credito alle imprese però non è delineato in maniera perfetta, ma vi è una fascia di casi in cui i due aspetti tendono a confondersi: l'area del micro-credito. In quest'ambito la figura della persona e quella dell'impresa tendono a coincidere e pongono non pochi problemi in relazione all'approccio da utilizzare per la concessione dei prestiti. A seguito di una lunga analisi sulla tipologia di informazioni che questa parte di sistema economico poteva fornire nella valutazione della rischiosità creditizia si è deciso di ricondurre tale valutazione alle tecniche del credito al consumo.

Quindi la transizione da parte delle tecniche di scoring da una valutazione del credito al consumo ad una valutazione del credito alle imprese è avvenuta in maniera graduale, anche perché la presenza di caratteristiche e informazioni molto diverse tra le due fattispecie poneva una serie di interrogativi relativi alle procedure da attuare allo scopo di migliorare in entrambi i settori la capacità diagnostica dei modelli utilizzati.

1.3 Gli obiettivi del Credit Scoring

Molteplici sono gli aspetti della gestione del rischio interessati dalle tecniche di credit scoring, e proprio per questo motivo lo scopo di tali tecniche non può essere circoscritto alla sola valutazione del rischio creditizio legato alla concessione o meno di un finanziamento.

La differenziazione tra i clienti meritevoli e quelli non meritevoli di credito costituisce l'obiettivo che in origine ha spinto numerosi studiosi a creare delle tecniche in grado di intervenire in tale ambito, ma sarebbe errato ricondurre la moltitudine dei modelli di scoring a quest'unico scopo.

I modelli normalmente utilizzati valutano ogni singolo richiedente lungo tutto l'arco di vita della posizione instaurata con l'intermediario; di conseguenza, interessando più fasi del rapporto, gli obiettivi abbracciati da tali modelli sono plurimi.

Ecco quindi il motivo per cui generalmente, nella gestione del credito, gli intermediari finanziari dispongono di una serie di modelli statistici a supporto delle decisioni creditizie. Infatti gli obiettivi da raggiungere sono diversi:

- obiettivi di marketing;
- obiettivi di credito;
- obiettivi di analisi comportamentale;
- obiettivi di gestione delle insolvenze.

Per la realizzazione degli obiettivi di marketing ci si affida a modelli di scoring in grado di stimare la capacità o meno da parte di un'iniziativa promozionale di attrarre potenziali clienti e di aumentare il grado di fidelizzazione nei confronti dei clienti già esistenti.

Il secondo gruppo di obiettivi può essere realizzato avvalendosi di modelli orientati alla valutazione dell'affidabilità creditizia della controparte, influenzando in questo modo la decisione sulla concessione del finanziamento.

Con riferimento al terzo punto, i modelli di scoring impiegati al fine di soddisfare tali obiettivi rivolgono l'attenzione al comportamento creditizio dei soggetti in relazione alla loro posizione debitoria in maniera da prevedere, in via anticipata, se il cliente potrà dimostrarsi insolvente in futuro causando così una situazione critica per l'intermediario.

Allo scopo di realizzare gli obiettivi delineati nell'ultimo punto, gli intermediari finanziari ricorrono ad una serie di modelli che descrivono le procedure mediante le quali è possibile agire sulle posizioni rilevatesi insolventi, in modo da limitare le eventuali perdite per l'intermediario stesso.

Sulla base di quanto detto possiamo affermare che al giorno d'oggi i modelli di credit scoring non si limitano ad identificare la probabilità che un soggetto finanziato entri in uno stato d'insolvenza, ma attraverso un'attività di marketing attuano una vera e propria selezione della clientela in modo da individuare i soggetti capaci di garantire il rimborso del prestito

concesso, agevolando così gli intermediari nel delineare un pricing coerente con il profilo di rischio del soggetto finanziato¹⁰.

Infatti il fine ultimo dei modelli di scoring è rappresentato dalla massimizzazione del profitto per l'intermediario; quest'ultimo li utilizza come strumenti di difesa in grado di analizzare le informazioni provenienti dalla controparte, le quali a volte sono soggette a notevoli distorsioni, e codificarle in modo da rendere il più possibile certo il profilo di rischio creditizio assunto dall'intermediario tramite l'opera di finanziamento.

1.4 Le fasi del Credit Scoring

Come delineato precedentemente, il risultato a cui si perviene attraverso l'applicazione delle tecniche di scoring è descritto da un punteggio (o score) in grado di classificare i soggetti analizzati sulla base della loro rischiosità. Il procedimento attraverso cui si giunge a tale risultato non è affatto semplice, ma richiede l'esecuzione di una serie di fasi specifiche allo scopo di evitare possibili errori connessi alla complessità della realtà da esaminare.

Generalmente, le fasi in cui il processo di scoring viene suddiviso sono le seguenti¹¹:

I° FASE: *Individuazione della popolazione di riferimento*; l'ipotesi fondamentale per l'applicabilità del metodo di scoring riguarda il fatto che la popolazione di riferimento risulti essere omogenea, ovvero che i membri di tale popolazione presentino caratteristiche simili (strutturali, economiche,...). Il venir meno di tale condizione pone gravi problemi in relazione all'efficacia e alla potenza del modello nel predire lo stato d'insolvenza, nel caso in cui ad esempio si stia trattando del credito alle imprese;

II° FASE: *Scelta della variabile di classificazione*; una volta terminata la delimitazione della popolazione di riferimento si passa alla scelta della variabile in base alla quale classificare il campione. La variabile considerata oggetto di discriminazione dipende dall'ambito in cui si sta operando; ad esempio se si opera nel credito al consumo la variabile sarà probabilmente legata alla capacità del soggetto di rimborsare il prestito concessogli, mentre nel caso in cui si operi all'interno del credito alle imprese la variabile di classificazione sarà rappresentata dallo stato d'insolvenza (o default) dell'impresa. Come si può notare, la variabile di classificazione

¹⁰ Thomas, LC (2000), *A Survey of Credit and Behavioural Scoring: Forecasting Financial Risk of Lending to Consumers*, 16(2) International Journal of Forecasting.

¹¹ Stanghellini E., *Introduzione ai metodi statistici per il credit scoring*, Springer-Verlag Italia 2009, Milano.

in entrambi i casi riflette il manifestarsi di un fenomeno negativo nei confronti dell'intermediario finanziario, e non può che essere così in quanto il ruolo chiave dei modelli di scoring è proprio quello di diagnosticare in anticipo le possibili situazioni di pericolo per l'intermediario.

III° FASE: *Selezione dei dati del campione*; in questa fase vengono selezionate le unità che andranno a formare il campione sul quale verrà effettuato lo studio tramite l'applicazione del modello di scoring. Il campione deve essere rappresentativo della popolazione di riferimento, nel senso che la selezione delle unità che lo compongono non viene effettuata sulla base di parametri specifici. Per ogni unità del campione vengono rilevate le informazioni considerate utili dal sistema al fine della classificazione. Tali informazioni si differenziano a seconda che il campione di riferimento sia composto da imprese o da consumatori. Nel primo caso le informazioni ritenute rilevanti solitamente possono essere di tipo quantitativo (bilanci d'esercizio), di tipo qualitativo (organizzazione aziendale, tipologia e qualità della gestione,...), di tipo settoriale (competitività all'interno del settore,...) e, in aggiunta, anche informazioni relative all'andamento del rapporto banca-impresa (analisi del rapporto con la Centrale dei Rischi). Mentre nel secondo caso le informazioni sono per lo più di natura socio-demografica e, in quanto tali, sono meno attendibili rispetto a quelle provenienti dal mondo aziendale.

Un aspetto fondamentale della selezione sta nel fatto che le unità vengano analizzate sulla base degli stessi parametri (orizzonte temporale costante, caratteristiche del finanziamento omogenee) in modo da non creare distorsioni nei risultati.

Resta inoltre da evidenziare che il campione d'analisi è nella maggior parte dei casi sbilanciato, ossia nel caso in cui ad esempio la variabile di classificazione sia rappresentata dallo stato d'insolvenza delle imprese si nota che il numero di imprese insolventi all'interno del campione è significativamente inferiore rispetto a quello delle imprese sane. Per questo motivo il più delle volte il campione viene strutturato in maniera da ottenere una situazione di quasi pareggio tra le due tipologie di imprese (si parla di campione "pareggiato").

IV° FASE: *Scelta della metodologia*; a seguito della definizione del campione avviene la scelta della metodologia da applicare per la costruzione del metodo di scoring. In primo luogo si attua una suddivisione del campione di riferimento in due sottogruppi; il primo, definito "training sample", costituisce la base di dati per un primo sviluppo del modello, mentre il secondo, definito "test sample", costituisce il campione di verifica al fine di valutarne

l'efficacia e la potenza. Particolarmente importanti in questa fase sono le modalità attraverso cui vengono selezionate le informazioni che costituiranno gli input della metodologia; infatti una scelta errata delle informazioni ritenute significative potrebbe portare il modello a commettere dei gravi errori di classificazione. Generalmente gli errori che si compiono nell'attuazione dei modelli di scoring sono di due tipologie: errori del I° tipo ed errori del II° tipo. La prima tipologia di errori conduce a selezionare come "sani" soggetti che in realtà non lo sono, mentre la seconda tipologia di errori porta a classificare come "non sani" soggetti che in realtà lo sono. Questo evidenzia come la scelta accurata delle informazioni e della metodologia, nello sviluppo delle tecniche di scoring, risulti fondamentale al fine di garantire la minor distorsione possibile dei risultati ottenuti. Una volta giunti alla stima del modello, si nota come alle singole variabili vengano attribuiti dei parametri o pesi in base al contributo da esse apportato alla corretta classificazione del soggetto in esame. Mantenendo costanti tali parametri e modificando il valore delle variabili in relazione ai diversi soggetti, si ottengono automaticamente i punteggi sulla base dei quali avverrà la classificazione finale.

V° FASE: *Stima dei costi di classificazione*; un aspetto significativo riguarda la stima degli errori di classificazione, poiché una stima errata può portare ad eventi estremamente sfavorevoli per l'intermediario che attua il modello. Infatti, per esempio, l'errore di classificare come "sani" soggetti insolventi comporta gravi perdite nel bilancio dell'intermediario, e ciò sottolinea come vi siano dei costi intrinseci nella selezione dei soggetti.

Allo scopo di ridurre i possibili errori di classificazione, si definisce il livello dei costi che l'intermediario è in grado di sopportare e, sulla base di tale livello, viene stabilito un valore di soglia capace di discriminare al meglio i soggetti esaminati. A conferma di ciò vi è la necessità di una sostanziale separazione tra il campione di sviluppo e il campione di verifica del modello in quanto una loro coincidenza porterebbe, a causa del mancato confronto con l'ambiente esterno, ad una sottostima degli errori e quindi dei costi (le informazioni con cui si sviluppa il modello e se ne valuta la potenza sono le stesse).

VI° FASE: *Monitoraggio della validità del modello nel tempo*; come in molte altre attività finanziarie è necessario monitorare la performance e l'efficacia del modello di scoring nel tempo. Confrontando periodicamente i risultati ottenuti dal modello con le informazioni provenienti dal mondo reale si ottiene una valutazione relativa alla precisione del modello nel predire la probabilità d'insolvenza. Nel caso in cui si presenti un largo divario tra i risultati

ottenuti dal modello e quelli osservabili, allora sarà necessaria una modifica della metodologia al fine di migliorare le capacità diagnostiche.

Tali modifiche sono molto frequenti per il fatto che, soprattutto negli ultimi anni, la popolazione di riferimento non rimane costante, ma è soggetta a profondi mutamenti (strutturali, sociali ed economici) che minano la solidità degli output generati dai diversi modelli di scoring. Proprio per questo un monitoraggio costante e accurato di un modello di scoring si rende necessario al fine di garantire che tale strumento possa divenire prezioso per la gestione del portafoglio clienti di un intermediario finanziario (in ogni fase della relazione con il cliente).

CAPITOLO 2

Modelli di previsione delle insolvenze

Come si è visto nel Capitolo 1, il ricorso a modelli di previsione delle insolvenze nasce dalla necessità degli intermediari finanziari di gestire al meglio il rischio di credito in cui incorrono attraverso la concessione di finanziamenti. I primi modelli che si svilupparono in materia erano dei modelli statistici univariati, il più famoso dei quali fu proposto da Beaver¹² nel 1966, i quali cercavano di studiare il fenomeno dell'insolvenza basandosi sull'analisi di singole variabili economiche.

In seguito a tali studi si cercò di incrementare la capacità predittiva delle analisi condotte mediante una transizione da modelli univariati a modelli multivariati; questi ultimi, avendo a disposizione un numero maggiore di informazioni, furono capaci di migliorare la diagnosi posta in essere dai precedenti modelli.

A partire da una serie di input costituiti dai principali indici economico-finanziari in grado di rappresentare al meglio la situazione dell'impresa, tali modelli giungono a stimare la probabilità d'insolvenza attribuendo a ciascun indice un peso sulla base della loro rilevanza nel predire il fallimento.

Nei paragrafi successivi sono descritti i principali modelli che adottano un approccio deduttivo all'analisi del fenomeno dell'insolvenza: l'Analisi Discriminante Lineare, il Modello Lineare di Probabilità (Linear Probability Model), il Modello Logit (Logistic Probability Model) e il Modello Probit (Probit Analysis).

2.1 Analisi Discriminante Lineare

Questo modello venne proposto da R. A. Fisher¹³ nel 1936 e, come delineato in precedenza, cerca di predire il fallimento delle imprese mediante una discriminazione tra imprese sane e imprese insolventi sulla base di alcuni indici economico-finanziari.

¹² Beaver W. H. (1966), *Financial ratios as predictors of failure*, Empirical Research in Accounting: Selected Studies 1966, Vol. 4, 41 (71-111). In questo famoso lavoro l'autore ha esaminato la capacità produttiva di alcuni singoli indicatori economico-finanziari rispetto al fenomeno dell'insolvenza. In esso vengono già considerate una pluralità di problematiche quali la definizione del concetto di insolvenza, l'utilizzo di campioni con pari composizione di società sane e anomale, il riferimento ad una teoria dell'insolvenza aziendale, l'analisi degli errori di I° e II° tipo, l'approccio bayesiano, che saranno oggetto del dibattito scientifico negli anni seguenti nelle ricerche con metodologie multivariate.

¹³ Fischer R.A. (1936), *The Use Of Multiple Measurement In Taxonomic Problems*, Annals of Eugenics, V. 7, p. 179-188.

L'approccio attuato da Fisher consiste nella definizione di una combinazione lineare delle variabili osservate, però tale scelta non avviene in maniera casuale.

Infatti il principio cardine su cui si fonda l'intero approccio afferma che tra tutte le combinazioni lineari possibili, la scelta deve ricadere sulla combinazione lineare capace di discriminare al meglio le variabili dei due campioni (imprese sane ed imprese insolventi) soddisfacendo le seguenti esigenze:

- massimizzare la varianza tra i due campioni;
- minimizzare la varianza all'interno di ogni singolo campione;

Tale combinazione lineare, giovandosi di una serie di pesi definiti attraverso specifiche tecniche di ponderazione, permette di sintetizzare le variabili relative ad ogni impresa in un unico valore (score):

$$S_j = \sum_{i=1}^n a_i X_{i,j}$$

dove S_j = score dell'impresa j-esima

a_i = coefficiente di discriminazione della variabile X_i

X_i = variabile descrittiva della caratteristica i-esima dell'impresa

I coefficienti di discriminazione sono calcolati in maniera da agevolare la combinazione lineare nel soddisfare le due esigenze prima enunciate, inoltre una volta giunti alla stima definitiva del modello, tali coefficienti vengono mantenuti costanti per tutte le successive applicazioni nello stesso ambito di studio.

Notiamo quindi come questa combinazione lineare permetta di considerare il profilo dell'impresa in esame nel suo complesso piuttosto che svolgere un'analisi sequenziale delle singole caratteristiche rilevanti.

La fase in cui si effettua la selezione delle variabili descrittive del modello è particolarmente delicata e solitamente avviene sulla base di due modalità: un metodo simultaneo o diretto e un metodo a passi (stepwise).

Nel primo caso le variabili sono scelte con riferimento ad un modello teorico definito a priori, mentre i coefficienti discriminanti vengono stimati attraverso analisi empiriche. Nel secondo

caso, invece, la selezione delle variabili avviene in base alla capacità discriminante manifestata dalle stesse.

Il secondo metodo presenta una pluralità di sfaccettature in quanto la selezione delle variabili può avvenire in più forme¹⁴:

- *backward elimination*, questa tipologia di selezione parte dall'insieme complessivo delle variabili e successivamente elimina quelle caratterizzate da un minore potere discriminante;
- *forward selection*, in questo caso invece si parte da una singola variabile e, in maniera progressiva, si aggiungono quelle variabili in grado di incrementare il potenziale diagnostico del modello sottolineando così un'elevata capacità discriminatoria;
- *stepwise selection*, quest'ultima procedura è una combinazione della *backward elimination* e della *forward selection*. Tale tecnica prevede che venga selezionata una variabile alla volta e che quest'ultima rimanga all'interno del modello solamente se il suo potere discriminatorio risulti elevato, ovvero verrà eliminata nel caso in cui venga introdotta un'altra variabile discriminante che ne faccia crollare il potere discriminatorio¹⁵.

Per semplificare la trattazione è riportato, in Figura 2.1, un esempio di Analisi Discriminante Lineare nel caso in cui siano esaminate due variabili (X_1 , X_2) riferite a due campioni distinti (A, B). Sul piano delineato dai due assi (X_1 e X_2) sono riportate le imprese appartenenti ai due campioni di riferimento assieme alle rispettive medie \bar{x}_A e \bar{x}_B (centroidi dei due campioni).

È possibile constatare dal grafico che i due campioni risultano sovrapposti in una certa area, quindi il ruolo dell'Analisi Discriminante è di riuscire ad individuare quella retta Y , nota come *Funzione Discriminante*, che separa al meglio i due campioni. Proprietà fondamentale della Funzione Discriminante è che, tracciata una retta S perpendicolare ad essa, permette di proiettare i punti del piano su tale retta S individuando due distribuzioni.

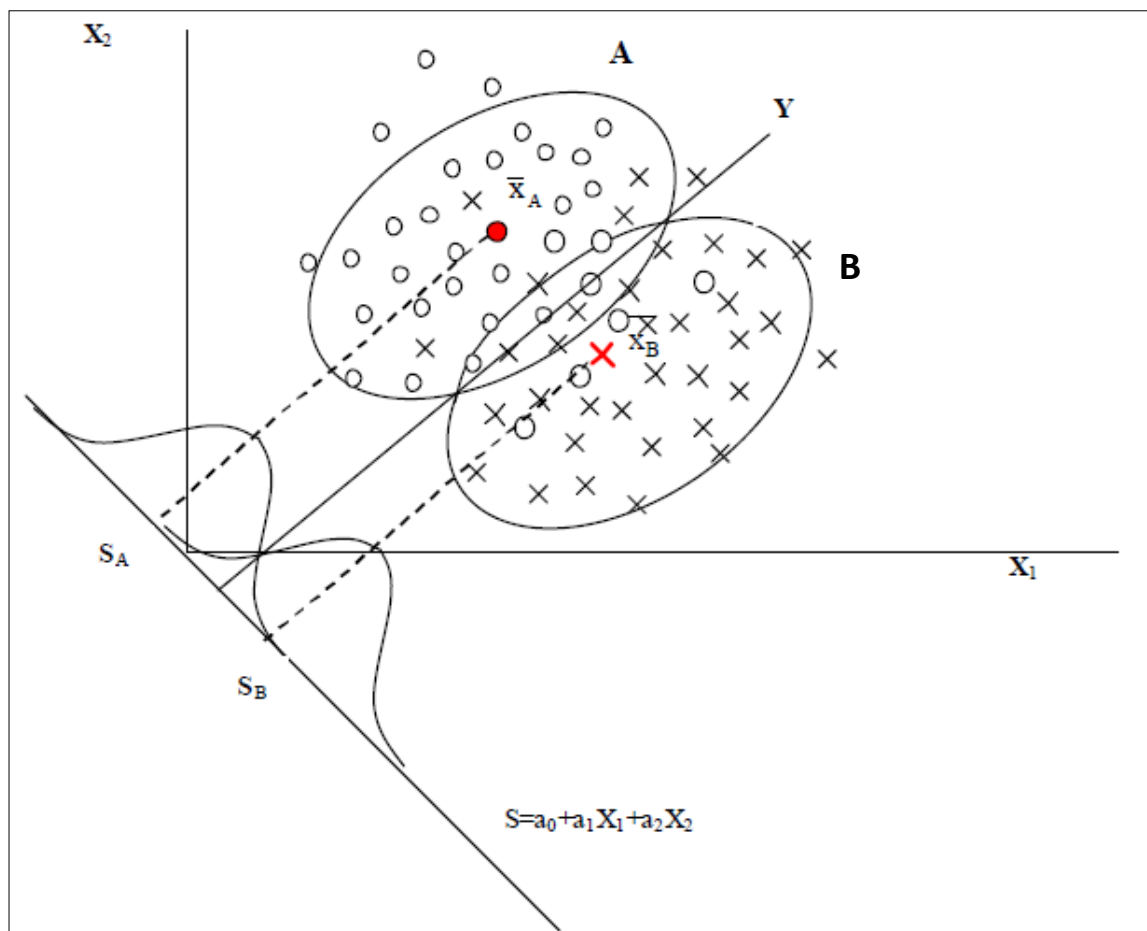
Le distribuzioni disegnate non sono due distribuzioni qualsiasi, ma tra tutte quelle possibili sono quelle che presentano la minor area di sovrapposizione; proprio per questo si può sostenere che S raffigura la migliore combinazione lineare tra tutte quelle realizzabili.

Uno dei maggiori vantaggi di questo modello risiede nella capacità di ridurre la dimensione delle caratteristiche osservate, infatti grazie alla combinazione lineare vi è il passaggio dalle n dimensioni delle variabili considerate ad una sola dimensione data dai punti della retta S .

¹⁴ Andrea Resti e Andrea Sironi(2008), *Rischio e valore nelle banche – Misura, regolamentazione, gestione*, EGEA, Milano.

¹⁵ Questo fenomeno può accadere nel momento in cui le variabili discriminanti selezionate siano espressione dello stesso evento economico.

Figura 2.1: Grafico dell'Analisi Discriminante Lineare



Fonte: G. Szego e F. Varetto (1999), *Il rischio Creditizio: misura e controllo*, ed. Utet, Torino

I punti della retta S rappresentano i punteggi (score) delle diverse imprese e consentono di effettuare una classificazione immediata mediante il confronto con un valore di soglia. Qualora lo score generato superi il valore soglia (cut-off point), valore che discrimina tra soggetti affidabili e soggetti non affidabili, allora l'impresa verrà classificata come sana, in caso contrario verrà inclusa nel campione delle imprese insolventi.

Questo modello non è però esente da errori nel processo di classificazione, infatti a seconda dell'estensione dell'area di sovrapposizione fra le due distribuzioni si può avere una maggiore o minore incertezza nella classificazione delle imprese. Ad esempio, nel caso in cui le due distribuzioni siano perfettamente coincidenti, l'incertezza classificatoria è massima in quanto le informazioni non permettono di identificare l'appartenenza delle singole imprese ad un campione o all'altro.

Gli errori di classificazione che possono essere commessi sono di due tipi:

- Classificare come sano un soggetto insolvente (Errore del I° tipo);
- Classificare come insolvente un soggetto sano (Errore del II° tipo).

Sulla base della probabilità di commettere uno dei due errori viene calibrato il valore soglia in quanto le conseguenze derivanti dai due errori non sono le stesse. In particolare l'intermediario finanziario subisce una perdita maggiore nel caso in cui commetta un errore del I° tipo per il fatto che non vi è solo la perdita del mancato guadagno nel caso in cui il soggetto fosse rimasto sano, ma anche la perdita del capitale e degli interessi a seguito dell'insolvenza. Ecco perché una fase cruciale del modello riguarda la stima dei costi provenienti dai possibili errori di classificazione, cosicché un'adeguata calibrazione del valore di soglia possa limitare al minimo le perdite per l'intermediario finanziario¹⁶.

Il modello, come tutti gli altri modelli di scoring, cerca di sottrarre la sfera soggettiva dalla valutazione del rischio, tanto è vero che la definizione della combinazione lineare e dei coefficienti di discriminazione è totalmente oggettiva, rilegando così la soggettività dell'analista solamente nella decisione delle variabili da osservare.

Inoltre, il modello di Fisher può essere considerato un modello non parametrico in quanto, a differenza dei modelli parametrici, non si fonda sull'ipotesi che le distribuzioni debbano necessariamente assumere una forma predeterminata¹⁷.

Ecco quindi che l'Analisi Discriminante si dimostra uno strumento di supporto molto utile per la valutazione della rischiosità del portafoglio crediti di un intermediario; attraverso la combinazione lineare delle diverse informazioni quest'ultimo riesce a classificare i soggetti riducendo al minimo i possibili errori legati a decisioni di carattere soggettivo.

2.2 Modello Z-Score

Il modello Z-Score venne proposto da Edward I. Altman nel 1968 e consisteva nell'applicazione del metodo dell'Analisi Discriminante Lineare sotto l'esistenza di due ipotesi¹⁸:

¹⁶ In generale, la determinazione del *cut-off point* può essere tale da rendere il costo atteso associato a errori del primo tipo pari a quello relativo a errori del secondo tipo. Analiticamente:

$$P_I \cdot C_I = P_{II} \cdot C_{II}$$

dove P_I e P_{II} rappresentano la probabilità rispettivamente di errori del primo e del secondo tipo e C_I e C_{II} rappresentano i costi medi connessi rispettivamente a errori del primo e del secondo tipo.

In presenza di $C_I > C_{II}$, il cut-off point dovrà essere fissato in modo che $P_I < P_{II}$.

¹⁷ I metodi non parametrici non necessitano di ipotesi a priori sulle caratteristiche della popolazione o comunque le ipotesi sono meno restrittive di quelle usuali nella statistica parametrica, in particolare non si assume l'ipotesi che i dati provengano da una popolazione normale o gaussiana.

¹⁸ Edward I. Altman, *Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy*, Journal of Finance, Vol. 23, No. 4, (September 1968), pp. 589-609.

- l'ipotesi che le variabili discriminanti siano caratterizzate da una distribuzione normale multivariata;
- l'ipotesi che le matrici varianze-covarianze per le variabili discriminanti siano uguali per i due gruppi d'impresе analizzate.

Quest'approccio al modello non fu esente da critiche per il fatto che, in relazione alla prima ipotesi, le esperienze empiriche dimostravano come la distribuzione normale, di per sé illimitata, male si adeguasse alla rappresentazione di variabili (indici economico-finanziari) che potevano assumere al massimo valori compresi tra 0 e 100. Tale carenza veniva principalmente riscontrata nella minore capacità diagnostica del modello rispetto a metodi che, in sostituzione alla distribuzione normale multivariata, si avvalevano di funzioni logistiche o probit¹⁹.

Altman applicò il suo modello ad un campione formato da 33 imprese industriali fallite nel periodo 1945-1965 e un campione "pareggiato" formato da 33 imprese sane. Quindi a seguito di un procedimento meticoloso Altman giunse alla stima del seguente modello²⁰:

$$Z = 0,012X_1 + 0,014X_2 + 0,033X_3 + 0,006X_4 + 0,999X_5$$

Dove X_1 = Capitale Circolante/Attivo Totale

X_2 = Riserve da Utili/ Attivo Totale

X_3 = Utile ante Interessi e Tasse/ Attivo Totale

X_4 = Valore di Mercato del Patrimonio Netto/ Debiti Totali

X_5 = Ricavi/ Attivo Totale

Il vettore dei coefficienti di discriminazione è dato da:

$$c = (mX_s - mX_{in}) \cdot S^{-1}$$

dove mX_s è il vettore delle medie del gruppo delle imprese sane

mX_{in} è il vettore delle medie del gruppo delle imprese insolventi

S^{-1} è la matrice delle varianze-covarianze tra le variabili X_i

¹⁹ Andrea Resti e Andrea Sironi(2008), *Rischio e valore nelle banche – Misura, regolamentazione, gestione*, EGEA, Milano.

²⁰ Questa formulazione del modello è stata elaborata per la valutazione delle società quotate.

Il modello sopra descritto abbraccia nella sua valutazione, più aspetti riferibili alla realtà in cui opera l'impresa: la liquidità, la redditività, la solidità finanziaria e l'efficienza complessiva.

Le aree di interesse indicate sono espresse dalle variabili del modello. Rivolgendo l'attenzione alla prima variabile si può notare come essa rappresenti un indicatore della liquidità dell'impresa. In effetti, essendo il capitale circolante definito come differenza tra le attività e le passività correnti, tale indicatore risulta essere molto utile nel predire la risposta dell'impresa in un periodo di crisi²¹.

La seconda variabile esprime la redditività dell'impresa e, in maniera implicita, fa riferimento anche all'età della stessa. Infatti è possibile dedurre che un'impresa giovane presenterà un valore dell'indice inferiore rispetto a quello di un'impresa più matura per il fatto che quest'ultima ha avuto il tempo di generare una riserva di utili. In aggiunta, questo indicatore esprime anche il grado di leva finanziaria dell'impresa, nel senso che maggiori riserve di utili rappresentano una maggiore capacità di autofinanziamento e di conseguenza un minor ricorso all'indebitamento.

La terza variabile rapporta l'EBIT (Earnings before Interest and Taxes) al totale dell'attivo; tale indicatore costituisce una vera e propria misura della produttività aziendale condivisa in campo internazionale.

La quarta variabile rappresenta un indicatore in grado di esprimere la struttura finanziaria dell'impresa in cui il Patrimonio Netto è valutato a valori di mercato. Tale indicatore è capace di identificare di quanto possono ridursi le attività aziendali prima che le passività eccedano il loro valore, permettendo così al modello di Altman di prevedere in anticipo i presupposti per l'insolvenza.

L'ultima variabile esprime la capacità dell'impresa di generare ricavi a partire da un determinato valore dell'attivo, quindi misura implicitamente anche l'abilità del management di operare in ambienti competitivi²². Lo stesso Altman sottolinea in un suo lavoro che, seppure su base individuale risulti il meno importante, posto in relazione agli altri indicatori considerati fornisce un contributo essenziale alla validità del modello²³.

²¹ Altri indicatori quali ad esempio il current ratio, ovvero il rapporto tra attività correnti e passività correnti, e il quick ratio, ovvero un rapporto che vede al numeratore la somma delle liquidità immediate e differite e al denominatore le passività correnti, si sono rivelati meno utili dell'indicatore in questione e quindi sono stati rifiutati durante lo sviluppo del modello.

²² Altman E., (2000), *Predicting financial distress of companies: revisiting the z-score and zeta® models*, Stern School of Business, New York University, pp. 9-12.

²³ Edward I. Altman, *Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy*, Journal of Finance, Vol. 23, No. 4, (September 1968), pp. 589-609.

Se si considera ora l'efficacia del modello di Altman si può evidenziare come, con riferimento all'anno immediatamente precedente all'insolvenza, esso riesca a classificare correttamente il 95% delle imprese esaminate, con una diversa influenza delle due tipologie di errore: l'errore del I° tipo è stato del 6% e l'errore del II° tipo è stato del 3% (Tabella 2.1).

Tabella 2.1: Capacità diagnostica del modello di Altman al periodo t-1(dati in %)²⁴

		Classificazione	
		<i>Insolventi</i>	<i>Sane</i>
Gruppo Effettivo	<i>Insolventi</i>	94	6 (Errore di I° tipo)
	<i>Sane</i>	3 (Errore di II° tipo)	97

Fonte: G. Szego e F. Varetto (1999), *Il rischio Creditizio: misura e controllo*, ed. Utet, Torino

La capacità diagnostica del modello di Altman risulta molto elevata al periodo t-1, però procedendo a ritroso nel tempo si nota un progressivo peggioramento dovuto al fatto che i due campioni di riferimento (imprese sane e imprese insolventi) non risultano più chiaramente distinti²⁵.

Ecco che il modello appare robusto e stabile nel valutare le imprese in periodi prossimi allo stato d'insolvenza, ma incrina la propria performance mano a mano che si procede a ritroso nel tempo.

Il modello classifica le imprese confrontando il valore ottenuto dall'equazione (Z) con un intervallo di cut-off caratterizzato da due estremi: uno inferiore pari a 1,81 ed uno superiore pari a 2,99. Infatti, nel calcolo dello score di cut-off come media dei centroidi delle imprese sane e di quelle insolventi, Altman ha individuato un intervallo noto come "grey area", ovvero zona grigia o zona d'ignoranza, in cui i due gruppi d'imprese non sono chiaramente distinti; ciò porta ad una possibilità maggiore di commettere degli errori di classificazione, non consentendo così di giungere a risultati soddisfacenti.

Sulla base di questa distinzione allora la probabilità di fallimento da parte dell'impresa risulta:

- alta se il valore dello Z-Score è minore di 1,81;
- media se il valore è compreso tra 1,81 e 2,99;

²⁴ Convenzionalmente sono indicati con t-1; t-2;...t-n gli anni precedenti al momento (anno t) dell'insolvenza o del fallimento.

²⁵ È stato dimostrato empiricamente che la capacità diagnostica del modello, con riferimento a due anni precedenti al momento dell'insolvenza, crolla dal 95% all'82%.

- bassa se il valore è maggiore di 2,99.

Adottando un'ottica di concessione del credito, tali range di valore di Z suggeriscono che un'impresa con uno score inferiore all'1,81 si presenta come inaffidabile poiché, con grande probabilità, si dimostrerà insolvente in futuro. In contrapposizione a ciò, un'impresa con uno score superiore a 2,99 si dimostra affidabile in quanto presenta una struttura finanziaria particolarmente solida. Punto critico rimane l'analisi delle imprese inglobate all'interno della zona d'ignoranza, per il fatto che il modello non riesce ad attuare una discriminazione efficace tra le due tipologie.

A seguito di questo lavoro, Altman apportò delle modifiche al modello originale allo scopo di migliorare la sua capacità diagnostica e di incrementare il numero dei campi in cui poteva essere utilizzato²⁶:

- affinché il modello possa essere applicato per la valutazione delle imprese non quotate, la quarta variabile è stata calcolata sostituendo al valore di mercato del Patrimonio Netto il suo valore contabile²⁷;
- per fare in modo che il modello sia uno strumento di analisi utile nei confronti delle imprese non industriali, la quinta variabile è stata eliminata in quanto è influenzata dall'appartenenza delle imprese al settore industriale;
- un ulteriore aggiustamento è stato attuato per adattare il modello ai rischi di credito dei nuovi paesi emergenti, con l'obiettivo di aumentare la correlazione tra gli score della funzione discriminante e le classi di rating definite sulle obbligazioni statunitensi (EM-Score)²⁸.

²⁶ Caouette J., E. Altman and P. Narayana (1998), *Managing Credit Risk: The Next Great Financial Challenge*, John Wiley & Sons, New York.

²⁷ Altman nel 1993 ritocca il modello originario giungendo alla formulazione dello Z' -Score riguardante la probabilità di fallimento delle aziende non quotate. Sempre con riferimento agli indicatori illustrati, la formula per questa variante del modello risulta essere:

$$Z' = 0,717X_1 + 0,847X_2 + 3,107X_3 + 0,420X_4 + 0,998X_5$$

In questo nuovo modello la precisione del primo tipo (ovvero la corretta identificazione di compagnie insolventi) è leggermente inferiore rispetto al modello classico, 91% contro 94%, mentre l'accuratezza del secondo tipo (la corretta identificazione delle compagnie sane) è pari al 97%.

²⁸ Edward I. Altman, *An emerging market credit scoring system for corporate bonds*, Emerging Markets Review, Volume 6, Issue 4, December 2005, Pages 311-323.

2.3 Modello Lineare di Probabilità (Liner Probability Model)

Questo modello è caratterizzato dal fatto che la probabilità d'insolvenza viene stimata attraverso una regressione lineare, e, seppur diversa, tale regressione presenta delle analogie con l'analisi discriminante lineare presentata nei paragrafi precedenti²⁹.

Attraverso un procedimento composto da una serie di fasi, simile a quello dell'analisi discriminante, il modello giunge a definire una probabilità d'insolvenza.

Vi è una prima fase in cui si procede a selezionare le imprese che formeranno il campione d'analisi. Come è già stato sottolineato nel primo capitolo, il punto cruciale del procedimento consiste nel creare un campione che sia bilanciato, ovvero in cui il numero delle imprese insolventi sia quasi uguale a quello delle imprese sane. In quest'ottica una banca dotata di un portafoglio ben nutrito di imprese insolventi perverrà ad una stima più precisa dei coefficienti della regressione lineare.

Fatto ciò si passa alla seconda fase in cui, una volta distinte le imprese sane da quelle insolventi, si individuano le informazioni che possono rappresentare le variabili della regressione. Tali variabili, come nell'analisi discriminante, sono solitamente degli indici economico-finanziari in grado di rispecchiare gli aspetti principali nell'andamento delle imprese (liquidità, profittabilità, ...).

La terza fase consiste nell'inclusione delle variabili indipendenti così selezionate in una regressione lineare in cui la variabile dipendente è costituita dalla variabile Y binaria che può assumere alternativamente valore nullo quando l'impresa è sana ($Y_i = 0$) o valore unitario quando l'impresa è insolvente ($Y_i = 1$). Analiticamente il modello è così presentato:

$$Y_i = F(\alpha + \beta X) = \alpha + \sum_{j=1}^n \beta_j X_{i,j} + \varepsilon_i$$

dove α indica il termine noto, $X_{i,j}$ sono gli indicatori economico-finanziari scelti come variabili casuali della regressione e β_j rappresenta l'influenza dell'indicatore j nel determinare lo stato d'insolvenza per l'impresa del campione.

²⁹ Si vedano Maddala G. (1983), *Limited-dependent and qualitative variables in econometrics*, ed. Cambridge Univ. Press, Cambridge e Maddala G. (1992), *Introduction to econometrics*, 2a edizione, ed. MacMillan, NY. In questi scritti viene presentata l'analogia principale tra i due modelli: è stato dimostrato che i coefficienti dell'analisi lineare sono pari a quelli della regressione con i minimi quadrati ordinari a meno di una costante.

La fase finale del procedimento consiste nell'estendere il campo di applicabilità del modello, ovvero nell'utilizzare la regressione stimata per predire la probabilità d'insolvenza di un'impresa, non appartenente al campione di riferimento, nel momento in cui si appresti a richiedere un finanziamento.

A questo punto un esempio esemplificativo può essere d'aiuto al fine di comprendere il funzionamento del modello; si può supporre che una banca, al termine del processo delineato in precedenza, sia giunta a stimare la seguente funzione lineare³⁰:

$$Y_i = \alpha + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \beta_3 X_{i,3} = 0,03 + 0,25 \cdot X_{i,1} - 5 \cdot X_{i,2} + 0,7 \cdot X_{i,3}$$

dove $X_{i,1}$ = Debito/ Capitale Proprio (D/E) dell'impresa i

$X_{i,2}$ = Margine Operativo Lordo/Attivo Totale dell'impresa i

$X_{i,3}$ = Oneri Finanziari/Margine Operativo Lordo (OF/MOL) dell'impresa i

Le variabili indipendenti incluse nel modello riflettono alcuni aspetti cardine delle imprese; infatti la prima variabile costituisce un indice del grado di leva finanziaria dell'impresa mentre la seconda rappresenta un indicatore della redditività dell'attivo. Ora si ipotizzi che l'impresa Alfa decida di richiedere un finanziamento alla banca; quest'ultima cercherà di stimare la sua probabilità d'insolvenza usufruendo del modello stimato. Se gli indicatori presentati per l'impresa i assumono i seguenti valori: $X_1 = 5$; $X_2 = 0,3$; $X_3 = 0,4$, allora la probabilità d'insolvenza per l'impresa Alfa risulta essere pari a:

$$Y_{Alfa} = 0,03 + 0,25 \cdot 5 - 5 \cdot 0,3 + 0,7 \cdot 0,4 = 0,03 + 1,25 - 1,5 + 0,28 = 0,06 = 6\%$$

Dal risultato si può dedurre che, pur avendo un indice di leva finanziaria molto elevato, simbolo di un significativo indebitamento, l'impresa Alfa non risulta insolvente.

Così presentato il modello sembra efficace e molto semplice da utilizzare, ma esso soffre di un problema essenziale, ovvero la mancanza di garanzia sul fatto che la probabilità d'insolvenza calcolata dal modello ricada sempre all'interno dell'intervallo (0;1). Infatti, sia nel caso in cui la funzione venga applicata al campione d'origine, sia che la stessa si riferisca

³⁰ Gli esempi riportati sono tratti da: Andrea Resti e Andrea Sironi (2008), *Rischio e valore nelle banche – Misura, regolamentazione, gestione*, EGEA, Milano.

ad un'impresa non appartenente a tale campione, non si esclude la possibilità che la variabile dipendente Y possa assumere valori esterni all'intervallo.

A conferma di quanto detto si può considerare un nuovo esempio in cui l'impresa Beta richieda un finanziamento alla banca presentando i seguenti valori delle variabili indipendenti: $X_1 = 4$; $X_2 = 0,3$; $X_3 = 0,5$. A questo punto la banca, utilizzando la funzione lineare stimata, giunge a definire la probabilità d'insolvenza per l'impresa Beta:

$$Y_{Beta} = 0,03 + 0,25 \cdot 4 - 5 \cdot 0,3 + 0,7 \cdot 0,5 = 0,03 + 1 - 1,5 + 0,35 = -0,15 = -15\%$$

La soluzione non può essere accettata in quanto la probabilità che si verifichi un evento nella realtà assume valori compresi tra 0 e 1, quindi l'esempio sopra descritto sottolinea una carenza del modello nel predire il fallimento di imprese che non appartengono al training sample (campione su cui si è sviluppato il modello).

Un ulteriore problema del modello è rappresentato dal fatto che la varianza dei residui non è costante ma è soggetta ad eteroschedasticità; ciò contribuisce a far venir meno l'ipotesi di fondo della regressione lineare, ossia l'indipendenza dei residui dalle variabili indipendenti (omoschedasticità), comportando così una distorsione nelle stime.

I problemi in questione verranno riproposti nel prossimo paragrafo attraverso la presentazione del modello logistico e del modello probit, tecniche prevalentemente utilizzate nelle diverse applicazioni in quanto permettono di risolvere le difficoltà incontrate dal Linear Probability Model.

2.4 Modello Logit (Logistic Probability Model) e Modello Probit (Probit Analysis)

Il ricorso al modello logistico è necessario al fine di risolvere i problemi incontrati nell'applicazione del modello lineare di probabilità; il modello logistico, alla pari del precedente, suppone l'esistenza di una relazione tra la probabilità d'insolvenza ed una serie di grandezze economiche legate al fallimento dell'impresa.

Per giungere ad una espressione analitica del modello in genere si è soliti indicare con y la probabilità d'insolvenza, con X il vettore delle variabili indipendenti, con α il termine costante e con β i coefficienti in modo tale da ottenere la seguente espressione:

$$y = F(\alpha + \beta X)$$

dove F identifica la funzione di ripartizione cumulata della variabile casuale logistica

$$F(\alpha + \beta X) = \int_{-\infty}^{\alpha + \beta X} f(h) dh = \frac{1}{1 + e^{-(\alpha + \beta X)}}$$

dove $f(h)$ indica la funzione di densità logistica³¹

$$f(h) = \frac{e^h}{(1 + e^h)^2}$$

In generale, la trasformazione descritta fa sì che F risulti pari a 0,5 per valori di $\alpha + \beta X$ pari a zero, per valori di $\alpha + \beta X$ inferiori a zero assume invece valori compresi tra 0 e 0,5 mentre per valori di $\alpha + \beta X$ maggiori di zero assume valori compresi tra 0,5 e 1.

Ciò che principalmente differenzia il modello logistico dal modello lineare di probabilità riguarda la forma assunta dalla distribuzione della probabilità d'insolvenza.

Per approfondire questo fattore si è soliti partire dalla probabilità d'insolvenza così delineata

$$y = \frac{1}{1 + e^{-(\alpha + \beta X)}}$$

e attraverso le opportune modifiche si ottiene l'equazione

$$e^{-(\alpha + \beta X)} = \frac{1 - y}{y}$$

in altre parole

$$e^{(\alpha + \beta X)} = \frac{y}{1 - y}$$

in cui il termine alla destra dell'uguaglianza rappresenta l'odd-ratio³², ovvero il rapporto tra la probabilità dell'evento e quella del suo complementare.

³¹ La distribuzione logistica è una funzione di densità cumulata definita su tutto \mathbb{R} alla pari di una distribuzione normale, ma presenta la caratteristica di essere più concentrata attorno alla propria media.

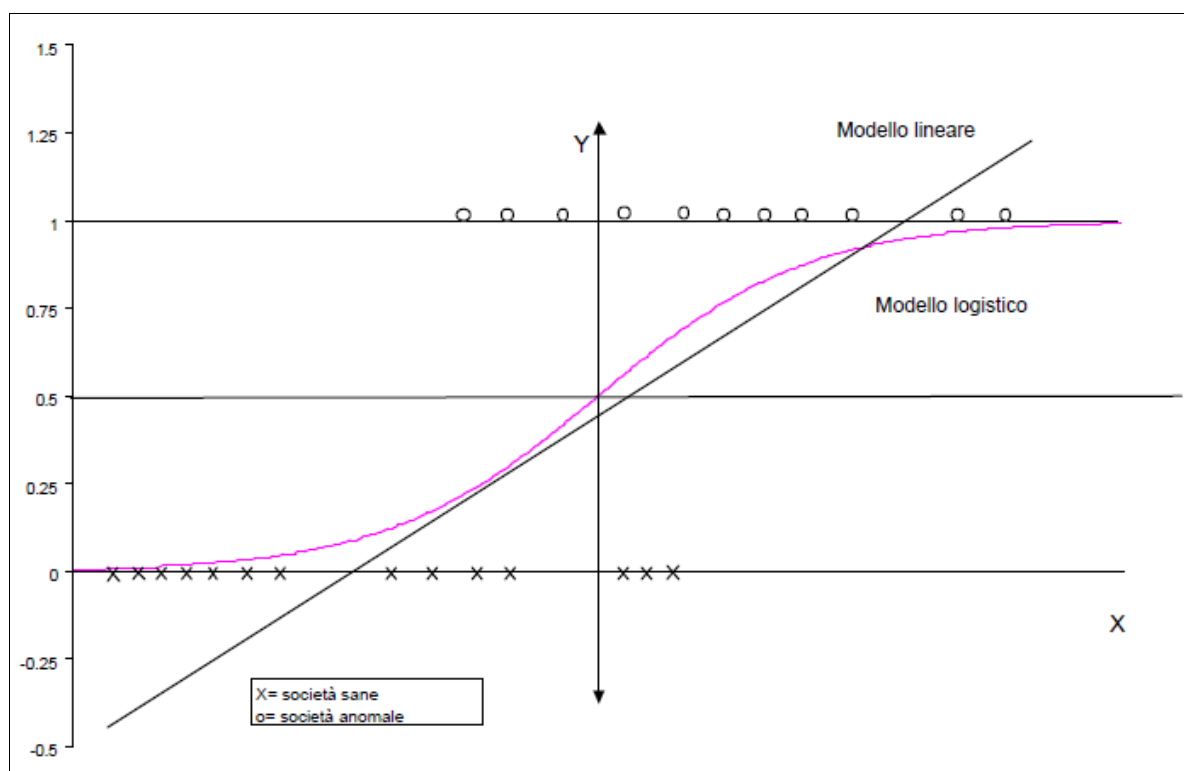
Attuando un'ulteriore trasformazione consistente nel considerare il logaritmo naturale di tale equazione si ottiene:

$$\ln\left(\frac{y}{1-y}\right) = \alpha + \beta X$$

Questa relazione esprime al meglio la differenza tra i due modelli; infatti, mentre nel Linear Probability Model ad essere messo in relazione con $\alpha + \beta X$ è y , nel modello logistico è il logaritmo dell'odd-ratio ad essere messo in relazione con $\alpha + \beta X$.

Per verificare la capacità del modello logistico di mantenere i valori delle stime all'interno dell'intervallo $[0;1]$ si può esaminare il grafico presentato nella Figura 2. In questo semplice grafico viene definita una relazione tra la variabile dipendente Y , la quale può assumere solamente valori interni all'intervallo, e un indicatore di bilancio X .

Figura 2.2: Modello Lineare e Modello Logistico



Fonte: G. Szego e F. Varetto (1999), *Il rischio Creditizio: misura e controllo*, ed. Utet, Torino

³² Con il termine odd-ratio si intende il rapporto tra la probabilità y di un evento e la probabilità che tale evento non accada, vale a dire la probabilità dell'evento complementare $1-y$. Il logaritmo naturale di tale rapporto, cioè $\ln\left(\frac{y}{1-y}\right)$, è detto logit.

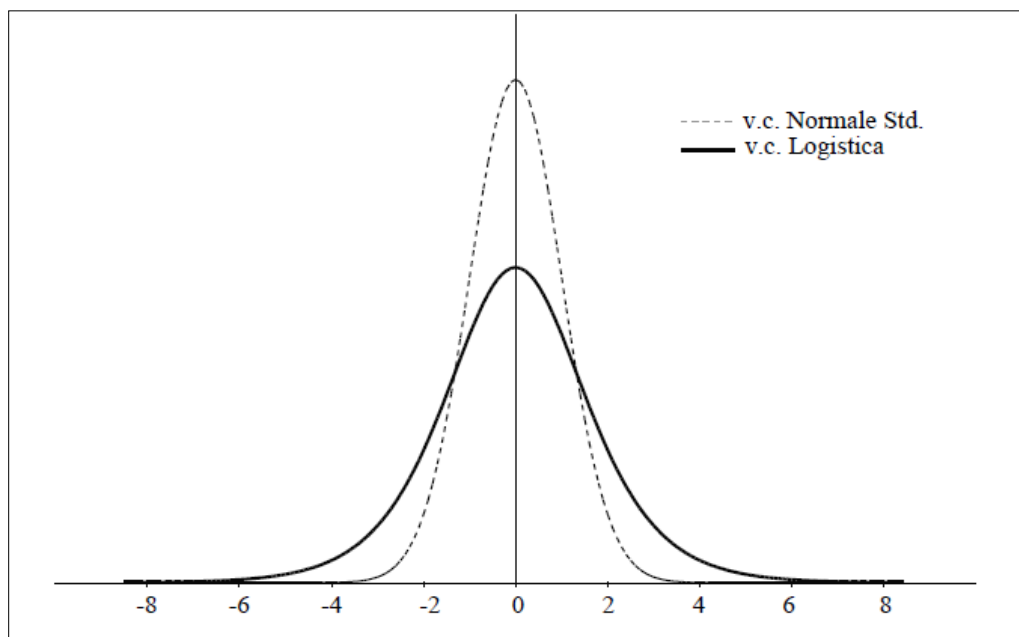
Osservando il grafico si può affermare che il modello logistico, a differenza del modello lineare, rappresenta in maniera più veritiera la realtà in quanto riesce a condurre la probabilità di tutte le possibili manifestazioni d'insolvenza all'intervallo compreso tra 0 e 1 (codominio della funzione logistica).

Un modello del tutto analogo a quello logistico è il modello probit, il quale presenta ugualmente un codominio compreso nell'intervallo (0;1), ma assume che la distribuzione F sia rappresentata da una normale standardizzata:

$$F(\alpha + \beta X) = \int_{-\infty}^{\alpha + \beta X} f(h) dh = \frac{1}{\sqrt{2\pi}} e^{-h^2/2}$$

I risultati ottenuti tramite l'applicazione dei due modelli sono per lo più simili; l'unica differenza è che la distribuzione logistica consente di facilitare i calcoli da un punto di vista matematico e, per questo motivo, risulta di più facile applicazione.

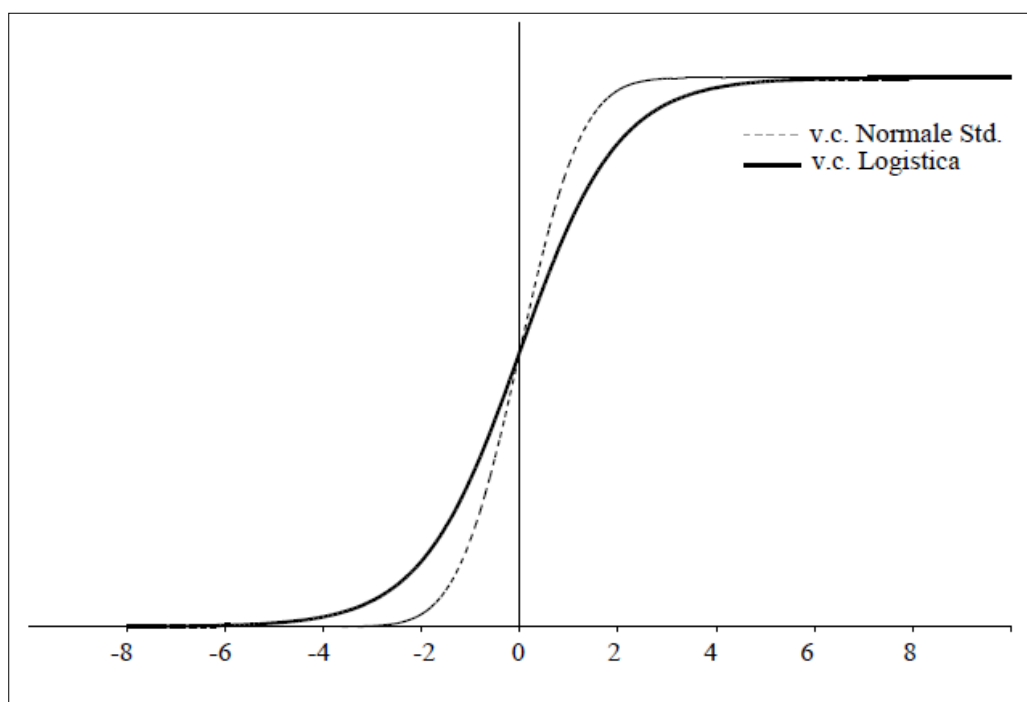
Figura 2.3: Confronto delle funzione di densità di probabilità del Modello Logit e del Modello Probit



Fonte: Giulio Palomba (2008), *Modelli a variabili dipendenti qualitative*

La funzione di densità logistica si caratterizza per code più spesse rispetto alla funzione di distribuzione normale standard per il modello probit, ma ciò non comporta molte differenze a meno che il campione di riferimento non includa parecchi valori estremi.

Figura 2.4: Funzioni di ripartizione del Modello Logit e del Modello Probit



Fonte: Giulio Palomba (2008), Modelli a variabili dipendenti qualitative

In quest'ottica il modello logistico, assieme al modello probit, tra i modelli di regressione sembra il più robusto in quanto mantiene la propria stabilità anche nel caso in cui venga applicato ad imprese al di fuori del campione di sviluppo.

A questo punto è significativo soffermarsi sulle differenze che caratterizzano i modelli di scoring presentati finora, in particolare ciò che differenzia l'analisi discriminante dal modello logistico (massima espressione dei modelli di regressione).

Il primo punto di divergenza tra i due modelli riguarda la selezione delle imprese per il fatto che, mentre l'analisi discriminante ipotizza che esse appartengano a due insiemi distinti, il modello logistico ne riconosce l'appartenenza ad un unico insieme comune.

Una seconda difformità tra i due modelli riguarda lo scopo a cui sono rivolti; infatti, mentre l'obiettivo dell'analisi discriminante risiede nell'individuare l'appartenenza dell'impresa ad uno dei due insiemi, l'obiettivo del modello logistico è quello di valutare lo stato di salute in cui versa l'impresa (entrambi si avvalgono di indicatori economico-finanziari per decretare tale giudizio).

Un'ulteriore disparità risiede nella modalità in cui sono concepite le variabili del modello: nell'analisi discriminante le variabili vengono valutate sulla base del loro contributo nella formulazione di un valore in grado di definire l'appartenenza delle imprese ad uno dei due

gruppi, nel modello logistico è come se svolgessero il ruolo di variabili esterne al modello, fondamentali però nell'offrire una panoramica dell'impresa dal punto di vista dell'intermediario finanziario.

Pur con la presenza di tutte queste differenze, le analisi empiriche dimostrano che i due modelli in generale non pervengono a risultati molto diversi. In particolare, per campioni bilanciati di numerosità contenuta è preferibile l'applicazione dell'analisi discriminante lineare, anche se le performance dei due modelli risultano molto simili. Invece, nel caso in cui il campione risulti sbilanciato o si abbia a che fare con distribuzioni non normali, gli studi empirici dimostrano che il modello logistico fornisce una migliore qualità delle stime previsionali.

Comunque, nonostante queste disparità in campo applicativo, i due modelli si presentano come dei validi strumenti mediante i quali analizzare l'affidabilità creditizia e lo stato di salute delle imprese oggetto di esame.

2.5 Campi di applicazione³³

Vale la pena ricordare che i modelli presentati finora hanno avuto una grande diffusione nell'ambito della gestione del rischio creditizio non solo grazie alla loro abilità nel predire il possibile fallimento di imprese finanziarie e non, ma anche grazie alla loro efficace applicabilità in una pluralità di campi.

Il campo in cui tali modelli trovano maggiore applicazione è quello dei finanziamenti bancari; in questo campo essi cercano di svolgere il ruolo di strumenti di supporto al fine di facilitare gli analisti del credito nelle classificazioni dei finanziamenti. Nel fare ciò, però, i modelli si trovano a dover fronteggiare una difficoltà, ossia il fatto che la classificazione avviene sulla base di varie classi di rischio³⁴ e non con la semplice distinzione in due insiemi (imprese sane e insolventi).

Un campo che presenta diverse analogie con quello dei finanziamenti bancari è quello dei prestiti personali, infatti anche in quest'ambito i modelli di scoring, utilizzando un set

³³ Osservazioni e cenni tratti da: G. Szego e F. Varetto (1999), *Il rischio Creditizio: misura e controllo*, ed. Utet, Torino.

³⁴ Un esempio di classificazione adottata dalle banche può essere il seguente in cui si individuano più classi di rischio: current (rischio accettabile), especially mentioned (vi sono elementi di debolezza nel debitore), substandard (rischio elevato), doubtful (rimborso assai problematico), loss (credito considerato inesigibile). Inoltre le banche possono avvalersi anche di sistemi di rating interno o di usufruire di modelli di classificazione offerti dalle agenzie di rating (Standard & Poor's e Moody's).

informativo diverso da quello precedente, tentano di classificare i diversi soggetti in più classi di rischio.

I modelli di scoring sono stati utilizzati anche per la valutazione del rischio di titoli obbligazionari e dei commercial papers, indipendentemente dal fatto che essi avessero un orizzonte di breve, medio o lungo periodo³⁵.

Ritornando poi al campo aziendale si può sottolineare come tali modelli siano stati impiegati non solo per la capacità di prevedere il fallimento o meno delle imprese, ma anche per individuare se queste fossero in grado di riprendere l'attività dopo aver subito una fase di ristrutturazione e fossero quindi potenzialmente competitive all'interno del settore di riferimento.

Inoltre è da segnalare come questi modelli predittivi siano stati adoperati anche allo scopo di individuare se esistessero delle relazioni tra l'andamento del mercato finanziario e lo stato di salute delle imprese, ovvero se da tale andamento si potessero trarre anticipatamente dei segnali d'insolvenza. Le ricerche empiriche in merito hanno dimostrato che un collegamento tra l'andamento dei prezzi del mercato finanziario e lo stato d'insolvenza delle imprese è possibile in quanto i prezzi riescono ad incorporare una parte delle informazioni relative al futuro stato di salute delle singole imprese.

Alla pari dei cambiamenti di mercato, i modelli di scoring si sono concentrati anche sulle modalità con cui i mutamenti di governo e di regime politico potessero influire sulla probabilità d'insolvenza, così da poter captare in anticipo i possibili eventi di crisi.

Questo alto potere predittivo dei modelli di scoring non viene sfruttato solamente dagli analisti del credito, ma anche dai revisori contabili nelle loro valutazioni. Infatti questi ultimi sono chiamati a segnalare, nelle loro relazioni di certificazione, se le imprese saranno capaci di sostenere la propria crescita o se saranno soggette a un possibile fallimento. I revisori contabili, nel dare questi giudizi, conducono un'analisi sui bilanci d'esercizio delle imprese e, proprio per questo, l'avvalersi di strumenti di supporto quali i modelli di scoring permette loro di dare maggiore validità alle proprie tesi in virtù dell'elevata capacità diagnostica di tali metodi.

Per concludere si può evidenziare come i modelli di scoring possano essere impiegati anche come strumenti di analisi per il Rischio Paese, così da generalizzare al massimo l'ottica della trattazione.

³⁵ Edward I. Altman, R. Avery, R. Eisenbeis, J. Sinkey (1981), *Application of classification techniques in business, banking and finance*, ed. Jai Press, NY.

Queste sono solo alcune delle aree di applicazione in cui si possono utilizzare tali modelli, grazie all'elevato grado di flessibilità di questi metodi nell'analizzare fenomeni diversi tra loro. Quindi, i modelli di scoring si prestano a numerosi processi di trasformazione e di miglioramento al fine di evolvere e di incrementare il loro potere predittivo in modo da divenire strumenti di risk management sempre più sofisticati.

2.6 Critiche e limiti dei modelli di scoring

I modelli di scoring, come osservato nel Capitolo 1, godono di una serie di vantaggi tra i quali è utile ricordare l'estrema flessibilità del modello, l'oggettività dei risultati e la capacità di adattarsi facilmente ai diversi scenari esaminati.

A fronte di questi vantaggi, la pluralità di casi concreti in cui sono stati impiegati ha messo in evidenza la presenza di alcuni limiti metodologici e operativi. Le problematiche solitamente riscontrate nei modelli di scoring sono le seguenti:

- la mancanza di una teoria di riferimento; ogni modello discende da un'analisi empirica e si nota l'assenza di un approccio teorico su cui basarsi. Infatti, nella maggior parte dei casi, la formulazione dei metodi differisce in base all'evento e ai dati su cui si sta effettuando l'analisi, e ciò simboleggia non solo l'elevata flessibilità ma anche la specificità degli stessi. Questo problema sottolinea come i modelli risultino estremamente legati al campione di riferimento e presentino delle difficoltà ad analizzare il mondo reale nel suo complesso;
- la definizione di stato d'insolvenza; la suddivisione tra imprese sane e insolventi avviene sulla base della definizione d'insolvenza usata dal modello. Il problema consiste nel fatto che esistono più "gradi d'insolvenza"³⁶ quindi, in base alla scelta della definizione, il modello classificherà come insolventi un numero di imprese più o meno grande influenzando così i risultati ottenuti;
- la stabilità del modello; mutamenti della sfera macroeconomica (cambiamenti del ciclo economico, variazioni strutturali dei mercati,...) possono influenzare sensibilmente la stabilità dei modelli nel tempo. Non vi è alcuna ragione per supporre che i rapporti tra le variabili indipendenti e l'evento dell'insolvenza rimangano immutati; inoltre i coefficienti di discriminazione o di regressione, a seconda del modello con cui si

³⁶ Esistono più "gradi di insolvenza": si può partire, ad esempio, da un semplice ritardo nel pagamento degli interessi fino al fallimento giuridico dell'impresa debitrice, passando attraverso una pluralità di stadi intermedi.

opera, non rimangono costanti nel tempo ma subiscono delle variazioni. Per questo motivo, al fine di garantire che la capacità diagnostica dei modelli rimanga inalterata nel tempo, essi sono soggetti a periodici interventi strutturali in modo da ridurre al minimo i possibili decrementi nella performance;

- la mancanza di variabili qualitative; i modelli di scoring nel condurre l'analisi si soffermano per lo più su dati statistici e di bilancio che rappresentano solamente in modo parziale la situazione dell'impresa, tralasciando così un ampio spettro informativo proveniente dai numerosi dati qualitativi. Il ricorso all'uso di variabili qualitative costituisce un tentativo di contrastare l'eventuale "window dressing"³⁷ che le imprese possono condurre sui dati di bilancio. Tra le principali variabili qualitative si possono segnalare la reputazione dell'impresa, la qualità del management, le notizie relative a particolari strategie adottate dall'impresa (pricing dei prodotti, apertura di nuove filiali), le prospettive del settore produttivo di appartenenza e mutamenti nell'ambiente esterno;
- la scarsa omogeneità del campione; il campione di riferimento deve essere il più omogeneo possibile, infatti l'appartenenza delle imprese a più settori economici porta ad una distorsione dei risultati in quanto il valore e il potere predittivo degli indicatori economico-finanziari possono cambiare da settore a settore.
- la presenza di campioni "sbilanciati"; è possibile che in molti casi il campione di riferimento risulti sbilanciato, in quanto il numero delle imprese sane è significativamente maggiore di quello delle imprese insolventi. È stato dimostrato che i modelli di scoring, all'aumentare del numero di imprese insolventi nel campione di riferimento, migliorano la capacità diagnostica e la precisione nei risultati. Per questo motivo lavorando con campioni poco sbilanciati è possibile ottenere delle stime più precise, incrementando il potere predittivo del modello.

³⁷ Con il termine *window dressing* si intendono una serie di politiche di bilancio poste in essere al fine di soddisfare formalmente obiettivi dimensionali, gestionali e di immagine, che rendono le evidenze contabili non conformi alla reale situazione aziendale e quindi scarsamente informativi.

CAPITOLO 3

Altri approcci: il contributo dell'intelligenza artificiale

Il seguente Capitolo, legato in parte al Capitolo 2, illustra come il processo di decisione nella concessione del credito venga affrontato anche mediante i procedimenti tipici dell'intelligenza artificiale, la quale ha subito un enorme sviluppo negli ultimi tempi che non poteva se non avere riflessi anche in questo campo.

A differenza dei modelli precedentemente analizzati, nei modelli di natura induttiva non vi è un approccio strutturale in cui, partendo dai ragionamenti di un analista, si cerca di trovare conferme alle ipotesi in un campione empirico di dati, ma si adotta un procedimento induttivo in cui la logica viene del tutto rovesciata, ovvero partendo da un campione di dati si cerca di trovare una regolarità (caratteristiche in comune tra le imprese insolventi) grazie alla quale si riesca a prevedere nel tempo il default di nuove imprese.

Ecco quindi che, mentre i modelli strutturati presentati in precedenza si fondano su una serie di algoritmi e parametri statistici, i modelli induttivi vengono spesso definiti delle scatole nere (black box), proprio per sottolineare la mancanza di un vero e proprio schema di riferimento. I paragrafi successivi tracciano una breve descrizione dei principali modelli di natura induttiva: i Sistemi Esperti, gli Alberi Decisionali, le Reti Neurali Artificiali e gli Algoritmi Genetici.

3.1 Sistemi Esperti

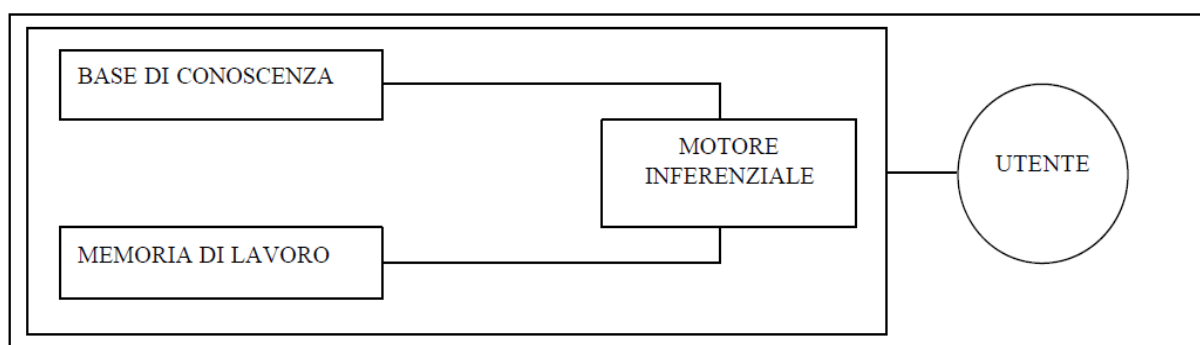
L'area in cui si concentrano maggiormente i sistemi esperti è rappresentata dall'area fidi; questo perché i procedimenti legati alla gestione dei crediti presentano dei caratteri facilmente standardizzabili e quindi possono essere integrati all'interno del sistema informativo della banca.

Un sistema esperto si differenzia notevolmente dai modelli descritti in precedenza per il fatto che non si limita a fornire una diagnosi del rischio di insolvenza, ma consente di effettuare una pluralità di servizi in supporto a tale azione (un supporto alla valutazione dell'impresa, la definizione dell'entità dei fidi accordati, le forme tecniche di erogazione dei fidi,...).

I vantaggi che possono essere ottenuti dall'applicazione di questi modelli sono molteplici ed includono³⁸:

- la realizzazione di una base di conoscenza a cui la banca può attingere qualora si concentri sulla concessione di finanziamenti;
- l'omogeneità del trattamento delle pratiche di credito;
- il risparmio di tempo da parte dell'analista, il quale può rivolgere la propria attenzione al trattamento dei clienti che richiedono approfondimenti in merito;
- il maggior grado di interazione tra la banca e i propri clienti.

Figura 3.1: Schema semplificato di un sistema esperto



Fonte: G. Szego e F. Varetto (1999), *Il rischio Creditizio: misura e controllo*, ed. Utet, Torino

In Figura 3.1 è rappresentata una semplificazione della struttura di un sistema esperto. L'insieme delle informazioni da cui si parte a costruire il modello, con riferimento ad un particolare dominio (campione di riferimento), è la base di conoscenza o memoria a lungo termine.

Generalmente la base di conoscenza viene realizzata da un "ingegnere³⁹" della conoscenza che assume informazioni dagli esperti del dominio⁴⁰, queste informazioni vengono poi strutturate in modo da essere rappresentate tramite un modello.

Nel caso in esame gli esperti del dominio altri non sono che gli analisti del credito della banca: è da questi che l'ingegnere della conoscenza trae informazioni sul dominio essenziali per lo sviluppo del modello.

³⁸ G. Szego e F. Varetto (1999), *Il rischio Creditizio: misura e controllo*, ed. Utet, Torino.

³⁹ L'ingegnere della conoscenza rappresenta l'individuo che codifica la conoscenza dell'esperto di dominio in una forma comprensibile al sistema esperto.

⁴⁰ L'esperto del dominio è l'attore che detiene la conoscenza del dominio e che è in grado di risolvere il problema in oggetto; l'obiettivo principale di un sistema esperto è imitare l'esperto di dominio umano.

Dallo schema notiamo che il motore inferenziale, ossia la componente che simula il processo di ragionamento mediante il quale è possibile trarre deduzioni logiche, è separata dalla base di conoscenza, cosicché l'ingegnere possa concentrare tutta la propria attenzione nell'estrazione di tale conoscenza dagli esperti del dominio. È infatti questa la fase più lunga e complessa dell'intero processo, ed occorre avere a disposizione dei buoni esperti in grado di trasmettere le proprie conoscenze in maniera efficace e di aiutare, qualora fosse necessario, l'ingegnere nella formalizzazione della base di conoscenza.

La memoria di lavoro o memoria a breve termine contiene i dati relativi allo specifico problema da risolvere; su questi si concentra il motore inferenziale che, attraverso una serie di procedure, mette in relazione la memoria di lavoro con gli elementi della base di conoscenza considerati rilevanti ai fini della soluzione del problema. Tale processo continua fino a che il motore inferenziale non individua tutti gli elementi conciliabili con il problema in esame.

Nei sistemi esperti esistono due metodologie per attuare una rappresentazione formale della conoscenza e sono: le regole di produzione e i frames⁴¹.

Le regole di produzione non sono altro che delle proposizioni in cui una premessa definisce le condizioni che devono essere rispettate per attivare la regola:

IF (premessa); THEN (conseguenza)

Questa struttura permette l'adozione di un approccio dinamico alla formazione della conoscenza, in quanto il sistema di regole può essere modificato o aggiornato in maniera progressiva.

Il motore inferenziale, come accennato in precedenza, per selezionare le regole ritenute rilevanti ai fini del problema può adottare due procedure contrapposte⁴²:

- *forward chaining*, ovvero una procedura in cui il motore inferenziale esamina le regole partendo dalle premesse, e tutte quelle che risultano compatibili con il problema specificato nella memoria di lavoro vengono selezionate (la selezione avviene in base ad un ordine di priorità o ad altri criteri);
- *backward chaining*, tramite questa procedura invece il motore inferenziale attiva le regole partendo dalle conseguenze piuttosto che dalle premesse. Questo approccio, del

⁴¹ G. Szego e F. Varetto (1999), *Il rischio Creditizio: misura e controllo*, ed. Utet, Torino.

⁴² Andrea Resti e Andrea Sironi (2008), *Rischio e valore nelle banche – Misura, regolamentazione, gestione*, EGEA, Milano.

tutto inverso rispetto al precedente, solitamente è più parsimonioso poiché si limita a lavorare sui fatti che sono strettamente rilevanti per la soluzione del problema.

La preposizione sopra riportata può essere completata aggiungendo un fattore di certezza alle conseguenze:

IF (premessa); THEN (conseguenza); CON PROBABILITÀ = (p)

dove p rappresenta la probabilità condizionale che, date le premesse, le conseguenze siano vere.

Un passo ulteriore consiste nell'aggiornare la struttura della regola considerando anche l'utilità da essa fornita al modello:

IF (premessa); THEN (conseguenza); CON PROBABILITÀ = (p); CON UTILITÀ = (u)

dove u è misurata in base alla quantità d'informazioni contenute nella regola⁴³.

In base a quanto affermato si può sottolineare come la struttura a regola presenti alcuni vantaggi, in particolar modo dal punto di vista del dinamismo, ma d'altro canto il numero delle regole, nel momento in cui diventa troppo elevato, pone degli interrogativi sull'efficacia del procedimento. Per di più va evidenziato che la coerenza interna del sistema di regole è difficile da garantire in quanto maggiore è il numero di regole presenti e maggiore è la possibilità che si verifichino contraddizioni o incoerenze all'interno della struttura.

Un approccio diverso rispetto a quello delle regole di produzione è attuato dal frame; tale metodologia è particolarmente statica e si fonda sull'individuazione di quelle caratteristiche chiave che descrivono in maniera compiuta lo specifico problema in esame. Ad esempio, ipotizzando che il nostro problema sia l'individuazione e la classificazione delle società insolventi, il frame costituirebbe quell'approccio teso ad identificare le caratteristiche chiave che differenziano tali imprese dal resto: redditività modesta o negativa, scarsa liquidità, struttura finanziaria fragile e così via.

Ecco quindi che il frame considera una pluralità di aspetti collegati al problema da studiare, in modo da assumere il maggior numero di informazioni possibili; compito del motore

⁴³ Il calcolo di *u* segue i principi della teoria dell'informazione.

inferenziale rimane quello di analizzare i frames ritenuti rilevanti per la soluzione del problema.

Entrambi i sistemi descritti, ossia il sistema a regole e quello a frame, oggi rivolgono la propria attenzione alla gestione dell'imprecisione, ossia dell'incertezza legata al procedimento, col ricorso alla teoria degli insiemi fuzzy⁴⁴.

I sistemi esperti sono molto utilizzati nell'area di analisi del merito creditizio, infatti ultimamente sono stati sviluppati dei sistemi esperti con l'obiettivo di migliorare la classificazione delle imprese appartenenti all'area d'ignoranza della funzione discriminante lineare.

I campi di applicabilità dei sistemi esperti non si limitano all'area del credito ma interessano una pluralità di ambiti; anch'essi però, alla pari di altri strumenti, presentano dei punti deboli. Il problema fondamentale è che per la realizzazione di un sistema esperto in grado di risolvere problemi particolarmente complessi è necessario molto tempo e un gruppo di analisti del credito di grande qualità. La scelta di compiere un progetto del genere impegna seriamente la struttura di una banca, sia in termini temporali sia in termini organizzativi, inoltre nella maggior parte dei casi non vi sono nemmeno le condizioni affinché tale progetto possa essere attuato.

In aggiunta, l'esperienza empirica dimostra che tali sistemi esperti sono meno precisi rispetto ad altri strumenti per il fatto che hanno come obiettivo la gestione della totalità delle fasi del rapporto creditizio, obiettivo che non può essere raggiunto se non diminuendo la specificità previsionale del modello in ogni singola fase.

Proprio a causa di queste limitazioni sono stati creati nuovi sistemi di apprendimento automatico, non influenzati dai problemi sopra descritti, i quali permettessero quindi di migliorare sia la velocità sia la precisione con cui giungere alla soluzione dello specifico problema esaminato.

⁴⁴ La teoria degli insiemi fuzzy è stata introdotta più di 40 anni fa da Lotfi Zadeh nell'opera "Fuzzy Sets". *Information and Control*, vol.8 (1965). La logica fuzzy, o logica sfumata, è una logica in cui si può attribuire a ciascuna proposizione un grado di verità compreso tra 0 e 1, intendendo con grado di verità quanto può essere vera una proprietà.

Infatti, a differenza della logica classica in cui una proprietà può essere vera (valore uguale a 1) o falsa (valore uguale a 0), nella logica fuzzy una proprietà può assumere un valore intermedio tra 0 e 1. Oggi la teoria degli insiemi fuzzy e le attività collegate sono applicate nei modelli di controllo ottimale, nei processi decisionali in condizioni di incertezza, negli studi comportamentali e nei modelli di intelligenza artificiale.

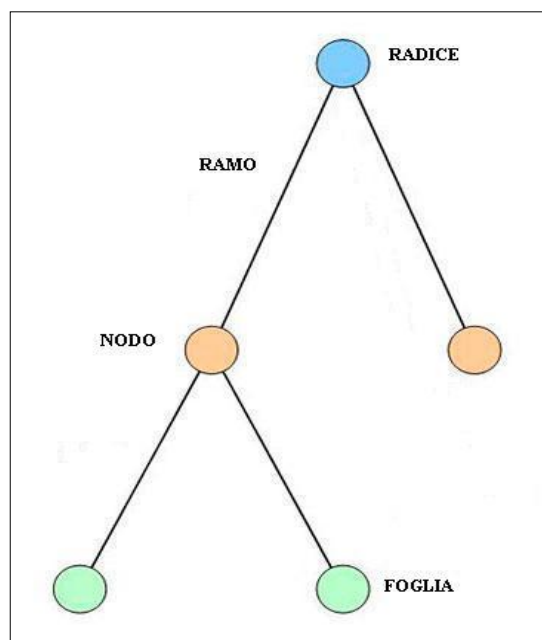
3.2 Alberi Decisionali⁴⁵

Gli Alberi Decisionali appartengono a quella branca dell'Intelligenza Artificiale nota con il nome di "Machine Learning" o Apprendimento Automatico, la quale contiene una serie di modelli che negli ultimi tempi, a fronte dei limiti riconosciuti ai sistemi esperti, hanno acquisito importanza.

Come descritto in precedenza, queste tecniche cercano di superare il principale punto debole riscontrato nel processo di attuazione dei sistemi esperti, ossia le distorsioni che possono crearsi nella fase di estrazione e formalizzazione della conoscenza. Infatti essi svolgono un ruolo di supporto all'azione dell'ingegnere della conoscenza e permettono quindi di velocizzare i tempi di messa a punto richiesti dai diversi modelli.

Gli alberi decisionali costituiscono delle tecniche che permettono di classificare un gruppo di oggetti sulla base di alcune caratteristiche osservate sugli stessi. Una volta rilevate le caratteristiche significative ai fini della classificazione, partendo dall'insieme generale degli oggetti, attraverso una serie di passaggi si giunge a comporre dei sottoinsiemi più "puri" rispetto ai sottoinsiemi degli stadi precedenti.

Figura 3.2: Albero Decisionale semplificato



Fonte: www.mokabyte.it

⁴⁵ Cenni tratti da G. Szego e F. Varetto (1999), *Il rischio Creditizio: misura e controllo*, ed. Utet, Torino.

Come si può notare dalla Figura 3.2, un albero decisionale presenta la struttura di una ramificazione discendente che, partendo dalla “radice” dell’albero, giunge alle “foglie”. Queste ultime costituiscono le classi o i sottoinsiemi specifici che distinguono l’oggetto classificato; esse rappresentano i punti terminali dell’albero e discendono dai diversi “nodi” decisionali che lo compongono.

Perciò, nel caso in cui si voglia classificare un oggetto, il procedimento è molto semplice: si parte dalla radice dell’albero in cui viene effettuato il primo test discriminatorio, in base al risultato conseguito si segue un apposito “ramo” che conduce ad un nodo decisionale (in tutti i nodi decisionali vengono condotti dei test su specifiche caratteristiche allo scopo di discriminare al meglio gli oggetti); a seconda del risultato ottenuto, dal nodo si diparte un ramo che conduce alternativamente ad un ulteriore nodo decisionale oppure ad una foglia; nel caso in cui si giunga ad una foglia il processo ha termine.

Questo procedimento presenta però un problema essenziale, infatti quando gli oggetti sono particolarmente complessi è necessario avvalersi per la classificazione di un elevato numero di rami e foglie, ma ciò costringe ad incrementare notevolmente la complessità degli Alberi Decisionali.

Per di più alberi troppo complessi tendono a generare fenomeni di overfitting⁴⁶ con conseguente limitazione della capacità di classificazione corretta di nuovi oggetti, non appartenenti al campione di partenza. Allo scopo di evitare tale fenomeno, ma con l’obiettivo di mantenere inalterata la capacità discriminatoria degli alberi decisionali, sono state sviluppate una serie di tecniche tese a semplificarne la struttura in quei casi in cui si rivelasse troppo complessa.

Esistono due procedimenti che possono essere attuati per semplificare la struttura di un albero⁴⁷:

- *la potatura in avanti*; con questa tecnica si cerca di valutare l’utilità proveniente dall’utilizzo di una nuova variabile nella classificazione. Nel caso in cui quest’utilità risulti essere troppo bassa può essere conveniente bloccare il procedimento all’interno di quel sotto-albero per non continuare in una classificazione che porterebbe

⁴⁶ In statistica si parla di overfitting (eccessivo adattamento) quando un modello statistico si adatta al campione osservato usando un numero eccessivo di parametri. Perciò, in presenza di un fenomeno di overfitting, le prestazioni (ovvero la capacità di prevedere) sui dati del campione aumenteranno, mentre le prestazioni sui dati esterni al campione di riferimento peggioreranno.

⁴⁷ G. Szego e F. Varetto (1999), *Il rischio Creditizio: misura e controllo*, ed. Utet, Torino.

solamente a delle foglie impure (foglie che incorporano una certa probabilità di distorsione della classificazione);

- *la potatura a ritroso*; questa tecnica, a differenza della precedente, parte dall'albero già completo e, successivamente ad una attenta analisi, cerca di eliminare quei sottoinsiemi che contribuiscono in maniera del tutto marginale alla capacità classificatoria dell'albero.

Le due tipologie di potatura si differenziano sostanzialmente per il fatto che mentre la prima opera in fase di costruzione dell'albero domandandosi se sia conveniente o meno effettuare ulteriori suddivisioni, la seconda agisce a costruzione ultimata cercando di capire se sia possibile migliorare la capacità classificatoria eliminando alcuni rami e sotto-alberi già realizzati. La fase di potatura è una fase molto delicata poiché la semplificazione deve essere effettuata mantenendo inalterata la validità dei risultati a cui il modello è giunto.

Un'ulteriore causa di distorsione nei risultati del modello potrebbe essere generata dalla presenza di informazioni non sempre accurate all'interno del campione di riferimento; per questo è necessaria un'attenta analisi dei casi durante la fase di apprendimento del sistema, al fine di evitare contraddizioni in grado di generare rumorosità all'interno dei dati⁴⁸.

Un altro svantaggio degli alberi decisionali è rappresentato dal fatto che il modello, nella classificazione delle imprese, fornisce una risposta dicotomica: 0 impresa sana, 1 impresa insolvente. In base a quanto delineato nel Capitolo 2, questo tipo di risposta non rispecchia la pluralità di situazioni in cui un'impresa può trovarsi e, proprio per questo, potrebbe essere più conveniente optare per l'utilizzo di uno score che assuma valori interni all'intervallo (0;1).

Forse però il punto di maggiore interesse di questo modello risiede nel fatto che un albero decisionale può essere convertito in un insieme di regole; tali regole non fanno altro che descrivere il percorso che è indispensabile seguire per raggiungere ogni foglia terminale, cercando così di semplificare l'approccio interpretativo.

È da sottolineare come gli alberi decisionali classifichino efficacemente gli oggetti appartenenti al campione di riferimento, ma, in base all'esperienza empirica, qualora ci si accinga a classificare degli oggetti esterni al campione è necessario ricorrere ad un nuovo addestramento del sistema ottenendo così un nuovo albero decisionale. Quindi gli alberi decisionali presentano una certa stabilità sia con riferimento al "training sample" sia con

⁴⁸ Con il termine "effetto rumore" si intendono gli errori che si possono creare durante la formalizzazione del contenuto informativo delle variabili.

riferimento agli outliers, questo anche in virtù del fatto che sono metodi non parametrici e non risentono delle limitazioni a cui sono soggetti i tradizionali approcci statistici.

Per concludere è d'obbligo evidenziare come, in relazione alla concessione del credito, gli alberi decisionali siano particolarmente utili non solo nell'individuare i fattori maggiormente significativi ai fini dell'analisi, ma anche nel fornire una visione d'insieme relativamente alla situazione aziendale in cui versa l'impresa.

3.3 Reti Neurali Artificiali

Nell'ampio ramo dell'Intelligenza Artificiale, le Reti Neurali Artificiali occupano un posto di rilievo tra i numerosi strumenti di apprendimento automatico. Il loro studio è iniziato nel 1949 con il lavoro di D.O.Hebb, il quale cercò di formalizzare in tale scritto il funzionamento del sistema nervoso biologico, ponendo così le basi per lo sviluppo di un algoritmo che potesse spiegarne la meccanica⁴⁹.

Le reti neurali artificiali, proprio per il fatto che simulano il funzionamento del sistema nervoso, hanno delle strutture non lineari che imitano relazioni molto complesse, garantendo un'operatività estremamente flessibile in più campi d'analisi.

Dal punto di vista strutturale, le reti neurali sono composte da un grande numero di neuroni o unità elaborative; ciascun neurone è interconnesso con gli altri ed è in grado di compiere delle semplici elaborazioni. I collegamenti tra i neuroni sono modificabili, infatti durante la fase di apprendimento, mediante un continuo scambio di informazioni tra la rete e il mondo esterno, il sistema modifica la propria struttura in modo da fornire un risultato più conforme alla realtà (uno schema semplice di rete neurale, rappresentante le singole unità e le connessioni che le legano, viene mostrato nella Figura 3.3).

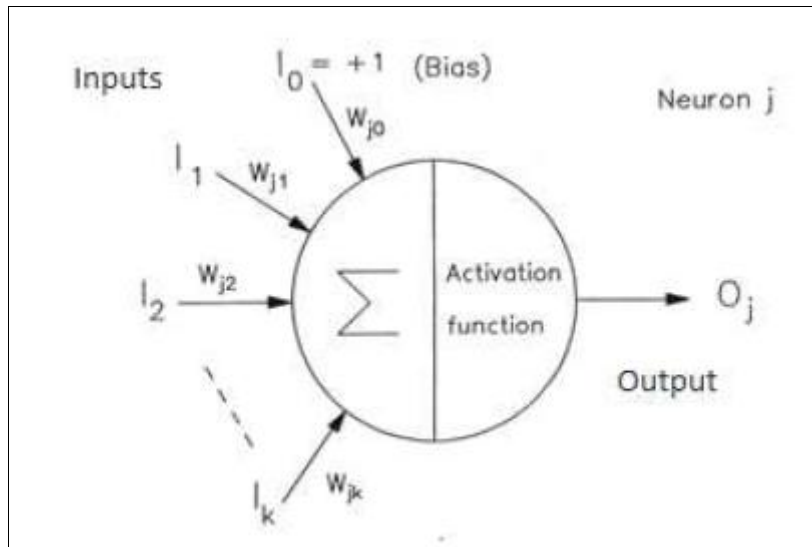
Ogni neurone riceve quindi degli input (I) o stimoli dall'esterno, o da altri neuroni a cui è collegato, con un'intensità o peso pari a $w_{j,i}$. L'input complessivo che il neurone j-esimo riceve è pari alla somma ponderata degli stimoli provenienti dall'esterno o dagli altri neuroni, espressa tramite un potenziale⁵⁰:

$$P_j = \sum_{i=1}^n w_{j,i} \cdot I_i$$

⁴⁹ Donald O. Hebb (1949), *The organization of behavior: a neuropsychological theory*, Wiley, New York.

⁵⁰ Si veda G. Szego e F. Varetto (1999), *Il rischio Creditizio: misura e controllo*, ed. Utet, Torino.

Figura 3.3: Schema di un neurone artificiale j-esimo



Fonte: A. Beltratti, S. Margarita and P. Terna (1996), "Neural Networks for economic and financial modeling", International Thomson Computer Press, London, UK.

L'output, o risposta, O_j generato dal neurone dipende dalla funzione di trasferimento, o funzione di attivazione, all'esterno del potenziale P_j .

Si può sottolineare come le informazioni in ingresso non abbiano tutte lo stesso effetto sul neurone; il ruolo svolto dai pesi $w_{j,i}$ è proprio quello di far valere il contributo delle informazioni più significative ai fini della classificazione. Infatti i valore assoluti assunti dai pesi si riferiscono all'importanza della connessione mentre il segno assunto indica la tipologia dello stimolo: in caso di segno positivo i collegamenti sono di tipo eccitatorio, in caso di segno negativo sono di tipo inibitorio.

Però è da evidenziare che la generazione dell'output da parte della rete non sempre avviene; ciò accade, ad esempio, quando la funzione di attivazione è una funzione di risposta "a salto" in cui si definisce una soglia di eccitazione; in questo caso l'output viene generato solamente se l'input complessivo risulta superiore alla soglia S_j stabilita:

$$P_j = \sum_{i=1}^n w_{j,i} \cdot I_i - S_j$$

Esistono una pluralità di funzioni di attivazione utilizzabili al fine di generare l'output del modello⁵¹, ma non esiste una vera e propria regola tecnica che definisca i casi in cui utilizzarle. L'algoritmo di apprendimento più diffuso necessita di una funzione di attivazione derivabile, per questo motivo una delle funzioni più utilizzate è la funzione logistica:

$$O_j = \frac{1}{1 + e^{-P_i}}$$

Le funzioni di attivazione generano dei valori interni ad un intervallo, nel nostro caso compresi tra 0 e 1, e gli estremi sono raggiunti solamente con pesi molto elevati. È importante rilevare come la risposta O_j del neurone j -esimo possa rappresentare sia l'output finale del modello, nel caso in cui il neurone sia quello d'uscita, sia un generico input per altri neuroni nel caso in cui il neurone sia connesso ad altri neuroni.

Esistono varie tipologie di reti neurali, infatti maggiore è la complessità del problema analizzato e maggiore è il grado di complessità della rete neurale impiegata per la soluzione:

- *Single-Layer Feed-forward*⁵² *Networks*, queste tipologie di reti sono le più semplici, sono costituite da un solo strato di neuroni (il caso limite è rappresentato da un solo neurone), ciascuno dei quali è a contatto diretto con gli stimoli provenienti dal mondo esterno e genera all'istante l'output della rete;
- *Multilayer Feed-forward Networks*, le reti neurali a due strati comprendono uno strato intermedio nascosto detto "hidden layer", il quale riceve gli stimoli dall'esterno, li elabora e, successivamente, li indirizza ad uno strato di output che formula le risposte della rete (Figura 3.4). Il numero di strati nascosti dipende dalla situazione in esame,

⁵¹ Esempi di funzioni di attivazione particolarmente utilizzate nelle reti neurali artificiali:

Funzione	Formula
A Gradino/A Salto	$f(x) = \begin{cases} 0 & \text{se } x < S_i \\ 1 & \text{se } x \geq S_i \end{cases}$
Lineare	$f(x) = x$
Logistica	$f(x) = \frac{1}{1 + e^{-P_i}}$
Arcotangente	$f(x) = \text{arctg}(x)$
Sinusoidale	$f(x) = \text{sen}(x)$
Gaussiana	$f(x) = e^{-x^2}$

⁵² Le reti neurali feed-forward sono definite anche come reti a funzionamento "in avanti" perché il flusso informativo scorre in un'unica direzione.

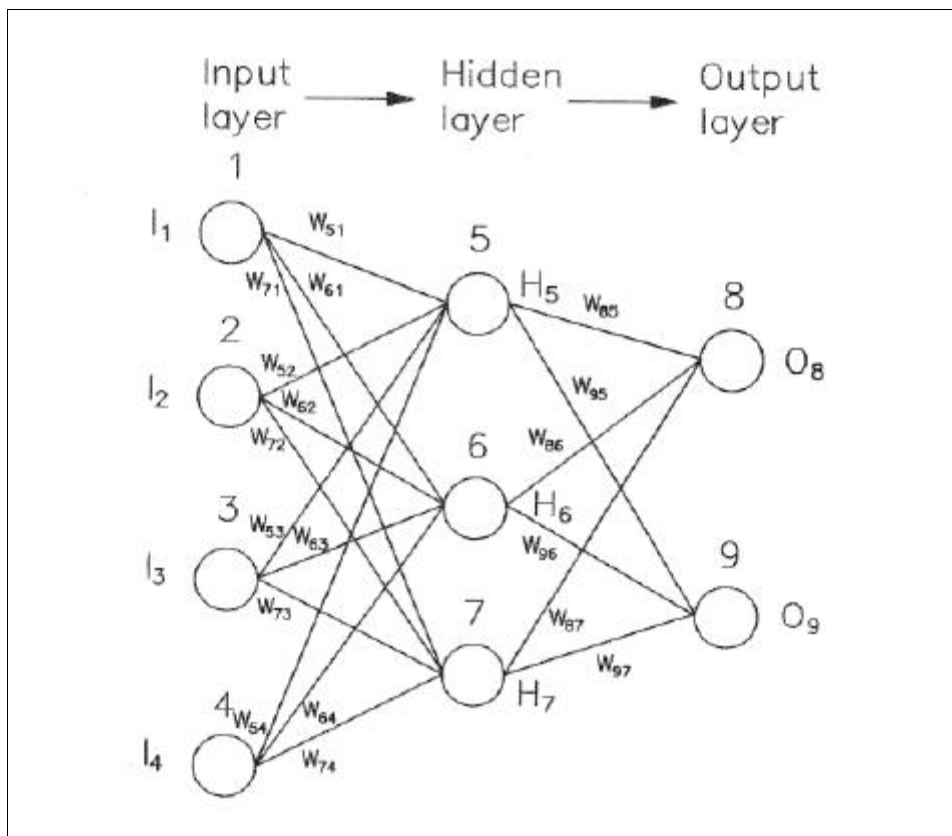
infatti quanto più complessa è la situazione tanto maggiore è il numero di livelli necessario per risolvere il problema;

- *Recurrent Neural Networks*, le reti neurali ricorrenti si differenziano dalle reti di tipo feed-forward per la presenza di almeno un nodo di feedback, ovvero l'esistenza di un neurone capace di generare un output che è anche input di se stesso o di un neurone dello stesso strato. La presenza dei nodi di feedback ha un profondo impatto sulla capacità di apprendimento della rete e sulle sue performance.

Per quanto riguarda la struttura della rete, rimane da evidenziare che non è univoca, ma le diverse modalità in cui si presenta dipendono da come i neuroni sono collegati tra di loro.

Una delle metodologie maggiormente utilizzate è la *fully connected neural network*, ovvero una rete neurale di piena connessione in cui ogni neurone di un livello è legato a tutti i neuroni del livello successivo (come riportato in Figura 3.4).

Figura 3.4: Esempio di rete neurale a due strati



Fonte: A. Beltratti, S. Margarita and P. Terna (1996), "Neural Networks for economic and financial modeling", International Thomson Computer Press, London, UK.

Come si evince dalla Figura 3.4, nelle reti feed-forward i neuroni appartenenti allo stesso livello non sono in connessione tra loro e il flusso di informazioni è unidirezionale, ovvero si muove dallo strato di input allo strato di output.

Punto di forza delle reti neurali, come delineato in precedenza, è rappresentato dall'estrema flessibilità della rete, che permette di modificare i pesi delle connessioni in modo che il sistema possa confrontare continuamente i risultati ottenuti con quelli desiderati originariamente. Questa è la caratteristica principale che differenzia le reti neurali dai tradizionali sistemi informatici in quanto, mentre questi ultimi definiscono una procedura standard attraverso cui analizzare il mondo esterno, le reti presentano un grado di dinamicità tale da rovesciare continuamente il meccanismo di apprendimento in modo da raggiungere, colmando i possibili divari di risposta, il risultato atteso.

Al fine di riassumere i punti cardine attraverso cui si snoda la risoluzione di un problema mediante l'uso delle reti neurali, si possono enunciare le seguenti fasi⁵³:

- la definizione di un meccanismo di apprendimento (comprendente anche la correzione dei pesi e le necessarie rettifiche);
- la preparazione di un training set rappresentativo del problema in esame;
- la definizione di un livello di accettabilità della risposta fornita dalla rete, cioè un livello di tolleranza per il possibile errore commesso dalla rete;
- la definizione dell'ordine in cui analizzare gli esempi e di una procedura di aggiustamento dei pesi (dopo ogni caso o dopo un certo numero di casi analizzati);
- l'analisi di una serie di casi di verifica con cui giudicare la capacità della rete di lavorare con input diversi da quelli ricevuti durante la fase di apprendimento.

A seguito di numerose iterazioni, la fase di apprendimento termina nel momento in cui il modello supera la fase di verifica; ciò significa che il possibile errore commesso dalla rete è stato ricondotto al livello di accettabilità e che non si giudicano più necessarie ulteriori rettifiche dei pesi. La rete neurale ha così raggiunto il suo stadio di stabilità ed è pronta per essere applicata, con un alto grado di efficienza, allo studio di casi reali.

Le reti neurali, qualora il campo di applicazione sia quello economico-finanziario, presentano una serie di vantaggi significativi; prima di tutto il modello ha un'ottima capacità di

⁵³ G. Szego e F. Varetto (1999), *Il rischio Creditizio: misura e controllo*, ed. Utet, Torino.

adattamento nel caso in cui si verificano dei mutamenti nell'ambiente circostante (tramite la rettifica dei pesi delle relazioni).

Un altro vantaggio riguarda la resistenza al rumore, ovvero la capacità di trattare con dati incompleti o influenzati da un grado di incertezza elevato; a questo si aggiunge poi la bassa sensibilità del modello alla presenza di outliers nel campione.

Un ulteriore vantaggio è riscontrato nella velocità computazionale del modello, ossia l'estrema celerità con cui avviene la fase di esecuzione poiché i calcoli eseguiti dai vari neuroni sono basilari (somme pesate e definizioni di valori di soglia).

Punto a sfavore del modello, invece, è costituito dal processo di apprendimento; in effetti per giungere alla formulazione finale sono necessari tempi lunghi e molti esempi in grado di calibrare il modello.

Lo svantaggio più significativo delle reti neurali è rappresentato dalla mancanza di una teoria rigorosa in grado di definire la struttura ottimale da utilizzare nell'affrontare i diversi problemi, il numero di strati e di neuroni per ogni strato, i valori iniziali dei pesi, la funzione di attivazione più consona e molti altri aspetti di funzionamento.

Per concludere si può affermare che le reti neurali possono costituire un modello molto utile nel valutare il rischio d'insolvenza delle imprese grazie alla loro estrema flessibilità nell'adattarsi a situazioni mutevoli o difficilmente riconducibili ad una rappresentazione matematica, tuttavia per minimizzare gli eventuali errori durante la classificazione è necessario che il numero di input e di esempi forniti alla rete sia il più grande possibile.

3.4 Algoritmi Genetici

Gli Algoritmi Genetici, similmente alle Reti Neurali Artificiali, simulano alcuni comportamenti biologici messi in atto dagli individui, in particolare focalizzano la loro attenzione sui meccanismi dell'evoluzione naturale.

Tali algoritmi vennero proposti da John H. Holland attorno agli anni '70⁵⁴ con l'obiettivo di creare, attraverso un'interpretazione artificiale dei principi base della selezione naturale di Darwin, un procedimento di ottimizzazione applicabile ad una pluralità di ambiti⁵⁵.

⁵⁴ John H. Holland (1975), *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control and artificial intelligence*, University of Michigan Press.

⁵⁵ Andrea Resti e Andrea Sironi (2008), *Rischio e valore nelle banche – Misura, regolamentazione, gestione*, EGEA, Milano.

Allora per poter comprendere al meglio il funzionamento degli Algoritmi Genetici è necessario indicare i principi su cui essi si basano; tali principi sono delineati nell'opera di Charles Darwin sul processo di evoluzione naturale⁵⁶: la selezione genetica e la riproduzione.

Il processo evolutivo si fonda quindi sull'adattamento della specie al mondo esterno; in effetti tutto questo trova affermazione nell'elemento tipico del meccanismo evolutivo, ovvero la ricombinazione genetica (cross-over). Questa tecnica consiste nel combinare i geni dei migliori individui al fine di ottenere soggetti con caratteristiche migliori (superfit), capaci di risultare vincenti nella lotta per la vita.

L'evoluzione, quindi, porta con sé un miglioramento progressivo della specie poiché tra le diverse generazioni hanno luogo dei cambiamenti volti a implementare le capacità di sopravvivenza dei singoli individui.

Gli Algoritmi Genetici si ispirano al naturale processo evolutivo; la principale differenza sta nel fatto che in questi gli individui da evolvere sono rappresentati dalle possibili soluzioni di un problema, con l'obiettivo finale di giungere ad una soluzione ottimale⁵⁷.

Infatti, alla pari del processo evolutivo, gli Algoritmi Genetici si caratterizzano per i seguenti principi⁵⁸: la selezione e la conseguente riproduzione dei soggetti migliori, la ricombinazione genetica e la mutazione casuale dei singoli geni.

Il processo di selezione di questi modelli si avvale di una funzione di valutazione o di fitness allo scopo di effettuare un'analisi sull'adeguatezza dei soggetti. Successivamente alla selezione avviene la riproduzione, la quale non si limita a replicare la popolazione iniziale ma, tramite la ricombinazione genetica, permette un'evoluzione della stessa.

In ultima istanza opera la mutazione casuale che, attraverso una ricombinazione dei geni, evita che la popolazione diventi troppo uniforme e, stabilizzandosi, perda la propria spinta evolutiva.

Un semplice esempio, al fine di ricondurre la trattazione al fenomeno dell'insolvenza, può essere rappresentato dalla costruzione di una funzione di valutazione, ipotizzata lineare per semplicità, che assegni valori più alti alle imprese sane e valori più bassi alle imprese insolventi⁵⁹:

⁵⁶ Charles Darwin (1859), *The Origin of Species by means of natural selection, or the preservation of favoured races in the struggle for life*, London.

⁵⁷ Andrea Resti e Andrea Sironi(2008), *Rischio e valore nelle banche – Misura, regolamentazione, gestione*, EGEA, Milano.

⁵⁸ G. Szego e F. Varetto (1999), *Il rischio Creditizio: misura e controllo*, ed. Utet, Torino.

⁵⁹ Esempio tratto da Andrea Resti e Andrea Sironi (2008), *Rischio e valore nelle banche – Misura, regolamentazione, gestione*, EGEA, Milano.

$$Z = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_i x_i + \dots + \alpha_m x_m$$

dove α_0 indica la costante

α_i indica il coefficiente i-esimo

x_i indica l'i-esimo indicatore di bilancio della popolazione

Ogni impresa costituisce una possibile soluzione del problema relativo al fenomeno dell'insolvenza; il ruolo della funzione di fitness è quello di individuare un valore numerico Z che possa rispecchiare l'adeguatezza di tali soluzioni al problema in esame.

La funzione di fitness così proposta mostra come ogni individuo, in questo caso si tratta di un'impresa, sia rappresentato da un vettore $\alpha = [\alpha_0 \dots \alpha_m]$. Tale vettore definisce il peso e il segno con cui i diversi indicatori influenzano lo score dell'impresa, infatti valori nulli di tali coefficienti indicano che i rispettivi indicatori economico-finanziari non vengono impiegati dall'impresa. La soluzione è rappresentata, quindi, dalla concessione di maggiori finanziamenti a quelle imprese che presentano valori di Z più alti, così da realizzare una selezione artificiale in base al grado di sopravvivenza dell'impresa stessa.

Il processo applicato dagli Algoritmi Genetici per giungere all'individuazione delle soluzioni migliori si compone di una serie di fasi⁶⁰:

- si genera in maniera casuale una popolazione iniziale di individui-soluzioni;
- si stima una funzione di fitness in grado di valutare la loro capacità di costituire una soluzione del problema in esame;
- si procede, tramite un algoritmo, alla selezione degli individui destinati a sopravvivere;
- si applica un secondo algoritmo, la ricombinazione genetica, che permette agli individui selezionati di riprodursi, generando così una nuova popolazione di soluzioni caratterizzata da un patrimonio genetico ereditato dalla popolazione precedente (i nuovi vettori α sono ottenuti mediante combinazioni della popolazione di soluzioni precedenti);
- si attua la mutazione casuale allo scopo di evidenziare improvvise modifiche nelle soluzioni della nuova generazione;

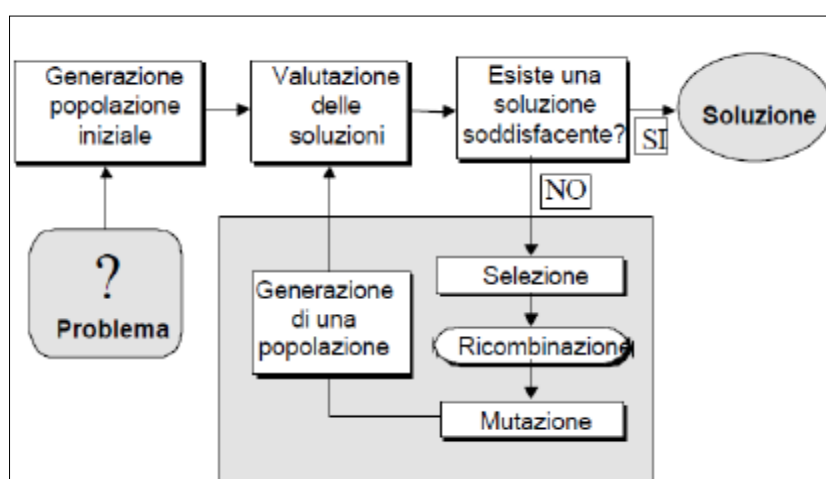
⁶⁰ Andrea Resti e Andrea Sironi(2008), *Rischio e valore nelle banche – Misura, regolamentazione, gestione*, EGEA, Milano.

- si misura, infine, l'adeguatezza (fitness) di ogni individuo della nuova generazione; se questa non soddisfa il problema allora è necessario ripetere ulteriormente tutto il procedimento fino a che non si perviene ad una soluzione accettabile per il sistema.

Il processo d'analisi termina con il raggiungimento delle soluzioni ottimali poiché, se si procede alla riproduzione di tali soluzioni, si nota che le popolazioni di soluzioni generate presentano una struttura del tutto simile a quella delle generazioni precedenti.

Le fasi del procedimento di individuazione delle soluzioni attuato dal sistema sono semplificate e riassunte nella figura sottostante (Figura 3.5).

Figura 3.5: La struttura di un Algoritmo Genetico



Fonte: U. Pomante (1998), *I modelli basati sugli algoritmi genetici*, in M. Marsella - A. Sironi (a cura di), *La misurazione e la gestione dei rischi di credito: modelli, strumenti e politiche*, Bancaria Editrice, Roma.

Gli Algoritmi Genetici presentano una serie di significativi vantaggi; prima di tutto non sono soggetti ad ipotesi statistiche restrittive come invece avviene per i metodi parametrici (in particolare l'Analisi Discriminante Lineare).

Inoltre, tramite l'applicazione della funzione di fitness, il modello permette di superare, ad esempio, il risultato dicotomico generato dagli Alberi Decisionali e di ottenere uno score numerico continuo che consente di separare in maniera più precisa e veritiera il gran numero di imprese.

Un ulteriore vantaggio riguarda la capacità di sondare con successo quei problemi caratterizzati da un ampio spazio di soluzioni; infatti, attraverso una ricerca adattiva, gli Algoritmi Genetici mettono a disposizione dell'analista finanziario l'intera popolazione di

soluzioni, consentendo così di approfondire la conoscenza dello spazio circostante alla soluzione ritenuta ottimale. Grazie a questa possibilità l'analista finanziario espande la propria prospettiva di analisi in modo da accrescere l'informazione e, di conseguenza, la comprensione relativa ai fenomeni economici che hanno dato luogo all'insolvenza⁶¹.

Rimane da evidenziare che questi modelli non possiedono solo vantaggi, ma presentano anche una serie di svantaggi, tra cui la mancanza di una garanzia relativamente al fatto che la soluzione delineata a seguito delle numerose operazioni sia effettivamente ottimale. In aggiunta, la maggior parte dei parametri (la numerosità della popolazione, la frequenza di riproduzione, la scelta dei coefficienti e degli indicatori di bilancio,...) non discendono dall'attuazione di un approccio teorico di riferimento, ma sono ottenuti tramite una serie di prove empiriche tenendo conto della qualità dei risultati ottenuti.

In conclusione, gli Algoritmi Genetici costituiscono dei validi strumenti di analisi per il fatto che offrono all'analista finanziario la possibilità di conoscere un ampio spettro di soluzioni ottimali e sub-ottimali. Però, la bontà del risultato risiede ancora una volta nell'abilità di quest'ultimo di selezionare gli indicatori economico-finanziari maggiormente rilevanti ai fini della classificazione, in modo da non creare distorsioni nei risultati, evitando così di giungere a delle soluzioni che potrebbero rivelarsi ingannevoli.

⁶¹ Andrea Resti e Andrea Sironi (2008), *Rischio e valore nelle banche – Misura, regolamentazione, gestione*, EGEA, Milano.

CAPITOLO 4

Data Envelopment Analysis: un metodo alternativo per la valutazione del Merito Creditizio

Nei capitoli precedenti si sono discussi i principali modelli che la letteratura, in materia di previsione delle insolvenze aziendali, è in grado di fornire all'analista finanziario. Però, a causa delle condizioni di mercato sempre più complesse, si è notato negli ultimi decenni un incremento nel numero dei fallimenti aziendali; ciò ha portato con sé la necessità per gli intermediari finanziari di identificare in anticipo i possibili sintomi dell'insolvenza aziendale. Proprio per il fatto che i modelli tradizionali, come visto in precedenza, non sono esenti dal presentare una serie di problemi nell'attuazione di tale obiettivo, vi è una continua ricerca rivolta all'esplorazione di metodi alternativi in grado di migliorarne la performance predittiva. In questo Capitolo viene presentato un modello alternativo a quelli tradizionali noto come Data Envelopment Analysis (DEA), il cui punto di forza giace nella capacità di comparare le imprese sulla base di un sistema di input e di output ad esse riferiti, ottenendo così un'unica misura capace di racchiudere la pluralità di aspetti della realtà aziendale.

I paragrafi successivi si sviluppano attorno alla presentazione matematica del modello e ai possibili vantaggi che tale tecnica può apportare ai modelli classici di previsione delle insolvenze.

4.1 Data Envelopment Analysis

La Data Envelopment Analysis fu introdotta per la prima volta da Charnes, Cooper e Rhodes nel 1978⁶², ed è una tecnica di programmazione lineare che permette di calcolare l'efficienza relativa di unità decisionali o unità produttive simili (Decision Making Units, DMU)⁶³ caratterizzate da un sistema di input e di output. L'obiettivo del loro lavoro era quello di estendere l'idea di Farrel⁶⁴ sulla possibilità di stimare una misura di efficienza tecnica

⁶² A. Charnes, W. W. Cooper, E. Rhodes (1978), *Measuring the Efficiency of Decision Making Units*, European Journal of Operational Research, Vol. 2, No. 6, pp. 429-444.

⁶³ Le *Decision Making Units*, DMU, sono unità produttive simili che operano nello stesso ambiente.

⁶⁴ M. J. Farrel (1957), *The Measurement of Productive Efficiency*, Journal of the Royal Statistical Society, Vol. 120, No. 3, pp. 253-290.

basandosi su una funzione di produzione, ovvero una funzione in grado di delineare il livello massimo di output ottenibile per un dato livello di input in maniera tecnicamente efficiente⁶⁵.

La DEA rientra tra i metodi non parametrici in quanto non richiede una specificazione della forma della funzione che lega tra loro gli input e gli output delle unità decisionali, e ciò non può se non rappresentare un vantaggio per la flessibilità del modello.

Questa tecnica analizza la prestazione delle unità decisionali in esame concentrandosi su tre concetti fondamentali: il processo produttivo, l'efficienza produttiva e l'efficacia produttiva.

Con il termine processo produttivo si intende la capacità di alcune unità decisionali di trasformare un dato numero di input in un livello desiderabile di output, quindi la produttività di un'unità decisionale discende dal rapporto tra l'output e l'input del processo produttivo.

Il calcolo di tale rapporto non crea particolari problemi nel caso in cui l'unità decisionale operi con un singolo input e un singolo output, ma diventa più complicato nel momento in cui il numero di input e di output aumenta per il fatto che le unità decisionali adottano un sistema di pesi diverso. Nell'approccio DEA i pesi da attribuire agli input e agli output non risultano predefiniti, ma per ogni unità decisionale si adottano i pesi che sono ad essa più favorevoli, massimizzando così la sua misura di efficienza.

L'efficienza produttiva, invece, misura di quanto si avvicina il valore osservato degli input e degli output di un'unità decisionale ad un valore ottimo teorico, ovvero la capacità di un'unità decisionale di ottimizzare le risorse disponibili. Il concetto di efficienza così enunciato è formato da due componenti⁶⁶:

- l'efficienza tecnica, che si riferisce all'abilità di un'unità decisionale di ottenere la maggior quantità di output per un dato livello di input o, al contrario, di impiegare il minor livello possibile di input per raggiungere il livello di output richiesto;
- l'efficienza allocativa, che si riferisce alla capacità di un'unità decisionale di combinare input e output in proporzioni ottimali tenendo conto dei loro prezzi di mercato e degli obiettivi del produttore (massimizzazione del profitto, minimizzazione dei costi).

Infine, con il termine di efficacia produttiva si intende la capacità di un'unità decisionale di assicurare con il livello di output raggiunto il soddisfacimento degli obiettivi prefissati.

⁶⁵ Paul C. Simak (1997), *DEA Based Analysis of Corporate Failure*, Working Paper, Department of Mechanical and Industrial Engineering, University of Toronto.

⁶⁶ M. J. Farrel (1957), *The Measurement of Productive Efficiency*, Journal of the Royal Statistical Society, Vol. 120, No. 3, pp. 253-290.

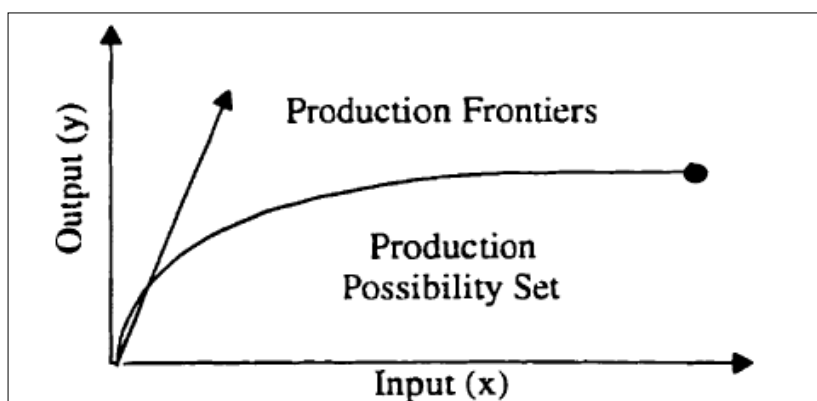
Quindi, è possibile affermare che un'unità decisionale risulta tecnicamente efficiente se un incremento di un qualsiasi output richiede un incremento in almeno un input o una riduzione di almeno un altro output, e se la diminuzione di un qualsiasi input richiede l'incremento di un qualsiasi altro input o la riduzione di almeno un output. Qualora si verifichi la situazione descritta, le condizioni di efficienza tecnica e allocativa sono soddisfatte in quanto l'unità decisionale utilizza l'esatta proporzione di input e output per il raggiungimento degli obiettivi prefissati.

4.2 Frontiere produttive: Teorica ed Empirica

Come detto in precedenza, la relazione tra la quantità di risorse impiegate e la quantità di output generati all'interno del processo produttivo è delineata da una funzione di produzione. Tale funzione, sulla base della relazione tra input e output, definisce una frontiera produttiva per i possibili gruppi di unità produttive ("Production Possibility Set"), ovvero una specie di barriera che indica il massimo ammontare di output che può essere generato a partire da un dato livello di input ed il livello minimo di input richiesto per ottenere un certo output.

La Figura 4.1 riporta un esempio di due frontiere produttive, in questo caso frontiere teoriche, la cui forma dipende principalmente dalle considerazioni e dalle assunzioni sottostanti.

Figura 4.1: Esempio di due Frontiere Produttive

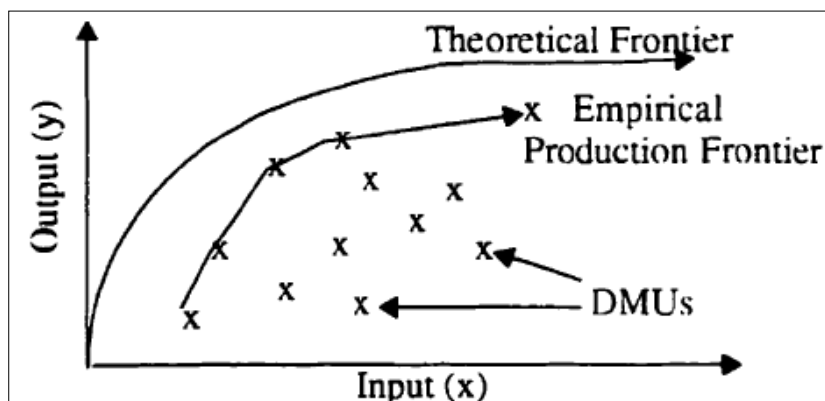


Fonte: Paul C. Simak (2000), Inverse and Negative DEA and their Application to Credit Risk Evaluation, Tesi di PhD, Department of Mechanical and Industrial Engineering, University of Toronto

Ecco quindi che l'efficienza di ogni unità produttiva potrebbe essere misurata rispetto alla frontiera teorica, ma in realtà ciò risulta difficile da realizzare per il fatto che, nella maggior parte dei casi, la frontiera teorica non è conosciuta. Infatti, solitamente, l'efficienza delle singole unità è misurata in relazione ad una frontiera produttiva empirica in quanto gli unici

dati disponibili, nel momento in cui si attua un'analisi del mondo reale, sono quelli direttamente osservabili. Attraverso l'uso di tali informazioni è possibile pervenire alla definizione di una frontiera produttiva empirica, ordinariamente inferiore alla frontiera teorica (Figura 4.2).

Figura 4.2: Frontiera Teorica e Frontiera Empirica



Fonte: Paul C. Simak (2000), *Inverse and Negative DEA and their Application to Credit Risk Evaluation*, Tesi di PhD, Department of Mechanical and Industrial Engineering, University of Toronto

La distinzione presentata in Figura 4.2 permette di identificare due misure di efficienza generalmente riferite ad ogni unità decisionale (DMU)⁶⁷:

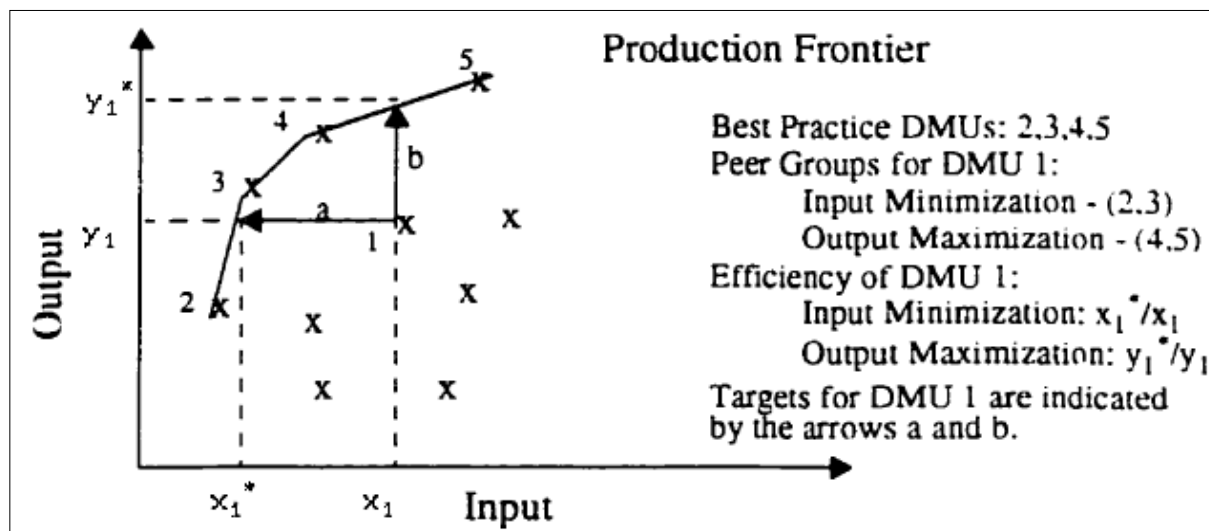
- *Efficienza Totale*: è misurata dalla distanza dell'unità decisionale dalla frontiera teorica ed indica la performance realizzata dall'unità rispetto alla migliore performance possibile;
- *Efficienza Relativa*: è misurata dalla distanza dell'unità decisionale dalla frontiera empirica ed evidenzia la performance dell'unità in esame rispetto alla migliore performance osservata.

In genere, la frontiera a cui si fa principalmente riferimento risulta la frontiera empirica; essa, oltre a rappresentare un benchmark di riferimento per il calcolo dell'efficienza di tutte le unità esaminate, costituisce anche il luogo di riferimento auspicabile per tutte le unità inefficienti ("efficiency reference set" o "peer group"). Infatti, come mostrato in Figura 4.3, le unità

⁶⁷ Paul C. Simak (2000), *Inverse and Negative DEA and their Application to Credit Risk Evaluation*, Tesi di PhD, Department of Mechanical and Industrial Engineering, University of Toronto.

produttive che formano tale frontiera e le loro combinazioni lineari costituiscono i diversi “peer groups” per le unità produttive inefficienti .

Figura 4.3: Un semplice esempio di ‘peer group’



Fonte: Paul C. Simak (2000), *Inverse and Negative DEA and their Application to Credit Risk Evaluation*, Tesi di PhD, Department of Mechanical and Industrial Engineering, University of Toronto

Per esempio, come si evince dalla Figura 4.3, i “peer groups” per la DMU#1 si differenziano a seconda del sentiero scelto per il raggiungimento dell’efficienza: in particolare le unità 2 – 3 rappresentano il “peer group” nel caso in cui si scelga di minimizzare l’input a parità di output, mentre le unità 4 – 5 costituiscono le unità di riferimento qualora ci si rivolga verso una massimizzazione degli output a parità di input.

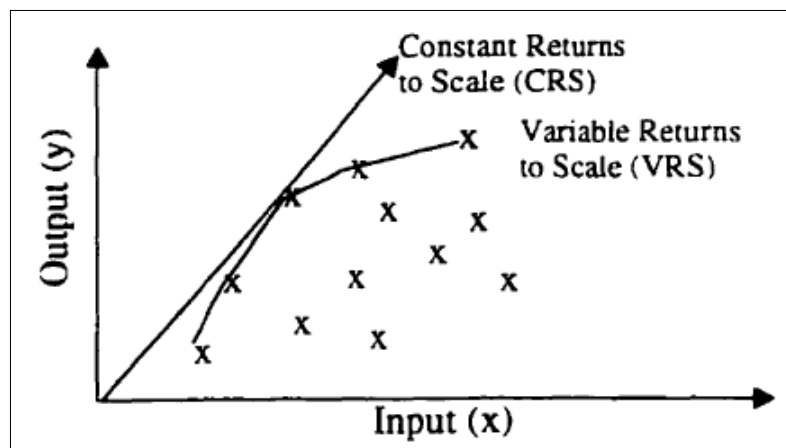
A questo punto è necessario sottolineare come esistano diversi modelli DEA distinti, per lo più, sulla base di due fattori: la tipologia della frontiera empirica impiegata nell’analisi e il sentiero scelto dalle unità decisionali per raggiungere la frontiera efficiente.

Per quanto riguarda la tipologia della frontiera si è soliti distinguere tra il modello CRS (Constant Returns to Scale), noto anche come modello CCR (Charnes, Cooper and Rhodes), e il modello VRS (Variable Returns to Scale) o modello BCC (Banker, Charnes and Cooper).

Il modello CRS opera ipotizzando che un dato aumento degli input impiegati, indipendentemente dall’operazione di scala, generi un proporzionale incremento negli output del modello. In maniera del tutto differente agisce, invece, il modello VRS, il quale non considera le restrizioni assunte dal modello CRS, evidenziando così come l’operazione di scala influisca notevolmente sulla relazione tra input e output.

Concentrando invece l'attenzione sui diversi sentieri di proiezione delle DMU inefficienti si possono individuare due modelli possibili⁶⁸:

Figura 4.4: Frontiere del Modello CRS e del Modello VRS



Fonte: Paul C. Simak (2000), *Inverse and Negative DEA and their Application to Credit Risk Evaluation*, Tesi di PhD, Department of Mechanical and Industrial Engineering, University of Toronto

- input-oriented, modello che proietta le DMU sulla frontiera efficiente riducendo tutti gli input e, allo stesso tempo, mantenendo costante il livello di output (sentiero rappresentato dalla freccia **a** in Figura 4.3);
- output-oriented, utilizza una logica inversa al precedente in quanto, mantenendo costante il livello di input, aumenta proporzionalmente gli output fino a raggiungere la frontiera efficiente (sentiero rappresentato dalla freccia **b** in Figura 4.3);

In aggiunta a questi due modelli vi è un altro approccio, definito dalla letteratura come Modello Additivo, il quale opera combinando tra loro una riduzione degli input simultaneamente ad un incremento degli output.

Tale approccio non costituisce un vero e proprio modello, ma rappresenta per lo più una variazione dei due modelli sopra descritti in quanto cerca di aggregare in un unico problema entrambi gli orientamenti (input-oriented e output-oriented)⁶⁹.

⁶⁸ A. Charnes, W. W. Cooper, E. Rhodes (1978), *Measuring the Efficiency of Decision Making Units*, European Journal of Operational Research, Vol. 2, No. 6, pp. 429-444 e Paul C. Simak (2000), *Inverse and Negative DEA and their Application to Credit Risk Evaluation*, Tesi di PhD, Department of Mechanical and Industrial Engineering, University of Toronto.

⁶⁹ Charnes A., Cooper W. W., Golany B., Seiford L. M., Stutz J. (1985), *Foundations of data envelopment analysis and Pareto-Koopmans empirical production functions*, Journal of Econometrics 30, 91-107 e Tone K. (2001), *A slacks - based measure of efficiency in Data Envelopment Analysis*, European Journal of Operational Research 130, pp. 498-509.

4.3 Il Modello CRS (Constant Returns to Scale)

Questo modello, capostipite dei modelli DEA, fu sviluppato e introdotto da Charnes, Cooper e Rhodes nel 1978. Al fine di presentarne la trattazione matematica è necessario utilizzare le seguenti notazioni:

$j = 1, \dots, n$ unità decisionali

$i = 1, \dots, m$ input

$r = 1, \dots, s$ output

x_{ij} quantità dell'input i per l'unità j

y_{rj} quantità dell'output r per l'unità j

u_r peso riferito all'output r

v_i peso riferito all'input i

Se i pesi associati, rispettivamente, agli output e agli input del modello sono noti, allora la teoria fondata sul rapporto tra i benefici e i costi consente di esprimere la misura di efficienza di una generica unità decisionale j come il rapporto tra la somma pesata degli output e la somma pesata degli input⁷⁰:

$$\frac{\sum_{r=1}^s u_r y_{rj}}{\sum_{i=1}^m v_i x_{ij}}$$

Però, come è stato definito in precedenza, i pesi associabili ai singoli input ed output generalmente non sono noti; per questo Charnes, Cooper e Rhodes proposero di derivare tali pesi risolvendo un particolare problema di programmazione non lineare. In particolare, la misura di efficienza tecnica per una generica unità decisionale j_0 (DMU_0) è data dalla soluzione del seguente problema di programmazione lineare frazionaria:

$$e_0 = \max \frac{\sum_{r=1}^s u_r y_{r0}}{\sum_{i=1}^m v_i x_{i0}} = \frac{u_1 y_{10} + u_2 y_{20} + \dots + u_s y_{s0}}{v_1 x_{10} + v_2 x_{20} + \dots + v_m x_{m0}} \quad (4.1)$$

⁷⁰ Le formule e gli esempi di questo capitolo sono tratti da: Wade D. Cook, Larry M. Seiford (2009), *Data envelopment analysis (DEA) – Thirty years on*, European Journal of Operational Research 192, pp. 1–17.

con rispetto dei seguenti vincoli⁷¹

$$\frac{\sum_{r=1}^s u_r y_{rj}}{\sum_{i=1}^m v_i x_{ij}} \leq 1 \quad j = 1, \dots, n \quad (4.2)$$

$$u_r > \varepsilon \quad r = 1, \dots, s \quad (4.3)$$

$$v_i > \varepsilon \quad i = 1, \dots, m \quad (4.4)$$

dove ε è un valore designato strettamente positivo.

Risolvendo il problema di massimizzazione (4.1) si ottiene la misura di efficienza per l'unità decisionale j_0 ; si può notare come il problema sia però sottoposto a dei vincoli, in particolare i vincoli (4.3) e (4.4) garantiscono che i pesi dell'unità decisionale assumano valori strettamente positivi. Inoltre la scelta dei pesi, come è già stato anticipato, non è casuale, infatti la funzione obiettivo (4.1) e il vincolo (4.2) impongono che i pesi scelti siano quelli ottimali per l'unità decisionale in esame e che il rapporto tra la somma pesata degli output e quella degli input risulti inferiore o al massimo pari ad 1 per tutte le unità decisionali. Il limite così definito, ovvero l'unità, viene raggiunto solamente dalle unità decisionali efficienti, ovvero quelle che giacciono sulla frontiera efficiente.

Inoltre, è possibile dimostrare che le misure di efficienza derivanti da tale problema di massimizzazione non risentono delle unità di misura adottate per calcolare gli input e gli output (però le unità di misura devono essere uguali per tutte le unità decisionali in esame affinché il problema abbia soluzione).

Il modello sopra descritto coinvolge il rapporto tra gli output e gli input ed è espressione di un modello input-oriented, ma, invertendo il rapporto e risolvendo il corrispondente problema di minimizzazione, si perviene ad un modello output-oriented.

Matematicamente il modello sopra descritto, rappresentante un problema di programmazione lineare frazionaria, può essere convertito, adottando precise condizioni⁷², in un più semplice problema di programmazione lineare:

$$e_0 = \max \sum_{r=1}^s u_r y_{r0} \quad (4.5)$$

⁷¹ Nel paper originale gli autori, tra i vincoli del modello, avevano imposto che i pesi fossero semplicemente non negativi ($\varepsilon = 0$); l'imposizione di un limite strettamente positivo fu introdotta in un paper successivo: A. Charnes, W. W. Cooper, E. Rhodes (1981), *Evaluating program and managerial efficiency: An application of data envelopment analysis to program follow through*, Management Science 27, pp. 668–697.

⁷² Le condizioni sottostanti la conversione del modello di programmazione lineare frazionaria sono le seguenti: $u_r = t u_r$ e $v_i = t v_i$, dove $t = (\sum_{i=1}^m v_i x_{i0})^{-1}$.

con i vincoli

$$\sum_{i=1}^m v_i x_{i0} = 1 \quad (4.6)$$

$$\sum_{r=1}^s u_r y_{rj} - \sum_{i=1}^m v_i x_{ij} \leq 0 \quad j = 1, \dots, n \quad (4.7)$$

$$u_r > \varepsilon \quad r = 1, \dots, s \quad (4.8)$$

$$v_i > \varepsilon \quad i = 1, \dots, m \quad (4.9)$$

Definito in questo modo il nuovo problema di massimizzazione, ed indicata con (θ^*, v^*, u^*) la soluzione ottima, allora l'unità decisionale in esame j_0 risulta CRS-efficiente se valgono le seguenti condizioni: $\theta^* = 1$, $v^* > 0$ e $u^* > 0$; in caso contrario risulta CRS-inefficiente.

Nel caso in cui l'unità decisionale j_0 risulti inefficiente, ovvero $\theta^* < 1$, allora è possibile incrementarne l'efficienza proiettando tale unità decisionale sulla frontiera efficiente.

Per fare questo è opportuno individuare, tra le DMU esaminate, quelle che soddisfano il vincolo (4.7) con il segno di uguaglianza; indicando con E_0 tale insieme risulta:

$$E_0 = \{j: \sum_{r=1}^s u_r^* y_{rj} = \sum_{i=1}^m v_i^* x_{ij}\} \quad j = 1, \dots, n \quad (4.10)$$

Tale insieme rappresenta, come detto in precedenza, l'efficiency reference set o peer group per l'unità inefficiente j_0 , e definisce la frontiera di efficienza raggiungibile attraverso una modifica nella quantità degli input o degli output.

L'insieme E_0 si differenzia dall'insieme delle possibilità produttive (production possibility set), infatti mentre il primo contiene solo le DMU riconosciute efficienti, il secondo contiene tutte le coppie di vettori (x, y) che possono essere ottenute dal processo produttivo.

Indicando con T l'insieme delle possibilità produttive, è possibile evidenziare le proprietà di cui gode tale insieme:

- tutte le coppie osservate (x_j, y_j) con $j = 1, \dots, n$ appartengono a T ;
- se (x, y) appartiene a T , allora anche (tx, ty) appartiene a T per ogni numero scalare $t > 0$;
- per una data coppia (x, y) in T , ogni coppia (x°, y°) in cui $x^\circ \geq x$ e $y^\circ \leq y$ appartiene a T ;
- tutte le combinazioni lineari positive delle coppie incluse in T appartiene a T .

Sulla base delle proprietà sopra elencate l'insieme delle possibilità produttive T può essere definito come:

$$T = \{(X, Y) \mid X \geq \sum_{j=1}^n \lambda_j X_j, Y \leq \sum_{j=1}^n \lambda_j Y_j, \lambda_j \geq 0\}$$

dove λ è un vettore non negativo di dimensione $N \times 1$, $\lambda = (\lambda_1, \dots, \lambda_n)$.

Al problema di programmazione lineare (4.5) - (4.9), definito anche problema primale, è associato un problema duale il quale, basandosi sulla definizione dell'insieme T appena fornita, è così specificato⁷³:

$$e_0 = \min \quad \vartheta_0 - \varepsilon (\sum_{r=1}^s s_r^+ + \sum_{i=1}^m s_i^-) \quad (4.11)$$

con i vincoli

$$\sum_{j=1}^n \lambda_j x_{ij} + s_i^- = \vartheta_0 x_{i0} \quad i = 1, \dots, m \quad (4.12)$$

$$\sum_{j=1}^n \lambda_j y_{rj} - s_r^+ = y_{r0} \quad r = 1, \dots, s \quad (4.13)$$

$$\lambda_j, s_i^-, s_r^+ \geq 0 \quad \forall i, j, r \quad (4.14)$$

dove ϑ_0 è una variabile reale compresa tra 0 e 1, s^- è il vettore che rappresenta gli eccessi di input, s^+ è il vettore che rappresenta i deficit di output e λ è un vettore non negativo di dimensione $N \times 1$ che connette input ed output.

L'obiettivo di questo problema duale è quello di identificare, all'interno dell'insieme delle possibilità produttive T, un'unità che sia in grado di garantire lo stesso livello di output dell'unità decisionale esaminata j_0 , riducendo il più possibile gli input impiegati.

A partire dai vincoli (4.12) – (4.13) è possibile ricavare la definizione dei vettori s^- e s^+ :

$$s^- = \vartheta x_{i0} - X\lambda \quad s^+ = Y\lambda - y_{r0} \quad (4.15)$$

dove X e Y sono le matrici che racchiudono gli input e gli output di tutte le unità decisionali in esame (X ha dimensione $m \times n$ e Y ha dimensione $s \times n$), $s^- \geq 0$ e $s^+ \geq 0$ per ogni soluzione (θ, λ) del problema.

⁷³ Wade D. Cook, Larry M. Seiford (2009), *Data envelopment analysis (DEA) – Thirty years on*, European Journal of Operational Research 192, pp. 1–17.

Si può dimostrare che la soluzione ottima θ^* del problema duale (4.11) – (4.14) è anche soluzione ottima del problema primale in virtù del teorema della dualità, quindi tale soluzione rappresenta la misura di efficienza, nell'ambito del modello CRS, per l'unità decisionale esaminata.

Una volta individuata la soluzione ottima θ^* , si può procedere all'identificazione degli eccessi di input e dei deficit di output tramite la risoluzione del seguente problema:

$$\max \rho = \sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+ \quad (4.16)$$

con i seguenti vincoli

$$s^- = \vartheta^* x_{i0} - X\lambda \quad (4.17)$$

$$s^+ = Y\lambda - y_{r0} \quad (4.18)$$

$$\lambda, s^-, s^+ \geq 0 \quad (4.19)$$

L'obiettivo del problema è quello di individuare una soluzione in grado di massimizzare la somma degli eccessi di input e deficit di output per la misura di efficienza del problema duale. La soluzione ottimale individuata dal problema (4.16) – (4.19) è chiamata soluzione di massimo slack, inoltre se tale soluzione soddisfa anche le condizioni per cui gli eccessi di input e deficit di output risultano nulli, $s^- = 0$ e $s^+ = 0$, allora viene definita soluzione zero – slack⁷⁴.

Quindi, si può dedurre che l'unità decisionale j_0 risulta CRS – efficiente se le soluzioni ottenute dalla risoluzione dei problemi (4.11) – (4.14) e (4.16) – (4.19) soddisfano le condizioni per cui $s^- = 0$, $s^+ = 0$ e $\theta^* = 1$, in caso contrario l'unità decisionale risulta inefficiente. Infatti se gli slack avessero valori positivi, ciò sottolineerebbe la possibilità di ottenere lo stesso livello di output riducendo la quantità di input o, viceversa, la possibilità di ottenere un output maggiore con la stessa quantità di input. La stessa situazione di inefficienza si verificherebbe per $\theta^* < 1$ in quanto vi sarebbe la possibilità di ottenere lo stesso livello di output con una significativa riduzione, in maniera proporzionale, di tutti gli input utilizzati nel processo.

⁷⁴ Qualora gli eccessi di input e deficit di output risultassero nulli, significa che l'unità decisionale in esame è collocata sulla frontiera efficiente in quanto non richiede modifiche delle quantità di input e di output impiegati nel processo produttivo.

Per semplificare la trattazione si può considerare il seguente esempio in cui viene presentato il modello CRS nel caso di un processo produttivo caratterizzato da un input e un output (la Tabella 4.1 presenta i dati delle sette DMU considerate).

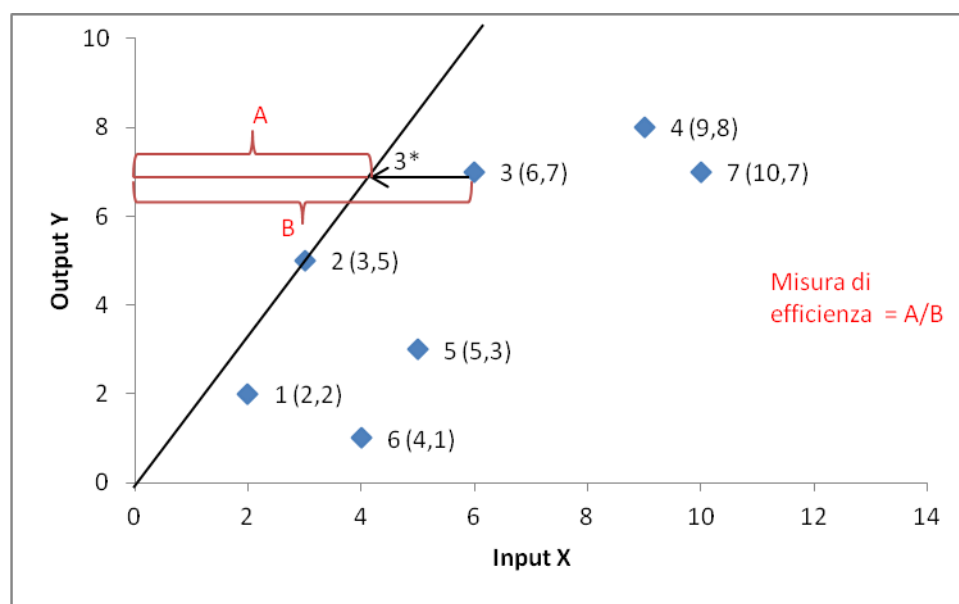
Tabella 4.1: Caso un input - un output

DMU	1	2	3	4	5	6	7
Input X	2	3	6	9	5	4	10
Output Y	2	5	7	8	3	1	7

Fonte: Wade D. Cook, Larry M. Seiford (2009), *Data envelopment analysis (DEA) – Thirty years on, European Journal of Operational Research* 192

La Figura 4.5 rappresenta la frontiera efficiente nel caso in esame; è possibile notare come tutte le unità decisionali facenti parte dell'insieme T appartengono alla frontiera o stanno al di sotto della stessa. Per risolvere il problema duale sopra descritto è necessario proiettare gli importi delle unità decisionali inefficienti sulla frontiera così da individuare le unità efficienti capaci di garantire lo stesso livello di output utilizzando un minor quantitativo di input. Soffermendosi nell'analizzare la DMU#3, ad esempio, la sua proiezione sulla frontiera efficiente, tenendo presente che la frontiera assume quella forma per il fatto che si sta analizzando il caso di rendimenti di scala costanti, è rappresentata dal punto 3*.

Figura 4.5: Modello CRS nel caso un input – un output



Fonte: Wade D. Cook, Larry M. Seiford (2009), *Data envelopment analysis (DEA) – Thirty years on, European Journal of Operational Research* 192

La misura di efficienza per la DMU#3 nei confronti del punto individuato sulla frontiera efficiente (3*) può essere calcolata come il rapporto $A/B = 4,2/6 = 0,7 = 70\%$. Il modello considerato nell'esempio è un modello input-oriented, ma per risolvere il corrispondente modello output-oriented è sufficiente proiettare verticalmente la DMU#3 in maniera da individuare quel punto sulla frontiera efficiente in grado di garantire un output maggiore mantenendo costante la quantità di input iniziale. Per trovare la misura di efficienza del modello output-oriented si inverte il rapporto, ottenendo così un valore pari ad 1,43.

La misura di efficienza per la DMU#3 si ritiene essere $1/1,43 = 0,7 = 70\%$, la stessa ottenuta per il modello input-oriented.

Un'ulteriore rappresentazione geometrica del problema duale è fornita dalla Figura 4.6, in cui si assiste all'analisi del caso due input – un output; la Tabella 4.2 riassume i dati relativi ai due input.

Si può notare come, a differenza del caso un input – un output, le combinazioni tecnicamente efficienti dei due input considerati formano un curva di livello o isoquanto. La forma di tale isoquanto, il quale rappresenta la frontiera efficiente, non è conosciuta, ma per poter applicare i concetti di efficienza descritti finora è necessario stimare tale frontiera attraverso due metodi alternativi: il metodo parametrico e il metodo non parametrico.

Tabella 4.2: Caso due input - un output

DMU	A	B	C	D	E	F	G
Input X ₁	3	5	8	12	6	8	10
Input X ₂	90	70	55	50	84	80	60

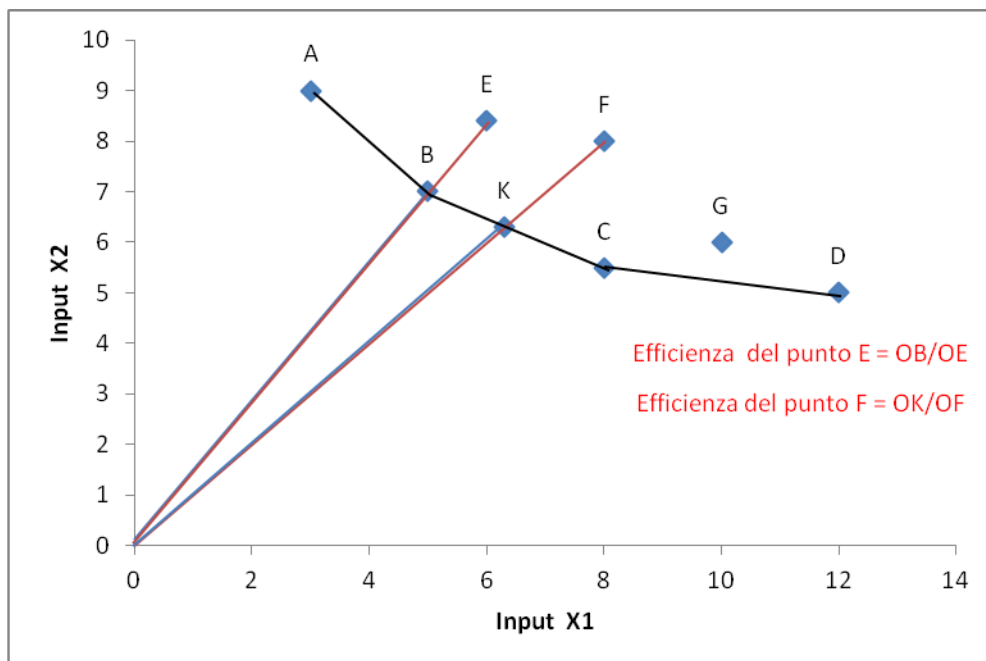
Fonte: Wade D. Cook, Larry M. Seiford (2009), Data envelopment analysis (DEA) – Thirty years on, European Journal of Operational Research 192

Il primo metodo, al fine di giungere ad una misura di efficienza per le DMU analizzate, richiede che la forma della frontiera efficiente sia stabilita a priori. Il metodo non parametrico, invece, determina la frontiera efficiente sulla base delle DMU che mostrano le migliori performance, senza dover necessariamente ipotizzare a priori la forma potenzialmente assumibile dalla frontiera.

Quindi, in base a quanto è stato detto, la curva di livello impiegata per l'analisi del caso due input – un output è costituita da un isoquanto non parametrico convesso e lineare a tratti⁷⁵.

⁷⁵ Tale tipologia di isoquanto, proposta da Farrell, è costruita in modo che nessuna delle DMU analizzate stia alla sinistra o al di sotto dell'isoquanto stesso.

Figura 4.6: Modello CRS nel caso due input – un output



Fonte: Wade D. Cook, Larry M. Seiford (2009), *Data envelopment analysis (DEA) – Thirty years on. European Journal of Operational Research* 192

È possibile affermare che le unità decisionali A, B, C e D, essendo le unità che compongono l'isoquante, sono tecnicamente efficienti ($\theta_A = \theta_B = \theta_C = \theta_D = 1$).

Le unità decisionali E ed F risultano, invece, tecnicamente inefficienti; per misurare tale inefficienza rispetto alle DMU che stanno sulla frontiera è necessario ricorrere ad un metodo formulato da Farrell⁷⁶ e Debreu⁷⁷. Tale metodologia determina il grado di inefficienza sulla base della distanza radiale, ovvero la distanza tra la DMU in esame e l'origine degli assi⁷⁸.

Ad esempio l'efficienza dell'unità decisionale E è pari a $\theta_E = 0,833 = 83,3\%$ e la sua proiezione sulla frontiera altro non è che l'unità decisionale B. La misura di efficienza è data dal rapporto OB/OE, il quale assume valori compresi tra 0 e 1, quindi tanto più E è vicina a B tanto più tale rapporto assume valori prossimi ad 1 e tanto più l'unità E sarà ritenuta efficiente. L'unità decisionale B riesce, quindi, a produrre lo stesso output dell'unità E utilizzando però una frazione esattamente pari all'83,3% di ciascun input impiegato nel processo produttivo.

⁷⁶ M. J. Farrell (1957), *The Measurement of Productive Efficiency*, Journal of the Royal Statistical Society, Vol. 120, No. 3, pp. 253-290.

⁷⁷ Debreu Gerard (1951), *The Coefficient of Resource Utilization*, Econometria, 19, 3, pp. 273-292.

⁷⁸ La distanza radiale è calcolata lungo la retta che collega le DMU in esame con l'origine degli assi, lungo tale retta il rapporto tra gli input rimane costante per il fatto che i loro importi variano in maniera proporzionale.

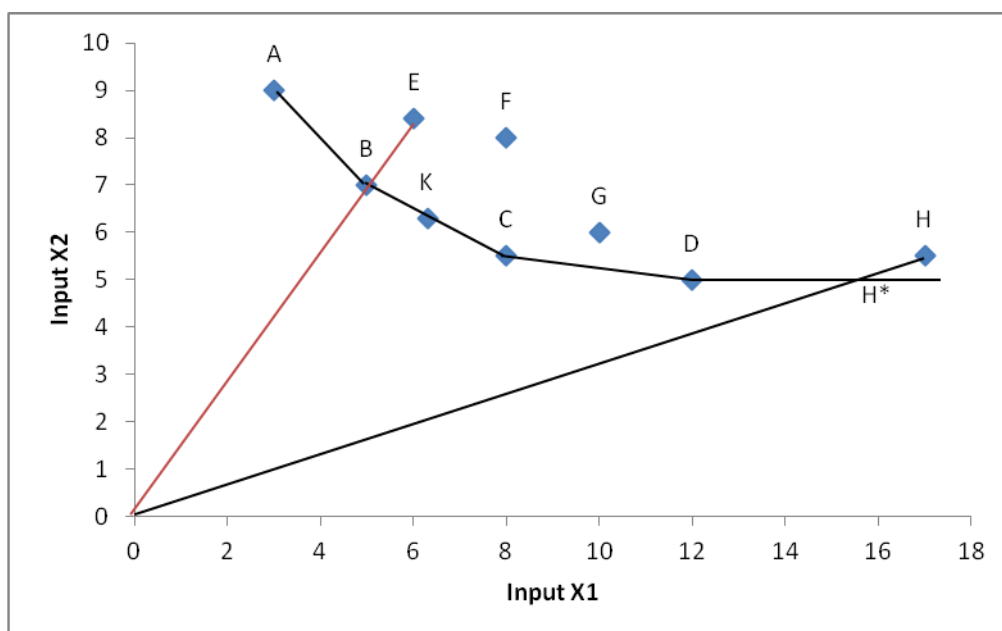
Considerando l'unità decisionale F si può notare come la sua proiezione sulla frontiera efficiente sia rappresentata dal punto K, quindi entrambe le unità decisionali B e C possono essere ritenuti dei benchmark di riferimento (peer group) per l'unità decisionale F.

Si può sottolineare come le unità decisionali A, B, C, D e K siano tutte soluzioni zero – slack per il problema esaminato in quanto soddisfano le uguaglianze $s_1^- = 0$ e $s_2^- = 0$, quindi tali unità risultano CRS – efficienti.

La figura 4.7 ridisegna il grafico riportato in figura 4.6 aggiungendo un'unità decisionale H la cui proiezione H^* ricade su un'estensione della frontiera e non sulla frontiera originale.

In questo caso lo slack corrispondente all'input X_1 (s_1^-) risulta essere positivo e pari alla distanza tra il punto H^* e il punto D.

Figura 4.7: Proiezione di un'unità decisionale impropriamente involuppata



Fonte: Wade D. Cook, Larry M. Seiford (2009), *Data envelopment analysis (DEA) – Thirty years on. European Journal of Operational Research 192*

A questo punto si dice che l'unità decisionale H è impropriamente involuppata⁷⁹, in quanto un'unità decisionale situata nel punto H^* risulta debolmente efficiente⁸⁰ rispetto alle unità A, B, C e D, le quali, essendo punti della frontiera produttiva empirica, risultano fortemente efficienti.

⁷⁹ La proprietà fondamentale della frontiera efficiente è quella di involuppare (to envelop) tutte le DMU esaminate, infatti il nome Data Envelopment Analysis deriva dallo sviluppo di questa proprietà della frontiera.

⁸⁰ Un'unità decisionale (DMU) risulta debolmente efficiente se e solo se: $\theta^* = 1$, $s_i^- \neq 0$ e/o $s_r^+ \neq 0$.

4.4 Il modello VRS (Variable Returns to Scale)

Il modello VRS fu introdotto da Banker, Charnes e Cooper nel 1984; l'obiettivo di tale lavoro fu quello di estendere il modello CRS in modo da poter valutare anche il caso di rendimenti di scala variabili.

Dal punto di vista matematico, considerando sempre una generica unità decisionale j_0 (DMU₀), il modello VRS differisce dal modello CRS in quanto nella formulazione del problema di programmazione lineare frazionaria vi è l'aggiunta di una variabile⁸¹:

$$e_0 = \max \frac{[\sum_{r=1}^s u_r y_{r0} - u_0]}{\sum_{i=1}^m v_i x_{i0}} \quad (4.20)$$

con i seguenti vincoli

$$\frac{\sum_{r=1}^s u_r y_{rj} - u_0}{\sum_{i=1}^m v_i x_{ij}} \leq 1 \quad j = 1, \dots, n \quad (4.21)$$

$$u_r > \varepsilon \quad r = 1, \dots, s \quad (4.22)$$

$$v_i > \varepsilon \quad i = 1, \dots, m \quad (4.23)$$

dove u_0 non ha vincoli per quanto riguarda il segno.

Il problema di programmazione lineare derivante dal problema (4.20) – (4.23) è così definito:

$$e_0 = \max \sum_{r=1}^s u_r y_{r0} - u_0 \quad (4.24)$$

con i vincoli

$$\sum_{i=1}^m v_i x_{i0} = 1 \quad (4.25)$$

$$\sum_{r=1}^s u_r y_{rj} - u_0 - \sum_{i=1}^m v_i x_{ij} \leq 0 \quad j = 1, \dots, n \quad (4.26)$$

$$u_r > \varepsilon \quad r = 1, \dots, s \quad (4.27)$$

$$v_i > \varepsilon \quad i = 1, \dots, m \quad (4.28)$$

e il corrispondente problema duale è pari a:

⁸¹ Wade D. Cook, Larry M. Seiford (2009), *Data envelopment analysis (DEA) – Thirty years on*, European Journal of Operational Research 192, pp. 1–17.

$$e_0 = \min \vartheta_0 - \varepsilon (\sum_{r=1}^s s_r^+ + \sum_{i=1}^m s_i^-) \quad (4.29)$$

con i vincoli

$$\sum_{j=1}^n \lambda_j x_{ij} + s_i^- = \vartheta_0 x_{i0} \quad i = 1, \dots, m \quad (4.30)$$

$$\sum_{j=1}^n \lambda_j y_{rj} - s_r^+ = y_{r0} \quad r = 1, \dots, s \quad (4.31)$$

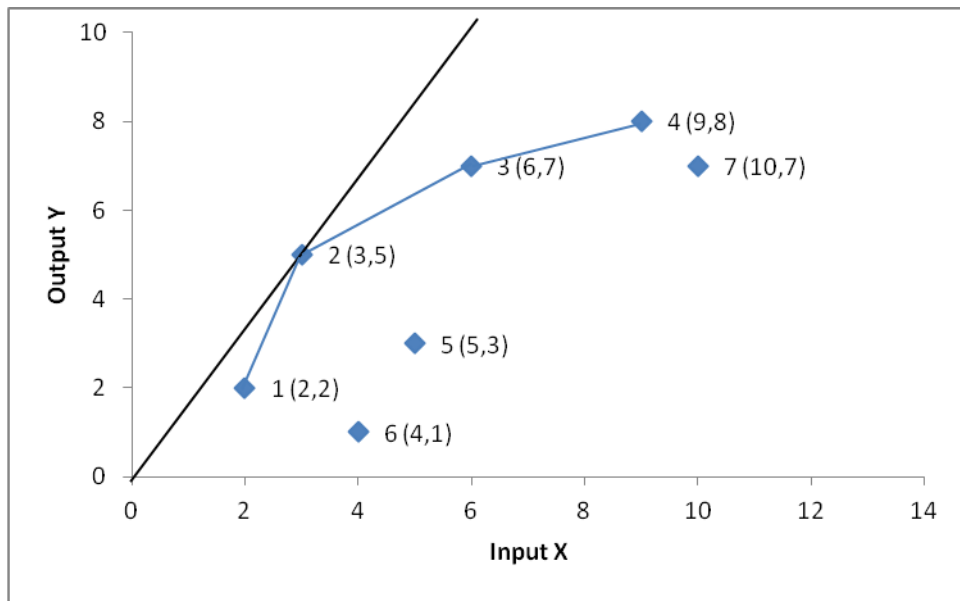
$$\sum_{j=1}^n \lambda_j = 1 \quad (4.32)$$

$$\lambda_j, s_i^-, s_r^+ \geq 0 \quad \forall i, j, r \quad (4.33)$$

Come si può notare, la differenza principale tra il modello CRS e il modello VRS è costituita dall'aggiunta, tra i vincoli del problema duale, del seguente vincolo di convessità:

$$\sum_{j=1}^n \lambda_j = 1.$$

Figura 4.8: Modello VRS nel caso un input – un output



Fonte: Wade D. Cook, Larry M. Seiford (2009), *Data envelopment analysis (DEA) – Thirty years on*, *European Journal of Operational Research* 192

Per una rappresentazione grafica del modello VRS è utile riprendere il caso un input – un output precedentemente esposto, infatti è possibile ridisegnare il grafico in modo da raffigurare la frontiera efficiente nel caso di rendimenti di scala variabili (Figura 4.8).

Si può notare come la frontiera efficiente del modello VRS sia composta dai segmenti 1 – 2, 2 – 3 e 3 – 4; in particolare, la porzione di frontiera che va dal punto 1 al punto 2 costituisce

quella parte di frontiera caratterizzata da rendimenti di scala crescenti. Il punto 2 però, essendo espressione di rendimenti di scala costanti, non è incluso nella frontiera del modello e svolge il ruolo di punto di separazione tra rendimenti di scala crescenti e rendimenti di scala decrescenti. Infatti tutti i punti della frontiera che stanno alla destra del punto 2, cioè i punti 3 e 4, costituiscono quella parte di frontiera caratterizzata da rendimenti di scala decrescenti. Alla pari del modello CRS, risolvendo il problema duale (4.29) – (4.33) per la DMU₀ si ottiene la soluzione ottima θ^* ; in base a tale risultato, l'unità decisionale j_0 è VRS – efficiente se la soluzione soddisfa le condizioni $s_i^{-*} = 0$, $s_r^{+*} = 0$ e $\theta^* = 1$, presentandosi come una soluzione zero – slack. Chiaramente, è possibile dimostrare che qualsiasi unità decisionale risulti efficiente per il modello CRS, è efficiente anche per il modello VRS⁸².

4.5 Il Modello Additivo o Modello Pareto – Koopmans (PK)

Il modello additivo, introdotto da Charnes et al.⁸³ nel 1985, si differenzia dai precedenti due modelli per il fatto che, mentre quest'ultimi sono costruiti basandosi su delle proiezioni radiali (in maniera specifica il modello input-oriented mantiene fisso il livello degli output e riduce proporzionalmente gli input e il modello output-oriented incrementa proporzionalmente gli output mantenendo fissi gli input), il modello in questione opera combinando entrambi gli orientamenti (riduce gli input ed incrementa gli output).

La Figura 4.9, raffigurante il caso un input – un output in ipotesi di rendimenti di scala costanti, presenta come nel modello additivo, allo scopo di raggiungere la frontiera efficiente, sia permessa qualsiasi direzione all'interno del triangolo A – B – C, ossia qualsiasi combinazione di input ed output in quell'area.

Questo modello, scelto tra la pluralità dei modelli DEA, verrà impiegato nel prossimo capitolo per valutare la performance delle aziende e discriminare le imprese sane da quelle maggiormente soggette ad insolvenza.

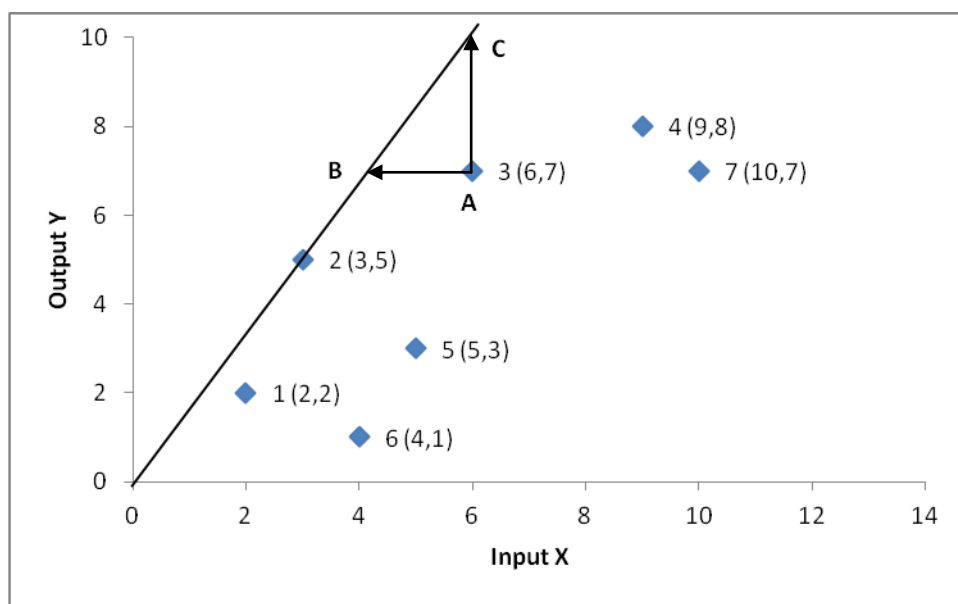
Esistono diverse versioni del modello additivo, ma la versione più comunemente utilizzata al fine di valutare la probabilità di fallimento da parte delle aziende è la seguente:

$$PK_0 = \max \sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+ \quad (4.34)$$

⁸² Wade D. Cook, Larry M. Seiford (2009), *Data envelopment analysis (DEA) – Thirty years on*, European Journal of Operational Research 192, pp. 1–17.

⁸³ A. Charnes, W. W. Cooper, B. Golany, L. M. Seiford, J. Stutz (1985), *Foundations of data envelopment analysis and Pareto–Koopmans empirical production functions*, Journal of Econometrics 30, 91–107.

Figura 4.9: Modello Additivo nel caso un input – un output



Fonte: Wade D. Cook, Larry M. Seiford (2009), *Data envelopment analysis (DEA) – Thirty years on*, *European Journal of Operational Research* 192

con i seguenti vincoli

$$\sum_{j=1}^n \lambda_j x_{ij} + s_i^- = x_{i0} \quad i = 1, \dots, m \quad (4.35)$$

$$\sum_{j=1}^n \lambda_j y_{rj} - s_r^+ = y_{r0} \quad r = 1, \dots, s \quad (4.36)$$

$$\sum_{j=1}^n \lambda_j = 1 \quad (4.37)$$

$$\lambda_j, s_i^-, s_r^+ \geq 0 \quad \forall i, j, r \quad (4.38)$$

La condizione di convessità del vincolo (4.37) implica che si sta utilizzando il modello VRS, infatti la frontiera dal modello (4.34) – (4.38) è del tutto uguale a quella ottenuta dal modello (4.29) – (4.33). Quindi è possibile affermare che un'unità decisionale j_0 è PK – efficiente, ovvero soddisfa le condizioni per cui $s_i^{-*} = 0$, $s_r^{+*} = 0$ e $\theta^* = 1$, se e solo se risulta VRS – efficiente⁸⁴.

Come si vedrà, il modello additivo determina l'inefficienza delle DMU rispetto ad ogni input ed ogni output, quindi, dal nostro punto di vista, identifica le imprese sane e quelle insolventi sulla base della presenza o meno di slack positivi.

⁸⁴ Wade D. Cook, Larry M. Seiford (2009), *Data envelopment analysis (DEA) – Thirty years on*, *European Journal of Operational Research* 192.

Tale modello presenta però un problema, infatti, a differenza dei modelli radiali descritti in precedenza (CRS e VRS), dalla risoluzione del problema (4.34) – (4.38) non si ottiene uno score di efficienza compreso nell'intervallo [0,1].

Charnes et al.⁸⁵ cercarono di trovare un rimedio a tale problema e proposero, come score di efficienza del modello additivo, il seguente valore:

$$Q_0 = \delta \left(\frac{\sum_{i=1}^m s_i^-}{x_{i0}} + \frac{\sum_{r=1}^s s_r^+}{y_{r0}} \right) \quad (4.39)$$

dove $\delta = 1/(m+s)$ e Q_0 è soggetto ai vincoli (4.35) – (4.38).

Grazie a questo espediente è possibile dimostrare che lo score di efficienza del modello additivo è compreso tra 0 e 1, inoltre la divisione degli slack per x_{i0} e y_{r0} è stata effettuata allo scopo di rendere gli slack univariati e monotoni decrescenti⁸⁶.

La scelta, compiuta da molti studiosi, di utilizzare il modello additivo per la valutazione della probabilità d'insolvenza delle imprese, è dovuta alla presenza di una serie di vantaggi rispetto ai modelli radiali precedentemente esposti.

In primo luogo, ad esempio, il modello additivo presenta il vantaggio, rispetto al modello CRS, di riuscire a trattare indicatori finanziari con valori negativi grazie alla proprietà definita *translation invariance*⁸⁷. Questa permette al modello di comprendere nella valutazione valori negativi di input e output e, qualora si verifichi una traslazione nei dati esaminati, la frontiera efficiente e la classificazione delle imprese come efficienti o inefficienti rimane invariata⁸⁸.

Quindi, nell'analizzare la possibilità di fallimento delle imprese, tale proprietà risulta estremamente utile in quanto molti degli indici considerati nei modelli di predizione dell'insolvenza assumono valori negativi.

In secondo luogo, una caratteristica singolare del modello additivo è costituita dal fatto che è sufficiente solo l'esame degli slack per determinare l'efficienza di una specifica unità decisionale. Ciò differenzia sensibilmente il modello additivo dai modelli radiali, infatti

⁸⁵ A. Charnes, W. W. Cooper, B. Golany, L. M. Seiford, J. Stutz (1985), *Foundations of data envelopment analysis and Pareto–Koopmans empirical production functions*, Journal of Econometrics 30, 91–107.

⁸⁶ I. M. Premachandra, Y. Chen, J. Watson (2011), *Dea as a tool for predicting corporate failure and success: A case of bankruptcy assessment*, Omega 39, pp.620-626.

⁸⁷ Il modello additivo presenta la proprietà *translation invariance* solamente quando viene aggiunto il vincolo di convessità al problema (4.34) – (4.38).

⁸⁸ I. Ali, L. Seiford (1990), *Translation invariance in data envelopment analysis*, Operations Research Letters 9, pp.403-405.

quest'ultimi necessitano di analizzare sia lo score di efficienza sia il valore degli slack per giungere a definire l'efficienza di una specifica DMU.

Inoltre nei modelli radiali, nel caso in cui lo score di efficienza sia rappresentato da un rapporto, i risultati dipendono dall'approccio adottato nell'analisi: la misura di efficienza dell'approccio input-oriented è diversa da quella fornita dall'approccio output-oriented.

Il modello additivo, invece, riesce ad evitare il problema associato allo score di efficienza in forma di rapporto (si sta facendo riferimento alla formula (4.1)) incorporando all'interno della misura di efficienza entrambi gli slack di input ed output.

In virtù di questi vantaggi, è possibile sostenere che l'impiego del modello additivo, per giungere alla definizione di uno stato di efficienza, risulta molto più semplice rispetto all'utilizzo dei precedenti modelli radiali⁸⁹.

4.6 Altri modelli DEA

I modelli CRS, VRS e Additivo, sono quelli più utilizzati in letteratura, ma con il passare del tempo sono stati introdotti modelli alternativi; questi ultimi altro non sono che delle estensioni realizzate al fine di colmare le lacune presentate dai modelli originari.

Nelle prossime righe riportiamo solo alcuni dei molteplici modelli alternativi formulati negli ultimi decenni⁹⁰.

Nel 1984 Deprins, Simar e Tulkens introducono il modello FDH (Free Disposal Hull), che opera ipotizzando che l'insieme delle possibilità produttive (production possibility set) sia non convesso.

Un ulteriore modello, noto con il nome di modello SBM (Slacks – Based Measures), fu introdotto da Tone nel 2001; particolarità di tale modello è che risulta invariato rispetto alle unità di misura e monotono crescente in ogni slack di input e in ogni slack di output⁹¹.

Nel 1999 Cooper propone un'estensione del modello additivo: il modello RAM (Range Adjusted Measure). Questo modello è del tutto simile al modello additivo, però presenta la caratteristica aggiuntiva che lo score di efficienza cade nell'intervallo [0,1].

⁸⁹ I. M. Premachandra, G. S. Bhabra, T. Sueyoshi (2009), *DEA as a tool for bankruptcy assessment: A comparative study with logistic regression technique*, European Journal of Operational Research 193, pp.412-424.

⁹⁰ Cenni tratti da Wade D. Cook, Larry M. Seiford (2009), *Data envelopment analysis (DEA) – Thirty years on*, European Journal of Operational Research 192, pp. 1–17.

⁹¹ K. Tone (2001), *A slacks – based measure of efficiency in Data Envelopment Analysis*, European Journal of Operational Research 130, pp. 498-509.

La varietà di modelli che caratterizza la metodologia DEA riflette la pluralità di campi in cui tale tecnica è stata impiegata. Infatti dalla prima applicazione relativa alla misurazione di efficienza di attività appartenenti al settore pubblico e di organizzazioni senza scopi di lucro, la DEA è stata applicata anche al campo della finanza (banche, imprese, fondi comuni d'investimento) e a molte altre aree d'interesse (sanità, trasporti, esercito,...).

4.7 I modelli di previsione dei fallimenti aziendali⁹²

Nei paragrafi precedenti sono stati presentati i tratti fondamentali della metodologia DEA, ma non è stato definito il motivo per cui lo studio che verrà condotto in seguito si fondi proprio su tale modello. Per poter spiegare il perché sia stato scelto il modello additivo per la valutazione del fenomeno dell'insolvenza aziendale, è necessario ripercorrere la letteratura in merito, al fine di scoprire come i modelli di previsione delle insolvenze si sono evoluti nel tempo.

4.7.1 Il modello Z-Score di Altman

Riprendendo quanto già visto nel Capitolo 2, il primo modello ad affrontare il problema di previsione dei fallimenti aziendali fu introdotto da Altman nel 1968. Basandosi sull'Analisi Discriminante Lineare l'obiettivo era quello di ottenere, tramite una combinazione lineare di variabili finanziarie, un punteggio per ogni impresa appartenente al campione (punteggio chiamato Z-Score) che permettesse di distinguere le imprese sane da quelle insolventi sulla base di un punteggio soglia.

A seguito di una serie di studi condotti sul modello di Altman, prima Eisenbeis⁹³ e poi Grice and Ingram⁹⁴, hanno dimostrato come tale modello presentasse alcuni problemi statistici nelle sue applicazioni econometriche. In modo particolare, come già affermato nel Capitolo 2, la capacità previsionale del modello calava notevolmente nel momento in cui l'analisi veniva condotta al di fuori del campione di riferimento portando così ad una distorsione dei risultati. Successivamente a tali studi, al fine di superare i difetti presentati dal modello di Altman, molti autori hanno fornito il loro contributo alla letteratura in materia, tentando di impiegare

⁹² Cenni tratti da I. M. Premachandra, Gurmeet Singh Bhabra, Toshiyuki Sueyoshi (2009), *DEA as a tool for bankruptcy assessment: A comparative study with logistic regression technique*, European Journal of Operational Research, 193, pp. 412-424 e I. M. Premachandra, Yao Chen, John Watson (2011), *DEA as a tool for predicting corporate failure and success: A case of bankruptcy assessment*, Omega 39, pp.620-626.

⁹³ Robert A. Eisenbeis (1977), *Pitfalls in the application of Discriminant Analysis in Business, Finance and Economics*, The Journal of Finance, Vol.32 no.3, pp.876-900.

⁹⁴ J. S. Grice, R. W. Ingram (2001), *Tests of the generalizability of Altman's bankruptcy prediction model*, Journal of Business Research 54, pp.53-61.

diversi tipi di Analisi Discriminante o di altre tecniche similari capaci di offrire un maggiore grado di affidabilità e correttezza nella classificazione delle imprese come sane ed insolventi. Come esempio si può considerare il lavoro di Freed and Glover⁹⁵ in cui si assiste alla trasformazione del problema di Analisi Discriminante Lineare in un problema di programmazione matematica risolvibile tramite un algoritmo di programmazione lineare.

4.7.2 Modelli di regressione: Logit e Probit

Con il passare degli anni, come già mostrato in maniera dettagliata nel Capitolo 2, vennero sviluppati nuovi modelli per superare le carenze presentate dal modello di Altman, e studiosi quali Ohlson⁹⁶, Zavgren⁹⁷ e Keasey e Watson⁹⁸, utilizzarono i modelli di regressioni logit e probit per la classificazione delle imprese.

In un ulteriore lavoro, Collins e Green⁹⁹ dimostrarono come la tecnica di regressione logistica non fosse superiore all'Analisi Discriminante solo per quanto riguarda la capacità diagnostica generale, ma anche nel determinare le potenziali imprese insolventi del campione.

Le tecniche di regressione però, pur essendo superiori all'Analisi Discriminante, non erano del tutto prive di difetti, infatti anch'esse non fornivano dei risultati robusti nel momento in cui l'analisi veniva condotta all'esterno del campione d'origine (come è stato evidenziato nel Capitolo 2).

I tre metodi sopra descritti dominarono la letteratura per molti anni, ma non sono gli unici metodi di cui un'analista finanziario può usufruire per una valutazione sul rischio d'insolvenza delle imprese. Negli ultimi tempi si sono sviluppati una serie di approcci tesi a svolgere un ruolo importante nella valutazione dei fallimenti aziendali, tra questi si possono includere la metodologia CUSUM (Cumulative sum), il metodo basato sull'utilizzo di tecniche dinamiche non lineari noto come approccio Caos ed infine la Data Envelopment Analysis.

⁹⁵ Ned Freed, Fred Glover (1981), *Simple but powerful goal programming models for discriminant problems*, European Journal of Operational Research 7, pp.44-60.

⁹⁶ James A. Ohlson (1980), *Financial Ratios and the Probabilistic Prediction of Bankruptcy*, Journal of Accounting Research, Vol.18 no.1, pp109-131.

⁹⁷ Christine V. Zavgren (1985), *Assessing the vulnerability to failure of American industrial firms: a Logistic Analysis*, Journal of Business Finance & Accounting 12, pp.19-45.

⁹⁸ K. Keasey, R. Watson (1987), *Non financial symptoms and the prediction of small company failure: a test of Argenti's Hypotheses*, Journal of Business Finance & Accounting 14(3), pp.335-354.

⁹⁹ R. A. Collins, R. D. Green (1982), *Statistical Methods for Bankruptcy Forecasting*, Journal of Economics and Business 34, pp.349-354.

4.7.3 Il metodo CUSUM

La metodologia CUSUM differisce dai modelli precedenti per il fatto che, a differenza di questi ultimi, non si limita ad utilizzare i dati di un singolo periodo per analizzare le condizioni finanziarie di un'impresa, ma considera il comportamento delle serie storiche relative alle variabili finanziarie nel loro complesso, in modo da non tralasciare informazioni passate significative per l'analisi.

Dal punto di vista matematico il modello CUSUM può essere così descritto¹⁰⁰:

$$C_{i,t} = \min(C_{i,t-1} + Z_{i,t} - K, 0) < -L \quad K, L > 0 \quad (4.40)$$

dove $C_{i,t}$ e $Z_{i,t}$ sono dei punteggi di performance relativi a serie storiche cumulative e annuali per una i -esima impresa al tempo t , mentre K e L sono dei parametri di sensibilità.

Il punteggio $Z_{i,t}$ è una funzione complessa delle variabili finanziarie considerate nell'analisi:

$$Z_{i,t} = \beta_0 + (X_{i,t} - A_h - X_{i,t-1}B_1 - \dots - X_{i,t-k}B_k)\beta_1 = \beta_0 + A_{f,s}\beta_1 + \varepsilon_{i,t}\beta_1 \quad (4.41)$$

$$\beta_0 = (1/2D)A_f\Sigma^{-1}A'_f = D/2 \quad (4.42)$$

$$\beta_1 = -(1/D)\Sigma^{-1}A'_f \quad (4.43)$$

$$D^2 = -A_f\Sigma^{-1}A'_f \quad (4.44)$$

dove β_0 e β_1 sono i parametri CUSUM, A_h è il vettore delle intercette per le imprese sane, $A_{f,s}$ sono le deviazioni dal vettore A_h associate ad un vettore di attributi estratto dalle imprese fallite s anni prima del fallimento (per semplicità $A_f \equiv A_{f,1}$), $\varepsilon_{i,t}$ è un vettore di errori indipendentemente distribuito con media zero e D è la distanza di Mahalanobis¹⁰¹ dei termini di errore delle variabili nelle popolazioni sane e fallite.

¹⁰⁰ Si veda E. Kahya, P. Theodossiou (1998), *Predicting corporate financial distress: A time-series CUSUM methodology*, Working Paper.

¹⁰¹ La distanza di Mahalanobis indica la distanza tra un campione incognito e uno noto tenendo in considerazione le correlazioni all'interno dell'insieme dei dati. Da un punto di vista formale la distanza di Mahalanobis di un vettore multivariato $x = (x_1, x_2, \dots, x_N)^T$ rispetto ad un insieme di valori di valore medio $\mu = (\mu_1, \mu_2, \dots, \mu_N)^T$ e matrice di covarianza S è definita come: $D_M(x) = \sqrt{(x - \mu)^T S^{-1} (x - \mu)}$.

Lo score $Z_{i,t}$ rappresenta quindi la performance annuale dell' i -esima impresa e, in accordo con il modello CUSUM, la performance complessiva dell' i -esima impresa al tempo t è misurata dal punteggio cumulativo $C_{i,t}$. Generalmente, per un'impresa sana i valori $Z_{i,t}$ sono positivi e maggiori di K , così da far risultare i punteggi $C_{i,t}$ pari a zero; contrariamente a questo, nel caso di un'impresa in difficoltà, i valori $Z_{i,t}$ sono minori di K così da far risultare negativi i punteggi $C_{i,t}$ ¹⁰².

Il segnale che permette di capire se le condizioni finanziarie dell'impresa siano mutate sensibilmente, aumentando così le possibilità di fallimento, è dato dal crollo del valore $C_{i,t}$ al di sotto della soglia $-L$.

Il modello CUSUM cerca quindi di determinare l'origine del deterioramento delle condizioni dell'impresa in modo da poter segnalare anticipatamente le possibili difficoltà finanziarie che l'impresa sarà chiamata a fronteggiare in un futuro prossimo.

4.7.4 L'approccio CAOS

A differenza del modello CUSUM, l'approccio Caos utilizza delle tecniche dinamiche non lineari nel predire il fallimento delle imprese. Questo aspetto, però, può indurre a pensare che i metodi che si avvalgono di dinamiche non lineari per predire i fallimenti possano dare risultati soddisfacenti solamente con un periodo di anticipo particolarmente ristretto a causa della forte sensibilità alle condizioni iniziali¹⁰³.

La metodologia si compone di più passi, in primo luogo si calcola, tramite un programma chiamato Chaos Data Analyzer, l'esponente di Lyapunov per ogni impresa, ovvero un indicatore che misura la dipendenza dalle condizioni iniziali delineando così la presenza o meno di dinamiche caotiche :

$$\lambda_i = \lim_{t \rightarrow \infty} \frac{1}{t} \log_2 \frac{p_i(t)}{p_i(0)} \quad (4.45)$$

In secondo luogo, avvalendosi di un test t per la verifica d'ipotesi, si calcolano le differenze tra i valori degli esponenti di Lyapunov in quanto i valori ottenuti saranno diversi per le imprese fallite rispetto alle altre.

¹⁰² Il punteggio CUSUM $C_{i,t}$ si comporta come un processo casuale (random walk) continuo a tempo discreto con una soglia di valore massima pari a zero.

¹⁰³ D. H. Lindsay, A. Campbell (1996), *A Chaos Approach to Bankruptcy Prediction*, Journal of Applied Business Research, Volume 12, Number 4.

Il neo più profondo di questo sistema è costituito dal fatto che l'alta sensibilità alle condizioni iniziali porta con sé un problema fondamentale, ovvero che una limitata quantità d'informazioni in partenza rende difficile prevedere i movimenti del sistema dopo un certo intervallo di tempo.

Tale limitazione temporale non deve però essere vista come uno svantaggio rispetto agli altri modelli in quanto, nella maggior parte dei casi, le richieste di previsione delle insolvenze riguardano il futuro prossimo, ovvero se l'impresa andrà in fallimento entro i prossimi due anni. Proprio in virtù di tale affermazione, e del fatto che vari studi hanno dimostrato la possibilità di ricondurre il comportamento delle imprese ai sistemi caotici, diversi lavori sono stati indirizzati all'applicazione di tali dinamiche non lineari ad una pluralità di campi tra cui l'analisi della probabilità d'insolvenza.

4.7.5 La metodologia DEA

La Data Envelopment Analysis, invece, è stata impiegata da Premachandra nei suoi studi più recenti¹⁰⁴; in queste ricerche vi è l'impiego del modello additivo in quanto, come già visto nel paragrafo 4.5, presenta una serie di vantaggi rispetto ai modelli input-oriented ed output-oriented, tali da migliorare la classificazione e la precisione con le quali imprese sane e insolventi vengono tra loro distinte.

Nel prossimo capitolo, partendo dagli studi condotti da Premachandra, si cercherà, tramite un campione casuale composto da imprese sane e fallite, di analizzare la capacità della metodologia DEA di predire lo stato d'insolvenza delle imprese.

Il modello DEA impiegato nell'analisi al fine di individuare l'efficienza relativa delle singole imprese è il modello additivo proposto da Charnes nel 1985¹⁰⁵:

$$\max es^- + es^+ \quad (4.46)$$

subordinatamente ai vincoli

¹⁰⁴ I. M. Premachandra, Gurmeet Singh Bhabra, Toshiyuki Sueyoshi (2009), *DEA as a tool for bankruptcy assessment: A comparative study with logistic regression technique*, European Journal of Operational Research, 193, pp. 412-424 e I. M. Premachandra, Yao Chen, John Watson (2011), *DEA as a tool for predicting corporate failure and success: A case of bankruptcy assessment*, Omega 39, pp.620-626.

¹⁰⁵ A. Charnes, W. W. Cooper, B. Golany, L. Seiford, J Stutz (1985), *Foundations of Data Envelopment Analysis for Pareto – Koopmans efficient empirical production functions*, Journal of Econometrics 30, pp. 91-107.

$$X\lambda + s^- = x_0 \quad (4.47)$$

$$Y\lambda - s^+ = y_0 \quad (4.48)$$

$$e\lambda = 1 \quad (4.49)$$

$$\lambda \geq 0, s^- \geq 0, s^+ \geq 0 \quad (4.50)$$

in cui n è il numero delle unità decisionali (imprese), k ed m sono rispettivamente il numero degli input e degli output, $X = (x_j) \in R^{k \times n}$ è una matrice di input $k \times n$, $Y = (y_j) \in R^{m \times n}$ è una matrice di output $m \times n$, e è un vettore riga con tutti gli elementi pari ad 1, s^- è il vettore degli slack di input, s^+ è il vettore degli slack di output, x_0 e y_0 sono rispettivamente i vettori colonna degli input e degli output per la 0 – esima impresa e $\lambda \in R^n$ è un vettore che connette gli input agli output.

Tale modello, paragonato con i metodi statistici precedenti, presenta delle caratteristiche uniche che lo rendono un ottimo strumento di previsione delle insolvenze¹⁰⁶:

- non richiede delle assunzioni a priori sulle relazioni tra gli input e gli output; il metodo ha la capacità di trattare una pluralità di input e di output mediante un modello matematico senza la necessità di spiegare i legami tra le singole misure in relazione alla performance dell'impresa a cui si riferiscono;
- rispetto ad alcuni approcci statistici classici (per esempio l'Analisi Discriminante Lineare), i quali assumono che le variabili esaminate siano distribuite come una variabile casuale normale multivariata, il modello DEA non ha bisogno di specificare la distribuzione che le variabili devono assumere al fine di non creare errori e distorsioni nei risultati;
- il modello analizza ogni singola impresa e genera il punteggio di efficienza, rappresentante la performance dell'impresa, in relazione all'intero campione d'analisi;
- non necessita, a differenza degli approcci statistici ed econometrici, di un campione particolarmente esteso per valutare il fenomeno dell'insolvenza.

¹⁰⁶ I. M. Premachandra, Yao Chen, John Watson (2011), *DEA as a tool for predicting corporate failure and success: A case of bankruptcy assessment*, Omega 39, pp.620-626.

Si può notare come la possibilità di operare con campioni numericamente piccoli di imprese costituisca un vero e proprio vantaggio in quanto permette di superare le difficoltà legate alla dimensione del campione.

In una recente ricerca sulla previsione delle insolvenze aziendali, Premachandra ha impiegato il modello additivo proposto da Charnes per identificare una frontiera dei fallimenti; i risultati così ottenuti sono stati confrontati con quelli ottenuti mediante la tecnica di regressione logistica ed è stato dimostrato come, rispetto a quest'ultima, il modello additivo ottenga dei risultati migliori nella previsione delle insolvenze¹⁰⁷.

Quindi, l'obiettivo dell'analisi che verrà condotta nel prossimo capitolo è quello di dimostrare come la Data Envelopment Analysis possa essere considerata un valido strumento nel diagnosticare il successo o il fallimento delle imprese all'interno del mondo economico – finanziario (in particolar modo l'attenzione è stata rivolta alle imprese di piccole e medie dimensioni).

¹⁰⁷ I. M. Premachandra, Gurmeet Singh Bhabra, Toshiyuki Sueyoshi (2009), *DEA as a tool for bankruptcy assessment: A comparative study with logistic regression technique*, European Journal of Operational Research, 193, pp. 412-424.

CAPITOLO 5

Caso applicativo: DEA come strumento d'analisi della probabilità d'insolvenza delle piccole e medie imprese

Sulla base di quanto esposto nel capitolo precedente, è possibile sostenere che la metodologia DEA, proprio per la sua capacità di non limitare il numero di input e output impiegati nel processo d'analisi, può essere utilizzata per la valutazione della probabilità d'insolvenza delle imprese.

Infatti, questa caratteristica permette di effettuare una valutazione del merito di credito delle diverse imprese tenendo conto di un'ampia varietà di indici e di informazioni provenienti dai bilanci aziendali e, di conseguenza, consente di realizzare una visione d'insieme dell'operato di ogni singola azienda.

In questo capitolo la metodologia DEA viene applicata ad un campione formato da piccole – medie imprese (PMI); la scelta di non concentrare l'analisi sulle grandi imprese rispecchia la convinzione che, nella maggior parte dei casi, i soggetti che si rivolgono agli intermediari finanziari per la concessione di finanziamenti sono per lo più soggetti appartenenti al mondo delle piccole e medie aziende.

Quindi, affinché l'approccio DEA possa essere considerato dagli istituti di credito un valido meccanismo per la valutazione del merito creditizio delle imprese, è necessario che presenti dei risultati classificatori soddisfacenti con riferimento alla porzione di clientela aziendale più ampia, ovvero quella costituita dalle piccole e medie imprese.

Nei paragrafi successivi vengono presentati i criteri in base ai quali è stata effettuata la selezione delle imprese facenti parte del campione e le variabili di input ed output impiegate nell'analisi.

5.1 Il campione di riferimento

L'analisi, come delineato in precedenza, è stata condotta su un campione casuale di 100 imprese di piccole e medie dimensioni non fallite, i cui dati di bilancio sono stati reperiti mediante il database AIDA Bureau Van Dijk.

A questo punto risulta utile, ai fini della trattazione, evidenziare i parametri in base ai quali un'impresa possa essere considerata una PMI. Proprio in merito a tale problema, la Commissione Europea ha provveduto a riformulare i criteri di definizione delle PMI tramite l'adozione della Raccomandazione n. 2003/361/Ce¹⁰⁸ che è stata recepita a livello nazionale dal Decreto Ministeriale 18 aprile 2005.

Sulla base di quanto esposto dalla normativa, affinché un'impresa possa essere qualificata come PMI deve soddisfare tre criteri: un criterio finanziario, uno legato al numero dei dipendenti e uno riguardante l'autonomia.

Noti i tre criteri di valutazione, i limiti posti dalla normativa sono i seguenti:

- *Numero dei dipendenti*: inferiore a 250 per le medie imprese, inferiore a 50 per le piccole imprese ed inferiore a 10 per le micro imprese;
- *Ammontare del Fatturato Annuo*: non superiore a 50 milioni di euro per le medie imprese, a 10 milioni di euro per le piccole imprese ed a 2 milioni di euro per le micro imprese;
- *Totale del Bilancio Annuo (Attivo Patrimoniale)*: non superiore a 43 milioni di euro per le medie imprese, a 10 milioni di euro per le piccole imprese ed a 2 milioni di euro per le micro imprese.

Per fatturato annuo si intende l'ammontare dei corrispettivi di competenza dell'esercizio derivanti dalla cessione di beni o dalla prestazione di servizi inerenti l'attività dell'impresa, nonché dalle cessioni di materie prime e sussidiarie, di semilavorati e di altri beni mobili, esclusi quelli strumentali, acquistati o prodotti per essere impiegati nella produzione.

Con il termine totale del bilancio annuo si indica il totale dell'attivo patrimoniale così come definito dall'articolo 2424 del codice civile (comprese successive modificazioni).

Fermo restando tali limiti, affinché un'impresa rientri nella relativa fascia dimensionale è necessario che sia verificato il criterio del numero di dipendenti e almeno uno dei restanti due parametri (fatturato o totale bilancio annuo).

Una volta enunciata la definizione di PMI, è necessario evidenziare le modalità in base alle quali è stato realizzato il campione d'analisi.

¹⁰⁸ Tale Raccomandazione è stata pubblicata in Gazzetta Ufficiale dell'Unione Europea L. 124, del 20 maggio 2003, con decorrenza dal 1° Gennaio 2005, essa fornisce indicazioni relative alla definizione delle dimensioni aziendali ai fini della concessione di aiuti alle attività produttive e si applica alle imprese operanti in tutti i settori produttivi.

Infatti, al fine di creare un campione che possa rappresentare nella maniera più rappresentativa possibile la realtà aziendale italiana, la selezione delle imprese è stata attuata tramite una procedura a più passi sulla base del contributo di ogni singola regione alla formazione del PIL (Prodotto Interno Lordo) italiano.

Tabella 5.1: Pil Regionale rispetto al Pil Totale (anno 2009, dati in milioni di euro)

Regione o Macroregione	Pil ai prezzi di mercato	% rispetto al Pil Totale
Piemonte	€ 120,696.00	7.92%
Valle d'Aosta	€ 4,347.60	0.29%
Liguria	€ 43,464.80	2.85%
Lombardia	€ 317,026.90	20.79%
Nord-Ovest	€ 485,535.30	31.85%
Trentino-Alto Adige	€ 33,427.20	2.19%
Veneto	€ 141,760.60	9.30%
Friuli-Venezia Giulia	€ 34,511.20	2.26%
Emilia-Romagna	€ 135,282.70	8.87%
Nord-Est	€ 344,981.70	22.63%
Marche	€ 40,297.80	2.64%
Toscana	€ 103,712.60	6.80%
Umbria	€ 21,204.30	1.39%
Lazio	€ 165,414.60	10.85%
Centro	€ 330,629.30	21.69%
Abruzzo	€ 28,395.60	1.86%
Molise	€ 6,531.00	0.43%
Campania	€ 97,093.70	6.37%
Puglia	€ 69,958.60	4.59%
Basilicata	€ 10,648.90	0.70%
Calabria	€ 33,216.10	2.18%
Sicilia	€ 84,852.60	5.57%
Sardegna	€ 32,781.10	2.15%
Sud	€ 363,477.60	23.84%
Totale Italia	€ 1,524,623.90	100.00%

Fonte: Istat

Come si può vedere dalla Tabella 5.1 le regioni italiane contribuiscono in maniera differente alla formazione del PIL, proprio per questo il primo passo è stato quello di pesare il numero di imprese del campione d'analisi facendo riferimento alla quota percentuale del PIL di ogni singola regione rispetto al PIL totale.

Quindi, per esempio, all'interno del campione il numero d'imprese provenienti dalla regione Piemonte è pari a otto in quanto la percentuale del PIL di tale regione rispetto al PIL italiano è

del 7,92%¹⁰⁹ (Tabella 5.1). Inoltre, è possibile notare come regioni quali la Valle d'Aosta e il Molise non vengano incluse nel modello in quanto, pur contribuendo alla formazione del PIL italiano, presentano una percentuale d'influenza particolarmente ridotta (rispettivamente 0.29% e 0.43%).

Tabella 5.2: Totale imprese del campione suddivise per macro – attività

Regione	% Agricoltura su Pil regionale	Numero imprese settore agricoltura	% Industria su Pil regionale	Numero imprese settore industria	% Servizi su Pil regionale	Numero imprese settore servizi	Numero imprese per regione
Piemonte	1.51%	0	29.52%	2	68.96%	6	8
Valle d'Aosta	1.33%	0	25.54%	0	73.13%	0	0
Liguria	1.67%	0	17.20%	1	81.13%	2	3
Lombardia	1.08%	0	33.07%	7	65.85%	14	21
Trentino-Alto Adige	3.00%	0	28.30%	1	68.70%	1	2
Veneto	1.84%	0	34.50%	3	63.66%	6	9
Friuli-Venezia Giulia	1.59%	0	26.02%	1	72.40%	1	2
Emilia-Romagna	2.31%	0	32.91%	3	64.78%	6	9
Marche	1.75%	0	32.09%	1	66.16%	2	3
Toscana	1.91%	0	26.47%	2	71.62%	5	7
Umbria	2.20%	0	27.71%	0	70.09%	1	1
Lazio	1.18%	0	14.58%	2	84.25%	9	11
Abruzzo	2.81%	0	30.62%	1	66.56%	1	2
Molise	3.72%	0	26.03%	0	70.25%	0	0
Campania	2.76%	0	18.72%	1	78.52%	5	6
Puglia	3.93%	0	23.27%	1	72.80%	4	5
Basilicata	4.93%	0	25.55%	0	69.52%	1	1
Calabria	4.67%	0	16.90%	0	78.43%	2	2
Sicilia	4.05%	0	17.06%	1	78.89%	5	6
Sardegna	3.55%	0	19.91%	0	76.54%	2	2
Totale Campione		0		27		73	100

¹⁰⁹ Il numero delle imprese facenti parte del campione è stato sottoposto ad approssimazione.

Dopo aver precisato l'apporto di ogni regione, alla formazione del campione, il secondo passo consiste nell'identificare come la suddivisione del PIL regionale tra le principali macro – attività economiche influisca nella selezione delle aziende (Tabella 5.2).

Analizzando la Tabella 5.2 si può notare quale sia l'apporto in termini percentuali dei settori di macro – attività alla formazione del PIL regionale e, rapportando tali dati alla percentuale ottenuta nella tabella 5.1 in riferimento al contributo di ogni regione alla costituzione del PIL italiano, si ottiene il numero di imprese selezionate operanti in ogni area di macro – attività.

Per esempio, al fine di chiarire il percorso seguito, le otto imprese rappresentative della regione Piemonte sono state selezionate in modo che due di esse appartengano al settore industriale e le restanti sei appartengano al settore dei servizi¹¹⁰.

A tali risultati si è giunti evidenziando l'apporto dei diversi settori alla percentuale di contribuzione di ogni regione per la formazione del PIL italiano, quindi, continuando l'esempio, il numero di imprese appartenenti al settore industriale per la regione Piemonte è stato individuato nella maniera seguente:

$$(8/100) \cdot 29,52 = 2,3616 \approx 2$$

L'applicazione di tale procedura rispecchia l'obiettivo di creare un campione in grado di riassumere la realtà delle PMI italiane nella maniera più rappresentativa possibile, in modo che per ogni impresa non venga considerata solo la regione d'appartenenza, ma anche il settore in cui essa opera.

Se il campione d'analisi così formulato fosse più ampio, allora si potrebbero confrontare tra loro dal punto di vista della performance le diverse regioni e allo stesso tempo, qualora il fenomeno d'insolvenza si verificasse con una certa regolarità o con caratteristiche simili in
Fonte: Elaborazione dati Confcommercio. zare un'analisi sulle possibili cause o sui possibili fenomeni sistemici colpevoli di tale manifestazione. Però, a causa del ristretto numero di imprese facenti parte del campione (122 imprese) ciò non risulta possibile.

¹¹⁰ I settori di macro – attività economica sono stati costruiti in modo da comprendere una pluralità di aree di attività:

- Il settore Agricolo è composto dalle attività relative all'Agricoltura, alla Silvicoltura e alla Pesca;
- Il settore Industriale è composto dalle attività relative all'Industria in senso stretto e alle Costruzioni;
- Il settore Servizi comprende le attività relative al Commercio, alle Riparazioni, Alberghi e Ristoranti, Trasporti e Comunicazioni, Intermediazione Monetaria e Finanziaria, Attività Immobiliari ed Imprenditoriali e altre attività di servizi.

Inoltre, nel momento in cui sono stati definiti i vincoli per la ricerca tramite il database AIDA, oltre alla definizione dei parametri validi per l'identificazione delle PMI visti in precedenza, ci si è concentrati sulla selezione di imprese non quotate in Borsa. Tale scelta è stata attuata, come già annunciato in precedenza, con l'obiettivo di riflettere al meglio la clientela del settore bancario regionale che, nella maggior parte dei casi, è composta da piccole e medie imprese generalmente non quotate in Borsa.

Successivamente, allo scopo di illustrare l'applicabilità del modello alla previsione delle insolvenze, al campione iniziale di 100 imprese è stato aggiunto un campione addizionale formato da 22 imprese fallite nel periodo 2007-2010, estratte dal database AIDA rispettando gli stessi parametri validi per il campione di imprese sane.

Una volta giunti alla formazione del campione finale, composto da 122 imprese, le variabili del modello sono state calcolate in modo da garantire il più alto grado di omogeneità possibile. Infatti per quanto riguarda le imprese non fallite, gli indicatori finanziari sono stati calcolati in riferimento al bilancio dell'anno 2010, mentre per le imprese fallite le variabili finanziarie sono state calcolate con riferimento all'ultimo bilancio disponibile precedente al fallimento.

Considerando che la maggior parte dei bilanci delle imprese fallite non appartiene all'anno 2010, al fine di garantire la massima omogeneità del campione si è scelto di creare un "matching" tra i due gruppi di imprese. In tal modo, ad ogni impresa fallita per cui l'ultimo bilancio disponibile risale ad un anno diverso dal 2010, corrisponde un'impresa sana il cui bilancio risale allo stesso anno.

Tale costruzione, ottenuta cercando di appianare le divergenze temporali tra i due campioni di imprese, permette di evitare che si creino distorsioni nei risultati in grado di falsare le capacità diagnostiche del modello.

5.2 Frontiera efficiente e Frontiera dei fallimenti

L'obiettivo principale dell'analisi riguarda la previsione delle insolvenze; quest'aspetto, in relazione anche alla selezione delle variabili di input e di output, è in parte differente dai convenzionali utilizzi della Data Envelopment Analysis.

Avvalendosi del modello additivo introdotto da Charnes nel 1985, come già delineato nel Capitolo precedente, e utilizzando le variabili di input e di output, l'analisi condotta ha identificato due frontiere: una frontiera dei fallimenti e una frontiera efficiente.

In particolare, allo scopo di individuare la frontiera dei fallimenti sono state considerate come variabili di input queglii indici finanziari i cui valori più bassi possono causare difficoltà finanziarie, mentre come variabili di output sono stati selezionati queglii indicatori per cui valori più elevati possono rappresentare sintomi di fallimento.

Tale classificazione identifica la frontiera dei fallimenti, ovvero quella barriera che individua le imprese in procinto di fallire e tendenti a riportare uno score di efficienza prossimo o pari ad 1.

Per giungere a definire la frontiera dei fallimenti e la corrispondente capacità del modello di predire l'appartenenza o meno delle imprese a tale barriera, è stata seguita una procedura in più passi¹¹¹:

- applicare il modello additivo al campione d'analisi nel suo insieme;
- classificare l'impresa basandosi sul fatto che abbia tutti gli slack nulli ed uno score pari ad 1. Se tutti gli slack sono nulli allora l'impresa è sulla frontiera dei fallimenti, in caso contrario l'impresa non si trova su tale frontiera;
- classificare allo stesso modo tutte le imprese del campione;
- classificare tutte le imprese nei seguenti gruppi: (a) imprese fallite appartenenti alla frontiera dei fallimenti, (b) imprese fallite non appartenenti alla frontiera dei fallimenti, (c) imprese non fallite appartenenti alla frontiera dei fallimenti e (d) imprese non fallite non appartenenti alla frontiera dei fallimenti;
- calcolare il numero di imprese appartenenti ad ogni singolo gruppo e identificare le rispettive probabilità:
 - (a) $P(BR/BR)$ = percentuale di imprese fallite predette come fallite
 - (b) $P(NBR/BR)$ = percentuale di imprese fallite predette come non fallite
 - (c) $P(BR/NBR)$ = percentuale di imprese non fallite predette come fallite
 - (d) $P(NBR/NBR)$ = percentuale di imprese non fallite predette come non fallite
 - (e) $P((BR/BR+NBR/NBR)/(BR+NBR))$ = percentuale di predizione corretta = il numero totale di previsioni corrette di imprese fallite e non fallite.

Ai fini di valutare la capacità del modello nel prevedere le insolvenze, l'attenzione è stata rivolta alle percentuali relative alle previsioni corrette, ovvero (a), (d), (e).

¹¹¹ I. M. Premachandra, Gurmeet Singh Bhabra, Toshiyuki Sueyoshi (2009), *DEA as a tool for bankruptcy assessment: A comparative study with logistic regression technique*, European Journal of Operational Research, 193, pp. 412-424.

L'individuazione delle precedenti percentuali riflette l'obiettivo per cui è stata attuata quest'analisi, ovvero quello di fornire uno strumento facile e veloce agli analisti finanziari per prevedere in maniera sufficientemente accurata il fenomeno delle insolvenze.

Utilizzando lo stesso processo, alterato solamente dall'inversione delle variabili di input e di output, è stata individuata la barriera efficiente. In questo caso l'impresa con tutti gli slack nulli e uno score pari ad 1 è stata classificata come efficiente e non come fallita.

Una volta definite le due frontiere, seguendo l'impronta di Premachandra¹¹², l'analisi è stata indirizzata alla combinazione delle due frontiere, allo scopo di giungere ad una migliore classificazione delle imprese.

Lo score di riferimento, in base al quale attuare la classificazione per questo modello DEA a due frontiere, è ottenuto tramite una combinazione degli score impiegati per l'identificazione delle due frontiere:

$$\lambda\theta_1 - (1 - \lambda)\theta_2$$

dove θ_1 è lo score associato al modello che si avvale della frontiera dei fallimenti, θ_2 è lo score associato al modello che si avvale della frontiera efficiente e λ è un peso che rappresenta la vicinanza o meno alle due frontiere (per $\lambda = 0$ si ottiene la frontiera efficiente, mentre per $\lambda = 1$ si ottiene la frontiera dei fallimenti).

Nei prossimi paragrafi sono presentati e discussi gli indici finanziari, fondamentali per calcolare gli score θ_1 e θ_2 e per l'individuazione delle rispettive frontiere, che verranno impiegati come variabili di input e di output nei diversi modelli esplorati.

5.3 La selezione delle variabili

Come già introdotto nel capitolo precedente, il modello DEA a cui si fa riferimento nell'analisi è il Modello Additivo in quanto, rispetto agli altri modelli DEA, riesce ad adempiere meglio alle esigenze richieste.

Punto focale dell'analisi riguarda la scelta delle variabili di input e di output da inserire nel modello, infatti una caratteristica della metodologia DEA è quella di lasciare ampia libertà all'analista nella scelta delle variabili da utilizzare.

¹¹² I. M. Premachandra, Yao Chen, John Watson (2011), *DEA as a tool for predicting corporate failure and success: A case of bankruptcy assessment*, Omega 39, pp.620-626.

Nell'analisi condotta sono stati impiegati degli indicatori che descrivono la forza o la debolezza finanziaria delle imprese esaminate, contribuendo così ad individuare la potenziale presenza o meno di una situazione d'insolvenza per le varie aziende.

In virtù del fatto che il metodo DEA non definisce un chiaro approccio nel determinare gli input e gli output del processo, le variabili d'interesse sono state selezionate attraverso una linea guida ricavata dagli studi passati.

In particolare, la scelta degli indicatori finanziari è avvenuta prendendo in considerazione quanto la letteratura era stata in grado di dimostrare in merito alla previsione dell'insolvenza; proprio per questo si è fatto riferimento in parte alle variabili utilizzate nel lavoro di Altman¹¹³ e in parte a quelle utilizzate nei più recenti studi di Cielen¹¹⁴ e di Premachandra¹¹⁵. Infatti, convinzione largamente condivisa dalla letteratura in merito, è costituita dal fatto che tali indici siano in grado di predire con maggiore efficienza rispetto a qualsiasi altro indice la probabilità d'insolvenza per le imprese esaminate.

Tabella 5.3: Variabili di Input e di Output utilizzate per il calcolo della frontiera dei fallimenti

Variabili di Input	Variabili di Output
Cash Flow/Total Assets	Total Debt/Total Assets
Net Income/Total Assets	Current Liabilities/Total Assets
Working Capital/Total Assets	
Current Assets/Total Assets	
Earnings Before Interest and Taxes/Total Assets	
Earnings Before Interest and Taxes/Interest Expense	

La Tabella 5.3 presenta le variabili di input e di output impiegate nell'analisi per identificare lo score θ_1 e la corrispondente frontiera dei fallimenti. In modo particolare, gli indici

¹¹³ Edward I. Altman, *Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy*, Journal of Finance, Vol. 23, No. 4, (September 1968), pp. 589-609.

¹¹⁴ A. Cielen, L. Peeters, K. Vanhoof (2004), *Bankruptcy prediction using a data envelopment analysis*, European Journal of Operational Research, 154, pp. 526-532.

¹¹⁵ I. M. Premachandra, Gurmeet Singh Bhabra, Toshiyuki Sueyoshi (2009) e I. M. Premachandra, Yao Chen, John Watson (2011), opere già citate.

finanziari i cui valori più bassi possono rappresentare sintomi di fallimento aziendale sono considerati variabili di input, mentre gli indici i cui valori più alti possono costituire causa di dissesto finanziario sono ritenuti variabili di output per il modello.

Al contrario, per ottenere lo score θ_2 ed identificare così la frontiera efficiente, le variabili di input sopra elencate vengono considerate variabili di output e, viceversa, le variabili di output vengono ritenute variabili di input.

In seguito si presenterà una descrizione dettagliata degli indicatori allo scopo di discutere la loro capacità nel valutare la presenza o meno di dissesti finanziari.

5.4 Le variabili di Input del modello per il calcolo della frontiera dei fallimenti

È possibile osservare che i seguenti indicatori finanziari, scelti come variabili di input, costituiscono per lo più misure di liquidità aziendale come l'utile netto, il capitale circolante, le attività correnti e altri indici, quindi risulta giustificata l'assunzione in base alla quale aziende che presentano valori piccoli per tali misure di liquidità siano più propense a versare in uno stato di default.

Le prossime righe presentano in dettaglio le variabili in questione¹¹⁶:

- Cash Flow¹¹⁷/Total Assets, rappresenta un indicatore capace di esprimere la liquidità dell'impresa concentrandosi sui flussi di cassa generati dalla gestione tipica, quindi costituisce un parametro estremamente importante al fine di giudicare l'operato della gestione aziendale;
- Net Income¹¹⁸/Total Assets, questo indice costituisce uno strumento in grado di valutare quanto efficacemente la gestione finanziaria impiega le risorse disponibili nell'attività dell'impresa, quindi indica la capacità delle attività in essere di generare profitto. Nel caso in cui l'indice presenti un valore ridotto, sintomo che i guadagni dell'impresa sono particolarmente limitati, confrontando tale valore con la media di settore è possibile evidenziare l'inefficienza da parte dell'impresa nell'utilizzo delle risorse aziendali. Infatti le imprese maggiormente esposte ad un rischio d'insolvenza

¹¹⁶ I. M. Premachandra, Gurmeet Singh Bhabra, Toshiyuki Sueyoshi (2009), *DEA as a tool for bankruptcy assessment: A comparative study with logistic regression technique*, European Journal of Operational Research, 193, pp. 412-424 e I. M. Premachandra, Yao Chen, John Watson (2011), *DEA as a tool for predicting corporate failure and success: A case of bankruptcy assessment*, Omega 39, pp.620-626.

¹¹⁷ Il Cash Flow è stato calcolato sommando all'utile o perdita d'esercizio le voci del conto economico relative agli ammortamenti e agli altri accantonamenti (fondi svalutazione,...).

¹¹⁸ Equivale all'utile netto o alla perdita d'esercizio.

sono quelle imprese per cui l'indice in questione assume dei valori particolarmente ridotti;

- Working Capital/Total Assets, questo rapporto rappresenta un indicatore di liquidità dell'impresa; il capitale circolante è calcolato come differenza tra le attività correnti e le passività correnti e, in quanto tale, fornisce informazioni relativamente alla capacità dell'impresa di affrontare difficoltà finanziarie che potrebbero sorgere in un futuro prossimo;
- Current Assets¹¹⁹/Total Assets, questo indice esprime la quota di risorse finanziarie attualmente investite a breve termine, per le quali ci si aspetta un ritorno in forma liquida entro dodici mesi, sul totale delle attività. Ciò permette, quindi, di valutare se l'impresa sia dotata di risorse finanziarie sufficienti per far fronte ai finanziamenti in scadenza, infatti un valore ridotto dell'indice può sottolineare difficoltà per l'impresa nell'adempiere alle proprie obbligazioni a breve termine e, di conseguenza, a una maggiore probabilità d'insolvenza per la stessa;
- EBIT/Total Assets, questa variabile costituisce una misura di produttività aziendale, infatti descrive la capacità dell'impresa di generare profitto indipendentemente dalle dimensioni della società stessa. In generale, più il valore di quest'indice è alto, maggiori sono i profitti potenziali dell'impresa e, di conseguenza, minore è la probabilità che l'impresa incorra in uno stato d'insolvenza;
- EBIT/Interest Expense, questo indice, conosciuto anche come "interest coverage ratio" (rapporto di copertura degli interessi), determina la facilità con cui un'impresa è in grado di rimborsare gli interessi sul debito in essere. Generalmente, più il valore di tale rapporto è basso e più l'impresa risulta gravata dalle spese del debito. Mano a mano che il valore dell'indicatore si riduce, la capacità dell'impresa di coprire gli interessi diventa discutibile e, nel caso in cui tale valore risulti inferiore a 1, è possibile sostenere che l'impresa non è in grado di generare profitti sufficienti a coprire gli interessi sul debito. Quest'indicatore assume una certa importanza in quanto consente di definire in maniera chiara la salute finanziaria di un'impresa con riferimento ad un orizzonte di breve termine.

¹¹⁹ Le Attività Correnti sono state calcolate considerando le voci inerenti all'Attivo Circolante dello Stato Patrimoniale e sommando ad esse i risconti attivi.

5.5 Le variabili di Output per la frontiera dei fallimenti

Contrariamente al paragrafo precedente, le variabili di output del modello sono state selezionate in modo che presentino dei valori significativamente elevati per le imprese maggiormente esposte al fenomeno dell'insolvenza.

Osservando la tabella 5.3 si può notare che gli indicatori selezionati ai fini della classificazione sono i seguenti¹²⁰:

- Total Debt/Total Assets, quest'indicatore finanziario, noto come "Debt Ratio", costituisce una misura di indebitamento ed esprime il rischio finanziario dell'impresa determinando il peso del debito totale (a breve e a lungo termine) sull'attivo totale, ovvero la percentuale dei beni dell'impresa finanziati dal debito. Un incremento nell'indebitamento da parte dell'impresa aumenta la probabilità di uno stress finanziario, infatti, qualora ciò si verificasse, potrebbero esserci momenti futuri in cui i fondi liquidi non sono in grado di rimborsare il debito in essere, conducendo così l'impresa al fallimento. Proprio per questo motivo valori elevati dell'indice sottolineano la presenza di deficit finanziari nelle imprese e palesano la possibilità che quest'ultime siano maggiormente esposte al fenomeno dell'insolvenza;
- Current Liabilities¹²¹/Total Assets, questa variabile esprime la capacità dell'impresa di soddisfare le proprie obbligazioni nel breve termine. Un valore elevato dell'indice evidenzia una mancanza di liquidità per finanziare le attività dell'impresa; infatti, in tale situazione, un'impresa avrebbe delle serie difficoltà nel gestire quotidianamente le proprie operazioni a causa di una ridotta flessibilità nell'ottenere il capitale circolante necessario all'attività. Essenzialmente è questa la ragione per cui imprese caratterizzate da una maggiore probabilità d'insolvenza presentano un valore dell'indice più elevato rispetto ad imprese ritenute finanziariamente sane.

Semplicemente invertendo le variabili di input con le variabili di output si ottengono gli indicatori necessari all'identificazione della frontiera efficiente e alla definizione delle rispettive percentuali di previsione corretta.

¹²⁰ I. M. Premachandra, Gurmeet Singh Bhabra, Toshiyuki Sueyoshi (2009), *DEA as a tool for bankruptcy assessment: A comparative study with logistic regression technique*, European Journal of Operational Research, 193, pp. 412-424.

¹²¹ Le passività Correnti sono state calcolate sommando ai debiti a breve termine la voce relativa ai ratei passivi.

La capacità di predire in maniera accurata la probabilità di fallimento di un'impresa è molto rilevante in quanto può fornire informazioni importanti a tutti quei soggetti (per esempio azionisti,creditori, impiegati) strettamente legati all'operatività dell'impresa stessa.

La scelta delle variabili di input e di output, ereditata dai precedenti studi di Premachandra, riflette l'idea che tali variabili siano le più adeguate per offrire una panoramica totale della situazione finanziaria dell'impresa, in modo che il modello non tralasci nessun aspetto operativo significativo ai fini della diagnosi.

5.6 Modelli e Risultati dell'analisi

In questo e nei successivi paragrafi vengono presentati i risultati dell'analisi condotta sul campione di PMI secondo il modello Additivo DEA presentato nel Capitolo precedente.

A questo punto è necessario sottolineare come l'indagine non si sia limitata ad analizzare il modello caratterizzato da sei variabili di input e due variabili di output descritto in precedenza, ma al fine di valutare la possibilità di implementarne la capacità predittiva sono state apportate delle estensioni a tale modello.

Una prima estensione è stata quella di aggiungere, tra le variabili di input, un indicatore finanziario aggiuntivo ereditato dal modello Z'-Score di Altman per le imprese non quotate; il nuovo indicatore è stato ottenuto rapportando il valore contabile del Patrimonio Netto ai debiti totali dell'impresa.

La seconda estensione è stata caratterizzata dalla costruzione di un nuovo indicatore definito in questo modo: Valore medio del Patrimonio Netto degli ultimi tre bilanci/Valore del Patrimonio Netto dell'ultimo bilancio.¹²²

Così facendo, la valutazione della capacità predittiva della metodologia DEA è stata effettuata avvalendosi di tre modelli leggermente differenti l'uno dall'altro, in modo da individuare quello che, sulla base della maggiore correttezza nelle percentuali di previsione, è in grado di rispondere più efficacemente alle esigenze degli analisti del credito.

I paragrafi successivi presentano i risultati dello studio per ognuno dei tre modelli. I relativi score delle imprese sono stati calcolati utilizzando il software MaxDEA Basic sviluppato da Gang Cheng (Peking University) e Zhenhua Qian (University of Science & Technology Beijing).

¹²² L'ultimo bilancio equivale all'anno 2010 per le imprese sane e all'ultimo disponibile prima del fallimento per quelle insolventi.

L'analisi, tra tutti i modelli disponibili all'interno del software, è stata effettuata tramite l'utilizzo del modello non radiale Slacks – Based Measures (SBM), introdotto da Tone nel 2001¹²³.

Inoltre, basandosi sui lavori di Premachandra, tutti i risultati sono stati ottenuti ipotizzando la presenza di rendimenti di scala variabili (VRS), i quali impongono un vincolo di convessità aggiuntivo ($\sum_{j=1}^n \lambda_j = 1$) al problema di programmazione lineare rispetto al caso di rendimenti di scala costanti.

Infine, per quanto riguarda l'orientazione, lo studio è stato condotto separatamente a seconda che il modello fosse input-oriented, output-oriented e non-oriented, confrontando poi i risultati allo scopo di evidenziare le differenze più significative, identificando così il modello più efficiente.

Alla fine del capitolo è riportata un'appendice in cui sono presentate le tabelle contenenti i risultati dettagliati dell'analisi che sono stati omessi nei prossimi paragrafi.

5.6.1 Modello a otto variabili

Sulla base di quanto anticipato, il primo modello di cui si presentano i risultati è il modello ad otto variabili, ovvero il modello presentato da Premachandra nei suoi lavori, caratterizzato però dall'esclusione della variabile relativa al valore di mercato del Patrimonio Netto in quanto, come già detto in precedenza, l'analisi è condotta su di un campione formato da imprese non quotate. Una volta ottenuti gli score θ_1 e θ_2 tramite il software MaxDEA Basic, in base alla formula vista in precedenza, è stato calcolato lo score di riferimento derivante dalla combinazione lineare delle due frontiere, quella dei fallimenti e quella efficiente.

Lo stadio successivo è stato quello di classificare le imprese come sane o insolventi sulla base di alcuni punteggi soglia (cut-off) e di calcolare le seguenti percentuali di previsioni corrette:

- $P(\text{BR}/\text{BR}) = \% \text{BR}$ = percentuale di imprese fallite predette come fallite;
- $P(\text{NBR}/\text{NBR}) = \% \text{NBR}$ = percentuale di imprese non fallite predette come non fallite;
- $P((\text{BR}/\text{BR} + \text{NBR}/\text{NBR})/(\text{BR} + \text{NBR})) = \text{Percentuale di previsione corretta} = \% \text{Totale} =$
numero totale di previsioni corrette di imprese fallite e non fallite.

¹²³ K. Tone (2001), *A slacks – based measure of efficiency in Data Envelopment Analysis*, European Journal of Operational Research 130, pp. 498-509.

In particolare, allo scopo di ottenere tali percentuali, le imprese il cui valore di score risultava maggiore o tutt'al più uguale al punteggio soglia stabilito sono state considerate come fallite, mentre le imprese il cui valore di score rimaneva al di sotto di tale soglia sono state ritenute sane.

Le tabelle seguenti riportano le percentuali di previsioni corrette ottenute a seconda dell'orientamento prescelto per il modello: input-oriented, output-oriented e non-oriented. Inoltre per ogni valore di λ assegnato, ricordando che per $\lambda = 0$ si ottiene la frontiera efficiente e per $\lambda = 1$ la frontiera dei fallimenti, è citato il punteggio soglia in relazione al quale si hanno le percentuali migliori. I cut-off sono stati selezionati in modo da mantenere una percentuale di previsione delle imprese non fallite accettabile in quanto la numerosità del campione è fortemente sbilanciata verso quest'ultime (100 imprese non fallite e 22 imprese fallite).

La tabella 5.4 presenta i risultati per il caso input-oriented. È possibile notare come all'aumentare del valore di λ si verifichi un calo nella percentuale di imprese fallite predette effettivamente come fallite, mentre la percentuale di imprese non fallite predette effettivamente come non fallite rimane per la maggior parte dei casi molto elevata tranne nel caso $\lambda = 0.9$ in cui assume un valore del 65%.

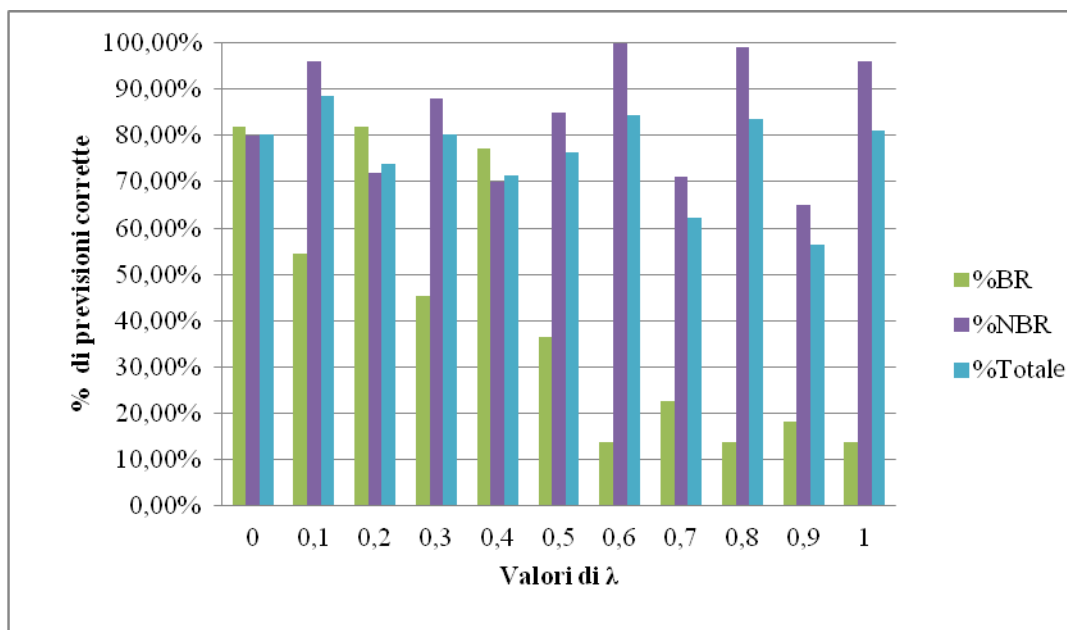
Ciò sottolinea come, ipotizzando un'ottica input-oriented, la scelta di λ influisca notevolmente sulle percentuali di previsione; infatti, nel caso in cui un analista fosse maggiormente interessato all'individuazione dei fallimenti allora dovrebbe optare per un

Tabella 5.4: Percentuali di previsioni corrette nel modello input-oriented a otto variabili

λ	Cut-off migliore	%BR	%NBR	%Totale
0	-0,2	81,82%	80,00%	80,33%
0,1	-0,1	54,55%	96,00%	88,52%
0,2	-0,1	81,82%	72,00%	73,77%
0,3	0	45,45%	88,00%	80,33%
0,4	0	77,27%	70,00%	71,31%
0,5	0,1	36,36%	85,00%	76,23%
0,6	0,5	13,64%	100,00%	84,43%
0,7	0,2	22,73%	71,00%	62,30%
0,8	0,7	13,64%	99,00%	83,61%
0,9	0,3	18,18%	65,00%	56,56%
1	0,3	13,64%	96,00%	81,15%

valore di λ più vicino a zero in quanto la crescita dello stesso parametro comporta un crollo drastico nella percentuale BR. Al contrario, qualora un analista fosse interessato ad identificare con una maggiore precisione le imprese sane, allora dovrebbe rivolgere la propria attenzione verso valori di λ più elevati, in corrispondenza dei quali vi è un aumento della percentuale NBR a scapito della percentuale BR.

Figura 5.1: Percentuali di previsioni corrette nel modello input-oriented a otto variabili



Il modello fornisce risultati più soddisfacenti nel caso in cui si adotti un'ottica output-oriented; infatti come si può osservare dalla Tabella 5.5 le percentuali di previsione corrette superano notevolmente le previsioni ottenute precedentemente.

Comparate con i risultati ottenuti con l'ottica input-oriented, si nota che all'aumentare di λ la percentuale di imprese fallite predette effettivamente come fallite scende da un apice del 77,27% ad un valore minimo del 50%, mentre la percentuale di imprese non fallite predette come non fallite rimane per lo più invariata.

Ciò si evince anche dalla Figura 5.2, in cui la percentuale BR subisce un leggero calo all'inizio per poi stabilizzarsi ad un valore costante pari al 50% senza influenzare in maniera significativa le restanti percentuali.

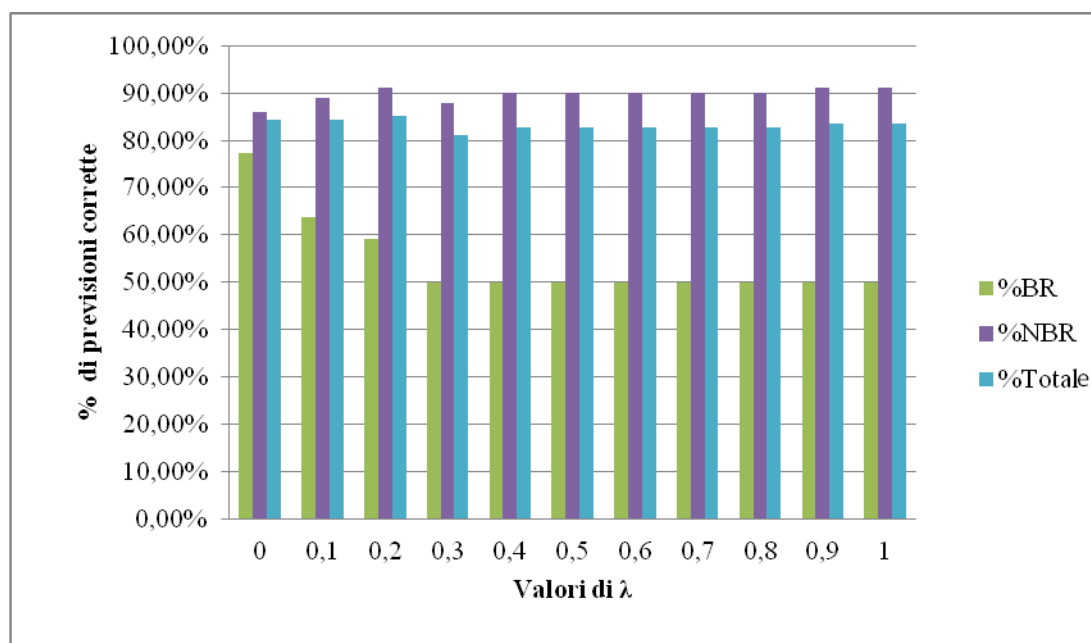
Osservando con attenzione la Tabella 5.5 è possibile notare che, per valori di λ che vanno da 0,3 a 0,8, le tre percentuali in esame rimangono immutate; l'unica cosa che si modifica è il cut-off ottimale che, ad ogni stadio, trasla costantemente di un valore pari all'incremento del valore di λ (ovvero di 0,1). Il risultato previsionale più soddisfacente, nell'ottica output-

oriented, è raggiunto per un valore di λ pari a 0,2, e infatti le percentuali corrispondenti al cut-off ottimale di -0,3 sono rispettivamente pari a 59,09%, 91% e 85,25%.

Tabella 5.5: Percentuali di previsioni corrette nel modello output-oriented a otto variabili

λ	Cut-off migliore	%BR	%NBR	%Totale
0	-0,5	77,27%	86,00%	84,43%
0,1	-0,4	63,64%	89,00%	84,43%
0,2	-0,3	59,09%	91,00%	85,25%
0,3	-0,2	50,00%	88,00%	81,15%
0,4	-0,1	50,00%	90,00%	82,79%
0,5	0	50,00%	90,00%	82,79%
0,6	0,1	50,00%	90,00%	82,79%
0,7	0,2	50,00%	90,00%	82,79%
0,8	0,3	50,00%	90,00%	82,79%
0,9	0,4	50,00%	91,00%	83,61%
1	0,5	50,00%	91,00%	83,61%

Figura 5.2: Percentuali di previsioni corrette nel modello output-oriented a otto variabili



Confrontando i risultati della Tabella 5.5 con quelli della Tabella 5.4 è possibile evidenziare come, in media, il modello output – oriented presenti una capacità diagnostica totale superiore

al modello input-oriented in quanto, rispetto a quest'ultimo, è in grado di prevedere con maggior precisione i casi di fallimento.

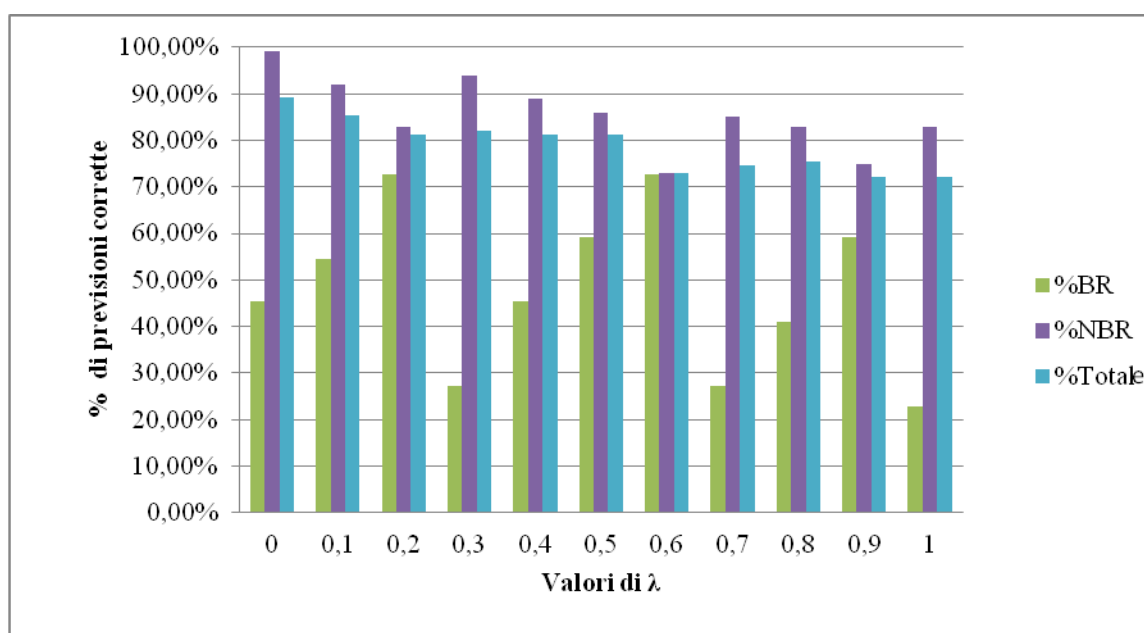
L'ultimo caso, ovvero l'applicazione del modello Additivo non oriented, differisce leggermente dai primi due in quanto al crescere del valore di λ presenta un andamento piuttosto altalenante (Figura 5.3).

Infatti, mentre le ultime due colonne della Tabella 5.6 presentano delle percentuali per lo più in linea con i casi precedenti, l'andamento altalenante è riconoscibile nella terza colonna, ovvero quella relativa alla previsione delle insolvenze.

Tabella 5.6: Percentuali di previsioni corrette nel modello non oriented a otto variabili

λ	Cut-off migliore	%BR	%NBR	%Totale
0	-0,1	45,45%	99,00%	89,34%
0,1	-0,1	54,55%	92,00%	85,25%
0,2	-0,1	72,73%	83,00%	81,15%
0,3	0	27,27%	94,00%	81,97%
0,4	0	45,45%	89,00%	81,15%
0,5	0	59,09%	86,00%	81,15%
0,6	0	72,73%	73,00%	72,95%
0,7	0,1	27,27%	85,00%	74,59%
0,8	0,1	40,91%	83,00%	75,41%
0,9	0,1	59,09%	75,00%	72,13%
1	0,2	22,73%	83,00%	72,13%

Figura 5.3: Percentuali di previsioni corrette nel modello non oriented a otto variabili



In questo caso, il risultato previsionale migliore lo si ottiene impiegando la barriera efficiente come cut-off ottimale, infatti per un valore di λ pari a 0 le percentuali di previsione sono rispettivamente 45,45%, 99% e 89,34%.

Quindi, tralasciando possibili eccezioni, è possibile affermare che i risultati numerici a cui si è pervenuti suggeriscono come il modello output-oriented, nel caso in cui ci si avvalga degli otto indicatori finanziari precedentemente descritti, risulti il più affidabile in quanto è in grado di garantire elevate percentuali previsionali, e questo indipendentemente dal valore assegnato al parametro λ .

Perciò, scelto il modello output-oriented come il modello di riferimento, si dimostra utile attuare un confronto con i risultati dei precedenti lavori di Premachandra¹²⁴ al fine di evidenziare possibili collegamenti con le prestazioni ottenute.

Partendo dallo studio del 2009, in cui Premachandra attua una classificazione delle imprese sulla base della frontiera dei fallimenti, le percentuali stimate rispettivamente per BR, NBR e Totale si attestano sui seguenti valori: 86%, 67,69% e 68,65%. Comparando queste percentuali con quelle ottenute dal modello output-oriented nel caso di un cut-off pari alla frontiera dei fallimenti ($\lambda = 0$), si nota che la percentuale di previsione delle imprese fallite è scesa al 50% a fronte, invece, di un sostanziale incremento delle restanti percentuali a livelli pari a 91% e 83,61%. Quindi, rapportato al lavoro di Premachandra, il modello realizzato in questo lavoro presenta una perdita di precisione nella previsione delle imprese fallite, ma acquista potere per quanto riguarda la capacità di identificazione delle imprese non fallite.

Passando ora al confronto con il lavoro di Premachandra del 2011, risulta conveniente riassumere nella Tabella 5.7 i diversi risultati.

Osservando la Tabella 5.7 si nota che il modello di Premachandra ottiene delle percentuali di previsione delle insolvenze leggermente superiori a quelle del modello analizzato, ad esclusione del caso in cui λ è pari a 0,1. Invece, per quanto riguarda la corretta identificazione delle imprese sane, il modello ad otto variabili supera le percentuali del modello di Premachandra nella maggior parte dei casi.

Infine, volgendo lo sguardo alla colonna delle percentuale totale di previsioni corrette, si può evidenziare come le percentuali non si discostino molto l'una dall'altra, infatti i range sono rispettivamente pari a 82,79-85,25% e 81,10-91,20%.

¹²⁴ I. M. Premachandra, Gurmeet Singh Bhabra, Toshiyuki Sueyoshi (2009), *DEA as a tool for bankruptcy assessment: A comparative study with logistic regression technique*, European Journal of Operational Research, 193, pp. 412-424 e I. M. Premachandra, Yao Chen, John Watson (2011), *DEA as a tool for predicting corporate failure and success: A case of bankruptcy assessment*, Omega 39, pp.620-626.

Tabella 5.7: Confronto tra i risultati del modello a otto variabili e quelli del modello Premachandra 2011

λ	Modello output-oriented caso otto variabili			Modello Premachandra 2011		
	%BR	%NBR	%Totale	%BR	%NBR	%Totale
0,1	63,64%	89,00%	84,43%	52,00%	93,30%	91,20%
0,2	59,09%	91,00%	85,25%	66,00%	88,20%	87,10%
0,5	50,00%	90,00%	82,79%	74,00%	81,50%	81,10%
0,9	50,00%	91,00%	83,61%	68,00%	86,10%	85,20%
1	50,00%	91,00%	83,61%	72,00%	83,20%	82,60%

Quindi, in virtù di quanto descritto, l'analista finanziario potrà avvalersi del modello output-oriented come strumento per diagnosticare nel miglior modo possibile la probabilità d'insolvenza delle imprese; nei prossimi paragrafi si cercherà di capire se l'introduzione di un indicatore aggiuntivo può in qualche modo migliorare la capacità predittiva del modello DEA.

5.6.2 Modello a nove variabili: prima estensione

Come già annunciato in precedenza, una volta ottenuti i risultati per il modello ad otto variabili, si è cercato di capire se l'aggiunta di un nuovo indicatore finanziario suggerito dal modello Z'-Score di Altman per le imprese non quotate fosse in grado di apportare o meno dei miglioramenti alla capacità previsionale del modello.

Anche in questo caso i risultati dell'analisi sono ripartiti in base all'orientamento designato (input-oriented, output-oriented, non oriented) allo scopo di creare una linea guida che faciliti i possibili confronti tra i diversi modelli.

La Tabella 5.8 presenta il caso input-oriented; confrontando tali risultati con quelli ottenuti dal modello ad otto variabili è possibile evidenziare come le percentuali non presentino differenze evidenti, anzi sovrapponendo i rispettivi grafici è possibile sostenere che vi sia quasi una totale coincidenza.

Infatti, anche in questo caso è possibile sottolineare come all'aumentare del valore di λ la percentuale di previsione dei fallimenti scenda drasticamente da picchi dell'81,82% al 13,64%.

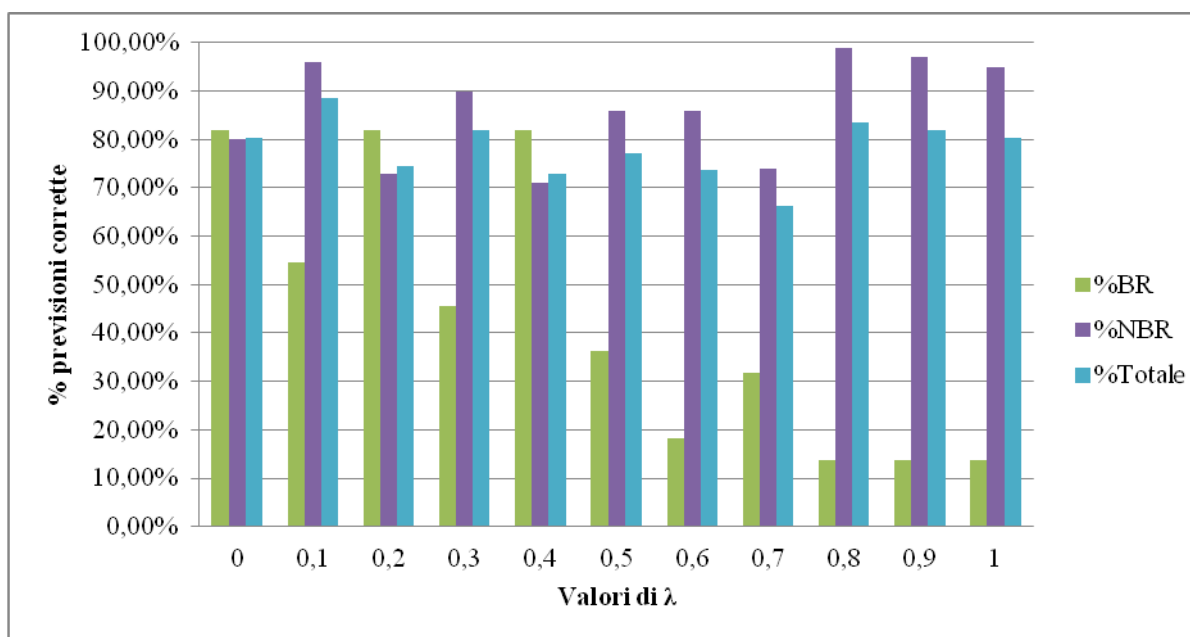
La differenza fondamentale tra i due modelli risiede nella diversità dei cut-off utilizzati, in effetti per $\lambda < 0,6$ essi rimangono inalterati, mentre per $\lambda \geq 0,6$ essi subiscono delle variazioni significative.

Tabella 5.8: Percentuali di previsioni corrette con il modello input-oriented a nove variabili

λ	Cut-off migliore	%BR	%NBR	%Totale
0	-0,2	81,82%	80,00%	80,33%
0,1	-0,1	54,55%	96,00%	88,52%
0,2	-0,1	81,82%	73,00%	74,59%
0,3	0	45,45%	90,00%	81,97%
0,4	0	81,82%	71,00%	72,95%
0,5	0,1	36,36%	86,00%	77,05%
0,6	0,2	18,18%	86,00%	73,77%
0,7	0,2	31,82%	74,00%	66,39%
0,8	0,7	13,64%	99,00%	83,61%
0,9	0,8	13,64%	97,00%	81,97%
1	0,8	13,64%	95,00%	80,33%

Il caso output-oriented, come in precedenza, incrementa le capacità previsionali del modello in quanto riesce ad individuare le imprese fallite con una precisione media attorno al 50% e le imprese non fallite con un'accuratezza che arriva anche al 98%.

Figura 5.4: Percentuali di previsioni corrette con il modello input-oriented a nove variabili

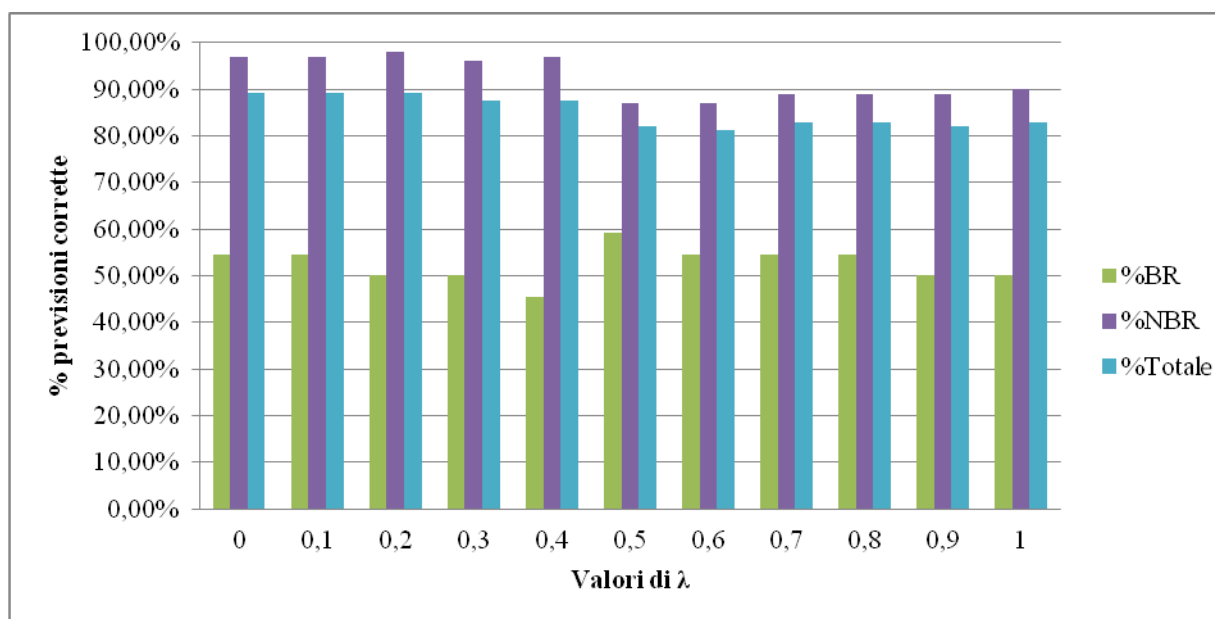


Confrontando tali risultati con quelli ottenuti dal modello ad otto variabili si nota che le differenze sono trascurabili: a scapito di percentuali BR leggermente inferiori per valori di λ compresi tra 0 e 0,3, il modello fornisce una maggiore capacità previsionale per le imprese sane.

Tabella 5.9: Percentuali di previsioni corrette con il modello output-oriented a nove variabili

λ	Cut-off migliore	%BR	%NBR	%Totale
0	-0,4	54,55%	97,00%	89,34%
0,1	-0,3	54,55%	97,00%	89,34%
0,2	-0,2	50,00%	98,00%	89,34%
0,3	-0,1	50,00%	96,00%	87,70%
0,4	0	45,45%	97,00%	87,70%
0,5	0	59,09%	87,00%	81,97%
0,6	0,1	54,55%	87,00%	81,15%
0,7	0,2	54,55%	89,00%	82,79%
0,8	0,3	54,55%	89,00%	82,79%
0,9	0,4	50,00%	89,00%	81,97%
1	0,5	50,00%	90,00%	82,79%

Figura 5.5: Percentuali di previsioni corrette con il modello output-oriented a nove variabili



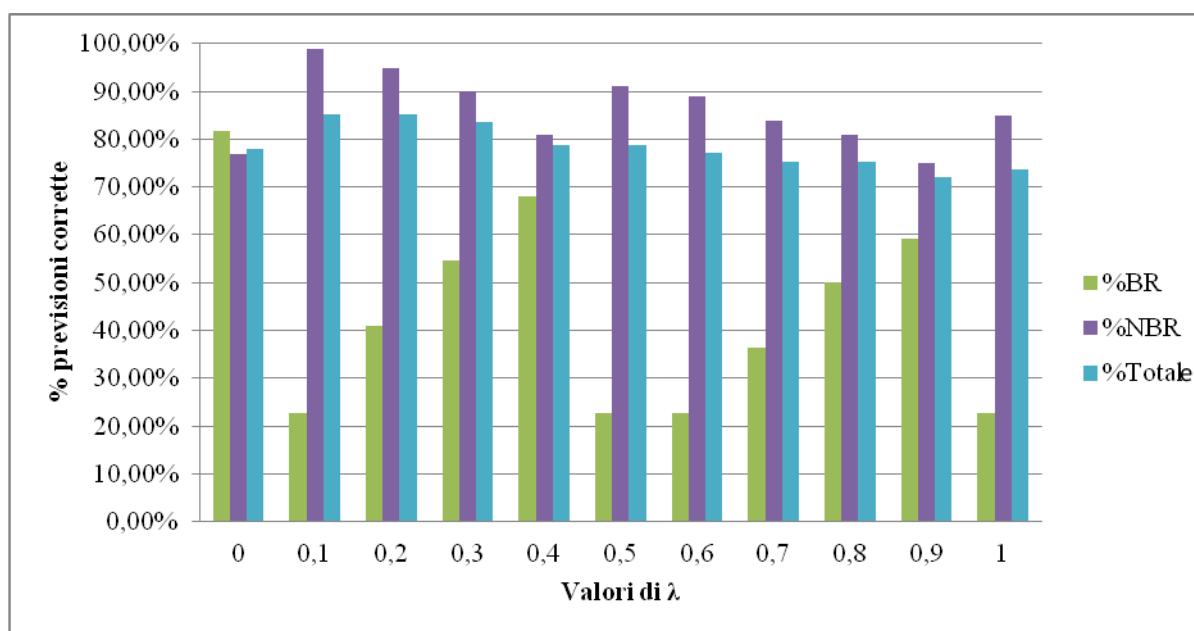
Quindi è possibile sostenere come nemmeno in questa situazione l'aggiunta di una variabile abbia apportato dei miglioramenti significativi al modello.

Infine, la Tabella 5.10 presenta i risultati per il caso non oriented, in cui si riscontra lo stesso andamento oscillatorio visto in precedenza nel modello ad otto variabili, senza particolari discrepanze degne di nota.

Tabella 5.10: Percentuali di previsioni corrette con il modello non oriented a nove variabili

λ	Cut-off migliore	%BR	%NBR	%Totale
0	-0,1	81,82%	77,00%	77,87%
0,1	0	22,73%	99,00%	85,25%
0,2	0	40,91%	95,00%	85,25%
0,3	0	54,55%	90,00%	83,61%
0,4	0	68,18%	81,00%	78,69%
0,5	0,1	22,73%	91,00%	78,69%
0,6	0,1	22,73%	89,00%	77,05%
0,7	0,1	36,36%	84,00%	75,41%
0,8	0,1	50,00%	81,00%	75,41%
0,9	0,1	59,09%	75,00%	72,13%
1	0,2	22,73%	85,00%	73,77%

Figura 5.6: Percentuali di previsioni corrette con il modello non oriented a nove variabili



Dalla Figura 5.6 è possibile osservare come l'andamento oscillatorio interessi anche la percentuale relativa alle imprese predette come non fallite, però tale moto segue una direzione opposta a quello delle imprese predette come fallite: all'aumentare della percentuale BR si nota un leggero decremento della percentuale NBR.

In generale, sulla base dei risultati ottenuti, l'aggiunta della variabile di Altman non ha modificato in maniera significativa le percentuali di previsione, e queste ultime sono rimaste per lo più immutate.

Il caso output-oriented si dimostra ancora l'orientamento più efficiente in quanto fornisce livelli di previsione delle insolvenze più elevati senza incidere negativamente sulle restanti percentuali.

Come nel paragrafo precedente, il confronto con i risultati ottenuti dai precedenti studi di Premachandra può fornire spunti per la convalida del modello quale efficace strumento di previsione dei fallimenti.

A fronte dello studio del 2009, nel quale le percentuali stimate da Premachandra per le imprese fallite, le imprese non fallite e il totale di predizioni corrette erano rispettivamente pari a 86%, 67,69% e 68,65%, la prima estensione del modello ad otto variabili fornisce una percentuale inferiore per la previsione delle imprese insolventi, pari a 50%, mentre si presenta molto più robusto nell'identificare le imprese sane, con una previsione corretta pari a al 90%.

La Tabella 5.11 compara i risultati del modello esaminato con quelli ottenuti dallo studio di Premachandra del 2011; anche in questo caso si osserva una maggiore capacità del modello di Premachandra di classificare le imprese insolventi a scapito, però, di percentuali inferiori per quanto riguarda la corretta identificazione delle imprese sane.

Tabella 5.11: Confronto tra i risultati del modello a nove variabili (prima estensione) e quelli del modello Premachandra 2011

λ	Modello output-oriented caso nove variabili (prima estensione)			Modello Premachandra 2011		
	%BR	%NBR	%Totale	%BR	%NBR	%Totale
0,1	54,55%	97,00%	89,34%	52,00%	93,30%	91,20%
0,2	50,00%	98,00%	89,34%	66,00%	88,20%	87,10%
0,5	59,09%	87,00%	81,97%	74,00%	81,50%	81,10%
0,9	50,00%	89,00%	81,97%	68,00%	86,10%	85,20%
1	50,00%	90,00%	82,79%	72,00%	83,20%	82,60%

Un ulteriore confronto interessante, dato dal fatto di aver incluso nell'analisi una variabile ereditata dallo studio di Altman del 1993, compara le prestazioni osservate con quelle del modello Z'-Score. Dal confronto si denota che la percentuale di corretta identificazione delle imprese sane si attesta su livelli simili (superiori al 90%), mentre la percentuale di corretta identificazione delle imprese insolventi è di molto superiore nel modello di Altman: 91% contro 59,09%. Tale differenza è dovuta alla diversa numerosità del campione di riferimento, in quanto il campione analizzato è composto da un numero ridotto d'impresе (122 imprese) rispetto al campione utilizzato da Altman e ciò limita la capacità previsionale del modello nel predire le imprese insolventi.

Comunque, prima di esporre i risultati ottenuti dalla seconda estensione del modello, è possibile affermare che il modello presentato, in virtù delle prestazioni conseguite nel caso output-oriented, può essere considerato un ottimo strumento di credit scoring per le piccole e medie imprese.

5.6.3 Modello a nove variabili: seconda estensione

L'ultimo modello considerato nell'analisi si differenzia dai precedenti per l'introduzione di un nuovo indicatore il quale risulta composto al denominatore dal valore del Patrimonio Netto nel 2010 e al numeratore dalla media dei valori assunti dal Patrimonio Netto negli ultimi tre anni di bilancio disponibili.

La Tabella 5.12 espone i valori delle percentuali di previsione corrette per il caso input-oriented; confrontando tali valori con quelli ottenuti dai due modelli a otto e a nove variabili precedenti è possibile affermare che quest'ultimo modello identifica con una maggiore precisione le imprese non fallite.

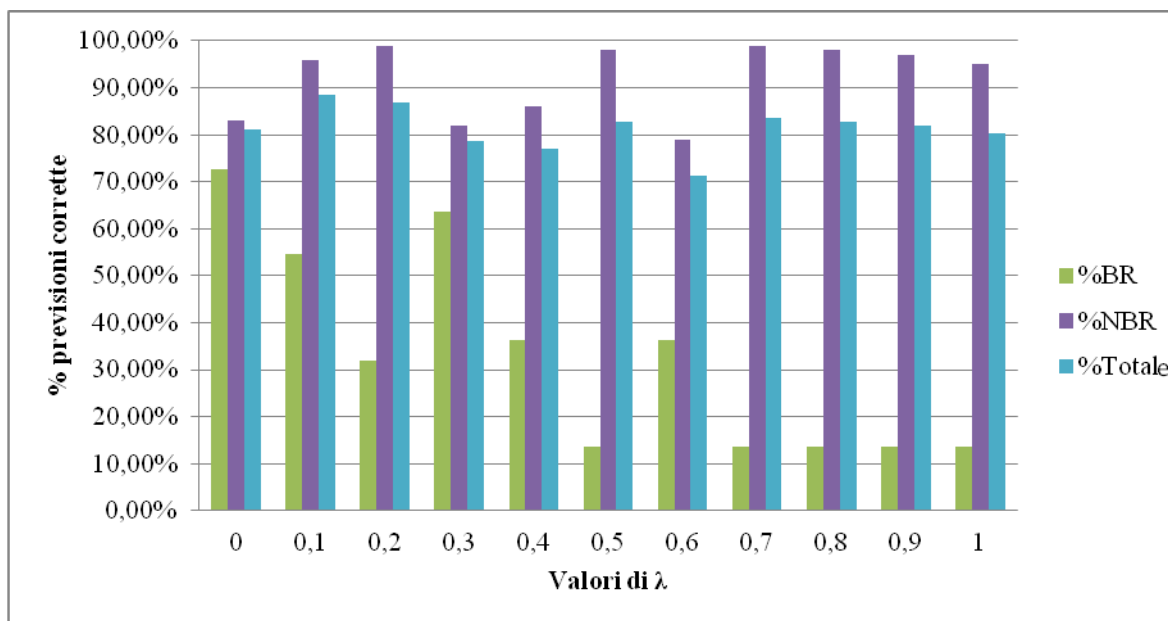
Infatti, comparando i tre modelli, notiamo che il range di percentuale del modello a otto variabili è pari a 65-100%, quello valido per il modello a nove variabili prima estensione risulta pari a 71-99%, mentre il range di percentuale relativo alla corretta previsione delle imprese non fallite per quest'ultimo modello è nettamente superiore e pari a 79-99%.

Questo incremento nelle percentuali di previsione delle imprese non fallite provoca però un deterioramento per quanto riguarda l'individuazione delle imprese fallite, infatti la percentuale massima a cui si perviene per un valore di λ pari a 0 è di 72,73% con la perdita di quasi dieci punti percentuali rispetto ai modelli precedenti.

Tabella 5.12: Percentuali di previsioni corrette con il modello input-oriented a nove variabili (seconda estensione)

λ	Cut-off migliore	%BR	%NBR	%Totale
0	-0,2	72,73%	83,00%	81,15%
0,1	-0,1	54,55%	96,00%	88,52%
0,2	0	31,82%	99,00%	86,89%
0,3	0	63,64%	82,00%	78,69%
0,4	0,1	36,36%	86,00%	77,05%
0,5	0,3	13,64%	98,00%	82,79%
0,6	0,2	36,36%	79,00%	71,31%
0,7	0,6	13,64%	99,00%	83,61%
0,8	0,7	13,64%	98,00%	82,79%
0,9	0,8	13,64%	97,00%	81,97%
1	0,9	13,64%	95,00%	80,33%

Figura 5.7: Percentuali di previsioni corrette con il modello input-oriented a nove variabili (seconda estensione)



Lo stesso si può dire per i risultati ottenuti nel caso output-oriented; infatti, rispetto ai due modelli precedentemente analizzati, si nota un incremento delle percentuali nel range previsionale relativo alle imprese non fallite, che però si verifica a scapito di una peggiore prestazione nell'individuazione dei fallimenti.

Infatti, mentre per il modello ad otto variabili e per la sua prima estensione a nove variabili i range di previsione dei fallimenti erano rispettivamente pari a 50-77,27% e 45,45-59,09%, per quest'ultimo modello il range si abbassa a 36,36-50%.

Invece, un discorso del tutto diverso può essere fatto per le percentuali di previsione relative alle imprese non fallite, infatti con un range pari a 90-97% sono superiori a quelle di tutti gli altri casi esaminati indipendentemente dall'orientamento prescelto per il modello (Figura 5.8). Anche in questo caso, come si era visto per il modello ad otto variabili, il cut-off migliore si sposta di un valore esattamente pari all'incremento del parametro λ e per i valori di λ compresi tra 0,7 e 1 le percentuali di previsione non cambiano.

Tabella 5.13: Percentuali di previsioni corrette con il modello output-oriented a nove variabili (seconda estensione)

λ	Cut-off migliore	%BR	%NBR	%Totale
0	-0,5	40,91%	97,00%	86,89%
0,1	-0,4	36,36%	95,00%	84,43%
0,2	-0,3	36,36%	94,00%	83,61%
0,3	-0,2	40,91%	93,00%	83,61%
0,4	-0,1	40,91%	93,00%	83,61%
0,5	0	40,91%	93,00%	83,61%
0,6	0,1	50,00%	91,00%	83,61%
0,7	0,2	50,00%	90,00%	82,79%
0,8	0,3	50,00%	90,00%	82,79%
0,9	0,4	50,00%	90,00%	82,79%
1	0,5	50,00%	90,00%	82,79%

Infine, il caso non oriented presenta lo stesso andamento altalenante già registrato nelle precedenti analisi, ma riporta delle differenze per quanto riguarda i range delle percentuali.

Dalla Tabella 5.14 è possibile evidenziare come la percentuale di previsione delle imprese fallite non scenda mai al di sotto del 36,36%, evento che invece accade per tre volte nel modello ad otto variabili e per quattro volte nella prima estensione di tale modello a nove variabili.

Figura 5.8: Percentuali di previsioni corrette con il modello output-oriented a nove variabili (seconda estensione)

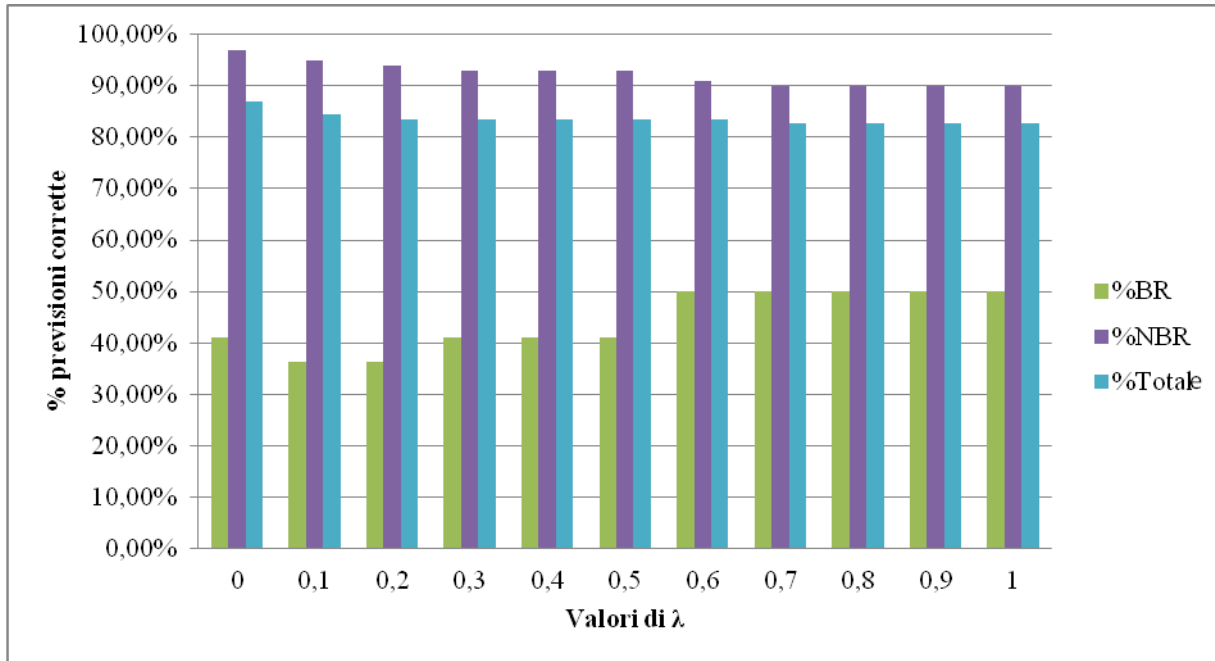
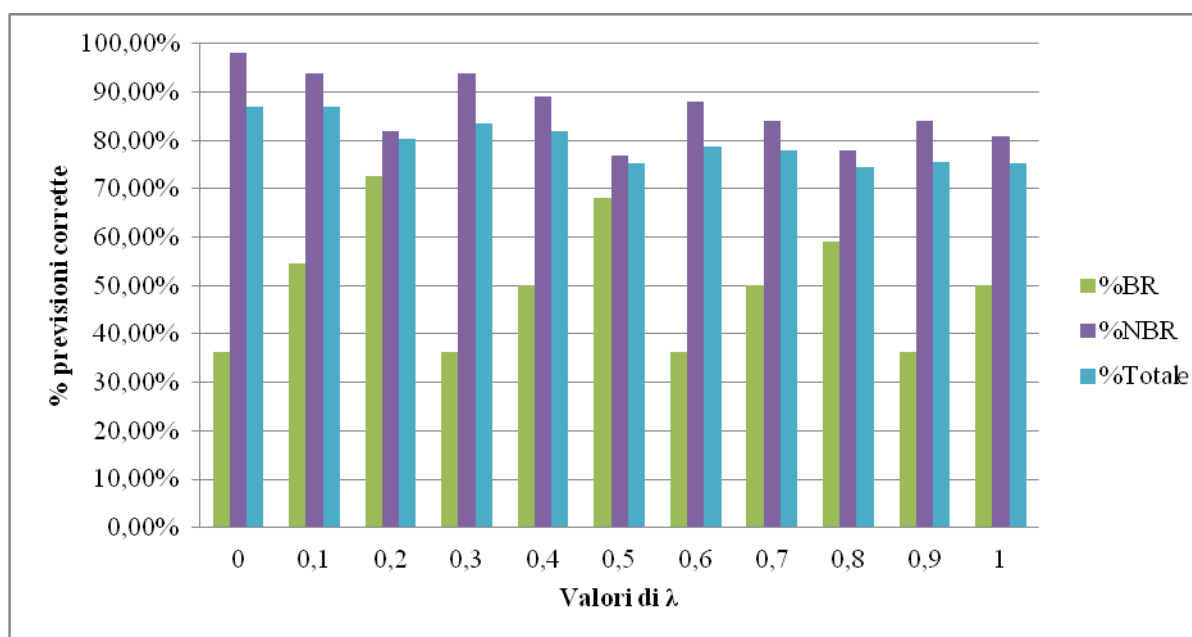


Tabella 5.14: Percentuali di previsioni corrette con il modello non oriented a nove variabili (seconda estensione)

λ	Cut-off migliore	%BR	%NBR	%Totale
0	-0,1	36,36%	98,00%	86,89%
0,1	-0,1	54,55%	94,00%	86,89%
0,2	-0,1	72,73%	82,00%	80,33%
0,3	0	36,36%	94,00%	83,61%
0,4	0	50,00%	89,00%	81,97%
0,5	0	68,18%	77,00%	75,41%
0,6	0,1	36,36%	88,00%	78,69%
0,7	0,1	50,00%	84,00%	77,87%
0,8	0,1	59,09%	78,00%	74,59%
0,9	0,2	36,36%	84,00%	75,49%
1	0,2	50,00%	81,00%	75,41%

Figura 5.9: Percentuali di previsioni corrette con il modello non oriented a nove variabili (seconda estensione)



Per quanto riguarda la percentuale di previsione corretta delle imprese non fallite e la percentuale di previsione totale, si osserva che si assestano per lo più sui valori già conseguiti dai precedenti modelli.

Alla luce di quanto riportato, l'aggiunta di quest'ultimo indicatore non comporta un miglioramento delle capacità diagnostiche rispetto al modello ad otto variabili, tuttavia ancora una volta il modello output-oriented presenta le più alte percentuali di previsione, candidandosi come valido strumento per l'identificazione delle insolvenze.

Anche in questo caso, un riscontro con i lavori di Premachandra su cui si è basato lo sviluppo dei modelli esaminati può fornire indicazioni essenziali riguardo alla capacità della metodologia DEA nel valutare la probabilità d'insolvenza delle imprese.

Metro di confronto con i lavori di Premachandra, lo stesso considerato nei paragrafi precedenti, è ancora una volta il modello output-oriented in virtù dei migliori risultati ottenuti.

Quest'ultimo, rispetto allo studio del 2009, presenta una percentuale più bassa per quanto riguarda l'identificazione delle imprese insolventi, ovvero un 50% rispetto all'86% realizzato da Premachandra, ma evidenzia una percentuale di corretta classificazione delle imprese sane pari al 90%, nettamente superiore al 67,69% conseguito da Premachandra.

La Tabella 5.15 riassume le principali differenze, per alcuni valori di λ , tra i risultati del modello proposto e quelli dello studio condotto da Premachandra nel 2011.

Tabella 5.15: Confronto tra i risultati del modello a nove variabili (seconda estensione) e quelli del modello Premachandra 2011

λ	Modello output-oriented caso nove variabili (seconda estensione)			Modello Premachandra 2011		
	%BR	%NBR	%Totale	%BR	%NBR	%Totale
0,1	36,36%	95,00%	84,43%	52,00%	93,30%	91,20%
0,2	36,36%	94,00%	83,61%	66,00%	88,20%	87,10%
0,5	40,91%	93,00%	83,61%	74,00%	81,50%	81,10%
0,9	50,00%	90,00%	82,79%	68,00%	86,10%	85,20%
1	50,00%	90,00%	82,79%	72,00%	83,20%	82,60%

Come si può vedere dalla Tabella, la seconda estensione del modello ad otto variabili non raggiunge i livelli previsionali conseguiti dal modello di Premachandra per quanto concerne le imprese insolventi; infatti mentre quest'ultimo presenta una percentuale minima di previsione del fallimento del 52%, il modello proposto perviene ad una percentuale massima di previsione del 50%.

Invece, un discorso inverso può essere fatto per quanto riguarda l'abilità previsionale relativa alle imprese reputate sane, infatti il modello proposto classifica correttamente le imprese sane con percentuali che vanno dal 90% al 95%, percentuali maggiori rispetto al range 81,50-93,30% ottenuto dal lavoro di Premachandra.

5.6.4 Confronto tra i modelli

Sulla base dei risultati descritti in maniera dettagliata nei paragrafi precedenti, è possibile affermare che l'ottica output-oriented prevale su tutti gli altri orientamenti. Infatti, tale orientamento è l'unico che riesce a garantire percentuali previsionali soddisfacenti sia per le imprese fallite che per quelle non fallite (la percentuale BR si aggira attorno al 50% e la percentuale NBR ha un range pari a 86-98%).

Inoltre, si può sottolineare come il modello si presti a molteplici usi da parte dell'analista finanziario; difatti, come già accennato in precedenza, a seconda del valore di λ selezionato è possibile utilizzare il modello per predire con una maggiore cura la percentuale di imprese insolventi oppure la percentuale di imprese solventi.

In particolare, nel caso in cui il modello fosse rivolto alla previsione delle insolvenze, l'analista dovrebbe rivolgere la propria attenzione al modello ad otto variabili output-oriented nel quale, per un valore di λ pari a 0 ed un cut-off pari a $-0,5$, si ottiene una percentuale predittiva per le imprese fallite pari al 77,27%. Al contrario, nel caso in cui l'esigenza fosse quella di identificare le imprese sane e meritevoli di credito, il modello a nove variabili (prima estensione) ricoprirebbe il ruolo di candidato ideale per la classificazione in quanto, per un valore di λ pari a 0,2 ed un cut-off pari a $-0,2$, consente di ottenere una percentuale di previsione per le imprese non fallite pari al 98%.

Il modello pertanto, a seconda dell'esigenza, può essere indirizzato dai diversi istituti di credito alla previsione dei fallimenti o all'individuazione delle imprese meritevoli di credito.

CONCLUSIONI

Questo lavoro si è posto l'obiettivo di verificare se la metodologia DEA possa essere considerata un strumento di credit scoring anche per le piccole e medie imprese.

Tramite tale metodologia è stato possibile analizzare le diverse unità decisionali (imprese) tenendo in considerazione una pluralità di input e di output, soprattutto grazie all'abilità dimostrata dal modello stesso nel maneggiare un'ampia quantità di dati ed informazioni.

Sulla base dei risultati ottenuti è stato possibile constatare che, tra i modelli proposti, il modello ad otto variabili e la sua prima estensione (caratterizzata dall'aggiunta di un indicatore finanziario suggerito dal modello Z'-Score di Altman) presentano le prestazioni previsionali più interessanti ai fini dell'indagine.

In primo luogo il confronto con lo studio condotto da Premachandra nel 2009, nel quale i risultati sono stati ottenuti mediante l'impiego di una sola frontiera (la frontiera dei fallimenti), ha mostrato come i modelli proposti accusino in parte un divario percentuale nella previsione delle insolvenze, ma siano in grado di identificare in maniera molto più corretta le imprese non fallite (90% contro 67,69%).

Gli stessi modelli, rispetto ai risultati del più recente lavoro di Premachandra, presentano un range di percentuali di previsione delle insolvenze leggermente inferiore, infatti, in opposizione al range pari a 52-74% realizzato da Premachandra, i modelli proposti manifestano dei livelli previsionali che vanno dal 50% al 63,64%.

D'altra parte, invece, i modelli sopra descritti si comportano molto meglio nell'identificare correttamente le imprese non fallite. Infatti mentre Premachandra ottiene delle percentuali di corretta identificazione che vanno dall'81,50% al 93,30%, l'analisi condotta ha permesso di classificare in maniera corretta le imprese ritenute sane con una precisione che va dall'87% al 98%.

In generale la metodologia proposta è in grado di classificare correttamente circa il 50% delle imprese insolventi, con picchi del 77,27%, e circa il 90% delle imprese sane; questo risultato, confrontato con la letteratura in merito, permette di considerare la metodologia DEA come un robusto strumento di valutazione dei fallimenti aziendali.

Inoltre, è necessario sottolineare che le prestazioni ottenute si riferiscono ad un campione di numerosità ridotta, pari a 122 imprese, e come già accennato in precedenza, selezionando un

campione di numerosità più elevata si potrebbero migliorare sensibilmente i risultati previsionali del modello.

Alla luce delle prestazioni ottenute, si è convinti che la metodologia DEA possa essere considerata da un generico analista finanziario un valido strumento di credit scoring per le PMI, soprattutto per i numerosi vantaggi derivanti dal fatto di non richiedere assunzioni a priori sulle caratteristiche delle variabili, dalla capacità di trattare una pluralità di input e di output e dal fatto di non necessitare di un campione particolarmente esteso per la valutazione del fenomeno dell'insolvenza.

In aggiunta, un vantaggio molto importante è costituito dal fatto che a seconda delle esigenze, tramite la selezione di specifici valori di λ , è possibile indirizzare il modello o alla previsione delle insolvenze o, d'altro canto, all'identificazione delle imprese meritevoli di finanziamenti.

APPENDICE A

Variabili di input ed output impiegate nell'indagine

Come già anticipato nel Capitolo 5, di seguito sono riportati i valori degli indicatori finanziari impiegati come variabili di input e di output per i tre modelli precedentemente esposti allo scopo di prevedere la probabilità d'insolvenza delle imprese in esame.

Nella Tabella seguente gli indicatori finanziari scelti come variabili di input sono evidenziati dalla lettera (i), mentre gli indici riconosciuti come variabili di output sono segnalati dalla lettera (o). Sulla base di tale distinzione, è possibile suddividere le variabili utilizzate in questo modo:

- variabili di input:

- cash flow/total assets
- net income/total assets
- working capital/total assets
- current assets/total assets
- ebit/ total assets
- ebit/interest expense
- average Equity last 3 years/Equity
- book value of Equity/total debt;

- variabili di output:

- total debt/total assets
- current liabilities/total assets.

Tabella A.1: Valori degli indici finanziari utilizzati come variabili di input e di output nei tre modelli

IMPRESA	(i) cash flow/total assets	(i) net income/total assets	(i) working capital/total assets	(i) current assets/total assets	(i) ebit/total assets	(i) ebit/interest expense	(i) book value of Equity/total debt	(i) average Equity last 3 years/Equity	(o) total debt/total assets	(o) current liabilities/total assets
A.I.S.A. S.P.A. AREZZO IMPIANTI E SERVIZI AMBIENTALI	0,08541	0,00531	0,21704	0,50242	0,01976	1,35780	0,82114	0,88535	0,48258	0,28538
A.T.C. AIR TRANSPORTATION CONSULTANTS S.P.A.	0,04658	0,03462	0,27104	0,91321	0,07406	16,84902	0,48097	0,99829	0,64217	0,64217
ACCIAI RIVESTITI VALDARNO S.R.L.	0,05964	0,04407	0,28072	0,96045	0,09625	5,40789	0,45375	0,81351	0,68051	0,67973
ACHITEX SRL	0,02004	-0,03419	0,15196	0,56146	-0,01517	-0,82714	0,37676	0,97535	0,67160	0,40950
ACQUE ALBULE S.P.A	0,06018	0,00699	0,04525	0,26388	0,03478	2,19066	0,68036	0,99466	0,59214	0,21863
ACQUE DI CALTANISSETTA SPA	0,00984	-0,03227	-0,32894	0,63082	-0,01775	-1,03254	0,01375	2,41362	0,97228	0,95976
ADAMANTE SRL	0,02557	0,00087	0,00136	0,71597	0,00643	0,77469	0,05214	0,51963	0,93538	0,71461
ADLER SPA	0,01886	-0,01768	-0,05160	0,36719	0,02175	0,72246	0,44178	0,99268	0,66750	0,41879
AEMME LINEA DISTRIBUZIONE S.R.L.	0,04808	0,00165	-0,33802	0,29226	0,02066	6,97874	0,47177	0,97256	0,63028	0,63028
AEPI COSTRUZIONI ELETTROMECCANICHE SRL	0,02415	0,00423	-0,00886	0,60762	0,01781	1,87286	0,10089	0,97757	0,85246	0,61648
AGENZIA ROGNONI SRL	0,01578	0,00043	-0,18774	0,38041	0,04133	1,49501	0,03477	0,98525	0,91901	0,56815
ALCHEMIA ITALIA S.R.L.	-0,24804	-0,30888	-0,35436	0,20412	-0,20534	-4,96345	0,16087	1,69933	0,85717	0,55848
ALHAMBRA'S GARDEN SRL	-1,17813	-1,21220	-1,60320	0,59811	-1,09080	-8,90176	-0,54865	0,05884	2,20131	2,20131
AMAGA ENERGIA E SERVIZI SRL	0,04331	0,01025	-0,06978	0,64517	0,02889	10,64630	0,18484	0,61124	0,76153	0,71495
ANDREOTTI FOTOINCISIONI S.R.L.	0,06481	0,00129	-0,05765	0,32754	0,02241	2,91533	0,70245	1,16410	0,54588	0,38519
AUTOSERVIZI MORETTI SRL	0,13166	0,05309	0,38706	0,76167	0,04977	195,44695	0,74701	0,70975	0,46694	0,37461
AZIENDA PROVINCIALE TRASPORTI SPA	0,12178	0,06460	0,51964	0,65631	0,09031	71,11843	5,71727	0,91173	0,13667	0,13667

B.R.B. S.R.L.	0,09054	0,00079	-0,59715	0,22051	0,02187	1,05386	0,08262	-0,55484	0,82636	0,81766
BBG S.P.A.	-0,60466	-0,88967	-0,73444	0,72909	-0,79874	-8,72746	-0,51716	-0,08063	1,60833	1,46353
BIT MARKET SERVICES S.P.A.	0,00620	0,00367	0,03009	0,26692	0,00661	20,83228	3,18588	1,00000	0,23683	0,23683
CALCESTRUZZI DELLA VAL ROJA - S.R.L.	0,11452	0,07348	0,36368	0,71246	0,12442	61,44254	1,57184	0,83306	0,36108	0,34878
CARPEDIL - S.P.A.	0,02805	0,01460	0,25470	0,55761	0,04941	1,89357	1,02398	0,56486	0,44810	0,30291
CASA DEL SOLE SPA	-0,00372	-0,02046	-0,02968	0,37457	-0,00407	-0,31204	0,55231	0,85387	0,51955	0,40425
CLEAN SERVICE S.R.L.	0,05224	0,00728	0,19579	0,91663	0,13535	13,15873	0,24535	0,60043	0,72084	0,72084
COMPAGNIA PORTUALE T GULLI SRL	0,16009	0,09140	-0,06032	0,43243	-0,00698	-1,84071	0,82022	0,59281	0,49275	0,49275
COSMO HOTEL S.P.A.	0,01714	0,00667	0,43332	0,70411	0,01840	4,37134	0,90951	0,39763	0,51710	0,27079
CPS COMPAGNIA PORTUALI SERVIZI SOCIETA COOPERATIVA	0,02398	-0,03345	0,31831	0,48509	-0,00237	0,83639	1,38223	0,89933	0,32632	0,16678
DATA PROCESSING	0,03659	0,03286	0,38927	0,94165	0,05322	76,73063	0,53671	0,65009	0,55238	0,55238
DIMENSIONE TURISMO S.R.L.	0,06982	-0,00788	-0,11376	0,48744	-0,00557	-0,32878	0,31463	0,20451	0,72262	0,60120
DOMUS VESTRA S.P.A.	0,06640	0,03423	0,02321	0,14313	0,06157	6,74701	2,11816	0,72548	0,31230	0,11992
DRIVE SERVICE SPA	0,05778	0,01917	0,04367	0,87557	-0,01160	-1,53578	0,10801	0,69207	0,88359	0,83190
ECOLOGICA SRL	0,11114	0,08213	0,31130	0,94034	0,09890	6,87398	0,45280	0,46026	0,62904	0,62904
ECOLOGISTICA SERVIZI SRL	0,11320	0,03226	-0,09332	0,70350	0,06843	3,90836	0,25004	0,96197	0,79682	0,79682
EQUITALIA GIUSTIZIA S.P.A.	-0,05660	-0,05671	0,86090	0,98592	-0,08204	-14,49460	6,97923	1,00000	0,12502	0,12502
ESDRA COSTRUZIONI S.R.L.	0,02650	0,00282	-0,05571	0,68943	0,04059	4,07692	0,06144	0,75072	0,91229	0,74514
EUGANEA MEDICA S.R.L	0,22482	0,15572	0,43253	0,75616	0,21722	47,00872	1,82572	0,59687	0,32363	0,32363
FERRARI - S.P.A.	0,02423	0,00676	0,22008	0,58244	0,06564	17,79359	0,51669	0,89220	0,62805	0,36236
FINARGE ARMAMENTO GENOVESE SRL	0,17691	0,10177	0,06711	0,16494	0,11549	5,96653	0,66627	0,62628	0,59888	0,09783
FONDERIA BOCCACCI S.P.A.	0,11902	-0,01964	0,23917	0,51000	-0,01596	-3,53687	1,09994	1,13056	0,46316	0,27083
FONDERIE PONTINE BREVETTI CAVIS SRL	0,09229	0,04957	-0,00875	0,52121	0,08157	6,02451	0,71572	0,84072	0,52996	0,52996
FRIULCO S.P.A.	-0,01411	-0,04482	0,05117	0,77806	-0,02298	-1,10088	0,15949	1,25801	0,85050	0,72689
FRIULSIDER S.P.A.	0,03053	0,02412	0,52871	0,81165	0,03818	73,54535	2,11967	1,07548	0,29772	0,28294
GENNARO COSTRUZIONI	-0,01469	-0,02007	0,04704	0,97826	-0,02371	1,18922	0,03808	1,51857	0,93122	0,93122

GEOVEST SRL	0,07693	0,00191	0,10572	0,55074	0,04428	1,26194	0,05873	0,92169	0,92115	0,44502
GESTIONE ESERCIZI TURISTICI ALBERGHIERI G.E.T.A. S.P.A.	0,01342	-0,01354	0,00993	0,13331	-0,00816	-4,57282	3,18699	0,86263	0,23175	0,12338
GESTIONE SERVIZI INFORMATICI S.P.A.	0,01761	0,00032	0,69607	0,79684	-0,02666	95,73195	0,03346	0,97764	0,94836	0,10077
GESTIONI AEROPORTI SARDI GEASAR SPA	0,08647	0,03534	0,04296	0,43196	0,06524	15,77137	1,44594	0,91030	0,38900	0,38900
HEINLOTH TRANSPORT GMBH & CO. KG	0,10438	0,07586	0,22171	0,80596	0,09190	41,95558	0,58909	0,97347	0,58425	0,58425
I VIAGGI DEL TURCHESE S.R.L.	0,13204	0,10806	0,55295	0,95758	0,06274	53,08629	1,38922	0,61101	0,40491	0,40463
IL DELFINO RISTORAZIONE S.R.L.	-0,14820	-0,19275	0,71678	0,71807	0,51147	22,73451	-0,70499	0,26838	1,71030	0,00129
IMPRESA VERDE FOGGIA SRL	0,00531	-0,04234	0,54939	0,86518	-0,00132	-131,43750	0,13359	0,65305	0,61559	0,31579
IMPRESA VERDE ROMAGNA SRL	0,05585	0,02562	0,48323	0,94663	0,09575	140,20650	0,17052	0,11335	0,55808	0,46340
ISTITUTO ANALISI MEDICHE SAN LORENZO - S.R.L.	0,09442	0,04092	0,22819	0,89770	0,11033	4,81417	0,21507	0,32568	0,75721	0,66951
ITACA RISTORAZIONE E SERVIZI S.R.L.	0,02957	0,00219	-0,17172	0,77085	0,04876	8,18378	0,03722	1,00000	0,94257	0,94257
ITALIANA COSTRUZIONI S.R.L.	-0,07846	-0,08973	-0,26207	0,91401	-0,05773	-3,70308	-0,16500	0,11188	1,17608	1,17608
J.D.S. SPA	0,07620	0,05176	0,19477	0,33771	0,09888	17,43275	1,46834	0,63374	0,35264	0,14294
JESOLO TURISMO SPA	0,09231	0,00077	-0,16693	0,25316	0,04723	3,95727	0,51315	0,98398	0,63092	0,42009
LA PERLA S.R.L.	0,00162	0,00197	0,54036	0,99558	0,05495	2,53166	0,79415	0,95811	0,50231	0,45522
LA TITANO EDILIZIA S.R.L.	0,02908	0,00600	-0,05187	0,83764	0,01541	0,74969	0,04009	0,70337	0,95182	0,88951
LAVORO E SERVIZI VALSUGANA SOCIETA' COOPERATIVA	0,09779	0,05960	0,44886	0,75807	0,08659	5,15790	1,20921	0,74406	0,42337	0,30921
LOGISERVICES SOCIETA COOPERATIVA A RL	-0,03153	-0,03588	0,10168	0,93504	0,00666	0,70890	0,10225	1,72838	0,83336	0,83336
LOGITRANS SRL	0,11377	0,03735	0,02775	0,86727	0,02371	8,82919	0,12554	0,26877	0,83952	0,83952
LUMAN C.M.L. S.R.L.	-0,71161	-0,86868	-0,77007	0,92044	-0,81520	-10,25730	-0,45960	-0,07486	1,76570	1,69051
M.A. INTERNATIONAL	-0,17752	-0,18607	-0,18274	0,96020	-0,14328	-10,97850	-0,18509	0,03016	1,14294	1,14294

CATERING SRL										
M.E. MARITTIMA EMILIANA S.P.A.	0,03564	-0,01310	0,13960	0,20673	0,00987	0,79019	0,87269	1,13151	0,51498	0,06713
MAINA ORGANI DI TRASMISSIONE S.P.A.	0,10405	0,05552	0,47379	0,84480	0,09262	21,48322	0,86851	0,76603	0,50452	0,37101
MEDITERRANEAN SHIPPING COMPANY S.R.L.	0,01881	-0,02592	0,02884	0,65594	0,04504	2,01016	0,08854	1,07156	0,72551	0,62710
METRONOTTE ISTITUTO DI VIGILANZA PRIVATA SRL	-0,01960	-0,03173	-0,05635	0,92311	0,01251	2,85982	-0,00032	-31,72387	0,99104	0,97946
MINOSSE S.R.L.	0,38127	0,36909	0,19860	0,94968	0,57673	614,59450	0,29006	0,56433	0,75188	0,75108
NAVITAL S.P.A.	0,02129	-0,01730	-0,05475	0,67217	0,04647	1,18180	0,25267	0,78423	0,75744	0,72692
NO PROBLEM PARKING S.P.A.	0,04936	0,00123	-0,26185	0,28435	0,02791	1,32478	0,29440	0,86832	0,75235	0,54620
NUOVA FONDERIA DI CASTENEDOLO S.P.A.	0,14606	0,09885	0,06646	0,43408	0,16153	19,58937	1,31832	0,32290	0,41905	0,36762
NUOVI CANTIERI APUANIA SPA	-0,15395	-0,21860	0,51087	0,73554	0,19119	7,64292	1,56504	1,02193	0,22467	0,22467
ORGANIZZAZIONE STRUTTURE CEMENTO ARMATO OSCA PREFABBRICATI SRL	0,00026	-0,01359	0,35906	0,77885	-0,00153	-0,09982	0,25779	0,83783	0,72393	0,41979
OTTAVI PREFABBRICATI SRL	-0,00789	0,01605	0,36991	0,94420	0,00425	0,21435	0,28621	1,06206	0,75531	0,57429
PAGLIERANI SRL	0,01346	0,00197	0,20780	0,82967	0,02241	1,90814	0,02314	3,45492	0,95080	0,62187
PARADIVI SERVIZI SRL	0,09474	0,06567	0,15021	0,90250	0,11326	56,91933	0,28115	0,69613	0,75229	0,75229
PIONEER HIBRED ITALIA SERVIZI AGRONOMICI SRL	0,11983	0,09646	0,21370	0,81526	0,14797	185,29074	0,63244	0,52050	0,60156	0,60156
PROGEMA IMPIANTI S.R.L.	0,02444	-0,02741	-0,29670	0,96354	-0,04793	-0,24571	-0,21921	0,55398	1,26024	1,26024
PUBBLICITA' S.R.L.	0,16172	0,02904	-0,11682	0,45298	0,09054	14,15293	0,42582	0,49147	0,67524	0,56980
REFITALIA S.R.L.	-0,22685	-0,22685	-0,16024	0,83492	-0,21463	-17,69870	-0,14600	-0,41353	1,12727	0,99516
RMG COSTRUZIONI SRL	0,02322	0,01845	0,09808	0,98657	0,07812	2,24433	0,11594	0,49011	0,88849	0,88849
RUSPA OFFICINE SPA	0,05745	0,00578	0,10407	0,76015	0,03570	4,65479	0,23867	1,08937	0,74331	0,65608
S.I.TEC. S.R.L.	-0,07291	-0,09436	-0,17047	0,86793	-0,05779	-2,51605	-0,08193	-0,05664	1,03840	1,03840
S.L.R. SERVICE	0,03150	0,00957	-0,06291	0,68129	0,07021	2,47089	0,01845	0,37810	0,93975	0,74420

SATER SERVIZI AMBIENTALI TERRITORIALI S.P.A.	0,12220	0,00509	0,22681	0,70085	0,06233	22,74618	0,96857	0,98033	0,47404	0,47404
SC SALDATURA CARPENTERIA SRL	0,03452	0,00217	0,06109	0,69258	0,02339	3,57341	0,16217	0,92504	0,75991	0,63149
SEA SOLUZIONI ECOLOGICHE AMBIENTALI SRL	0,03294	-0,01218	-0,36186	0,52364	0,05438	1,32035	0,06297	0,96412	0,91462	0,88550
SEA SRL SERVIZI ECOLOGICI AMBIENTALI	0,08854	0,00105	-0,20236	0,58850	0,06317	4,22687	0,05514	0,76949	0,90113	0,79086
SECURITY SERVICE SISTEMI S.R.L.	0,03290	0,00115	0,11056	0,95257	0,03806	1,84845	0,10342	0,83516	0,84201	0,84201
SEGHERIA PIAVE S.R.L.	0,03011	0,01434	0,24155	0,96185	0,10891	2,63129	0,07497	0,80297	0,73942	0,72030
SERVECO SRL	0,05391	0,02487	0,01167	0,68984	0,05998	5,03976	0,19223	0,85646	0,80699	0,67817
SERVIZI ALLA STRADA SPA	0,05916	0,00154	0,02272	0,68110	0,03726	138,56506	0,32964	0,77686	0,65838	0,65838
SICILIA POLICE S.R.L.	0,06217	0,02327	0,10960	0,85179	0,12316	5,04463	0,17306	0,57005	0,74219	0,74219
SIDOTI COSTRUZIONI SRL	0,00945	0,00455	0,04786	0,96569	0,00502	0,73460	0,05956	0,84159	0,92965	0,91783
SIMONT S.P.A.	0,00899	0,00345	0,29435	0,95886	0,06172	1,32647	0,15625	0,93925	0,84232	0,66451
SOCIETA ALGERESE TURISTICO ALBERGHIERA SRL	0,05020	0,02038	0,52945	0,61910	0,06514	3,47156	0,08012	0,59045	0,91686	0,08965
SOCIETA ANONIMA CUPRENSE SERVIZI AUTOMOBILISTICI SRL	0,11906	-0,01193	0,11895	0,44075	0,05730	9,99555	0,12163	1,24681	0,32180	0,32180
SOLIDARIETA CINQUE SOC COOP SOCIALE	0,03910	0,00685	0,01771	0,83727	0,03894	5,39377	0,12543	1,09496	0,85114	0,81956
SORA S.R.L.	-0,12411	-0,14032	0,30275	0,97040	-0,08164	-4,89210	0,01392	17,92914	0,85334	0,66765
SPREA EDITORI S.P.A.	0,08964	0,01831	0,07717	0,72260	-0,01263	-3,00537	0,45708	1,39836	0,64770	0,64543
SPRING S.R.L.	0,00787	-0,01099	-0,03143	0,75266	-0,14967	-3,78106	0,00985	0,86147	0,98215	0,78409
SRB COSTRUZIONI SRL	0,15198	0,01274	-0,13750	0,56229	0,07642	5,02581	0,12905	0,71021	0,84383	0,69979
STA SERVIZI TECNICI AMMINISTRATIVI SOCARL	0,09362	0,00537	0,38528	0,79076	0,05136	311,55556	0,69942	1,03423	0,40548	0,40548
STANADYNE S.P.A.	0,02909	-0,02503	0,22352	0,72054	0,01316	0,61302	0,15356	1,12978	0,63037	0,49702
SYMANTEC SRL	0,12105	0,09918	0,45680	0,97953	0,19656	281,48170	0,69321	0,19774	0,52273	0,52273
TEKNOSERVICE S.R.L.	0,02774	-0,01398	-0,13015	0,72494	0,00545	0,94618	0,11678	0,88997	0,85844	0,85509

TIPOGRAFIA TOSCANA S.R.L. INDUSTRIA GRAFICA	-0,36716	-0,44351	-0,29570	0,75113	-0,42527	-13,51100	-0,17638	-0,86685	1,12438	1,04683
TRANSPOBANK SRL	0,01523	-0,03157	-0,38408	0,35881	-0,02576	-2,57563	0,03027	2,87258	0,80124	0,74289
TRASIMENO SERVIZI AMBIENTALI S.P.A.	0,03166	0,00851	0,10171	0,80570	0,03136	9,37644	0,31054	0,90033	0,70399	0,70399
TRASPORTI ROMA RX S.R.L.	0,01564	0,01008	0,14329	0,97804	0,03718	2,69230	0,14362	0,96812	0,83475	0,83475
TREVISO SINERGIE S.R.L.	-0,06872	-0,07602	0,47038	0,95541	0,01964	2,88233	0,49664	1,45984	0,49632	0,48503
UNICOOP SERVIZI LIVORNO SOCIETA COOPERATIVA	-0,06723	-0,08433	-0,19936	0,52141	-0,07059	3,90346	-0,04493	-1,90595	0,99837	0,72077
USI ITALIA SRL	0,02651	-0,03958	0,20510	0,85885	-0,05858	-29,27151	0,25641	1,55391	0,74976	0,65375
VACCARI GIOVANNI S.R.L.	0,01814	-0,02636	-0,28283	0,64620	0,00196	0,08457	0,01659	0,32818	0,93804	0,92903
VENEZIA INFORMATICA E SISTEMI VENIS SPA	0,01320	0,00113	0,41403	0,97237	0,01673	1,20758	0,12583	0,93763	0,83928	0,55834
VETRICERAMICI S.P.A.	0,10235	0,00080	0,17916	0,41012	0,04653	2,31444	0,39308	0,69180	0,70471	0,23096
VETROBALSAMO SPA	0,06350	-0,00022	-0,13122	0,24334	-0,00422	-0,93042	0,57693	0,57404	0,62651	0,37456
VILLAIRIS SRL	0,03142	0,00569	0,19797	0,41728	0,02607	6,91113	3,43937	0,96782	0,21931	0,21931
WORLD STARTEL COMMUNICATIONS S.R.L.	-0,48442	-0,49799	-0,48250	0,69684	-0,26868	-4,97936	-0,25472	-0,50076	1,30050	1,17934
ZETA FASHION GROUP S.R.L.	0,03171	0,01448	0,03455	0,95087	0,01969	0,46206	0,04415	0,65325	0,94018	0,91632
ZETA-FPA ENGINEERING	-0,28319	0,29141	0,12140	0,91293	-0,24192	-12,07620	-0,30785	-0,06426	0,88121	0,79153

APPENDICE B

Risultati

Le tabelle seguenti presentano i risultati ottenuti dall'indagine per ognuno dei tre modelli precedentemente esposti; l'analisi è stata focalizzata sul caso output-oriented in quanto, come si è già visto nel Capitolo 5, si è dimostrato l'orientamento in grado di fornire le prestazioni più efficienti.

In ogni tabella è stato riportato il nome dell'impresa, il punteggio di score per diversi valori di λ (i valori ritenuti più significativi ai fini dell'analisi) e il numero di imprese classificate come fallite per un determinato livello di cut-off.

I valori di λ per i quali sono stati riportati i punteggi di score rappresentano rispettivamente la frontiera efficiente ($\lambda = 0$) e la frontiera dei fallimenti ($\lambda = 1$), inoltre è stato aggiunto un caso intermedio, in cui il valore di λ è compreso tra i due estremi, al fine di constatare i progressi ottenuti dall'utilizzo di entrambe le frontiere per la classificazione delle imprese.

Per ogni impresa viene indicato il punteggio di score per i diversi valori di λ e, nel caso in cui l'impresa venga classificata come fallita dal modello, vi è un'indicazione del fallimento nella rispettiva colonna tramite la lettera F.

Tabella B.1: Score di efficienza DEA e numero di imprese fallite per il modello output-oriented ad otto variabili, per diversi valori di λ

Impresa	$\lambda=0$	Imprese classificate fallite (cut-off -0,5)	$\lambda=0,2$	Imprese classificate fallite (cut-off -0,3)	$\lambda=1$	Imprese classificate fallite (cut-off 0,5)	Imprese realmente fallite
A.I.S.A. S.P.A. AREZZO IMPIANTI E SERVIZI AMBIENTALI	-0,698568		-0,52011		0,193715		
A.T.C. AIR TRANSPORTATION CONSULTANTS S.P.A.	-0,591281		-0,41468		0,291722		
ACCIAI RIVESTITI VALDARNO S.R.L.	-0,586604		-0,40749		0,308961		
ACHITEX SRL	-0,612999		-0,44118		0,246104		
ACQUE ALBULE S.P.A	-0,6122		-0,4245		0,326314		
ACQUE DI CALTANISSETTA SPA	-0,467461		-0,2862		0,43882		
ADAMANTE SRL	-0,502636		-0,3285		0,368065		
ADLER SPA	-0,558698		-0,37111		0,379219		
AEMME LINEA DISTRIBUZIONE S.R.L.	-0,457303		-0,24957		0,581386		
AEPI COSTRUZIONI ELETTROMECCANICHE SRL	-0,527482		-0,35698		0,325041		
AGENZIA ROGNONI SRL	-0,497682		-0,29823		0,499602		
ALCHEMIA ITALIA S.R.L.	-0,38611	F	-0,10889	F	1	F	F
ALHAMBRA'S GARDEN SRL	-0,00402	F	0,196788	F	1	F	F
AMAGA ENERGIA E SERVIZI SRL	-0,513411		-0,34372		0,33503		
ANDREOTTI FOTOINCISIONI S.R.L.	-0,566792		-0,37893		0,372531		
AUTOSERVIZI MORETTI SRL	-0,897382		-0,68014		0,188847		
AZIENDA PROVINCIALE TRASPORTI SPA	-1		-0,78758		0,062086		
B.R.B. S.R.L.	-0,38105	F	-0,10484	F	1	F	F
BBG S.P.A.	-0,26625	F	-0,07376	F	0,696184	F	F
BIT MARKET SERVICES S.P.A.	-0,621093		-0,4491		0,238851		
CALCESTRUZZI DELLA VAL ROJA - S.R.L.	-0,81376		-0,61877		0,161187		
CARPEDIL - S.P.A.	-0,703229		-0,52737		0,176045		
CASA DEL SOLE SPA	-0,570068		-0,39037		0,328439		
CLEAN SERVICE S.R.L.	-0,551744		-0,3759		0,32746		
COMPAGNIA PORTUALE T GULLI SRL	-0,559934		-0,3862		0,308719		
COSMO HOTEL S.P.A.	-0,755373		-0,572		0,161469		
CPS COMPAGNIA PORTUALI SERVIZI SOCIETA COOPERATIVA	-0,786928		-0,60485		0,123453		

DATA PROCESSING	-0,725863		-0,5305		0,250932		
DIMENSIONE TURISMO S.R.L.	-0,512516		-0,33696		0,365256		
DOMUS VESTRA S.P.A.	-0,582233		-0,31066		0,775625		
DRIVE SERVICE SPA	-0,499859		-0,32203		0,389298		
ECOLOGICA SRL	-0,591975		-0,41643		0,285757		
ECOLOGISTICA SERVIZI SRL	-0,501519		-0,32882		0,361975		
EQUITALIA GIUSTIZIA S.P.A.	-1		-0,78822		0,058914		
ESDRA COSTRUZIONI S.R.L.	-0,497279		-0,3233		0,372636		
EUGANEA MEDICA S.R.L.	-1		-0,7706		0,147017		
FERRARI - S.P.A.	-0,685749		-0,50573		0,214348		
FINARGE ARMAMENTO GENOVESE SRL	-1		-0,6		1		
FONDERIA BOCCACCI S.P.A.	-0,701687		-0,52497		0,181885		
FONDERIE PONTINE BREVETTI CAVIS SRL	-0,57638		-0,40591		0,275962		
FRIULCO S.P.A.	-0,49582	F	-0,32544		0,356084		F
FRIULSIDER S.P.A.	-0,882229		-0,67942		0,131804		
GENNARO COSTRUZIONI SRL	-0,673062		-0,45384		0,42303		
GEOVEST SRL	-0,595573		-0,41728		0,295903		
GESTIONE ESERCIZI TURISTICI ALBERGHIERI G.E.T.A. S.P.A.	-0,501508		-0,20121		1		
GESTIONE SERVIZI INFORMATICI S.P.A.	-1		-0,78345		0,082761		
GESTIONI AEROPORTI SARDI GEASAR SPA	-0,639634		-0,46291		0,24398		
HEINLOTH TRANSPORT GMBH & CO. KG	-0,653576		-0,46978		0,26541		
I VIAGGI DEL TURCHESE S.R.L.	-0,817858		-0,61751		0,183877		
IL DELFINO RISTORAZIONE S.R.L.	-1		-0,79977		0,001171		
IMPRESA VERDE FOGGIA SRL	-0,001022		0,199182		1		
IMPRESA VERDE ROMAGNA SRL	-0,842188		-0,62775		0,230023		
ISTITUTO ANALISI MEDICHE SAN LORENZO - S.R.L.	-0,55686		-0,38092		0,322837		
ITACA RISTORAZIONE E SERVIZI S.R.L.	-0,504994		-0,31836		0,428186		
ITALIANA COSTRUZIONI S.R.L.	-0,47427	F	-0,27257	F	0,534264	F	F
J.D.S. SPA	-0,77455		-0,58706		0,162889		
JESOLO TURISMO SPA	-0,509625		-0,30044		0,53628		
LA PERLA S.R.L.	-1		-0,75661		0,216965		
LA TITANO EDILIZIA S.R.L.	-0,498888		-0,31556		0,417756		
LAVORO E SERVIZI VALSUGANA SOCIETA' COOPERATIVA	-0,759618		-0,57522		0,162356		

LOGISERVICES SOCIETA COOPERATIVA A RL	-0,502843		-0,32656		0,378575		
LOGITRANS SRL	-0,522702		-0,34189		0,381373		
LUMAN C.M.L. S.R.L.	-0,25994	F	-0,04963	F	0,791581	F	F
M.A. INTERNATIONAL CATERING SRL	-0,49707	F	-0,29241	F	0,526241	F	F
M.E. MARITTIMA EMILIANA S.P.A.	-1		-0,76596		0,170185		
MAINA ORGANI DI TRASMISSIONE S.P.A.	-0,757619		-0,56725		0,194241		
MEDITERRANEAN SHIPPING COMPANY S.R.L.	-0,533345		-0,36556		0,305602		
METRONOTTE ISTITUTO DI VIGILANZA PRIVATA SRL	-0,502691		-0,31264		0,447559		
MINOSSE S.R.L.	-1		-0,73172		0,341378		
NAVITAL S.P.A.	-0,49598	F	-0,32938		0,337011		F
NO PROBLEM PARKING S.P.A.	-0,472961		-0,25834		0,600126		
NUOVA FONDERIA DI CASTENEDOLO S.P.A.	-0,670318		-0,48736		0,244476		
NUOVI CANTIERI APUANIA SPA	-1		-0,77959		0,102062		
ORGANIZZAZIONE STRUTTURE CEMENTO ARMATO OSCA PREFABBRICATI SRL	-0,643167		-0,46625		0,241411		
OTTAVI PREFABBRICATI SRL	-0,589073		-0,41198		0,296404		
PAGLIERANI SRL	-0,550185		-0,37183		0,341586		
PARADIVI SERVIZI SRL	-0,607146		-0,41737		0,341747		
PIONEER HIBRED ITALIA SERVIZI AGRONOMICI SRL	-0,789455		-0,57691		0,273274		
PROGEMA IMPIANTI S.R.L.	-0,55434		-0,32897		0,572495	F	F
PUBBLICITA' S.R.L.	-0,550833		-0,36669		0,369863		
REFITALIA S.R.L.	-0,42635	F	-0,23924	F	0,509193	F	F
RMG COSTRUZIONI SRL	-0,79779		-0,55751		0,403619		F
RUSPA OFFICINE SPA	-0,541154		-0,3696		0,316619		
S.I.TEC. S.R.L.	-0,47749	F	-0,28765	F	0,471719		F
S.L.R. SERVICE	-0,496	F	-0,32134		0,37733		F
SATER SERVIZI AMBIENTALI TERRITORIALI S.P.A.	-0,664985		-0,48892		0,215344		
SC SALDATURA CARPENTERIA SRL	-0,538811		-0,36838		0,313347		
SEA SOLUZIONI ECOLOGICHE AMBIENTALI SRL	-0,463847		-0,2778		0,466406		
SEA SRL SERVIZI ECOLOGICI AMBIENTALI	-0,485072		-0,31028		0,388885		
SECURITY SERVICE SISTEMI S.R.L.	-0,526027		-0,34432		0,382504		
SEGHIERA PIAVE S.R.L.	-0,585761		-0,40231		0,3315		

SERVECO SRL	-0,525626		-0,35354		0,334798		
SERVIZI ALLA STRADA SPA	-0,679518		-0,4838		0,299086		
SICILIA POLICE S.R.L.	-0,523649		-0,35149		0,337158		
SIDOTI COSTRUZIONI SRL	-0,589789		-0,38791		0,419615		
SIMONT S.P.A.	-0,56767		-0,38664		0,337492		F
SOCIETA ALGERESE TURISTICO ALBERGHIERA SRL	-1		-0,78516		0,074197		
SOCIETA ANONIMA CUPRENSE SERVIZI AUTOMOBILISTICI SRL	-0,679225		-0,50381		0,197848		
SOLIDARIETA CINQUE SOC COOP SOCIALE	-0,510913		-0,33286		0,379343		
SORA S.R.L.	-0,59411		-0,40722		0,340325		F
SPREA EDITORI S.P.A.	-0,529569		-0,36491		0,293717		
SPRING S.R.L.	-0,4746	F	-0,30045		0,396135		F
SRB COSTRUZIONI SRL	-0,503959		-0,32926		0,369532		
STA SERVIZI TECNICI AMMINISTRATIVI SOCARL	-1		-0,76316		0,184199		
STANADYNE S.P.A.	-0,600483		-0,42989		0,25249		
SYMANTEC SRL	-1		-0,75251		0,237463		
TEKNOSERVICE S.R.L.	-0,488103		-0,31264		0,389205		
TIPOGRAFIA TOSCANA S.R.L. INDUSTRIA GRAFICA	-0,3892	F	-0,20983	F	0,507636	F	F
TRANSPOBANK SRL	-0,427043		-0,22545		0,580943		
TRASIMENO SERVIZI AMBIENTALI S.P.A.	-0,53514		-0,36415		0,319805		
TRASPORTI ROMA RX S.R.L.	-0,688692		-0,47511		0,379206		
TREVISIO SINERGIE S.R.L.	-0,644125		-0,47073		0,222872		
UNICOOP SERVIZI LIVORNO SOCIETA COOPERATIVA	-0,467903		-0,28716		0,435793		
USI ITALIA SRL	-0,472302		-0,30474		0,365484		
VACCARI GIOVANNI S.R.L.	-0,47399	F	-0,29438	F	0,424072		F
VENEZIA INFORMATICA E SISTEMI VENIS SPA	-0,657893		-0,46539		0,304625		
VETRICERAMICI S.P.A.	-0,693613		-0,50892		0,229861		
VETROBALSAMO SPA	-0,512674		-0,30418		0,529798		
VILLAIRIS SRL	-0,735533		-0,55996		0,142338		
WORLD STARTEL COMMUNICATIONS S.R.L.	-0,39148	F	-0,2008	F	0,56192	F	F
ZETA FASHION GROUP S.R.L.	-0,51275		-0,32588		0,421611		F
ZETA-FPA ENGINEERING	-0,45876	F	-0,28965	F	0,386787		F
TOTALE IMPRESE CLASSIFICATE FALLITE		17		13		11	22

Tabella B.2: Score di efficienza DEA e numero di imprese fallite per il modello output-oriented (prima estensione), per diversi valori di λ

Impresa	$\lambda=0$	Imprese classificate fallite (cut-off -0,4)	$\lambda=0,5$	Imprese classificate fallite (cut-off 0)	$\lambda=1$	Imprese classificate fallite (cut-off 0,5)	Imprese realmente fallite
A.I.S.A. S.P.A. AREZZO IMPIANTI E SERVIZI AMBIENTALI	-0,57901		-0,19265		0,193715		
A.T.C. AIR TRANSPORTATION CONSULTANTS S.P.A.	-0,57883		-0,14355		0,291722		
ACCIAI RIVESTITI VALDARNO S.R.L.	-0,571108		-0,13107		0,308961		
ACHITEX SRL	-0,509891		-0,13189		0,246104		
ACQUE ALBULE S.P.A	-0,500808		-0,08725		0,326314		
ACQUE DI CALTANISSETTA SPA	-0,414423		0,012199		0,43882		
ADAMANTE SRL	-0,445242		-0,03859		0,368065		
ADLER SPA	-0,484744		-0,05276		0,379219		
AEMME LINEA DISTRIBUZIONE S.R.L.	-0,460688		0,060349		0,581386		
AEPI COSTRUZIONI ELETTROMECCANIC HE SRL	-0,451808		-0,06338		0,325041		
AGENZIA ROGNONI SRL	-0,405352		0,047125		0,499602		
ALCHEMIA ITALIA S.R.L.	-0,35341	F	0,323294	F	1	F	F
ALHAMBRA'S GARDEN SRL	-0,00467	F	0,497666	F	1	F	F
AMAGA ENERGIA E SERVIZI SRL	-0,481128		-0,07305		0,33503		
ANDREOTTI FOTOINCISIONI S.R.L.	-0,503653		-0,06556		0,372531		
AUTOSERVIZI MORETTI SRL	-0,70172		-0,25644		0,188847		
AZIENDA PROVINCIALE TRASPORTI SPA	-1		-0,46896		0,062086		
B.R.B. S.R.L.	-0,37093	F	0,314534	F	1	F	F
BBG S.P.A.	-0,14896	F	0,27361	F	0,696184	F	F
BIT MARKET SERVICES S.P.A.	-0,618318		-0,18973		0,238851		
CALCESTRUZZI DELLA VAL ROJA - S.R.L.	-0,73184		-0,28533		0,161187		
CARPEDIL - S.P.A.	-0,603541		-0,21375		0,176045		
CASA DEL SOLE SPA	-0,496929		-0,08425		0,328439		
CLEAN SERVICE S.R.L.	-0,530598		-0,10157		0,32746		
COMPAGNIA PORTUALE T GULLI SRL	-0,532768		-0,11202		0,308719		
COSMO HOTEL S.P.A.	-0,62055		-0,22954		0,161469		

CPS COMPAGNIA PORTUALI SERVIZI SOCIETA COOPERATIVA	-0,654271		-0,26541		0,123453		
DATA PROCESSING	-0,61866		-0,18386		0,250932		
DIMENSIONE TURISMO S.R.L.	-0,48569		-0,06022		0,365256		
DOMUS VESTRA S.P.A.	-0,570701		0,21465		1		
DRIVE SERVICE SPA	-0,474151		-0,04243		0,389298		
ECOLOGICA SRL	-0,576194		-0,14522		0,285757		
ECOLOGISTICA SERVIZI SRL	-0,507515		-0,07277		0,361975		
EQUITALIA GIUSTIZIA S.P.A.	-1		-0,47054		0,058914		
ESDRA COSTRUZIONI S.R.L.	-0,447201		-0,03728		0,372636		
EUGANEA MEDICA S.R.L.	-1		-0,42649		0,147017		
FERRARI - S.P.A.	-0,554029		-0,16984		0,214348		
FINARGE ARMAMENTO GENOVESE SRL	-1		0		1		
FONDERIA BOCCACCI S.P.A.	-0,61244		-0,21528		0,181885		
FONDERIE PONTINE BREVETTI CAVIS SRL	-0,550595		-0,13732		0,275962		
FRIULCO S.P.A.	-0,47008		-0,057		0,356084		F
FRIULSIDER S.P.A.	-0,794405		-0,3313		0,131804		
GENNARO COSTRUZIONI SRL	-0,443127		-0,01005		0,42303		
GEOVEST SRL	-0,449373		-0,07674		0,295903		
GESTIONE ESERCIZI TURISTICI ALBERGHIERI G.E.T.A. S.P.A.	-0,512514		0,243743		1		
GESTIONE SERVIZI INFORMATICI S.P.A.	-1		-0,45862		0,082761		
GESTIONI AEROPORTI SARDI GEASAR SPA	-0,602546		-0,17928		0,24398		
HEINLOTH TRANSPORT GMBH & CO. KG	-0,611782		-0,17319		0,26541		
I VIAGGI DEL TURCHESE S.R.L.	-0,733471		-0,2748		0,183877		
IL DELFINO RISTORAZIONE S.R.L.	-1		0		1		
IMPRESA VERDE FOGGIA SRL	-0,001191		0,499405		1		
IMPRESA VERDE ROMAGNA SRL	-0,55345		-0,16171		0,230023		
ISTITUTO ANALISI MEDICHE SAN LORENZO - S.R.L.	-0,520811		-0,09899		0,322837		
ITACA RISTORAZIONE E SERVIZI S.R.L.	-0,443574		-0,00769		0,428186		
ITALIANA	-0,35251	F	0,090876	F	0,534264	F	F

COSTRUZIONI S.R.L.							
J.D.S. SPA	-0,669346		-0,25323		0,162889		
JESOLO TURISMO SPA	-0,456945		0,039668		0,53628		
LA PERLA S.R.L.	-1		-0,39152		0,216965		
LA TITANO EDILIZIA S.R.L.	-0,446959		-0,0146		0,417756		
LAVORO E SERVIZI VALSUGANA SOCIETA' COOPERATIVA	-0,666046		-0,25185		0,162356		
LOGISERVICES SOCIETA COOPERATIVA A RL	-0,465454		-0,04344		0,378575		
LOGITRANS SRL	-0,496587		-0,05761		0,381373		
LUMAN C.M.L. S.R.L.	-0,17164	F	0,309969	F	0,791581	F	F
M.A. INTERNATIONAL CATERING SRL	-0,33833	F	0,093954	F	0,526241	F	F
M.E. MARITTIMA EMILIANA S.P.A.	-1		-0,41491		0,170185		
MAINA ORGANI DI TRASMISSIONE S.P.A.	-0,644217		-0,22499		0,194241		
MEDITERRANEAN SHIPPING COMPANY S.R.L.	-0,453465		-0,07393		0,305602		
METRONOTTE ISTITUTO DI VIGILANZA PRIVATA SRL	-0,431658		0,007951		0,447559		
MINOSSE S.R.L.	-1		-0,32931		0,341378		
NAVITAL S.P.A.	-0,49232		-0,07765		0,337011		F
NO PROBLEM PARKING S.P.A.	-0,434542		0,082792		0,600126		
NUOVA FONDERIA DI CASTENEDOLO S.P.A.	-0,619887		-0,18771		0,244476		
NUOVI CANTIERI APUANIA SPA	-1		-0,44897		0,102062		
ORGANIZZAZIONE STRUTTURE CEMENTO ARMATO OSCA PREFABBRICATI SRL	-0,507048		-0,13282		0,241411		
OTTAVI PREFABBRICATI SRL	-0,524933		-0,11426		0,296404		
PAGLIERANI SRL	-0,444841		-0,05163		0,341586		
PARADIVI SERVIZI SRL	-0,563258		-0,11076		0,341747		
PIIONEER HIBRED ITALIA SERVIZI AGRONOMICI SRL	-0,682714		-0,20472		0,273274		
PROGEMA IMPIANTI S.R.L.	-0,35028	F	0,111108	F	0,572495	F	F
PUBBLICITA' S.R.L.	-0,51866		-0,0744		0,369863		
REFITALIA S.R.L.	-0,34637	F	0,081413	F	0,509193	F	F
RMG COSTRUZIONI SRL	-0,64442		-0,1204		0,403619		F
RUSPA OFFICINE SPA	-0,506119		-0,09475		0,316619		
S.I.TEC. S.R.L.	-0,38465	F	0,043533	F	0,471719		F

S.L.R. SERVICE	-0,43819		-0,03043		0,37733		F
SATER SERVIZI AMBIENTALI TERRITORIALI S.P.A.	-0,621313		-0,20298		0,215344		
SC SALDATURA CARPENTERIA SRL	-0,478007		-0,08233		0,313347		
SEA SOLUZIONI ECOLOGICHE AMBIENTALI SRL	-0,425775		0,020316		0,466406		
SEA SRL SERVIZI ECOLOGICI AMBIENTALI	-0,447861		-0,02949		0,388885		
SECURITY SERVICE SISTEMI S.R.L.	-0,475617		-0,04656		0,382504		
SEGHIERA PIAVE S.R.L.	-0,4804		-0,07445		0,3315		
SERVECO SRL	-0,48919		-0,0772		0,334798		
SERVIZI ALLA STRADA SPA	-0,559576		-0,13025		0,299086		
SICILIA POLICE S.R.L.	-0,502661		-0,08275		0,337158		
SIDOTI COSTRUZIONI SRL	-0,455629		-0,01801		0,419615		
SIMONT S.P.A.	-0,49554		-0,07902		0,337492		F
SOCIETA ALGHERESE TURISTICO ALBERGHIERA SRL	-1		-0,4629		0,074197		
SOCIETA ANONIMA CUPRENSE SERVIZI AUTOMOBILISTICI SRL	-0,466693		-0,13442		0,197848		
SOLIDARIETA CINQUE SOC COOP SOCIALE	-0,477504		-0,04908		0,379343		
SORA S.R.L.	-0,42555		-0,04261		0,340325		F
SPREA EDITORI S.P.A.	-0,535394		-0,12084		0,293717		
SPRING S.R.L.	-0,42396		-0,01391		0,396135		F
SRB COSTRUZIONI SRL	-0,477927		-0,0542		0,369532		
STA SERVIZI TECNICI AMMINISTRATIVI SOCA RL	-1		-0,4079		0,184199		
STANADYNE S.P.A.	-0,478409		-0,11296		0,25249		
SYMANTEC SRL	-1		-0,38127		0,237463		
TEKNOSERVICE S.R.L.	-0,459586		-0,03519		0,389205		
TIPOGRAFIA TOSCANA S.R.L. INDUSTRIA GRAFICA	-0,3176	F	0,095017	F	0,507636	F	F
TRANSPOBANK SRL	-0,389941		0,095501		0,580943		
TRASIMENO SERVIZI AMBIENTALI S.P.A.	-0,523898		-0,10205		0,319805		
TRASPORTI ROMA RX S.R.L	-0,487076		-0,05394		0,379206		
TREVISO SINERGIE S.R.L	-0,564158		-0,17064		0,222872		
UNICOOP SERVIZI LIVORNO SOCIETA COOPERATIVA	-0,380941		0,027426		0,435793		
USI ITALIA SRL	-0,479017		-0,05677		0,365484		

VACCARI GIOVANNI S.R.L.	-0,42036		0,001857	F	0,424072		F
VENEZIA INFORMATICA E SISTEMI VENIS SPA	-0,484922		-0,09015		0,304625		
VETRICERAMICI S.P.A.	-0,51662		-0,14338		0,229861		
VETROBALSAMO SPA	-0,453898		0,03795		0,529798		
VILLAIRIS SRL	-0,719661		-0,28866		0,142338		
WORLD STARTEL COMMUNICATIONS S.R.L	-0,28596	F	0,137982	F	0,56192	F	F
ZETA FASHION GROUP S.R.L.	-0,45525		-0,01682		0,421611		F
ZETA-FPA ENGINEERING	-0,41022		-0,01172		0,386787		F
TOTALE IMPRESE CLASSIFICATE FALLITE		12		13		11	22

Tabella B.3: Score di efficienza DEA e numero di imprese fallite per il modello output-oriented (seconda estensione), per diversi valori di λ

Impresa	$\lambda=0$	Imprese classificate fallite (cut-off -0,5)	$\lambda=0,6$	Imprese classificate fallite (cut-off 0,1)	$\lambda=1$	Imprese classificate fallite (cut-off 0,5)	Imprese realmente fallite
A.I.S.A. S.P.A. AREZZO IMPIANTI E SERVIZI AMBIENTALI	-0,731118		-0,1762182		0,193715		
A.T.C. AIR TRANSPORTATION CONSULTANTS S.P.A.	-0,635504		-0,0791684		0,291722		
ACCIAI RIVESTITI VALDARNO S.R.L.	-0,627928		-0,0657946		0,308961		
ACHITEX SRL	-0,65279		-0,1134536		0,246104		
ACQUE ALBULE S.P.A	-0,650042		-0,0642284		0,326314		
ACQUE DI CALTANISSETTA SPA	-0,530863		0,0509468		0,43882		
ADAMANTE SRL	-0,541042		0,0044222		0,368065		
ADLER SPA	-0,599603		-0,0123098		0,379219		
AEMME LINEA DISTRIBUZIONE S.R.L.	-0,500061		0,1488072		0,581386		
AEPI COSTRUZIONI ELETTROMECCANICHE SRL	-0,571229		-0,033467		0,325041		
AGENZIA ROGNONI SRL	-0,541361		0,0832168		0,499602		

ALCHEMIA ITALIA S.R.L.	-0,432363	F	0,4270548	F	1	F	F
ALHAMBRA'S GARDEN SRL	-0,004681	F	0,5981276	F	1	F	F
AMAGA ENERGIA E SERVIZI SRL	-0,552075		-0,019812		0,33503		
ANDREOTTI FOTOINCISIONI S.R.L.	-0,609836		-0,0204158		0,372531		
AUTOSERVIZI MORETTI SRL	-0,910493		-0,250889		0,188847		
AZIENDA PROVINCIALE TRASPORTI SPA	-1		-0,3627484		0,062086		
B.R.B. S.R.L.	-0,417115	F	0,433154	F	1	F	F
BBG S.P.A.	-0,29717	F	0,2998576	F	0,697876	F	F
BIT MARKET SERVICES S.P.A.	-0,658331		-0,1200218		0,238851		
CALCESTRUZZI DELLA VAL ROJA - S.R.L.	-0,836943		-0,238065		0,161187		
CARPEDIL - S.P.A.	-0,733743		-0,1878702		0,176045		
CASA DEL SOLE SPA	-0,60826		-0,0462406		0,328439		
CLEAN SERVICE S.R.L.	-0,589829		-0,0394556		0,32746		
COMPAGNIA PORTUALE T GULLI SRL	-0,597258		-0,0536718		0,308719		
COSMO HOTEL S.P.A.	-0,781501		-0,215719		0,161469		
CPS COMPAGNIA PORTUALI SERVIZI SOCIETA COOPERATIVA	-0,811621		-0,2505766		0,123453		
DATA PROCESSING	-0,755327		-0,1515716		0,250932		
DIMENSIONE TURISMO S.R.L.	-0,550269		-0,000954		0,365256		
DOMUS VESTRA S.P.A.	-0,618966		0,2425896		0,81696		
DRIVE SERVICE SPA	-0,540187		0,017504		0,389298		
ECOLOGICA SRL	-0,628292		-0,0798626		0,285757		
ECOLOGISTICA SERVIZI SRL	-0,545597		-0,0010538		0,361975		
EQUITALIA GIUSTIZIA S.P.A.	-1		-0,3646516		0,058914		
ESDRA COSTRUZIONI S.R.L.	-0,53812		0,0083336		0,372636		
EUGANEA MEDICA S.R.L.	-1		-0,3117898		0,147017		
FERRARI - S.P.A.	-0,719949		-0,1593708		0,214348		
FINARGE ARMAMENTO GENOVESE SRL	-1		0,2		1		
FONDERIA BOCCACCI S.P.A.	-0,739989		-0,1868646		0,181885		
FONDERIE PONTINE BREVETTI CAVIS SRL	-0,616605		-0,0810648		0,275962		
FRIULCO S.P.A.	-0,543573		-0,0037788		0,356084		F
FRIULSIDER S.P.A.	-0,903221		-0,282206		0,131804		
GENNARO COSTRUZIONI SRL	-0,727145		-0,03704		0,42303		

GEOVEST SRL	-0,636401		-0,0770186		0,295903		
GESTIONE ESERCIZI TURISTICI ALBERGHIERI G.E.T.A. S.P.A.	-0,53987		0,384052		1		
GESTIONE SERVIZI INFORMATICI S.P.A.	-1		-0,3503434		0,082761		
GESTIONI AEROPORTI SARDI GEASAR SPA	-0,676551		-0,1242324		0,24398		
HEINLOTH TRANSPORT GMBH & CO. KG	-0,694395		-0,118512		0,26541		
I VIAGGI DEL TURCHESE S.R.L.	-0,839113		-0,225319		0,183877		
IL DELFINO RISTORAZIONE S.R.L.	-1		-0,3992974		0,001171		
IMPRESA VERDE FOGGIA SRL	-0,001192		0,5995232		1		
IMPRESA VERDE ROMAGNA SRL	-0,859501		-0,2057866		0,230023		
ISTITUTO ANALISI MEDICHE SAN LORENZO - S.R.L.	-0,59404		-0,0439138		0,322837		
ITACA RISTORAZIONE E SERVIZI S.R.L.	-0,54957		0,0370836		0,428186		
ITALIANA COSTRUZIONI S.R.L.	-0,512253		0,1156572	F	0,534264	F	F
J.D.S. SPA	-0,799606		-0,222109		0,162889		
JESOLO TURISMO SPA	-0,550576		0,1015376		0,53628		
LA PERLA S.R.L.	-1		-0,269821		0,216965		
LA TITANO EDILIZIA S.R.L.	-0,53935		0,0349136		0,417756		
LAVORO E SERVIZI VALSUGANA SOCIETA' COOPERATIVA	-0,786424		-0,217156		0,162356		
LOGISERVICES SOCIETA COOPERATIVA A RL	-0,558865		0,003599		0,378575		
LOGITRANS SRL	-0,560536		0,0046094		0,381373		
LUMAN C.M.L. S.R.L.	-0,290423	F	0,3601138	F	0,793805	F	F
M.A. INTERNATIONAL CATERING SRL	-0,535024		0,101981	F	0,526651	F	F
M.E. MARITTIMA EMILIANA S.P.A.	-1		-0,297889		0,170185		
MAINA ORGANI DI TRASMISSIONE S.P.A.	-0,784704		-0,197337		0,194241		
MEDITERRANEAN SHIPPING COMPANY S.R.L.	-0,577776		-0,0477492		0,305602		
METRONOTTE ISTITUTO DI VIGILANZA PRIVATA SRL	-0,011441		0,5954236		1		
MINOSSE S.R.L.	-1		-0,1951732		0,341378		

NAVITAL S.P.A.	-0,537087		-0,0126282		0,337011		F
NO PROBLEM PARKING S.P.A.	-0,51343		0,1547036		0,600126		
NUOVA FONDERIA DI CASTENEDOLO S.P.A.	-0,702365		-0,1342604		0,244476		
NUOVI CANTIERI APUANIA SPA	-1		-0,3387628		0,102062		
ORGANIZZAZIONE STRUTTURE CEMENTO ARMATO OSCA PREFABBRICATI SRL	-0,679169		-0,126821		0,241411		
OTTAVI PREFABBRICATI SRL	-0,63245		-0,0751376		0,296404		
PAGLIERANI SRL	-0,636688		-0,0497236		0,341586		
PARADIVI SERVIZI SRL	-0,645421		-0,0531202		0,341747		
PIONEER HIBRED ITALIA SERVIZI AGRONOMICI SRL	-0,813502		-0,1614364		0,273274		
PROGEMA IMPIANTI S.R.L.	-0,593825		0,105967	F	0,572495	F	F
PUBBLICITA' S.R.L.	-0,588304		-0,0134038		0,369863		
REFITALIA S.R.L.	-0,46345	F	0,1234778	F	0,514763	F	F
RMG COSTRUZIONI SRL	-0,821396		-0,086387		0,403619		F
RUSPA OFFICINE SPA	-0,586031		-0,044441		0,316619		
S.I.TEC. S.R.L.	-0,515262		0,077496		0,472668		F
S.L.R. SERVICE	-0,534253		0,0126968		0,37733		F
SATER SERVIZI AMBIENTALI TERRITORIALI S.P.A.	-0,704753		-0,1526948		0,215344		
SC SALDATURA CARPENTERIA SRL	-0,581036		-0,0444062		0,313347		
SEA SOLUZIONI ECOLOGICHE AMBIENTALI SRL	-0,507308		0,0769204		0,466406		
SEA SRL SERVIZI ECOLOGICI AMBIENTALI	-0,526302		0,0228102		0,388885		
SECURITY SERVICE SISTEMI S.R.L.	-0,567279		0,0025908		0,382504		
SEGHERIA PIAVE S.R.L.	-0,627264		-0,0520056		0,3315		
SERVECO SRL	-0,567702		-0,026202		0,334798		
SERVIZI ALLA STRADA SPA	-0,714652		-0,1064092		0,299086		
SICILIA POLICE S.R.L.	-0,561889		-0,0224608		0,337158		
SIDOTI COSTRUZIONI SRL	-0,632401		-0,0011914		0,419615		
SIMONT S.P.A.	-0,610646		-0,0417632		0,337492		F
SOCIETA ALGHERESE TURISTICO ALBERGHIERA SRL	-1		-0,3554818		0,074197		
SOCIETA ANONIMA CUPRENSE SERVIZI AUTOMOBILISTICI	-0,723325		-0,1706212		0,197848		

SRL							
SOLIDARIETA CINQUE SOC COOP SOCIALE	-0,557099		0,0047662		0,379343		
SORA S.R.L.	-1		-0,195805		0,340325		F
SPREA EDITORI S.P.A.	-0,581869		-0,0565174		0,293717		
SPRING S.R.L.	-0,517196		0,0308026		0,396135		F
SRB COSTRUZIONI SRL	-0,544486		0,0039248		0,369532		
STA SERVIZI TECNICI AMMINISTRATIVI SOCA RL	-1		-0,2894806		0,184199		
STANADYNE S.P.A.	-0,644476		-0,1062964		0,25249		
SYMANTEC SRL	-1		-0,2575222		0,237463		
TEKNOSERVICE S.R.L.	-0,53112		0,021075		0,389205		
TIPOGRAFIA TOSCANA S.R.L. INDUSTRIA GRAFICA	-0,425211	F	0,1405086	F	0,517655	F	F
TRANSPOBANK SRL	-0,489546		0,1527474		0,580943		
TRASIMENO SERVIZI AMBIENTALI S.P.A.	-0,5787		-0,039597		0,319805		
TRASPORTI ROMA RX S.R.L.	-0,733041		-0,0656928		0,379206		
TREVISIO SINERGIE S.R.L.	-0,693567		-0,1437036		0,222872		
UNICOOP SERVIZI LIVORNO SOCIETA COOPERATIVA	-0,503402		0,0801874		0,469247		
USI ITALIA SRL	-0,524636		0,009436		0,365484		
VACCARI GIOVANNI S.R.L.	-0,512228		0,049552		0,424072		F
VENEZIA INFORMATICA E SISTEMI VENIS SPA	-0,7017		-0,097905		0,304625		
VETRICERAMICI S.P.A.	-0,724907		-0,1520462		0,229861		
VETROBALSAMO SPA	-0,550714		0,0975932		0,529798		
VILLAIRIS SRL	-0,765878		-0,2209484		0,142338		
WORLD STARTEL COMMUNICATIONS S.R.L .	-0,427862	F	0,1693198	F	0,567441	F	F
ZETA FASHIONGROU P S.R.L.	-0,552097		0,0321278		0,421611		F
ZETA-FPA ENGINEERING	-0,496508	F	0,0341212		0,387874		F
TOTALE IMPRESE CLASSIFICATE FALLITE		9		11		11	22

RIFERIMENTI BIBLIOGRAFICI

Ali I., Seiford L. M. (1990), *Translation invariance in data envelopment analysis*, Operations Research Letters 9, pp.403-405.

Altman Edward I. (1968), *Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy*, Journal of Finance, Vol. 23, No. 4, September, pp. 589-609.

Altman Edward I., Avery R., Eisenbeis R., Sinkey J. (1981), *Application of classification techniques in business, banking and finance*, ed. Jai Press, NY.

Altman Edward I., (2000), *Predicting financial distress of companies: revisiting the z-score and zeta® models*, Stern School of Business, New York University, pp. 9-12.

Altman Edward I. (2005), *An emerging market credit scoring system for corporate bonds*, Emerging Markets Review, Volume 6, Issue 4, December, Pages 311-323.

Beaver W. H. (1966), *Financial ratios as predictors of failure*, Empirical Research in Accounting: Selected Studies 1966, Vol. 4, 41 (71-111).

Beltratti A., Margarita S. and Terna P. (1996), *Neural Networks for economic and financial modeling*, International Thomson Computer Press, London, UK.

Caouette J., Altman E. I. and Narayana P. (1998), *Managing Credit Risk: The Next Great Financial Challenge*, John Wiley & Sons, New York.

Charnes A., Cooper W. W., Rhodes E. (1978), *Measuring the Efficiency of Decision Making Units*, European Journal of Operational Research, Vol. 2, No. 6, pp. 429-444.

Charnes A., Cooper W. W., Rhodes E. (1981), *Evaluating program and managerial efficiency: An application of data envelopment analysis to program follow through*, Management Science 27, pp. 668–697.

Charnes A., Cooper W. W., Golany B., Seiford L. M., Stutz J. (1985), *Foundations of data envelopment analysis and Pareto–Koopmans empirical production functions*, Journal of Econometrics 30, 91–107.

Cielen A., Peeters L., Vanhoof K. (2004), *Bankruptcy prediction using a data envelopment analysis*, European Journal of Operational Research, 154, pp. 526-532.

Collins R. A., Green R. D. (1982), *Statistical Methods for Bankruptcy Forecasting*, Journal of Economics and Business 34, pp.349-354.

Cook Wade D., Seiford L. M. (2009), *Data envelopment analysis (DEA) – Thirty years on*, European Journal of Operational Research 192, pp. 1–17.

Darwin Charles (1859), *The Origin of Species by means of natural selection, or the preservation of favoured races in the struggle for life*, London.

Debreu Gerard (1951), *The Coefficient of Resource Utilization*, Econometria, 19, 3, pp. 273-292.

Eisenbeis Robert A. (1977), *Pitfalls in the application of Discriminant Analysis in Business, Finance and Economics*, The Journal of Finance, Vol.32 no.3, pp.876-900.

Farrel M. J. (1957), *The Measurement of Productive Efficiency*, Journal of the Royal Statistical Society, Vol. 120, No. 3, pp. 253-290.

Fischer R. A. (1936), *The Use Of Multiple Measurement In Taxonomic Problems*, Annals of Eugenics, V. 7, p. 179-188.

Freed Ned, Glover Fred (1981), *Simple but powerful goal programming models for discriminant problems*, European Journal of Operational Research 7, pp.44-60.

Grice J. S., Ingram R. W. (2001), *Tests of the generalizability of Altman's bankruptcy prediction model*, Journal of Business Research 54, pp.53-61.

Hebb Donald O. (1949), *The organization of behavior: a neuropsychological theory*, Wiley, New York.

Holland John H. (1975), *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control and artificial intelligence*, University of Michigan Press.

Kahya E., Theodossiou P. (1998), *Predicting corporate financial distress: A time-series CUSUM methodology*, Working Paper.

Keasey K., Watson R. (1987), *Non financial symptoms and the prediction of small company failure: a test of Argenti's Hypotheses*, Journal of Business Finance & Accounting 14(3), pp.335-354.

Lindsay D. H., Campbell A. (1996), *A Chaos Approach to Bankruptcy Prediction*, Journal of Applied Business Research, Volume 12, Number 4.

Maddala G. (1983), *Limited-dependent and qualitative variables in econometrics*, ed. Cambridge Univ. Press, Cambridge.

Maddala G. (1992), *Introduction to econometrics*, 2a edizione, ed. MacMillan, NY.

Ohlson James A. (1980), *Financial Ratios and the Probabilistic Prediction of Bankruptcy*, Journal of Accounting Research, Vol.18 no.1, pp109-131.

Orgler Y. E. (1970), *A Credit Scoring Model for Commercial Loans*, Journal of Money, Credit & Banking (Ohio State University Press) 2, 435-445.

Pomante U. (1998), *I modelli basati sugli algoritmi genetici*, in M. Marsella - A. Sironi (a cura di), *La misurazione e la gestione dei rischi di credito: modelli, strumenti e politiche*, Bancaria Editrice, Roma.

Premachandra I. M., Bhabra G. S., Sueyoshi T. (2009), *DEA as a tool for bankruptcy assessment: A comparative study with logistic regression technique*, European Journal of Operational Research 193, pp.412-424.

Premachandra I. M., Chen Y., Watson J. (2011), *Dea as a tool for predicting corporate failure and success: A case of bankruptcy assessment*, Omega 39, pp.620-626.

Resti Andrea e Sironi Andrea (2008), *Rischio e valore nelle banche – Misura, regolamentazione, gestione*, EGEA, Milano.

Simak Paul C. (1997), *DEA Based Analysis of Corporate Failure*, Working Paper, Department of Mechanical and Industrial Engineering, University of Toronto.

Simak Paul C. (2000), *Inverse and Negative DEA and their Application to Credit Risk Evaluation*, Tesi di PhD, Department of Mechanical and Industrial Engineering, University of Toronto.

Stanghellini E., *Introduzione ai metodi statistici per il credit scoring*, Springer-Verlag Italia 2009, Milano.

Szego G. e Varetto F. (1999), *Il rischio Creditizio: misura e controllo*, ed. Utet, Torino.

Thomas Lyn C. (2000), *A Survey of Credit and Behavioural Scoring: Forecasting Financial Risk of Lending to Consumers*, 16(2) International Journal of Forecasting, 149-172.

Tone K. (2001), *A slacks – based measure of efficiency in Data Envelopment Analysis*, European Journal of Operational Research 130, pp. 498-509.

Vitalie B. & Ashta A., *The conceptual framework of credit scoring from its origins to microfinance*, Second European Research Conference on Microfinance, Groningen, The Netherlands, June 2011.

Zavgren Christine V. (1985), *Assessing the vulnerability to failure of American industrial firms: a Logistic Analysis*, Journal of Business Finance & Accounting 12, pp.19-45.

SITOGRAFIA

<http://www.bancaditalia.it>

<http://www.bancaditalia.it>, *Questioni di Economia e Finanza*, Occasional papers, Aprile 2008

<http://www.bis.org>, Bank for International Settlements

<http://www.confcommercio.it>

<http://aida.bvdep.com>

<http://www.istat.it>

<http://www.mokabyte.it>

RINGRAZIAMENTI

Giunto alla conclusione del mio percorso universitario, vorrei ringraziare la prof.ssa Antonella Basso per l'attenzione e il supporto concessomi durante l'intera stesura di questo lavoro.

Un grande ringraziamento è rivolto alla mia famiglia e a Carla, punti cardine della mia vita. È grazie a loro che sono quello che sono.

Infine, ringrazio tutte le persone che mi hanno accompagnato e sostenuto in questo lungo cammino. Grazie di cuore a tutti.