



Università
Ca' Foscari
Venezia

Corso di Laurea magistrale
in Economia e Finanza

Tesi di Laurea

L'intelligenza artificiale per la previsione degli indici azionari

Relatore

Prof. Claudio Pizzi

Laureanda

Giada Schiochet

Matricola 848355

Anno Accademico

2018 / 2019

*Tu ci hai insegnato tante cose,
ma non ci hai insegnato
cosa faremo senza di te.*

A mio nonno Bruno.

Sommario

Introduzione	ii
CAPITOLO I – LA STORIA DELL’INTELLIGENZA ARTIFICIALE	1
1.1 Dal sillogismo di Aristotele al sogno di Leibniz	1
1.2 Dal linguaggio proposizionale a quello dei predicati	5
1.3 La nascita dell’intelligenza artificiale	8
1.4 L’intelligenza artificiale contemporanea	16
1.5 Dall’automazione all’intelligenza artificiale debole	23
1.6 Dall’intelligenza artificiale debole all’intelligenza artificiale forte	24
CAPITOLO II – FUNZIONAMENTO DELL’INTELLIGENZA ARTIFICIALE	29
2.1 Modalità di apprendimento supervisionato	31
2.1.1 Tipologie di apprendimento supervisionato	31
2.2 Modalità di apprendimento non supervisionato	45
2.2.1 Tipologie di apprendimento non supervisionato	45
2.3 Modalità di apprendimento per rinforzo	53
CAPITOLO III – LE RETI NEURALI ARTIFICIALI: LO STATO DELL’ARTE	57
3.1 Le reti neurali artificiali	57
3.2 La relativa letteratura: lo stato dell’arte	68
CAPITOLO IV – GLI AMBITI APPLICATIVI DELL’INTELLIGENZA ARTIFICIALE	81
CAPITOLO V – IMPATTI DELL’INTELLIGENZA ARTIFICIALE ED EVOLUZIONE FUTURA	99
CAPITOLO VI – IMPLEMENTAZIONE DI UNA RETE NEURALE ARTIFICIALE PER LA PREVISIONE FINANZIARIA	109
6.1 L’indice FTSE MIB. Cenni storici	109
6.1.1 Composizione e metodologia di calcolo dell’indice FTSE MIB	112
6.1.2 Le reti neurali artificiali in R	115
6.2 L’indice Hang Seng. Cenni storici.	125
Conclusioni	135
APPENDICE A. SCRIPT RETE NEURALE FEED – FORWARD FTSE MIB	137
APPENDICE B. SCRIPT RETE NEURALE FEED – FORWARD FTSE MIB – HANG SENG	143
APPENDICE C. SCRIPT RETE NEURALE FEED – FORWARD FTSE MIB – HANG SENG CON RIDUZIONE DELLE VARIABILI DI INPUT	151
Bibliografia e sitografia	159
Ringraziamenti	169

Introduzione

La capacità di prevedere le dinamiche degli strumenti finanziari ha rappresentato, e rappresenta tuttora, una delle principali sfide nella comunità dell'intelligenza artificiale.

La previsione è quel processo che permette di stimare qualcosa di futuro sulla base di dati passati e presenti, o di analisi delle tendenze, guidando così molte decisioni importanti nei più svariati ambiti tra cui quello economico – finanziario.

Per molto tempo, a fine predittivo, sono stati utilizzati modelli lineari ovvero strumenti efficaci ed efficienti per elaborare e risolvere dei problemi che richiedevano la semplice ripetizione di una serie di operazioni ben definite, ma, in un contesto quale il mercato finanziario definito come un ambiente caratterizzato da un'ampia mole di informazioni eterogenee e influenzato da molteplici fattori, si è presentata la necessità di utilizzare dei modelli in grado di cogliere le diverse correlazioni tra i dati che costituiscono problemi non lineari: proprio per questo, le tecniche di apprendimento automatico sono divenute nel tempo, molto importanti nella modellizzazione finanziaria.

Tale studio, inizia presentando la nascita e l'evoluzione dell'intelligenza artificiale: dal filosofo greco Aristotele, il quale formulò un insieme preciso di leggi che governavano la parte razionale della mente e con il quale riuscì a spiegare il ragionamento logico deduttivo tramite un sistema informale di sillogismi, fino ad arrivare al più recente successo in tale ambito, ovvero l'algoritmo *Deep Mind* ideato da Google e applicato al software *Alphago*. Questo algoritmo permette alla macchina di imparare da sé stessa attraverso l'esperienza.

Successivamente tale elaborato presenta il funzionamento dell'intelligenza artificiale, vale a dire il modo in cui una macchina può apprendere e di conseguenza risolvere un determinato problema; sono state analizzate le principali modalità di apprendimento, cioè l'apprendimento per rinforzo, non supervisionato e supervisionato, esplicitando per ognuno le principali tipologie utilizzate con maggiore frequenza.

Successivamente il lavoro si concentra sull'utilizzo delle reti neurali artificiali in ambito finanziario. Dopo aver presentato una panoramica generale sulla loro struttura e sul loro funzionamento, è stata analizzata la letteratura attinente al loro utilizzo a fine previsivo in ambito finanziario degli ultimi due decenni: essa, restituendo dei risultati molto

promettenti in merito, rende ottimisti sul fatto che, grazie alla continua evoluzione tecnologia e disponibilità crescente di dati, il trend degli indici azionari possa essere previsto ex -ante grazie alle informazioni storiche.

In particolar modo, l'oggetto dello studio è rappresentato dall'utilizzo delle reti neurali artificiali *feed – forward* per la previsione a un giorno dell'indice azionario FTSE MIB italiano sulla base dei dati storici dell'ultimo decennio i quali comprendono il prezzo di chiusura, il prezzo di apertura, il prezzo massimo e minimo verificatosi giornalmente, i volumi negoziati e la relativa variazione percentuale. L'implementazione e l'addestramento della rete neurale scelta, nonché la previsione dell'indice azionario selezionato, sono avvenute con il l'utilizzo di R, un linguaggio di programmazione per l'elaborazione statistica e grafica.

Dopo aver evidenziato i vari ambiti nei quali, attualmente, si ricorre all'utilizzo dell'intelligenza artificiale e quali impatti ha avuto e può avere in futuro la sua continua evoluzione, si è proceduto all'implementazione di una seconda rete neurale *feed – forward* al fine di verificare l'eventuale presenza, e la sua relativa incidenza, della correlazione tra l'andamento dell'indice italiano FTSE MIB e l'indice asiatico HANG SENG; in particolar modo il seguente elaborato mira a verificare se il prezzo di apertura dell'indice asiatico al tempo t , unitamente al prezzo di chiusura, di apertura, il prezzo massimo e minimo al tempo $t-3$ dell'indice italiano, hanno qualche influenza sul prezzo di chiusura al tempo t di quest'ultimo.

CAPITOLO I – LA STORIA DELL’INTELLIGENZA ARTIFICIALE

1.1 Dal sillogismo di Aristotele al sogno di Leibniz

L’intelligenza artificiale è una disciplina di studio, divenuta una vera e propria scienza nel tempo, le cui origini sono molto antiche, infatti i primi studi attinenti alla materia, anche se indiretti, furono compiuti dai filosofi dell’antica Grecia.

Benché già Platone e i pre - socratici utilizzassero tecniche di ragionamento come le dimostrazioni per assurdo o la dialettica, solo nel IV secolo a. C. il filosofo greco Aristotele formulò un insieme preciso di leggi che governavano la parte razionale della mente, riuscendo così a spiegare il ragionamento logico deduttivo tramite un sistema informale di sillogismi che consentiva a chiunque, date le premesse iniziali, di generare meccanicamente le conclusioni¹. È un tipo di inferenza costituita da almeno due premesse, accomunate da una proprietà e una conclusione: “gli uomini sono animali, gli animali sono mortali, quindi gli uomini sono mortali”².

Non è scontato, tuttavia, che la veridicità delle premesse implichi quella della conclusione, dal momento che le inferenze possono risultare corrette anche qualora le conclusioni siano false.

Tale tipologia di inferenza divenne oggetto di interesse in quanto venne studiata all’interno della disciplina della logica, la quale si occupa del funzionamento delle tecniche di ragionamento, nonché delle facoltà deduttive degli esseri umani.

I sillogismi si organizzano in quattro forme fondamentali: l’universale affermativa “Tutti i P sono Q”, l’universale negativa “Tutti i P non sono Q”, la particolare affermativa “Qualche P è Q” e la particolare negativa “Qualche P non è Q”, nelle quali P e Q sono due elementi detti proprietà³.

La correttezza, nonché la veridicità dei sillogismi, può essere verificata tramite tecniche specifiche che permettono di esplicitare i legami che intercorrono tra il meccanismo

¹ Cordeschi R. e Tamburrini G., (2001), *L’intelligenza artificiale: la storia e le idee*, Carocci, Roma.

² Russell S. e Norving P., (2010), *Intelligenza artificiale. Un approccio moderno*, Pearson, Milano.

³ Aristotele, (2016), “*Analitici Primi*”, Giorgio Colli (a cura di), *Organon*, Giulio Einaudi Editore, Torino, pp. 105 – 110.

inferenziale sillogistico e la teoria degli insiemi, la quale considera che l'insieme dei P è sottoinsieme dell'insieme dei Q per la prima proposizione, l'intersezione dell'insieme dei P e dell'insieme dei Q è vuota per la seconda proposizione, l'intersezione dell'insieme dei P e dell'insieme dei Q non è vuota per la terza proposizione e infine che l'insieme dei P non è incluso nell'insieme dei Q per l'ultima proposizione⁴.

Utilizzando il famoso quadrilatero aristotelico per evidenziare i rapporti esistenti tra le diverse proposizioni, si rappresenta la verifica apportata dalla teoria degli insiemi:

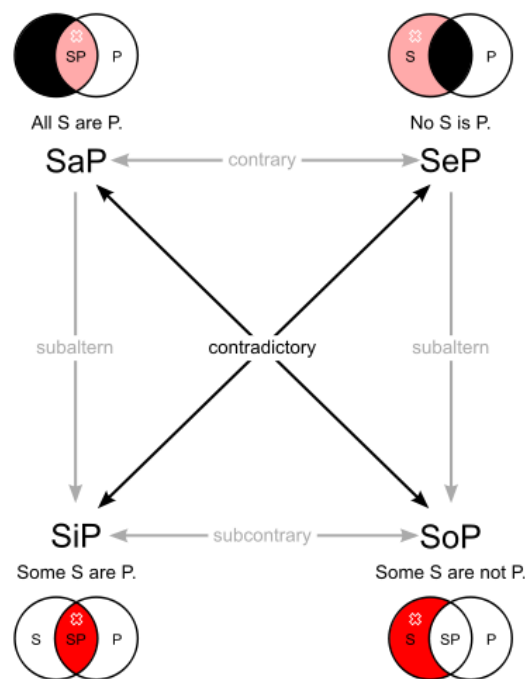


FIGURA 1: RAPPRESENTAZIONE GRAFICA DEL QUADRILATERO ARISTOTELICO. FONTE: [HTTPS://IT.WIKIPEDIA.ORG/WIKI/QUADRATO DELLE OPPOSIZIONI](https://it.wikipedia.org/wiki/Quadrato_delle_opposizioni)

Aristotele ha costituito così, la prima forma di organizzazione categoriale che porta a una sistematizzazione della disciplina logica, detta anche logica analitica. Egli considerava innanzitutto la struttura del ragionamento in quanto tale, motivo per cui, come esposto pocanzi, la conclusione dei sillogismi può risultare anche falsa. Da un punto di vista di ricerca del vero, la questione cambia in quanto si manifesta l'assoluta necessità di premesse vere, "mentre ciò non è necessario che si verifichi nel sillogismo come tale, poiché in quest'ultimo interessa determinare solo se un certo conseguente

⁴ Aristotele, (2016), "Analitici Primi", Giorgio Colli (a cura di), *Organon*, Giulio Einaudi Editore, Torino, pp. 105 – 110.

segue o non segue dalle premesse poste, per il solo fatto che esse sono poste, indipendentemente dal valore di verità che possono avere”⁵.

La trattazione del sillogismo è contenuta interamente negli *Analitici* e più specificatamente negli *Analitici Primi*, opera dello stesso Aristotele. Egli analizza in maniera ordinata e specifica la definizione di premessa e termine per poi concludere con la definizione di sillogismo.

Per “premesse” egli intende il discorso, la proposizione che si può presentare in forma universale o particolare, positiva o negativa, e può essere dimostrativa o dialettica, cioè che può considerare qualcosa vero o porsi una domanda in termini di contraddizione.

Il termine, è “ciò in cui si scompone la premessa” e infine definisce il sillogismo come: *“... un discorso in cui, poste certe cose, qualcosa di diverso rispetto ai dati risulta di necessità per il fatto che sono questi. Quando dico “per il fatto che sono questi” intendo che <esso> risulta a causa di quelli, e quando dico “risulta a causa di quelli” intendo che non c’è bisogno di alcun termine preso dall’esterno perché la necessità <del risultato> venga ad esserci. Quindi, da un lato chiamo sillogismo perfetto quello che non ha bisogno di null’altro oltre agli assunti perché la necessità <del risultato> si manifesti; dall’altro, chiamo imperfetto quello che invece ha bisogno di una o più cose che sono sì necessarie in ragione dei termini dati di base, ma che non sono espressamente assunte con le premesse. Infine, dire che una cosa è in un’altra come in un intero e dire che l’una “è predicata di ogni” rispetto all’altra è la stessa cosa. E diciamo “... è predicato di ogni ...” quando non è possibile assumere nulla del soggetto di cui non sarà detto l’altro termine; lo stesso vale per “... non è predicato di nessun ...”⁶.*

Benché Aristotele non cercasse di automatizzare il ragionamento, contribuì in modo considerevole alla nascita di una riflessione sui principi della validità, strumento che permette di creare nuova conoscenza anticipando il funzionamento dei procedimenti matematici moderni⁷.

La teoria aristotelica stimolò così l’interesse di molti studiosi tra cui quello del giovane Leibniz, il quale dopo aver conseguito il titolo di baccelliere con una tesi sulla metafisica

⁵ M. Mignucci, *La teoria aristotelica della scienza*, Sansoni, Firenze 1965, pp. 110 - 111.

⁶ Aristotele, *Analitici Primi in Organon*, Maurizio Migliori (a cura di), Bompiani, Milano 2016, pp. 375 - 377.

⁷ Davis M., (2000), *The Universal Computer. The Road from Leibniz to Turing*, Adelphi Edizioni, Milano.

aristotelica, durante una delle sue visite a Londra nel 1673, presentò un modello di macchina calcolatrice in grado di eseguire le quattro operazioni aritmetiche fondamentali. Essa si caratterizzava da un funzionamento acuto utilizzato fino alla metà del Novecento nelle macchine calcolatrici, ovvero, la ruota di Leibniz, un meccanismo composto da nove denti di lunghezza crescente abbinato a una ruota dentata che poteva scorrere lungo la direzione del suo asse trovandosi in una posizione relativa differente rispetto al numero di denti ingranati, il quale ovviamente era variabile.

Risultato considerevole per l'epoca, in quanto una macchina ora, era in grado di compiere le operazioni elementari di aritmetica ordinaria, ma l'ambizione di Leibniz lo aveva portato a capire che tale meccanizzazione poteva comportare dei risultati dal significato molto più ampio. L'anno dopo infatti, egli descrisse una macchina in grado di portare alla risoluzione di equazioni algebriche a seguito della quale, fece un raffronto tra ragionamento logico e meccanismo: ridurre ogni ragionamento a una sorta di calcolo e costruire da ultimo, una macchina capace di calcolare in questo senso generale⁸.

Il suo contributo più importante, purtroppo lasciato incompiuto a causa di tutte le possibili metodologie arbitrarie di studio per prevedere e spiegare i risultati correlati al comportamento umano, fu il progetto che avrebbe portato a definire un insieme di regole per formalizzare il ragionamento logico deduttivo, infatti per Leibniz la cosa più importante era riuscire a identificare dei simboli adatti a trovare regole che ne governassero la manipolazione, cioè un vero e proprio alfabeto del pensiero umano e relative regole: il cosiddetto sogno del *calculus ratiocinator* o più comunemente chiamato logica simbolica, che anticipa la tecnologia predittiva attuale⁹.

Egli, dopo aver svolto numerose considerazioni sull'importanza e sui vantaggi che portavano i simboli utilizzati in discipline limitrofe quali l'algebra, l'aritmetica o perfino la chimica e l'astronomia, nelle quali ogni singolo simbolo esprimeva un concetto ben definito, compì uno dei passi più importanti: rappresentò i numeri binari utilizzando un alfabeto idoneo il quale comprendeva le cifre 0 e 1 esplicitando inoltre le regole dell'aritmetica binaria.

⁸ Davis M., (2000), *The Universal Computer. The Road from Leibniz to Turing*, Adelphi Edizioni, Milano.

⁹ Ibidem.

1.2 Dal linguaggio proposizionale a quello dei predicati

La maggior parte delle astrazioni riguardanti l'intelligenza artificiale sono state studiate ed elaborate da filosofi, ma il passaggio a una scienza formale richiedeva un maggior livello di formalizzazione matematica, in particolar modo in tre aree fondamentali¹⁰.

Come già esplicitato, l'idea di logica formale può essere fatta risalire ai filosofi dell'antica Grecia, ma il suo sviluppo in termini matematici ebbe effettivamente inizio nel XIX secolo con lo studio di George Boole, matematico e logico britannico considerato il fondatore della logica matematica, il quale cercò di ideare un sistema deduttivo basato sulla logica proposizionale, utilizzabile nella pratica e avente la stessa impostazione di quella leibniziana. Egli infatti, nella maggior parte dei suoi scritti giovanili dimostra ammirazione nei confronti di Leibniz sostenendo anch'egli la capacità, di un ben definito formalismo matematico, di produrre automaticamente la risposta corretta a ogni problema¹¹.

La logica di Aristotele studiava enunciati come: tutte le piante sono viventi, nessun ippopotamo è intelligente o alcune persone parlano inglese; Boole considerava il modello aristotelico, purché caratterizzato da un importante fascino, incapace di sorreggere l'impalcatura della logica stessa e comprese che, ai fini del ragionamento logico, l'aspetto significativo di parole come "vivente", "ippopotamo" o "persona" era la classe o collezione di tutti gli individui descritti dalla parola stessa, ovvero la classe degli esseri viventi, la classe degli ippopotami o la classe delle persone e così ricorse all'applicazione del calcolo simbolico alla logica con un metodo in grado di formulare tali ragionamenti per mezzo di un'algebra di tali classi¹².

Egli indicava le classi mediante lettere: << ...se come termine di una descrizione usiamo un aggettivo, per esempio buono, con una lettera y, rappresenteremo tutte quelle cose alle quali può applicarsi la descrizione buono, vale a dire tutte le cose buone, o la classe cose buone. Conveniamo inoltre di rappresentare con la combinazione xy la classe di

¹⁰Russell S. e Norving P., (2005), *Intelligenza artificiale. Un approccio moderno*, Pearson Edition Italia, Milano.

¹¹ Davis M., (2000), *The Universal Computer. The Road from Leibniz to Turing*, Adelphi Edizioni, Milano.

¹² Ibidem.

cose a cui sono applicabili, simultaneamente, i nomi o le descrizioni rappresentati da x e da y . Così se x da solo sta per cose bianche, e y sta per animali, xy starà per animali bianchi...>> (Boole, 1854-1958).

Nelle intenzioni di Boole la notazione xy doveva suggerire un'analogia con la moltiplicazione dell'algebra ordinaria, oggi denominata intersezione di x e y . Un ulteriore progresso ottenuto da Boole fu dimostrare che l'equazione $xx=x$ è sempre vera quando x è uguale a 0 o a 1 arrivando al principio che l'algebra della logica è esattamente ciò che sarebbe l'algebra ordinaria se fosse limitata a due soli valori, 0 e 1 i quali possono essere reinterpretati come classi: $0x$ sarà identica a 0 per ogni x , si interpreta 0 come una classe alla quale non appartiene nulla e $1x$ sarà invece identica a x per ogni x se 1 contiene tutti gli oggetti che si considerano¹³, dimostrando così, con grande soddisfazione, che la deduzione logica poteva essere considerata e studiata come un ramo della matematica. Proprio a tal fine, Boole, fautore della teoria delle algebre, conosciuta ancora oggi come algebra booleana, attribuì alla logica un ambito matematico algebrico, dedicandosi inoltre alle equazioni differenziali, differenze finite e a problemi generali del calcolo delle probabilità.

Circa un decennio dopo, Gottlob Frege, matematico, logico e filosofo tedesco, padre della moderna logica matematica e poi proclamato padre della filosofia analitica, estese la logica di Boole in modo da includere tutte le inferenze deduttive applicate nella pratica matematica, ideando così, per la prima volta nella storia, la logica di primo ordine. Detta anche dei predicati, essa si basa sulla logica booleana la quale, a sua volta, poggiava sull'algebra ordinaria: Frege, che voleva reinterpretare tanto l'algebra quanto gli altri segmenti della matematica come sovrastrutture fondate sulla sua logica, riteneva indispensabile invece l'introduzione di simboli speciali che rappresentassero le relazioni logiche¹⁴.

Mentre per Boole le correlazioni fra proposizioni potevano essere espresse a loro volta da proposizioni secondarie, Frege apprese ed esplicitò che tali correlazioni possono essere utilizzate anche per studiare la struttura di una singola proposizione: fondamento alla base della sua logica e della logica moderna. Frege, per esempio, analizza

¹³Davis M., (2000), *The Universal Computer. The Road from Leibniz to Turing*, Adelphi Edizioni, Milano.

¹⁴ Ibidem.

l'enunciato: "tutti i cavalli sono mammiferi" usando la relazione logica se...allora...: se x è un cavallo, allora x è un mammifero. L'enunciato "alcuni cavalli sono purosangue" viene analizzato, analogamente, usando la relazione logica...e...: x è un cavallo e x è un purosangue¹⁵.

Pur essendo che in entrambi gli enunciati viene utilizzata la lettera x , essa viene utilizzata in modo diverso; nel primo viene utilizzata per far capire che la cosa sostenuta è vera qualunque sia x , cioè per ogni x , mentre nel secondo viene utilizzata solo per rappresentare qualche x . Nel simbolismo corrente il simbolo \forall esprime "per ogni x " ovvero la quantificazione universale, mentre il simbolo \exists indica "per qualche x " quindi la quantificazione esistenziale: i due enunciati assumeranno rispettivamente la forma $\forall x$ e $\exists x$. Per la relazione se...allora, si usa in genere il simbolo \supset , mentre per la relazione ...e... il simbolo \wedge . I due enunciati ora potrebbero essere rappresentati molto più stenograficamente come: $(\forall x) (c(x) \supset m(x))$ e $(\exists x) (c(x) \wedge p(x))$.

Sembrerebbe quindi che Frege non stesse solamente studiando un trattamento matematico della logica, ma cercasse di elaborare un nuovo linguaggio dimostrando che tutta la matematica poteva basarsi sulla logica e affinché tale dimostrazione fosse convincente, egli doveva trovare il modo di sviluppare la sua logica senza usare la logica stessa. La soluzione si manifestò con un linguaggio artificiale caratterizzato da precise e rigorose regole sintattiche, detto *Begriffsschrift* e grazie al quale divenne possibile rappresentare le inferenze logiche come operazioni meccaniche guidate da regole di inferenza sulle quali si basavano le configurazioni con cui erano disposti i simboli. Fu il primo esempio di linguaggio formale artificiale¹⁶, benché avesse un limite: mentre Leibniz aveva sognato un linguaggio rappresentato da uno strumento di calcolo efficiente che permettesse di risolvere inferenze logiche solo attraverso la manipolazione dei simboli, tale sistema implicava a tutte le deduzioni, tranne ovviamente quelle elementari, una complessità insostenibile.

La seconda formalizzazione matematica, che permise il passaggio dell'intelligenza artificiale a una scienza formale, riguardò la logica della computazione: l'atto del computare presuppone sempre un'interpretazione dell'identità e ciò implica qualcosa

¹⁵ Davis M., (2000), *The Universal Computer. The Road from Leibniz to Turing*, Adelphi Edizioni, Milano.

¹⁶ Ibidem.

di interpretabile; computabile è anzitutto un tipo di oggetto inesistente e tecnico costruito sull'interazione. Tale oggetto è il numero¹⁷.

Il terzo contributo, oltre alla logica e alla teoria della computazione, è stata la teoria della probabilità. Gerolamo Cardano fu il primo a formulare il concetto di probabilità, il quale divenne parte integrante di tutte le scienze quantitative, rendendo più facilmente gestibili misurazioni incerte e teorie incomplete. Basandosi su tale concetto, circa 200 anni dopo, Thomas Bayes propose una regola capace di aggiornare i valori di probabilità in base ai dati via via raccolti. Ciò che ne risultò, vale a dire l'analisi bayesiana, è tuttora alla base della maggior parte degli approcci moderni del ragionamento in condizioni di incertezza¹⁸.

Ci furono, in questi anni, anche dei tentativi concreti di costruzione di macchine intelligenti, ma la tecnologia ottocentesca fondata ancora sulla meccanica risultò inadeguata, quindi furono questi approcci che influenzarono in modo rilevante gli studi nel XX secolo.

1.3 La nascita dell'intelligenza artificiale

Come esposto nei paragrafi precedenti, molti studiosi tra cui Aristotele, Leibniz, Boole e Frege posero le basi teoriche e applicative dell'intelligenza artificiale: Leibniz aspirava a una ragione umana ricondotta al puro calcolo e grandi macchine meccaniche che fossero in grado di eseguire i calcoli mentre Frege ideò un sistema di regole in grado di dimostrare tutti i ragionamenti deduttivi umani estendendo la logica booleana. Tutto ciò aveva come unico scopo quello della ricerca di un algoritmo, ovvero di una procedura in grado, non solo di manipolare le semplici espressioni algebriche e risolvere le equazioni, ma di minimizzare tutti i ragionamenti deduttivi umani a calcolo primitivo¹⁹. Nel 1950 Alan Turing cominciò a interrogarsi sul modo con cui si poteva dimostrare l'irragionevolezza dell'esistenza di una simile tipologia di algoritmo. Egli infatti sosteneva

¹⁷Possati L., (2017), *Osservazioni sul numero e sulla computabilità*, Hal, 2017.

¹⁸Russell S. e Norving P., (2005), *Intelligenza artificiale. Un approccio moderno*, Pearson Edition Italia, Milano.

¹⁹ Davis M., (2000), *The Universal Computer. The Road from Leibniz to Turing*, Adelphi Edizioni, Milano.

che un algoritmo fosse una procedura definita da una serie di regole che una persona potesse seguire in modo minuzioso e meccanico, ma eliminando conseguentemente i dettagli non rilevanti, riuscì a dimostrare che una persona era limitata, ovvero era in grado di compiere solamente poche azioni di base senza che il risultato finale cambiasse: da ciò comprese che una macchina era in grado di sostituire un essere umano e svolgere quelle stesse azioni basilari. Egli, in questi stessi anni, pubblicò uno dei primi articoli attinenti tale disciplina, introducendo un metodo di verifica del grado di intelligenza di una macchina, l'apprendimento automatico, gli algoritmi genetici e l'apprendimento per rinforzo. Inizialmente Turing aveva definito un modello generale astratto di calcolatore, il quale riceveva il suo input sottoforma di una stringa di simboli su un nastro, e le uniche operazioni che era in grado di attuare erano quelle di far scorrere il nastro in entrambe le direzioni, leggere la cella del nastro su cui era posizionata, e scrivere tale cella. Non erano stati imposti limiti alla lunghezza del nastro, ma il limite molto importante che presentava era che permetteva di capire solamente se la stringa in input avesse una certa forma. Per superare tale limite, era necessario costruire una macchina intelligente, cioè avente un comportamento che, visto da un umano, fosse giudicato simile a quello di un altro umano²⁰. Idea che esemplificò nel 1950 tramite il cosiddetto Test di Turing.



FIGURA 2: UN ESEMPIO DI MACCHINA DI TURING. FONTE: WIKIPEDIA, L'ENCICLOPEDIA LIBERA.

L'idea di base del test è che un osservatore impegnato in una conversazione “non faccia a faccia”, sia con il sistema di elaborazione dei simboli sia con un essere umano, interagendo solo per mezzo di un testo, può distinguere l'uno dall'altro. Come teorizzato

²⁰ A. Turing, *Computing Machinery and Intelligence*, Mind 1950, pp. 433 - 460.

da Turing, se il sistema è in grado di simulare in modo così efficace la risposta umana da impedire al 30% degli interlocutori di discernere in modo affidabile, durante un test di cinque minuti, se hanno a che fare con un computer o una persona, allora il sistema supera il test. Secondo Turing, qualsiasi sistema artificiale in grado di superare il test dovrebbe essere considerato capace di intelligenza umana²¹. Un approccio che si potrebbe considerare behaviorista: egli esaminava il comportamento, l'esterno e le risposte osservabili dal sistema piuttosto che chiedersi cosa accadeva all'interno dello stesso.

Suddetto test è considerato ancora oggi un benchmark che permette di decretare l'avvenuto raggiungimento di un livello intellettuale umano da parte di una macchina. Con ciò, Turing non implicava l'idea della macchina divenuta umano, ma molto più realisticamente quella della macchina agente come se fosse un umano²². Il computer infatti avrebbe dovuto possedere delle capacità come l'interpretazione del linguaggio naturale al fine di comunicare con l'esaminatore nel suo linguaggio umano, rappresentare la conoscenza per memorizzarla, ragionare in maniera automatica per utilizzare la conoscenza acquisita in modo da rispondere alle domande e giungere a nuove conclusioni e apprendere per adattarsi a nuove situazioni individuando ed estrapolando così nuovi schemi e modelli²³.

La data della nascita dell'intelligenza artificiale, ufficialmente riconosciuta e coincidente con il raggiungimento di un ragionevole sviluppo delle tecnologie disponibili, parallelamente alle basi teoriche, è il 1956, anno in cui si tenne presso il Dartmouth College di Hannover nel New Hampshire, un famoso seminario durante il quale la nuova disciplina venne fondata programmaticamente e identificata come *Artificial Intelligence* (A. I.) da Marvin Minsky, insieme a Claude Shannon, John McCarthy, Herbert Simon e Nathaniel Rochester: primo gruppo di ricerca ufficiale sulle A. I.

Si partiva dall'assunto che: *“ogni aspetto dell'apprendimento e delle capacità cognitive in generale, sarebbe stato suscettibile di una descrizione accurata; sulla base di questa,*

²¹ D. Matsumoto, *The Cambridge Dictionary of Psychology*, Cambridge 2009, p. 555.

²² B. Lenzi, Passato, (2012), “Presente e futuro dell'intelligenza artificiale”, *Divenire*, 5, p. 235.

²³ Russell S. e Norving P., (2005), *Intelligenza artificiale. Un approccio moderno*, Pearson Edition Italia, Milano.

*una macchina capace di svolgere i medesimi compiti, sarebbe stata progettabile e realizzabile*²⁴.

Minsky, economista statunitense e noto per la sua teoria dell'instabilità finanziaria e delle cause delle crisi dei mercati, orientò la ricerca verso le cosiddette reti neurali, modelli, inizialmente astrattamente matematici e solo successivamente concretamente elettronici, sia di singoli neuroni che di connessioni tra i medesimi. Egli, con grande spirito critico, realizzò lo SNARC (*Stochastic Neuro-Analog Reinforcement Computer*), primo computer neurale, il quale fu molto criticato per la sua funzionalità imprecisa e mancanza di rigore, caratteristiche che lo rendevano uno strumento inadatto a svolgere compiti complessi²⁵.

Tali critiche permisero a Minsky di cogliere una questione di fondamentale importanza: la necessità di arrivare alla rappresentazione di un Tutto, partendo da frammenti.

Egli incanalò tale consapevolezza nelle sue successive ricerche, le quali lo portarono alla realizzabilità concreta della *Register Machine*, equivalente alla Macchina di Turing quanto a computabilità, fino a quel momento considerata solo un modello teorico. Superò il problema del tradurre in un dispositivo reale il concetto di infinità, propria del nastro infinito che alimenta la Macchina di Turing; nello SNARC tutto si riduceva all'essere i singoli registri non vuoti seppure virtualmente infiniti. Tutto questo ricondusse Minsky al concetto di *Frame Representation of Knowledge* che persegue l'obiettivo di contestualizzare la conoscenza, così come le possibili risposte della macchina a una situazione reale. Traendo ispirazione da ciò, egli ideò tre Schemi Rappresentativi: il Dichiarazionale basato sulla Logica Formale, il Procedurale basato sulle preesistenti nozioni di Programmazione e l'Associativo di ispirazione linguistica²⁶.

Parallelamente, McCarthy, informatico statunitense, intraprese degli studi volti a colmare retroattivamente la carenza della Logica Formale classica, la quale appariva come un processo razionale parziale, non deterministico e sempre aperto a correzioni nel caso in cui si fossero presentati fattori ambientali nuovi e non previsti ex-ante. Egli nel 1958 sviluppò il linguaggio di programmazione dell'A. I., il famoso LISP e l'*Advice*

²⁴ B. Lenzi, Passato, (2012), "Presente e futuro dell'intelligenza artificiale", *Divenire*, 5, p. 236.

²⁵ Russell S. e Norving P., (2005), *Intelligenza artificiale. Un approccio moderno*, Pearson Edition Italia, Milano.

²⁶ *Ibidem*.

Taker, i quali delineavano l'approccio simbolico adottato da egli stesso. Esso consisteva nel rappresentare, utilizzando un opportuno linguaggio formale, tutto quello che un agente conosce su un dato mondo; alla base vi era la rappresentazione esplicita della conoscenza e proprio da questa si traevano le conclusioni necessarie al fine di far agire l'agente in modo "intelligente".²⁷ Approccio che può essere ricondotto alle teorie assiomatiche in matematica: la conoscenza su un dato pezzo di mondo viene rappresentata in forma di assiomi da cui si traggono conclusioni. Al contempo ne differisce in quanto, mentre in matematica la conoscenza è di tipo statico, nell'intelligenza artificiale si vuole rappresentare una realtà dinamica, in cui le proprietà possono passare dal vero al falso, semplicemente perché è cambiato il contesto di riferimento e/o sono state eseguite azioni che hanno falsificato cose che prima erano vere, o viceversa²⁸. Ad esempio, l'*Advice Taker* sarebbe stato in grado di generare un piano per guidare fino all'aeroporto e prendere un aereo; inoltre, fu progettato al fine di poter accettare anche nuovi assiomi durante la normale esecuzione, acquisendo così competenze in nuove aree senza dover essere riprogrammato. Rappresenta, ancora oggi, il principio fondamentale della rappresentazione della conoscenza e del ragionamento, portando a una rappresentazione formale ed esplicita del mondo e del modo in cui le azioni di un agente lo modificano²⁹.

La conferenza avvenuta nel Dartmouth fu seguita da un periodo, di quasi vent'anni caratterizzato da un successo rilevante nel campo dell'intelligenza artificiale. Oltre ai successi conseguiti da McCarthy e Minsky nelle rispettive aree di ricerca, tra il 1964 e il 1966, venne creato, da Joseph Weizenbaum, un programma di elaborazione del linguaggio naturale. Il programma, noto con il nome ELIZA, rappresenta ancora oggi una delle pietre miliari dell'intelligenza artificiale. La macchina era in grado di simulare la conversazione con un essere umano utilizzando semplici sostituzioni e regole di ricerche

²⁷ Russell S. e Norving P., (2005), *Intelligenza artificiale. Un approccio moderno*, Pearson Education Italia, Milano.

²⁸ L. Carlucci e M. Aiello, (2004), "Intelligenza Artificiale: i primi 50 anni", *Mondo Digitale*, 2, p. 10.

²⁹ Russell S. e Norving P., (2005), *Intelligenza artificiale. Un approccio moderno*, Pearson Education Italia, Milano.

per corrispondenza (*pattern matching*). Essa rappresenta la prima iterazione uomo-macchina³⁰.

Un ulteriore successo di quegli anni fu il programma *General Problem Solver*, sviluppato dal vincitore del premio Nobel Herbert Simon e gli studiosi della *RAND Corporation*, quali Cliff Shaw e Allen Newell, in grado di risolvere automaticamente alcuni tipi di problemi semplici come quello della torre di Hanoi, gioco matematico che prevede la presenza di tre aste e un numero di dischi (da 3 a 7) di dimensione diversa. Il gioco inizia con i dischi disposti in ordine crescente nell'asta centrale, così da formare appunto una torre. L'obiettivo del gioco consiste nello spostare l'intera torre dall'asta centrale a una delle due aste laterali mantenendo inalterato l'ordine crescente originario³¹.

Conseguentemente a questi importanti risultati, lo studio dell'intelligenza artificiale fu stimolato da sostanziali finanziamenti, i quali portarono a un numero sempre maggiore di progetti e così anche l'ottimismo degli studiosi venne alimentato; a tal proposito nel 1970 Marvin Minsky rilasciava un'intervista a *Life Magazine* in cui affermava che una macchina con intelligenza generale di un essere umano medio, poteva essere sviluppata entro 3-8 anni. Fin dagli inizi infatti, i ricercatori di intelligenza artificiale non ebbero remore a fare predizioni sui loro futuri successi:

<< Il mio scopo non è stupire o sbalordire – ma il modo più semplice in cui posso esprimermi è dicendo che ora nel mondo esistono macchine che possono pensare, imparare e creare. Inoltre, la loro abilità nel fare queste cose aumenterà rapidamente finché, in un futuro vicino, il campo dei problemi che potremmo gestire avrà la stessa estensione di quello a cui si è applicata la mente umana>>. (Simon H., 1957)³².

Sfortunatamente però, il termine “futuro vicino” può essere interpretato in vari modi e le predizioni fatte da Simon, come quella che un computer sarebbe stato campione mondiale di scacchi e che una macchina avrebbe dimostrato un teorema matematico di importanza rilevante, si verificheranno dopo quarant'anni anziché nel decennio successivo.

³⁰ J. Weizenbaum, (1966), *Joseph Weizenbaum's ELIZA: Communications of the ACM*, Cambridge.

³¹ L. Carlucci e M. Aiello, (2004), “Intelligenza Artificiale: i primi 50 anni”, *Mondo Digitale*, 2, p. 10.

³² Russell S. e Norving P., (2005), *Intelligenza artificiale. Un approccio moderno*, Pearson Education Italia, Milano.

Ciò conseguentemente al fatto che si manifestarono delle difficoltà e dei limiti che non erano stati previsti. I primi sistemi contenevano una quantità pressoché insignificante di informazioni specifiche attinenti all'area delle loro elaborazioni in quanto il loro agire era basato su semplici manipolazioni sintattiche: questo rappresentò il primo ostacolo³³. Il secondo ostacolo era caratterizzato dal fatto che i primi programmi giungevano alla soluzione provando varie combinazioni di passi, ma ciò chiaramente, funzionava con un numero relativamente ridotto di problemi; si verificò quindi l'intrattabilità di questi nel caso si fossero presentati in numero più elevato. Tale problema fu studiato, ma purtroppo non risolto, da Friedberg nel 1958 tramite esperimenti di evoluzione automatica (oggi chiamati algoritmi genetici) i quali apportavano una serie appropriata di piccole modifiche a un programma per generarne un altro con prestazioni migliori per una qualsiasi specifica attività. Ciò non comportò nessun progresso rilevante³⁴.

La mancanza di una soluzione al problema dell'esplosione combinatoria rappresentò una delle critiche principali contenuto nel rapporto Lighthill nel 1973; egli affermò che le macchine avrebbero raggiunto il livello di "dilettanti esperti" solo in giochi come gli scacchi e che il ragionamento del buonsenso sarebbe stato sempre al di là delle loro capacità. Conseguentemente a ciò, il governo britannico, seguito da quello statunitense, interruppe il supporto finanziario all'intelligenza artificiale³⁵.

Le aspettative e le speranze che avevano nutrito per anni gli scienziati, si rivelarono troppo ambiziose; proprio nelle mansioni che sembravano più semplici, come la traduzione automatica o la riproduzione del linguaggio naturale, l'A. I. aveva maggiori difficoltà. A ciò si aggiungevano alcuni limiti fondamentali delle strutture di base utilizzate per generare comportamento intelligente: benché delle semplici forme di reti neurale potessero apprendere tutto ciò che erano in grado di rappresentare, potevano effettivamente rappresentare ben poco; ad esempio, una rete neurale basica a due

³³ Russell S. e Norving P., (2005), *Intelligenza artificiale. Un approccio moderno*, Pearson Education Italia, Milano.

³⁴ L. Carlucci e M. Aiello, (2004), "Intelligenza Artificiale: i primi 50 anni", *Mondo Digitale*, 2, p. 10.

³⁵ M. Haenlein e A. Kaplan, *A brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence*, University of California.

input non era in grado, e di imparare, a distinguere quando i suoi due input erano diversi³⁶.

La riduzione di interesse che inevitabilmente ne conseguì, si protrarrà fino alla prima metà degli anni Ottanta quando la ricerca riprese vigore.

I progressi e i risultati sull'intelligenza artificiale conseguiti fino a questo momento, tesero a differenziarsi dagli avanzamenti e dai risultati conseguiti dalle discipline limitrofe quali matematica, informatica, teoria del controllo e statistica.

Solamente, a partire dagli inizi degli anni Ottanta, l'A. I. ha creato una maggiore integrazione con tali discipline; la comunità scientifica, superata l'era pionieristica, cominciò a pretendere un maggior rigore, i risultati dovevano essere caratterizzati, alla base, da solide teorie oppure derivare da chiare sperimentazioni³⁷.

In questi anni si "riscoprono" le reti neurali, il cui principale contributo, è attribuito ai risultati del fisico J. J. Hopfield il quale è riuscito ad allargarne, in modo considerevole, il campo di applicazione: egli utilizzava tecniche di meccanica statistica per studiare le proprietà di memorizzazione e l'ottimizzazione delle reti. Psicologi come David Rumhart e Geoff Hinton continuavano ad analizzare i modelli della memoria basati su reti neurali.

Ma il vero ritorno della disciplina dell'intelligenza artificiale si ebbe intorno alla metà degli anni Ottanta, quando quattro gruppi differenti di studiosi reinventarono l'algoritmo di autoapprendimento con retro propagazione scoperto nel 1969 da Bryson e Ho. L'algoritmo fu applicato in informatica e psicologi, e la pubblicazione delle conclusioni tratte, nella collezione *Parallel Distributed Processing* suscitò molto stupore e turbamento: questi modelli di sistemi intelligenti, chiamati connessionisti, furono considerati diametralmente opposti ai modelli simbolici di Newell e Simon, ma anche all'approccio logistico di McCarthy³⁸.

³⁶ Russell S., Norving P., (2005), *Intelligenza artificiale. Un approccio moderno*. Pearson Education Italia, Milano.

³⁷Carlucci Aiello L., (2004), "Intelligenza artificiale: i primi 50 anni", *Mondo Digitale*, 2, p. 9.

³⁸ Russell S., Norving P., (2005), *Intelligenza artificiale. Un approccio moderno*. Pearson Education Italia, Milano.

1.4 L'intelligenza artificiale contemporanea

In anni recenti si è assistito a una vera e propria rivoluzione nel campo dell'intelligenza artificiale; attualmente è più comune basarsi su teorie esistenti che avanzarne di nuove ovvero ideare le proprie supposizioni su rigorosi teoremi o prove sperimentali proponendo applicazioni reali anziché esempi limitati.

Oggi l'intelligenza artificiale si basa completamente nell'ambito del metodo scientifico, cioè le ipotesi devono essere analizzate da rigorosi esperimenti e i relativi risultati convalidati da un'analisi statistica rigorosa. In riferimento a ciò vengono utilizzati approcci basati sui modelli nascosti di Markov (HMM, *hidden Markov models*) i quali si fondano su una rigorosa teoria matematica e vengono generati mediante un processo di apprendimento basato su una grande quantità di dati reali dando prova di un'affidabilità crescente³⁹.

Essi sono dei processi doppiamente stocastici, ovvero caratterizzati da un processo stocastico sottostante non osservabile, e proprio per questo, sono detti nascosti. Questo può essere osservato solo attraverso un'ulteriore serie di processi stocastici che produce la sequenza di simboli osservati⁴⁰. Per comprendere tali modelli è possibile rifarsi a un classico esempio: un individuo è in una stanza nella quale è presente una tenda e attraverso di essa non è possibile vedere cosa sta accadendo dall'altro lato. Nella parte opposta di tale barriera, è presente un ulteriore individuo che sta eseguendo un esperimento di lancio di monete o monete multiple. I due individui non comunicano tra di loro se non per esplicitare il risultato di ogni lancio della moneta: si esegue così una sequenza di esperimenti di lancio di monete nascoste, dei quali si osservano solamente i risultati, cioè dove sta testa e dove sta croce⁴¹.

Il problema sta nel costruire un modello nascosto di Markov per spiegare la sequenza osservata di teste e croci. Un possibile modello è quello della "*moneta equa*": si verificano due stati nel modello ognuno dei quali è associato in modo univoco con teste

³⁹ Russell S. e Norving P., (2005), *Intelligenza artificiale. Un approccio moderno*, Pearson Education Italia, Milano.

⁴⁰ Rabiner L. R. e Juang B. H., (1986), "An Introduction to Hidden Markov Models", *ASSP MAGAZINE*, 86, pp. 5 - 6.

⁴¹ Ibidem.

(stato 1) o con croci (stato 2)⁴². Questo non è un modello nascosto in quanto la sequenza di osservazione definisce in modo univoco lo stato e rappresenta la moneta equa conseguentemente al fatto che la probabilità di generare una testa (o una croce) dopo una testa (o una croce) è 0,5: non si presenta parzialità sull'attuale osservazione. Tale esempio dimostra che esperimenti indipendenti possano essere interpretati come un insieme di eventi sequenziali; se l'individuo dietro la tenda sta effettivamente lanciando solo una moneta, suddetto modello dovrebbe spiegare molto bene i risultati⁴³.

Un modello nascosto di Markov alternativo è denominato "*moneta da due monete*" e prevede ancora due stati, ma diversamente dal precedente, nessuno dei due è associato in modo univoco a testa o croce. La probabilità di testa (o croce) in entrambi gli stati è 0,5 e la medesima si presenta nel caso di rimanere o lasciare. In questo caso è possibile associare a ogni stato una moneta (imparziale) equa: le statistiche delle sequenze di output osservabili del modello quindi, sono indipendenti dalle transazioni di stato. La ragione di ciò è che tale modello è nascosto, cioè non è possibile conoscere con esattezza quale moneta equa ha portato alle teste o alle croci osservate, ma è essenzialmente indistinguibile, in senso statistico, dal primo modello citato⁴⁴.

Poi c'è essenzialmente un altro modello nascosto di Markov, denominato "*modello a due monete non equa*" ed è appropriato nel caso in cui la persona dietro la barriera ha tre monete, una equa e le rimanenti distorte. Queste ultime sono associate rispettivamente alle due facce della moneta equa e per esplicitare il risultato di ogni lancio di una moneta, l'individuo dietro la barriera lancia prima la moneta equa per decidere quale moneta distorta utilizzare in seguito e ottenere così il risultato. Con tale modello si è quindi in grado di spiegare i sottili cambiamenti delle monete distorte.

Per essere in grado di modellare i risultati dell'esperimento tramite gli HMM è necessario innanzitutto decidere la dimensione, ovvero il numero di stati del modello senza nessuna informazione *ex-ante* e le dimensioni della sequenza di osservazioni; infatti prendendo in considerazione una sequenza estremamente esigua, si potrebbe non essere in grado di stimare in modo ottimale i parametri del modello.

⁴² Elliot, Robert J.; Mamon Rogemar S., (2007), *Hidden Markov models in finance.*, Springer, New York.

⁴³ Ibidem.

⁴⁴ Ibidem.

In un certo senso quindi, a seconda della quantità di dati considerati dal modello, alcuni HMM potrebbero non essere statisticamente attendibili⁴⁵.

Anche le reti neurali sono state soggette a una rivoluzione importante che le ha condotte a un elevato grado di maturazione; oggi esse sono in grado di utilizzare le tecniche di riconoscimento dei pattern e di apprendimento automatico, in modo da applicare a ogni problema la soluzione più efficiente e attinente.

In seguito a tali sviluppi, è nata una nuova industria basata su una tecnologia del tutto nuova, il *data mining* ovvero un processo assistito da computer per esplorare e analizzare set di dati di dimensioni rilevanti al fine di estrarre informazioni e conoscenze utili⁴⁶. L'approccio utilizzato da tale tecnologia si basa sul rilevamento di modelli e altri algoritmi utilizzati per determinare le relazioni chiave in dati inesplorati.

Contemporaneamente si sono verificati dei progressi anche nel campo della robotica⁴⁷, della visione e della rappresentazione della conoscenza che, uniti alle miglione

⁴⁵Elliot, Robert J.; Mamon Rogemar S., (2007), *Hidden Markov models in finance.*, Springer, New York.

⁴⁶ Il processo di ricerca dei dati per scoprire connessioni nascoste e prevedere le tendenze future ha una lunga storia. A volte indicato come "scoperta della conoscenza nei database", il termine "data mining" non è stato coniato fino agli anni '90. Ma la sua fondazione comprende tre discipline scientifiche intrecciate: la statistica (lo studio numerico delle relazioni di dati), l'intelligenza artificiale (intelligenza simile a quella umana visualizzata da software e / o macchine) e l'apprendimento automatico (algoritmi che possono apprendere dai dati per fare previsioni). Ciò che era vecchio è nuovo di nuovo, poiché la tecnologia di data mining continua a evolversi per stare al passo con il potenziale illimitato dei big data e della potenza di elaborazione accessibile.

Nell'ultimo decennio, i progressi nella potenza e nella velocità di elaborazione ci hanno permesso di andare oltre le pratiche manuali, noiose e che richiedono tempo per l'analisi dei dati rapida, semplice e automatizzata. Più sono complessi i set di dati raccolti, maggiore è il potenziale per scoprire approfondimenti pertinenti. I rivenditori, le banche, i produttori, i fornitori di servizi di telecomunicazione e gli assicuratori, tra gli altri, stanno utilizzando il data mining per scoprire relazioni tra tutto, dall'ottimizzazione dei prezzi, promozioni e dati demografici a come l'economia, il rischio, la concorrenza e i social media influenzano i loro modelli di business, i ricavi, le operazioni e relazioni con i clienti. Fonte: Schrage M., (2020), *Data Mining. What it is and why it matters*, SAS Insights.

⁴⁷ La definizione di "robot" è stata sempre difficile da rintracciare, spesso confusa anche a livello di lingua di origine. La parola apparve per la prima volta nel 1921, nella commedia di Karel Capek in R.U.R, Rossum's Universal Robots. "Robot" viene dal ceco dove il significato è "lavoro forzato". Nella commedia, questi robot erano robot più nello spirito che nella forma, però, sembravano a tutti gli effetti esseri umani e invece di essere fatti di metallo, erano fatti di una sorta di plastica lucida. Nella commedia i robot erano molto più efficienti delle loro controparti umane ma anche molto più cattivi, sviluppando un istinto omicida e una importante dose di violenza. Fonte: Redazione, (2020), *Robotica. Storia, evoluzioni e ultime novità*, Impactschool. The future is opensource.

apportate agli strumenti matematici, hanno comportato un irrobustimento della ricerca e dello sviluppo di metodi più efficienti ed efficaci.

Questi settori possono essere riunificati nella visione comune dell'intelligenza artificiale come attività di progettazione di agenti razionali i quali sono stati presi in considerazione, nella loro interezza alla fine degli anni Novanta⁴⁸.

Un agente è qualsiasi cosa che può essere vista come un sistema che percepisce l'ambiente in cui è collocato attraverso dei sensori e agisce su di esso mediante attuatori. Gli input percettivi dell'agente in un dato momento vengono detti percezioni e la loro storia completa, ovvero ciò che l'agente ha percepito nell'arco della sua esistenza, è detta sequenza percettiva; quest'ultima può influenzare la scelta di un'azione di un agente in un qualsiasi istante, la quale sarà implementata da un programma agente⁴⁹.

Data la definizione di agente, è possibile affermare che un agente razionale è l'agente che fa la cosa giusta, cioè che gli permette di massimizzare la misura di prestazione che definisce il grado di successo, la conoscenza pregressa dell'ambiente, le azioni che può effettuare e la sequenza percettiva fino a quell'istante.

L'ambiente che circonda l'agente e nel quale questo agisce può essere fisico, quindi reale o artificiale, computazionale o virtuale. Nel primo caso quindi l'agente sarà un robot, nel secondo caso sarà un agente software e nell'ultimo caso un agente artificiale⁵⁰.

Esso inoltre dev'essere completamente o parzialmente osservabile, deterministico se lo stato successivo dell'ambiente è completamente determinato dallo stato corrente e dall'azione eseguita dall'agente, mentre in caso contrario sarà stocastico⁵¹. Può essere un ambiente operativo episodico se la scelta dell'azione dipende solo dall'episodio corrente, o sequenziale se ogni decisione può influenzare tutte quelle successive. Può

⁴⁸ Russell S. e Norving P., (2005), *Intelligenza artificiale. Un approccio moderno*. Pearson Education Italia, Milano.

⁴⁹ Dopico J. R. R., De la Calle J. D. e Sierra A. P., (2009), *Encyclopedia of Artificial Intelligence*. Information sci, New York.

⁵⁰ Goy A., Torre I., (2009), "Agenti artificiali e agenti intelligenti: paradigmi, applicazioni e prospettive", *LEXIA*, 03/04, pp. 297 - 313.

⁵¹ Russell S. e Norving P., (2005), *Intelligenza artificiale. Un approccio moderno*. Pearson Education Italia, Milano.

essere statico se l'agente non deve continuamente osservare il mondo mentre decide l'azione successiva, oppure dinamico se esso cambia mentre un agente sta pensando⁵².

Può essere discreto nel caso in cui l'ambiente sia caratterizzato da un numero finito di stati distinti o continuo se caratterizzato da azioni continue al passare del tempo.

Infine è possibile che l'ambiente sia caratterizzato da un singolo agente (come nel caso in cui quest'ultimo risolve da solo un cruciverba) o da due o più agenti (come nel caso in cui un agente giochi a scacchi)⁵³.

Uno dei maggiori successi dell'intelligenza artificiale, ottenuto dal lavoro di migliaia di ricercatori che iniziò ufficialmente alla fine degli 50', fu l'evento emblematico che si verificò l'11 maggio 1997 quando Garry Kasparov, l'allora campione mondiale di scacchi, venne battuto da un avversario piuttosto particolare, ovvero una macchina chiamata *Deep Blue*. Esso era un computer, progettato espressamente per giocare a scacchi, che combinava gli algoritmi di ricerca, proposti inizialmente da Claude Shannon e successivamente perfezionati da numerosi scienziati informatici.



FIGURA 3: DEEP BLUE VS. KASPAROV SCACCO MATTO ALL'INTELLIGENZA UMANA. FONTE: LA STAMPA.

Esso fu progettato da un team di specialisti ai quali faceva capo Feng-Hsiung Hsu. Il "tallone d'Achille" della macchina però, stava proprio nel fatto che nessuno dei

⁵² Russell S. e Norving P., (2005), *Intelligenza artificiale. Un approccio moderno*. Pearson Education Italia, Milano.

⁵³ Ibidem.

progettisti era un esperto di scacchi; essi stessi così, per ovviare a tale problema, optarono per l'implementazione di un programma basato sostanzialmente su un algoritmo che prevedeva come input la posizione sulla scacchiera e restituiva come output un valore che quantificava il vantaggio o svantaggio rispetto all'avversario: poteva consistere in un vantaggio o svantaggio di materiale (come conseguenza del guadagno o della perdita di uno o più pezzi) o di vantaggi/svantaggi posizionali (conseguenti al posizionare un pezzo in una casella chiave)⁵⁴. Tale soluzione risultò ottimale, infatti *Deep Blue* era in grado di analizzare rapidamente un numero considerevole di varianti: era in grado di valutare 200 milioni di posizioni al secondo, il che gli permetteva di analizzare una profondità di 6/8 mosse successive se sulla scacchiera erano presenti molti pezzi, ma anche più di 20 mosse successive in presenza di pochi pezzi⁵⁵.

Un più recente successo, ma non meno importante, ottenuto sempre nell'ambito dell'intelligenza artificiale nel 2016, fu l'algoritmo *Deep Mind* di Google applicato al software *Alphago* per il gioco del Go⁵⁶; tale algoritmo permette alla macchina di imparare da sé stessa con l'esperienza dimostrando di riuscire ad apprendere ed adattarsi, nel tempo, a diverse tipologie di giochi, infatti esso non fu creato appositamente per il gioco Go, diversamente dal *Deep Blue*, creato esclusivamente per il gioco degli scacchi.

⁵⁴ Campbell M., Hoane J. A., Hsu F. H., (2002), "Deep Blue", *Elsevier*, 134, pp. 57 - 83.

⁵⁵ *Ibidem*.

⁵⁶ Il Go è un gioco da tavolo strategico per due giocatori. Ebbe origine in Cina, dove è giocato da almeno 2500 anni; è molto popolare in Asia orientale e si è diffuso nel resto del mondo negli anni recenti. È un gioco molto complesso strategicamente malgrado le sue regole semplici; un proverbio coreano dice che nessuna partita di go è mai stata giocata due volte, il che è verosimile se si pensa che ci sono $2,08 \times 10^{170}$ diverse posizioni possibili.

Il Go è giocato da due giocatori che collocano alternativamente pedine (dette pietre) nere e bianche sulle intersezioni vuote di una "scacchiera" (detta goban) formata da una griglia 19 × 19. Lo scopo del gioco è il controllo di una zona del goban maggiore di quella controllata dall'avversario; a questo scopo i giocatori cercano di disporre le proprie pietre in modo che non possano essere catturate, ritagliandosi allo stesso tempo dei territori che l'avversario non possa invadere senza essere catturato.

A parte la dimensione del goban e delle posizioni di partenza, le regole sono state mantenute nei secoli, cosicché può essere considerato il gioco più antico ancora praticato.

Fonte: Autori di Wikipedia, "Go (gioco)", Wikipedia, l'Enciclopedia libera, Tratto il 12 dicembre 2019 da [https://it.wikipedia.org/wiki/Go_\(gioco\)](https://it.wikipedia.org/wiki/Go_(gioco)).

Alphago è stato allenato in 49 giochi differenti superando non solo algoritmi concorrenti, ma anche gli umani, con una potenza di calcolo in grado di considerare 10^{170} situazioni differenti e ad avere contemporaneamente la possibilità di ragionare sulle insidie poste dal gioco.

Oggi l'A. I. replica in qualche modo un modello di mente che potremmo definire modulare, ciascun sistema sa svolgere molto bene un compito specifico (riconoscimento visivo o vocale). Quello che veramente manca, non solo a livello applicativo ma anche a livello neuroscientifico, è capire come l'intelligenza generale possa emergere dall'interazione complessa tra tutti questi sistemi modulari e da altri sistemi che sono dominio-generalisti, cioè che si occupano già di diversi compiti assieme. In neuroscienze si parla sempre di più di rete e di dinamiche *whole brain* che coinvolgono l'intero cervello, dell'interazione tra aree che non sono indipendenti ma che comunicano e partecipano a una dinamica estremamente più ricca e complessa rispetto a quella che si aveva in mente qualche anno fa⁵⁷; basti pensare che alla fine dell'ottobre del 2019, durante il *Future Investment Initiative*, l'Arabia Saudita ha concesso la cittadinanza al robot donna Sophia, la quale interagisce con i giornalisti e riconosce le emozioni umane; a livello di programmazione non possiede pensiero, ma strategie avanzate di *problem solving* e non solo algoritmi⁵⁸.

Ovviamente tali risultati sono stati possibili a partire dal nuovo millennio, grazie a un ambiente tecnico – scientifico più favorevole caratterizzato da una maggiore capacità di calcolo e teorie sull'apprendimento automatico più evolute⁵⁹.

⁵⁷ Suman F., (2019), *Dove sta andando oggi l'intelligenza artificiale?*, Il Bo live Università di Padova, 11/03/2019.

⁵⁸ Proverbio D., (2018), *Intelligenza artificiale e robotica: le sfide del nuovo millennio*, Rivista di affari europei, 11/01/2018.

⁵⁹ Ibidem.

1.5 Dall'automazione all'intelligenza artificiale debole

Per decenni, la ragione che ha spinto ininterrottamente l'uomo a migliorare ed implementare i propri studi e sperimentazioni riguardanti l'intelligenza artificiale, si identifica nel desiderio di creare una connessione inscindibile tra automazione e ragionamento. I primi tentativi si basavano su algoritmi matematici che, con il tempo, si sono evoluti e sono divenuti sempre più complessi ed affidabili.

L'automazione ha così permesso l'ideazione di due teorie fondamentali: l'Intelligenza Artificiale debole e l'Intelligenza Artificiale forte, così denominate dallo studioso John Searle. Entrambe si basano e si sviluppano sulla visione della mente umana come un programma: offrendo stimoli al cervello umano, viene a crearsi un ragionamento, il quale a sua volta produce un certo comportamento.

Scopo dell'intelligenza artificiale quindi, è sempre stato quello di riuscire a creare una macchina o un calcolatore capace di simulare il ragionamento umano⁶⁰.

Durante il primo decennio di studi attinenti all'intelligenza artificiale, aveva preso forma una metodologia per la risoluzione dei problemi che consisteva in un meccanismo di ricerca di uso generale che cercava di mettere assieme dei passi elementari di ragionamento, uno dopo l'altro fino al raggiungimento di soluzioni complete. Tali approcci vennero denominati metodi deboli e questo perché, pur essendo generali, non erano in grado di raggiungere la risoluzione di problemi molto grandi e/o difficili. L'intelligenza artificiale debole si identifica in sistemi tecnologici in grado di agire e ragionare come se possedessero un cervello senza però, raggiungere le effettive funzionalità cognitive e capacità intellettuali proprie dell'essere umano, infatti una macchina non è in grado di ragionare in modo autonomo, ma necessita della presenza dell'uomo⁶¹.

Il programma DENDRAL, sviluppato a Stanford da Ed Feigenbaum, Bruce Buchanan e Joshua Lederberg, fu una delle prime applicazioni di un metodo debole.

⁶⁰Boldrini N., (2019), *Cos'è l'intelligenza artificiale, perché tutti ne parlano e quali sono gli ambienti applicativi*, AI4 BUSINESS, 01/03/2019.

⁶¹ Russell S. e Norving P., (2005), *Intelligenza artificiale. Un approccio moderno*. Pearson Education Italia, Milano.

Inizialmente sviluppato per risolvere il problema della ricostruzione dell'informazione molecolare partendo dai dati forniti da uno spettrometro di massa, esso fu il primo sistema a conoscenza intensiva di successo e la sua capacità era dovuta al fatto che fosse limitato da un numero elevato di regole speciali; esso applica un modello di elaborazione delle informazioni di ragionamento scientifico a un problema specifico. Sostanzialmente raccoglie esperienza analizzando casi simili, li confronta tra di loro, verifica in maniera empirica diverse ipotesi fino ad essere in grado di elaborare diverse soluzioni, optando infine per quella maggiormente razionale e congrua: si occupa esclusivamente del *problem solving*, ovvero la risoluzione di problemi⁶². Una macchina del genere sa fare molto bene una cosa, ma sa fare solo quella, come *Deep Blue*, spiegato nel paragrafo precedente. Successivamente, lo stesso Kasparov affermò che dall'altra parte della scacchiera aveva percepito "un nuovo tipo di intelligenza"⁶³.

1.6 Dall'intelligenza artificiale debole all'intelligenza artificiale forte

La seconda teoria idealizzata da Searle sosteneva che la macchina non era solamente uno strumento, e da qui, la definizione di intelligenza artificiale forte, ovvero dei sistemi esperti che possono sviluppare una propria intelligenza senza emulare processi di pensiero o capacità cognitive e intellettuali proprie dell'uomo, ma sviluppandone una propria autonomamente, pressoché indistinguibile da quella umana. Essa si basa sulla logica matematica, sul ragionamento e la dimostrazione automatica del problema, sull'analisi del linguaggio che permette alla macchina di comprendere le richieste dell'uomo e infine, sulla pianificazione ad opera degli algoritmi.

Un'intelligenza di questo tipo, è in grado di automigliorarsi in un processo continuo che potrebbe portarla ad evolversi molto di più della mente umana, fino ad essere in grado di risolvere complessi problemi di natura etica e giuridica, quando le soluzioni proposte dal sistema automatico possono danneggiare persone e cose o favorire servizi segreti,

⁶² Buchanan B. G., Lederberg J., (1971), *Il programma Dendral per spiegare i dati empirici*, Dipartimento di scienze del computer, Università di Standford, no. CS203.

⁶³ Campbell M., Hoane J. A., Hsu F. H., (2002), "Deep Blue", *Elsevier*, 134, pp. 57 - 83.

criminalità organizzate e così via, fino a comportare responsabilità morali e sociali della macchina.

Attualmente, l'unica differenza che intercorre tra computer e cervello umano, è che quest'ultimo svolge più operazioni contemporaneamente, diversamente da una macchina che ne svolge una dopo l'altra⁶⁴.

Un esempio di sistema esperto, ovvero un programma di intelligenza artificiale forte, è *Mycin* il quale si occupa del trattamento delle infezioni del sangue. Esso, sviluppato da Feigenbaum, Buchanan e il Dott. Edward Shortliffe, era in grado di offrire prestazioni pari a quelle di molti esperti. Diversamente dal programma DENDRAL, esso non conteneva alcun modello teorico generale da cui poter dedurre le regole, le quali a loro volta, dovevano essere acquisite da una serie di interviste agli esperti che le avevano apprese dai libri e dall'esperienza personale. Inoltre, la seconda differenza derivava dal fatto che le regole dovevano rispecchiare intrinsecamente l'incertezza associata alla conoscenza medica per cui, incorporato in *Mycin* era presente un metodo di calcolo di incertezza, denominato fattore di certezza e che rispecchiava le modalità con le quali i dottori utilizzavano le informazioni disponibili durante il processo diagnostico⁶⁵.

Queste diverse teorie, nel corso degli anni, hanno creato un dibattito tra i diversi studiosi della materia; alla base si trova il teorema fondamentale: se il cervello è una macchina, in linea teorica è possibile costruire una macchina che svolga esattamente le funzioni del cervello⁶⁶.

I sostenitori della concezione debole, come Dreyfus, affermarono che l'intelligenza artificiale non è una vera e propria intelligenza; egli in particolar modo contestò il concetto di A. I. forte sostenendo che il fondamento sul quale si basava era completamente errato. Un calcolatore utilizza dati attraverso una serie di dettami ben specificati e di conseguenza è in grado di cogliere solamente alcuni aspetti della realtà,

⁶⁴Guerini M., (2017), *Intelligenza Artificiale, forte e debole, tra mito e realtà*, Corriere della sera Blog, 10/11/2017.

⁶⁵ Russell S. e Norving P., (2005), *Intelligenza artificiale. Un approccio moderno*. Pearson Education Italia, Milano.

⁶⁶ Guerini M., (2017), *Intelligenza Artificiale, forte e debole, tra mito e realtà*, Corriere della sera Blog, 10/11/2017.

al contrario della mente umana che ragiona cogliendo i vari elementi nella loro totalità⁶⁷. A tal proposito quindi Searle ideò un test, denominato “test della stanza cinese”, in netta contrapposizione con il famoso Test di Turing. Tale test consisteva nel costruire un computer che si comportasse come se capisse il cinese: il computer considerava degli ideogrammi cinesi in ingresso, eseguiva un programma con il quale, in base alla forma degli ideogrammi, forniva tutte le regole di combinazione e associazione tra di essi e produceva degli ideogrammi in uscita. Tale macchina sarebbe in grado di rispondere ad un interlocutore umano cinese, inducendolo però a credere di tenere una conversazione con un soggetto umano che conosce e capisce correttamente la sua lingua, così ingannandolo⁶⁸.

Con tale test, Searle dimostrò che una macchina, pur avendo superato con successo il Test di Turing, non può essere definita intelligente: la macchina ha svolto un’azione su dei dati che gli sono stati affidati senza alcuna comprensione del significato del linguaggio.

Quindi, se da una parte John Searle nega sostanzialmente che un cervello computerizzato possa raggiungere il livello di funzionamento di quello di un essere umano, l’opinione speculare dei sostenitori dell’intelligenza artificiale forte, come Daniel Dennett, sostiene che se un’entità si comporta in maniera tale da dimostrare una determinata intelligenza, ciò basta per affermare che tale entità è dotata di una mente equivalente a quella dell’uomo, in quanto gli stati mentali non solo sono riconducibili a puri comportamenti osservabili, ma quest’ultimi sono gli unici rilevanti⁶⁹.

Con riguardo al test della “stanza cinese” di Searle, i sostenitori di tale concezione definirebbero che il computer capisce il cinese al pari di un individuo umano, in quanto non sussiste alcuna differenza tra il comportamento della macchina e quello di un uomo che conosce il cinese.

Sebbene le enormi differenze tra le due concezioni fossero molto evidenti, entrambe erano accomunate da un denominatore comune: nel condurre il dibattito

⁶⁷ Dreyfus H. L., (2002), *Phenomenology and the Cognitive Sciences*, Dipartimento di Filosofia edell’Università della California a Berkeley, No. 1.

⁶⁸ Wang J., (2011), *The Turing Test, Searle’s chinese room, and E.T.’s chinese room*, School of Business, Richard Stockton College of New Jersey, No. 4628.

⁶⁹ Wright R., (1996), “Can machines think?”, *Time Magazine*, Volume. 147, p. 50.

sull'intelligenza artificiale, sia Dennett che Searle dovettero mettere in discussione prima di tutto la stessa natura umana, le basi del suo modo di pensare e di agire.

Negli anni più recenti, la dicotomia debole/forte è stata superata, giungendo così alla conclusione comune che una macchina potrà essere definita realmente intelligente soltanto quando sarà in grado di riprodurre il funzionamento del cervello umano a livello cellulare⁷⁰.

⁷⁰Wright R., (1996), "Can machines think?", *Time Magazine*, Volume. 147, p. 50.

CAPITOLO II – FUNZIONAMENTO DELL’INTELLIGENZA ARTIFICIALE

Sin dalla loro evoluzione, gli esseri umani hanno utilizzato molteplici tipologie di strumenti per svolgere vari compiti e proprio la creatività del cervello umano ha portato l’evoluzione di diverse macchine, le quali hanno consentito il soddisfacimento di esigenze riguardanti i più svariati campi, quali i viaggi, le industrie, le costruzioni e l’informatica⁷¹.

Nonostante i rapidi sviluppi di settore, la pigrizia è rimasta la differenza fondamentale tra umani e macchine nell’esecuzione dei loro compiti; un essere umano utilizza i suoi sensi per raccogliere informazioni dall’ambiente che lo circonda e, dopo averle analizzate, assume le opportune decisioni.

Le macchine però, al contrario degli esseri umani, non sono intelligenti per natura, non hanno le capacità di analizzare dati e prendere decisioni o perlomeno questo si sostenne fino alla metà del ventunesimo secolo quando Alan Turing ritenne fosse possibile che le macchine pensassero⁷².

La vera svolta si verificò quando vennero ricreati degli algoritmi specifici in grado di perfezionare la capacità di agire e di decidere di una macchina.

L’apprendimento automatico o *machine learning* venne definito inizialmente da Hebert Simon nel 1959 come quella branca dell’informatica che permette ad una macchina di imparare ad eseguire un compito senza essere stata esplicitamente programmata per farlo⁷³. Un’ulteriore definizione che gli venne attribuita dal professor Tom Mitchell, in un articolo pubblicato nel 2006 e intitolato “*The Discipline of Machine Learning*” fu quella di crescita naturale dell’intersezione tra l’informatica e la statistica⁷⁴; esso allena il sistema ad imparare dall’esperienza, infatti modifica sé stesso ovvero allo stesso input, a distanza di tempo, non produce necessariamente lo stesso output in quanto l’input ricevuto in precedenza ha rappresentato nuova esperienza per l’agente; esso si basa

⁷¹ Mohammed M., Khan B. M., e Bashier M.B.E., (2017), *Machine Learning. Algorithms and Applications*, CRC Press, Boca Raton.

⁷² Nilsson N. J., (1998), *Introduction to Machine Learning. An early draft of a proposed textbook*, Stanford University, Stanford.

⁷³ Ibidem.

⁷⁴ Mohammed M., Khan B. M., e Bashier M.B.E., (2017), *Machine Learning. Algorithms and Applications*, CRC Press, Boca Raton.

quindi, su algoritmi le cui prestazioni migliorano in modo adattivo con l'aumentare del numero dei campioni disponibili per l'apprendimento: l'esperienza è il dato utilizzato e l'allievo è l'algoritmo⁷⁵.

Attualmente l'apprendimento automatico ricopre un ruolo rilevante nella vita degli esseri umani per diverse ragioni. Innanzitutto, ci sono numerosi compiti che gli esseri umani svolgono abitualmente, ma l'introspezione con cui questi vengono svolti non è sufficientemente elaborata per estrarne un programma ben congegnato; esempi di tali compiti si identificano nella guida, il riconoscimento vocale e la comprensione di immagini. Inoltre possono presentarsi attività da svolgere che vanno al di là delle capacità umane, come l'analisi di insiemi di dati dimensionalmente elevati e complessi oppure delle mansioni che cambiano nel tempo con la conseguente necessità di adattabilità che contraddistingue gli agenti intelligenti dalle macchine utensili ad esempio. È importante quindi la loro flessibilità rispetto ai cambiamenti che si verificano all'ambiente che le circonda⁷⁶.

Proprio per tale ragione, alla macchina viene fornito un algoritmo di apprendimento, il quale si identifica come la descrizione di uno specifico modello di apprendimento e dei dati, i quali possono essere strutturati ovvero organizzati in tabelle, chiave-valore oppure non strutturati se si presentano come testi, audio, immagini ecc..... I tipi di problemi ai quali viene applicato il Machine Learning sono molteplici:

- classificazione quando è necessario decidere a quale categoria appartiene un determinato dato;
- regressione per prevedere il valore futuro di un dato avendo a disposizione il suo valore attuale e
- raggruppamento nel momento in cui si necessita di raggruppare i dati che presentano caratteristiche simili.

In base al tipo di algoritmo utilizzato dalla macchina per apprendere, nonché accumulare dati e informazioni sui quali deve operare, si possono distinguere tre modalità diverse

⁷⁵ Mitchell T., (2006), *The Discipline of Machine Learning*, School of Computer Science Carnegie Mellon University, 07/2006.

⁷⁶ Nilsson N. J., (1998), *Introduction to Machine Learning. An early draft of a proposed textbook*, Stanford University, Stanford.

di apprendimento in modo da assicurare la massima performance e il risultato più efficiente ed efficace come risposta agli stimoli dell'ambiente esterno.

2.1 Modalità di apprendimento supervisionato

La terza e ultima modalità di apprendimento che viene analizzata, è quella supervisionata detta anche *Supervised Machine Learning*. È una tecnica di apprendimento utilizzata per addestrare degli strumenti fornendo una serie di nozioni specifiche, ossia modelli ed esempi che permettono di creare un database di informazioni e di esperienze a disposizione della macchina. L'obiettivo è trovare un'approssimazione sufficientemente buona $f'(x)$ della funzione $f(x)$, la quale è alla base della relazione tra input e output⁷⁷. Il processo di apprendimento, in un modello di apprendimento supervisionato, è composto principalmente da due fasi: la fase di addestramento, nella quale il set di dati viene appreso come funzione dall'algoritmo di apprendimento, e la fase test nella quale il modello addestrato utilizza l'apprendimento per effettuare la previsione dei dati⁷⁸. In altre parole, l'algoritmo di apprendimento utilizzato riceve un set input insieme agli output osservati e apprende grazie al confronto tra l'output stimato e quello osservato in modo da identificare eventuali errori che gli permetteranno di modificare e quindi migliorare ulteriormente il modello⁷⁹.

2.1.1 Tipologie di apprendimento supervisionato

A seconda della natura dell'obiettivo, l'apprendimento supervisionato può identificarsi con:

- l'attività di classificazione avente il compito di approssimare una funzione di mappatura f dalle variabili di input X , le quali possono avere valori reali o discreti, alle variabili di output discrete Y (chiamate più comunemente etichette o categorie); tale attività quindi

⁷⁷ Cangini N., *Quantum Supervised Learning: algoritmi e implementazione*, tesi di laurea, Università di Bologna, a. a. 2017-2018, relatore Sartori C.

⁷⁸ Nasteski V., (2018), *An Overview of the supervised machine learning methods*, Research Gate, 11/12/2018.

⁷⁹ Nasteski V., (2018), *An Overview of the supervised machine learning methods*, Research Gate, 11/12/2018.

prevede la classe o la categoria per una data osservazione⁸⁰. Questa tecnica di apprendimento supervisionato richiede che gli esempi siano classificati in una o più classi e prevede un valore continuo come probabilità, di un dato esempio, appartenente a ciascuna classe di output. È possibile calcolare l'accuratezza della classificazione come la percentuale di esempi correttamente classificati rispetto a tutte le previsioni fatte. Gli algoritmi in grado di apprendere un modello predittivo di classificazione sono detti appunto algoritmi di classificazione⁸¹.

- l'attività di regressione, diversamente da quella precedente, ha il compito di approssimare una funzione di mappatura f dalle variabili X , le quali possono avere valori reali o discreti, a una variabile di output continua Y . Tale tecnica richiede la previsione di una quantità, valutata tramite il calcolo del RMSE⁸². Gli algoritmi in grado di apprendere un modello predittivo di regressione sono chiamati quindi algoritmi di regressione.

A titolo esemplificativo, di seguito viene riportata la rappresentazione grafica della differenza che intercorre tra le due diverse modellizzazioni:

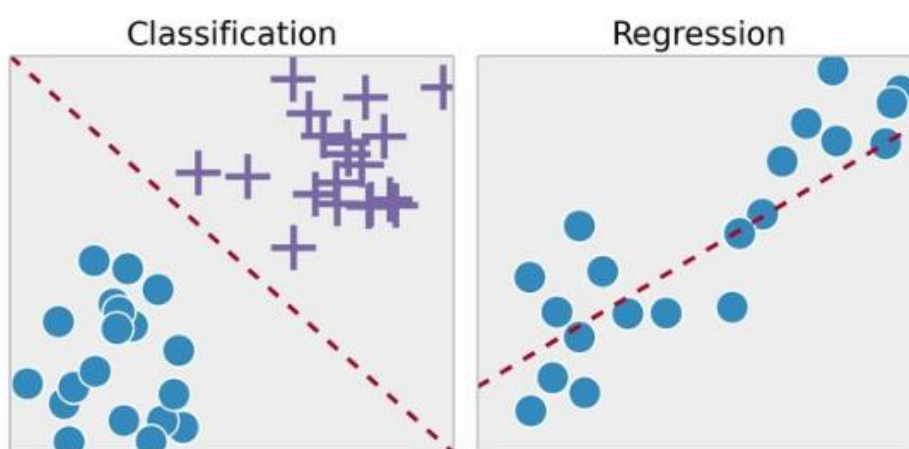


FIGURA 4: A SINISTRA VIENE RAPPRESENTATO GRAFICAMENTE UN PROBLEMA DI CLASSIFICAZIONE, MENTRE A DESTRA, UN PROBLEMA DI REGRESSIONE. FONTE: [HTTPS://BLOG.STATSBOT.CO/MACHINE-LEARNING-ALGORITHMS-183CC73197C](https://blog.statsbot.co/machine-learning-algorithms-183cc73197c).

⁸⁰ Brownlee J., (2017), *Difference Between Classification and Regression in Machine Learning*, Machine Learning Mastery, 11/12/2017.

⁸¹ Ibidem.

⁸² RMSE o errore quadratico medio di radice, è una regola di punteggio quadratico che misura l'entità media dell'errore. È la radice quadrata della media delle differenze al quadrato tra previsione e osservazione effettiva. Si identifica nella formula: $RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_i - \hat{y}_i)^2}$.

Fonte: J.J., (2016), *MAE and RMSE – Which Metric is Better?*, Medium, 23/03/2016.

Il primo algoritmo utilizzato nell'ambito dell'apprendimento automatico supervisionato, che si va ad analizzare, è la regressione lineare ovvero un modello che si occupa di prevedere, nonché specificare, la relazione tra una singola variabile dipendente di input (valore da prevedere Y) e una o più variabili indipendenti di output X (predittori) assumendo che tale relazione possa essere rappresentata da una linea retta e che i valori da prevedere si trovino all'interno di un intervallo continuo⁸³. Esistono principalmente due tipi di regressione lineare:

- nel caso in cui la variabile indipendente sia solamente una, si parla di regressione lineare semplice e in termini equazionali può essere scritta come:

$$y = \alpha + \beta x,$$

dove α indica l'intercetta della retta, mentre β indica la pendenza (coefficiente angolare). A titolo di esempio, si considera un set di dati che descrivono quando un'azienda sostiene, in termini di costo, ogni anno di pubblicità radiofonica e le rispettive vendite in termini di unità vendute; se si volesse sviluppare un'equazione per prevedere le unità vendute in base al costo sostenuto per la pubblicità radiofonica, l'equazione utilizzata risulterebbe:

$$Vendite = \text{Peso} \times \text{Radio} + \text{Bias}$$

dove il *Peso* rappresenta il coefficiente della variabile indipendente e il *Bias* rappresenta la distorsione che compensa le previsioni che vengono fatte, quindi gli eventuali errori⁸⁴.

⁸³ Lantz B., (2013), *Machine Learning with R*, Packt Publishing, Birmingham.

⁸⁴ Ibidem.

Tale equazione, alla fine dell'addestramento si avvicinerà alla linea di migliore adattamento e a livello grafico ne risulterà:

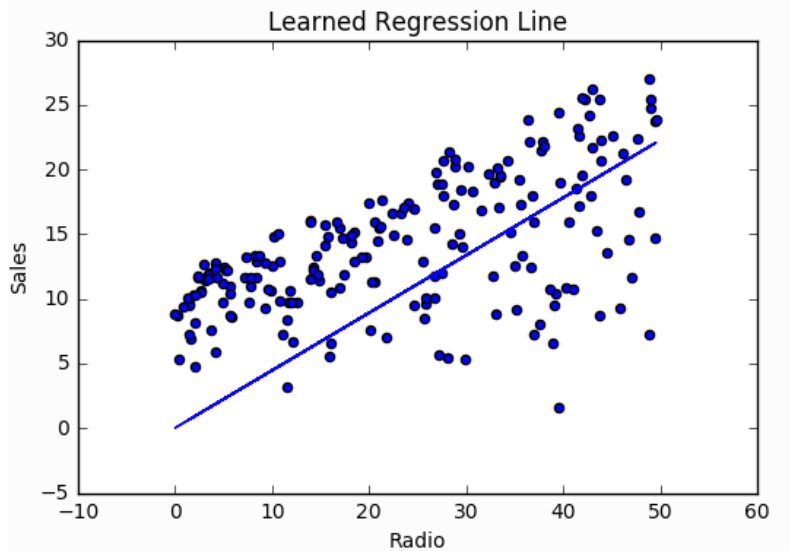


FIGURA 5: RAPPRESENTAZIONE GRAFICA DEL MODELLO DI REGRESSIONE LINEARE SEMPLICE. FONTE: LANTZ B., (2013), MACHINE LEARNING WITH R, PACKT PUBLISHING, BIRMINGHAM.

- la maggior parte delle analisi del mondo reale però, considera una molteplicità di variabili indipendenti pertanto viene utilizzata la regressione lineare multivariata, considerata un'estensione della precedente e rappresentata con un'equazione del tipo:

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i$$

Estendendo l'esempio fatto in precedenza, si consideri di avere a disposizione un set di dati sui costi sostenuti in pubblicità, in televisione, radio e giornali con l'obiettivo di prevedere le vendite in termini di unità vendute.

La funzione di previsione in questo caso sarà più complessa e assumerà la forma del tipo:

$$Vendite = W_1 Tv + W_2 Radio + W_3 Giornale$$

In altre parole, tale metodologia permette di identificare le eventuali relazioni tra una variabile dipendente e un determinato insieme di variabili indipendenti.

Un ulteriore strumento di machine learning supervisionato, facente parte della famiglia degli algoritmi di classificazione e divenuto particolarmente importante, è la

regressione logistica, il cui pioniere fu Ohlson nel 1980 che affrontò il problema di classificare alcune società distinguendole tra adempienti e inadempienti⁸⁵. La regressione logistica, utilizzando dei metodi statistici basati su dati storici, fornisce un risultato rappresentato da una probabilità che una determinata unità statistica appartenga a una determinata classe. Essendo che si tratta di un algoritmo di apprendimento automatico, chiaramente esso dovrebbe migliorarsi nel prevedere le diverse classificazioni ogniqualvolta viene utilizzato un set di dati più consistente⁸⁶.

Tali problemi sono detti anche problemi di regressione logistica binomiale in quanto la variabile che si cerca di prevedere è binaria, ovvero può assumere solamente due valori: il valore 1 che rappresenta la classe positiva con una certa probabilità (P), o il valore 0 che rappresenta la classe negativa con una probabilità diversa (1-P), stimate entrambe tramite una funzione logistica che assume la forma di

$$F(x) = \frac{1}{1 + e^{-(b_0 + b_1 x)}}$$

dove x è la variabile indipendente e b_0 e b_1 sono i parametri del modello⁸⁷.

Tali coefficienti nel metodo della regressione logistica vengono stimati tramite il metodo di massima verosomiglianza che permette di ottenere una serie di coefficienti di regressione per i quali la probabilità di ottenere i dati osservati è massima minimizzando l'errore nelle probabilità previste dal modello rispetto a quelle dei dati. Tale funzione trasforma i valori reali arbitrari in numeri compresi tra zero e uno; maggiore è il valore reale, maggiormente il risultato sarà prossimo a uno e viceversa.

⁸⁵ Zaidi M., Amirat A., (2016), "Forecasting stock market trends by logistic regression and neural networks", *International Journal of Economics, Commerce and Management*, 4, pp. 221-222.

⁸⁶ Govoni L., (2018), *Come funziona un algoritmo di regressione logistica*, Business e Tecnologia, 03/09/2018.

⁸⁷ Ibidem.

A livello grafico, essa può essere rappresentata come segue:

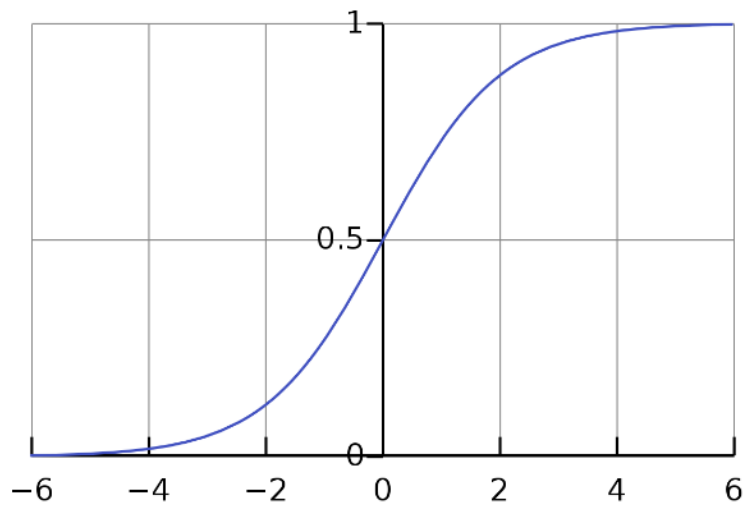


FIGURA 6: RAPPRESENTAZIONE GRAFICA DI UNA FUNZIONE LOGISTICA. FONTE: GOVONI L., (2018), COME FUNZIONA UN ALGORITMO DI REGRESSIONE LOGISTICA, BUSINESS E TECNOLOGIA, 03/09/2018.

Essa quindi è una tecnica che permette di modellizzare la probabilità che un evento si verifichi a seconda dei valori delle variabili indipendenti⁸⁸.

Se applicata alla previsione dei movimenti degli indici azionari, considera come variabili indipendenti il valore di apertura, quello di chiusura e il massimo e minimo valore assunto dall'indice analizzato e il prezzo di chiusura del giorno successivo definita come variabile dipendente dicotomica – *bull or bears*. Ricorrendo all'equazione di regressione ottenuta dal set dei dati di addestramento, è possibile formulare una funzione di risposta logistica per prevedere il valore della variabile dipendente: se superiore allo 0,5 il giorno considerato viene classificato in uno stato rialzista (*bull*), altrimenti in uno stato di ribasso (*bears*)⁸⁹.

Un'ulteriore tecnica di classificazione è rappresentata dai classificatori *Nearest Neighbor* il cui nome definisce la loro peculiarità ovvero la capacità, che li contraddistingue, di assegnare la classe più simile a dati non classificati in base a quelli che già compongono le classi. Essi sono adatti per le attività in cui le relazioni tra le caratteristiche e le classi target sono numerose, complicate o difficili da comprendere.

⁸⁸ Govoni L., (2018), *Come funziona un algoritmo di regressione logistica*, Business e Tecnologia, 03/09/2018.

⁸⁹ Subha M. V., Nambi S.T., (2012), *Classification of Stock Index movement using k-Nearest Neighbours (k-NN) algorithm*, Wseas Transaction on Information Science and applications, 09/09/2012.

Già utilizzati con successo per applicazioni di visione artificiale, identificazione di schemi nei dati genetici o per la previsione dei gusti cinematografici di un individuo per un film che è stato raccomandato⁹⁰.

L'algoritmo maggiormente utilizzato in tale tecnica è il kNN; un algoritmo semplice ed efficace che non formula ipotesi sulla distribuzione dei dati sottostanti e caratterizzato da una fase di allenamento celere, ma per contro non produce un modello che limita la capacità di trovare nuove intuizioni in relazioni con le caratteristiche. Esso è caratterizzato da una fase di classificazione lenta che richiede così una grande quantità di memoria e nel caso di funzionalità nominale e dati mancanti, richiede un'ulteriore elaborazione⁹¹. Esso presuppone che gli elementi simili siano vicini gli uni agli altri catturando l'idea di somiglianza, detta distanza⁹². Supponendo di avere un set di dati contenente informazioni non classificate, ma che hanno caratteristiche simili al dataset esistente, l'algoritmo kNN identifica k unità statistiche con caratteristiche più simili alle classi target determinando così la bontà del modello nel generalizzare i dati futuri: maggiore risulta tale parametro, minore sarà la varianza causata da eventuali errori e viceversa, la scelta di un k piccolo influenzerà indebitamente la classificazione. Inoltre l'individuazione dei *nearest neighbor* richiede una funzione di distanza che tradizionalmente si identifica nella distanza euclidea⁹³ e specificata dalla formula:

$$\text{dist}(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2}$$

Essa ricorda la classica formula per il calcolo di un segmento avente come estremi due punti, solo che in tale caso è estesa a uno spazio n – dimensionale⁹⁴.

⁹⁰ Lantz B., (2013), *Machine Learning with R*, Packt Publishing, Birmingham.

⁹¹ Ibidem.

⁹² Harrison O., (2018), *Machine Learning Basics with the K – Nearest Neighbors Algorithm*, Towards Data Science, 10/09/2018.

⁹³ Lantz B., (2013), *Machine Learning with R*, Packt Publishing, Birmingham.

⁹⁴ Govoni L., (2018), *Algoritmo K-Nearest Neighbors (KNN)*, Business e Tecnologia, 28/01/2019.

In termini grafici l'applicazione di tale algoritmo può essere rappresentata come segue:

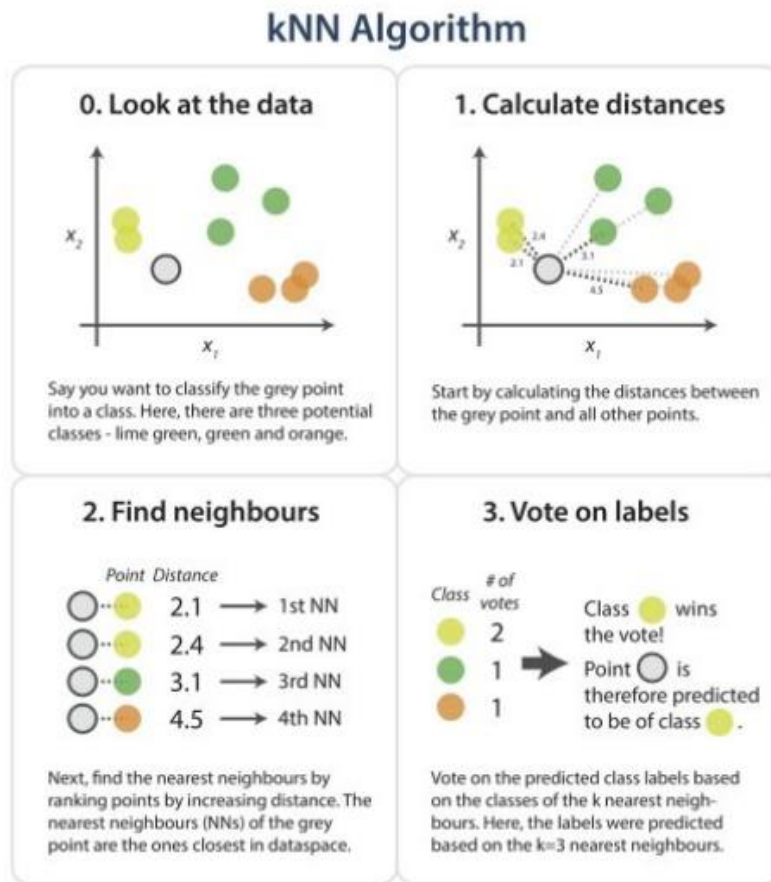


FIGURA 7: RAPPRESENTAZIONE GRAFICA DELL'APPLICAZIONE DELL'ALGORITMO KNN. FONTE: HARRISON O., (2018), MACHINE LEARNING BASICS WITH THE K – NEAREST NEIGHBORS ALGORITHM, TOWARDS DATA SCIENCE, 10/09/2018.

Recenti ricerche sulla previsione dei prezzi di mercato azionario hanno condotto a risultati interessanti: come si evince dal paper *“Stock Price Prediction Using K – Nearest Neighbor (Knn) Algorithm”*, pubblicato nell’*International Journal of Business, Humanities and Technology* nel 2013, tale algoritmo risulta stabile e robusto con un piccolo rapporto di errore, fornendo così risultati razionali e ragionevoli. Inoltre, a seconda dei dati, i risultati della previsione risultano prossimi ai prezzi osservati⁹⁵.

Gli alberi decisionali (o *Decision Tree*) sono un modello di apprendimento automatico supervisionato utilizzato con il fine di prevedere un certo obiettivo imparando da delle regole di decisione derivanti dalle funzionalità. Come lo stesso nome

⁹⁵ Alkhatib K., Najadat H., Hmeidi I. e Shatnawi M., (2013), *“Stock Price Prediction Using K – Nearest Neighbor (Knn) Algorithm”*, *International Journal of Business, Humanities and Technology*, 3, pp. 32 - 33.

suggerisce, è possibile pensare a tale modello come una scomposizione dell'insieme di dati: tale modello quindi, ad ogni punto, considera una serie di domande in grado di partizionare il dataset che si ha a disposizione; scegliendo una determinata domanda che fornisce la suddivisione migliore, si trovano nuovamente le domande migliori per le partizioni seguenti⁹⁶.

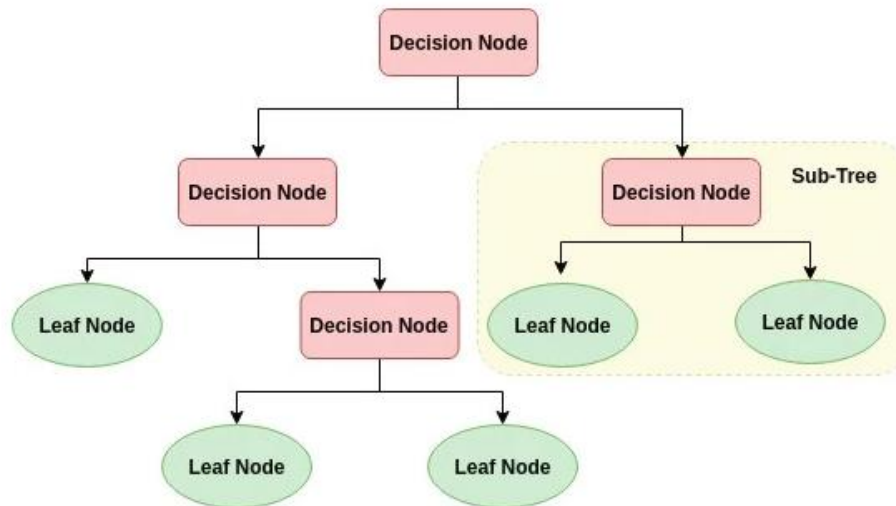


FIGURA 8: RAPPRESENTAZIONE GRAFICA DEL FUNZIONAMENTO DI UN ALBERO DECISIONALE. FONTE: LE J., (2018), DECISIONE TREES IN R, DATACAMP, 19/06/2018.

Essi quindi, sono costruiti da un partizionamento ricorsivo che assume la forma di un diagramma a flusso imitando facilmente il pensiero a livello umano e risultando così facile da comprendere ed interpretare, come emerge dalla rappresentazione grafica precedente. Come si può notare il *root node* (o nodo radice) rappresenta l'intero campione di dati il quale si divide in due o più insiemi omogenei detti *decision node* (o nodi decisionali), i quali a loro volta comportano ulteriori nodi detti *leaf* (o nodi terminali) che non prevedono ulteriori divisioni⁹⁷. In altre parole, ogni nodo rappresenta una singola variabile di input (X) e un punto di divisione su quella variabile supponendo che la variabile sia numerica, mentre i nodi terminali dell'albero contengono una variabile di output (Y): la creazione di un albero decisionale è sostanzialmente un processo di divisione dello spazio di input che si arresta nel momento in cui l'euristica,

⁹⁶ Le J., (2018), *Decisione Trees in R*, DataCamp, 19/06/2018.

⁹⁷ Ibidem.

utilizzata per selezionare il criterio di divisione, fornisce a ogni nodo terminale un numero di osservazioni inferiore a un minimo di osservazioni⁹⁸.

Infine possono essere utilizzati nella classificazione al fine di prevedere una risposta qualitativa derivante da un set di dati discreti: tutte le variabili di dati vengono assegnate al nodo principale, il quale poi viene suddiviso in ulteriori nodi in base al criterio della suddivisione che divide i dati caratterizzandoli attraverso una specifica domanda; a seconda della risposta, sì o no, i dati vengono posizionati nel nodo secondario sinistro o in quello destro. La divisione continua fino a che i nodi secondari risultino privi di errori o fino a quando il numero di dati all'interno di ciascuno di essi non raggiunge un numero predefinito⁹⁹. Gli alberi decisionali possono essere utilizzati anche nella regressione nella quale il processo di costituzione è il medesimo; si differenzia solo dal fatto che il risultato che si ottiene è di tipo quantitativo e deriva da un set di dati continuo.

Da una recente ricerca condotta al fine di valutare le capacità degli alberi decisionali di prevedere la direzione del movimento degli indici di mercato, è emerso che sebbene tale tecnica sia facilmente interpretabile e comprensibile, non presenta lo stesso livello di precisione predittiva di altri approcci in quanto questi non sono modelli robusti e una piccola modifica dei dati può causare importanti cambiamenti nell'albero finale stimato¹⁰⁰. Aggregando molti alberi decisionali, utilizzando tecniche come le *Random Forest* (o Foreste casuali), le prestazioni predittive possono essere sostanzialmente migliorate¹⁰¹.

La *Random Forest* è un tipo di algoritmo di apprendimento automatico supervisionato basato sull'apprendimento d'insieme in quanto, il modello di previsione, deriva dall'unione ripetitiva di diversi tipi di algoritmi o del medesimo algoritmo. Essa è costituita da molteplici alberi decisionali generalmente addestrati con il metodo del *bagging* (o aggregazione) la cui idea di base è che la combinazione di modelli di

⁹⁸ Brownlee J., (2016), *Classification and Regression Trees for Machine Learning*, Machine Learning Mastery, 08/04/2016.

⁹⁹ Ibidem.

¹⁰⁰ Imandoust S. B., Bolandraftar M., (2014), "Forecasting the direction of stock market index movement using three data mining techniques: the case of Tehran Stock Exchange", *International Journal of Engineering and Applications*, 4, p. 106-117.

¹⁰¹ Le J., (2018), *Decision Trees in R*, DataCamp, 19/06/2018.

apprendimento aumenta la precisione del risultato, nonché della previsione finale riducendone la varianza¹⁰². A livello grafico, una *Random Forest* apparirebbe come segue:

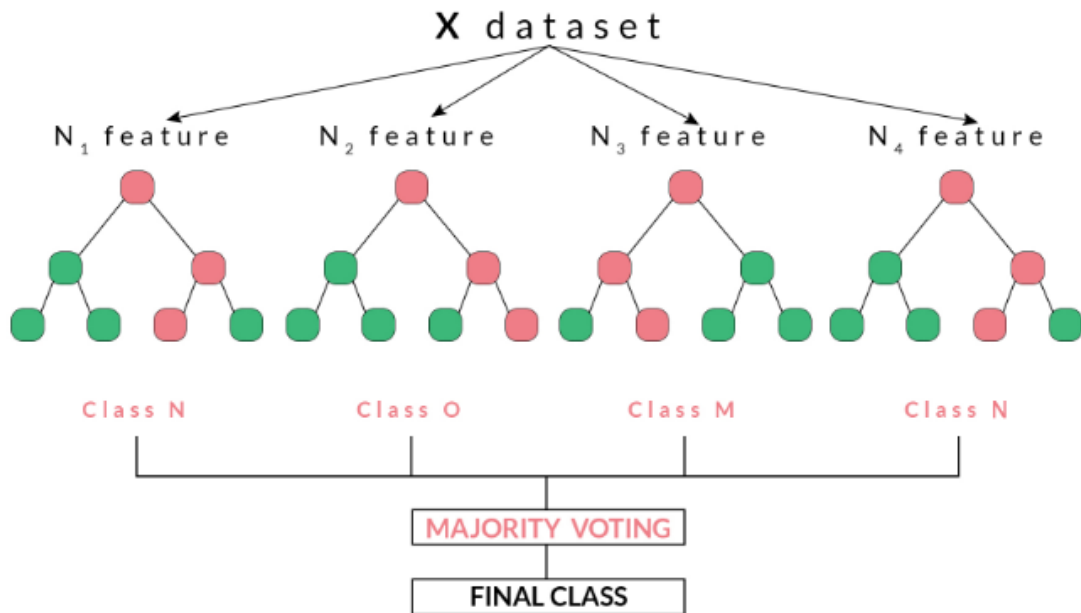


FIGURA 9: RAPPRESENTAZIONE GRAFICA DI UNA RANDOM FOREST COSTITUITA DA QUATTRO ALBERI DECISIONALI. FONTE: [HTTPS://BLOG.QUANTINSTI.COM/RANDOM-FOREST-ALGORITHM-IN-PYTHON/](https://blog.quantinsti.com/random-forest-algorithm-in-python/).

È un algoritmo utilizzabile sia nei problemi di classificazione che in quelli di regressione che costituiscono la maggior parte degli attuali sistemi di apprendimento automatico. Essa, come emerso da un recente progetto di ricerca attinente alla previsione per il marketing azionario, presenta notevoli vantaggi rispetto ad altre metodologie come la regressione logistica o gli alberi decisionali: è un algoritmo non distorto in quanto, ogni albero che lo compone, viene addestrato su un sottoinsieme di dati, è molto stabile infatti l’inserimento di un nuovo dato nel dataset originario non influenza l’algoritmo generale (potrebbe avere un notevole impatto su un albero, ma difficilmente causerebbe cambiamenti su tutti gli alberi che compongono l’algoritmo) e infine, funziona bene quando si hanno a disposizione dati categorici e/o numerici e perfino, anche in mancanza di dati¹⁰³.

¹⁰² Deng H., (2018), *An Introduction to Random Forest*, Towards Data Science, 07/12/2018.

¹⁰³ Jain S., Kain M., (2018), "Prediction for Stock Marketing Using Machine Learning", *International Journal on Recent and Innovation Trends in Computing and Communication*, 6, pp. 131-135.

Per contro però, tale algoritmo richiede molte risorse computazionali e tempistiche a causa dell'unione di un elevato numero di alberi decisionali e comporta una precisione minore, se non inaccettabile, se confrontata con altre tecniche¹⁰⁴.

I classificatori basati sui metodi bayesiani sono uno dei più efficaci ed efficienti algoritmi utilizzati per la risoluzione dei problemi di classificazione. Essi sono una raccolta di algoritmi che utilizzano un approccio probabilistico e che condividono un principio comune, ovvero il Teorema di Bayes. Tale approccio è basato sull'assunzione che tutte le grandezze considerate sono caratterizzate da distribuzioni di probabilità e conseguentemente, una decisione ottimale può essere assunta tramite l'analisi di tali probabilità e del dataset che si ha a disposizione¹⁰⁵. Esso fornisce quindi un metodo per il calcolo della probabilità di un'ipotesi conoscendo la sua probabilità a priori e la probabilità di osservare un certo insieme di dati, avendo a disposizione l'ipotesi e i dati osservati¹⁰⁶.

La formula del teorema di Bayes è la seguente:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

I metodi di apprendimento Bayesiano sono molteplici e accomunati da diverse caratteristiche. Ciascun esempio contenuto all'interno del *training set* contribuisce ad incrementare o decrementare la probabilità che una certa ipotesi sia corretta, la conoscenza a priori può essere combinata con l'analisi dei dati osservati per il calcolo della probabilità finale di un'ipotesi e la conoscenza a priori viene ricavata fornendo per ognuna delle possibili ipotesi, una probabilità a priori e una distribuzione di probabilità sui dati osservati. Essi inoltre, favoriscono ipotesi che effettuano predizioni

¹⁰⁴ Imandoust S. B., Bolandraftar M., (2014), "Forecasting the direction of stock market index movement using three data mining techniques: the case of Tehran Stock Exchange", *International Journal of Engineering and Applications*, 4, p. 106-117.

¹⁰⁵ Brownlee J., (2016), *Naive Bayes for Machine Learning*, Machine Learning Mastery, 11/04/2016.

¹⁰⁶ Stefenelli M., *Apprendimento automatico nei giochi di strategia*, tesi di laurea, Università di Bologna, a. a. 2013-2014, relatore Roli A.

probabilistiche e permettono di classificare nuovi dati combinando le predizioni di diverse ipotesi, pesate con le loro probabilità¹⁰⁷.

L'algoritmo più utilizzato è il *Naive Bayes*, il quale presuppone che tutte le caratteristiche nel set di dati siano ugualmente importanti e indipendenti, presupposti però raramente verificati nella maggior parte delle applicazioni del mondo reale. Essi sono semplici e veloci da utilizzare, funzionano bene anche in presenza di errori o dati mancanti, richiedono relativamente pochi esempi per il loro addestramento e stimare la probabilità di una previsione non comporta calcoli complessi¹⁰⁸ ma per contro, come evidenzia una recente ricerca attinente la previsione del movimento del mercato azionario utilizzando un modello *Naive Bayes*, essi non sono adatti a set di dati di grandi dimensioni e comportano una notevole riduzione di precisione rispetto ad altre tecniche, quali ad esempio le *Support Vector Machine*¹⁰⁹.

Un *Support Vector Machine* (o SVM) è un classificatore discriminante, utilizzabile sia per i problemi di classificazione che di regressione, e definito formalmente da un iperpiano di separazione. In altre parole, avendo un set di dati classificati di addestramento, l'algoritmo produce un iperpiano ottimale che permette di classificare nuovi dati. In uno spazio n – dimensionale, dove n rappresenta il numero di variabili, tale algoritmo classifica distintamente i dati che vengono forniti¹¹⁰. Il compito principale dell'algoritmo è quello di trovare l'iperpiano che abbia il margine massimo, ovvero la distanza massima tra esso e il punto più vicino su ciascun lato; questo consentirà di incrementare l'accuratezza e la sicurezza di classificazione per i dati futuri. La dimensione dell'iperpiano inoltre, dipende dal numero di funzioni presenti: in presenza di due funzioni esso è solo una linea, con tre funzioni assumerà la forma di un piano dimensionale e così via¹¹¹.

¹⁰⁷ Stefenelli M., *Apprendimento automatico nei giochi di strategia*, tesi di laurea, Università di Bologna, a. a. 2013-2014, relatore Roli A.

¹⁰⁸ Lantz B., (2013), *Machine Learning with R*, Packt Publishing, Birmingham.

¹⁰⁹ Mahajan S., Deshmukh K., Thite P., Bhavana S., (2016), *Stock Market Prediction and Analysis Using Naïve Bayes*, International Journal on Recent and Innovation Trends in Computing and Communication, Working Paper No.4.

¹¹⁰ Patel S., (2017), *Support Vector Machine – Theory*, Medium, 03/05/2017.

¹¹¹ Gandhi R., (2018), *Support Vector Machine – Introduction to Machine Learning Algorithms*, Towards Data Science, 07/06/2018.

A titolo esemplificativo, il processo di una SVM è rappresentato come segue:

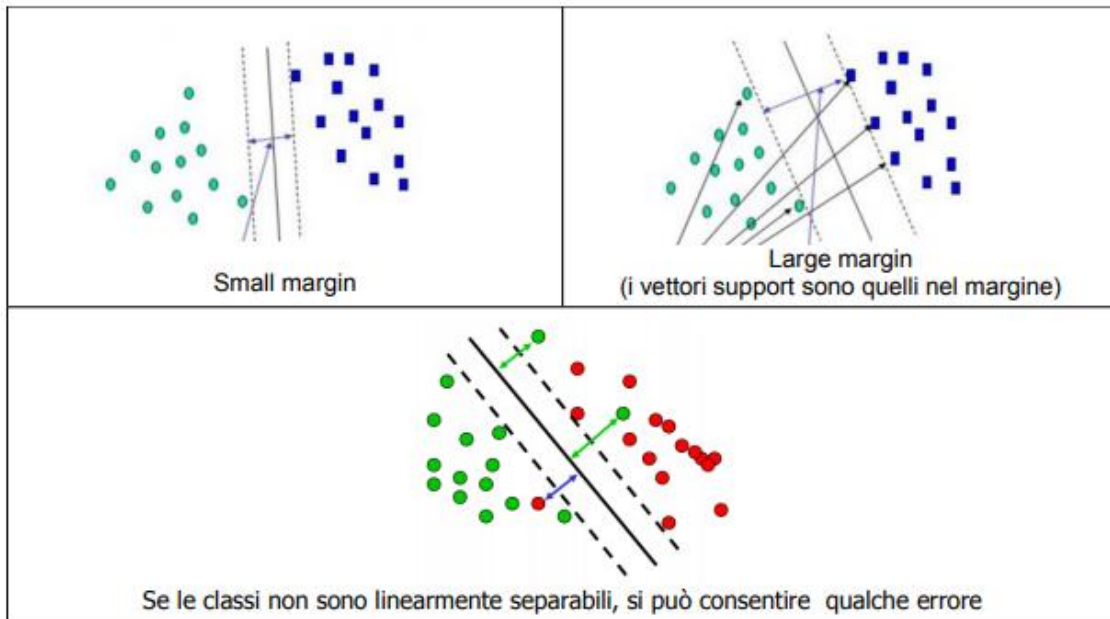


FIGURA 10: RAPPRESENTAZIONE GRAFICA DEL CLASSIFICATORE A MARGINE MASSIMO. FONTE: GANDHI R., (2018), SUPPORT VECTOR MACHINE – INTRODUCTION TO MACHINE LEARNING ALGORITHMS, TOWARDS DATA SCIENCE, 07/06/2018.

Un'ulteriore metodologia di apprendimento automatico è costituita dalle Reti Neurali Artificiali (o RNA), alle quali, essendo il principale oggetto di studio della presente tesi, verranno dedicati i successivi due capitoli.

2.2 Modalità di apprendimento non supervisionato

La prima metodologia di apprendimento automatico che viene analizzata è quella non supervisionata, detta anche *Unsupervised Machine Learning*. In tale approccio, ogni problema che viene analizzato comincia con il dover osservare le caratteristiche dei dati senza avere a disposizione informazioni o esempi che costituirebbero esperienza pregressa¹¹².

2.2.1 Tipologie di apprendimento non supervisionato

Alcune applicazioni di tecniche di apprendimento automatico senza supervisione includono:

- il clustering: tecnica che consente il raggruppamento automatico del set di dati in gruppi (detti cluster) in base alla somiglianza ovvero basandosi sulle informazioni che li descrivono e sulle relazioni esistenti tra di essi. I dati appartenenti allo stesso gruppo dovranno essere caratterizzati dalle medesime caratteristiche e/o proprietà, mentre quelli appartenenti a gruppi diversi avranno caratteristiche e/o proprietà diametralmente opposte¹¹³.

Sono tecniche semplici ed efficaci e ne esistono diverse tipologie:

- esclusivo: metodo che raggruppa i dati in modo tale che una singola informazione possa appartenere a un solo cluster o gruppo. Un esempio di algoritmo basato su tale tipologia di cluster è il *K-means*. Esso è un algoritmo, progettato nel 1967 da MacQueen, basato sul centroide o punto medio, il cui obiettivo è determinare gruppi di dati generati da distribuzioni gaussiane i cui attributi sono rappresentati come vettori. Esso si propone di minimizzare la varianza intra-cluster tenendo presente che ogni cluster viene identificato mediante un punto medio¹¹⁴. Esso segue una procedura iterativa che consente inizialmente di creare k partizioni e assegnando ad ognuna di esse i punti di

¹¹² Mohammed M., Khan B. M., e Bashier M.B.E., (2017), *Machine Learning. Algorithms and Applications*, CRC Press, Boca Raton

¹¹³ Seif G., (2019), *Una facile introduzione all'apprendimento senza supervisione con 4 tecniche di base*, Towards Data Science, 19/08/2019.

¹¹⁴ Govoni L., (2019), *Algoritmo k – means: cose 'è e come funziona?*, Business e Tecnologia, 02/07/2018.

ingresso casualmente o utilizzando informazioni euristiche, successivamente è necessario calcolare il centroide di ogni gruppo con la costituzione conseguente di una nuova partizione che associa ogni punto di ingresso al cluster il cui centroide è più vicino ad esso e infine, ricalcola i centroidi per i nuovi cluster affinché l'algoritmo non converga¹¹⁵. A titolo esemplificativo, la dinamica utilizzata dal clustering esclusivo, viene riportata di seguito:

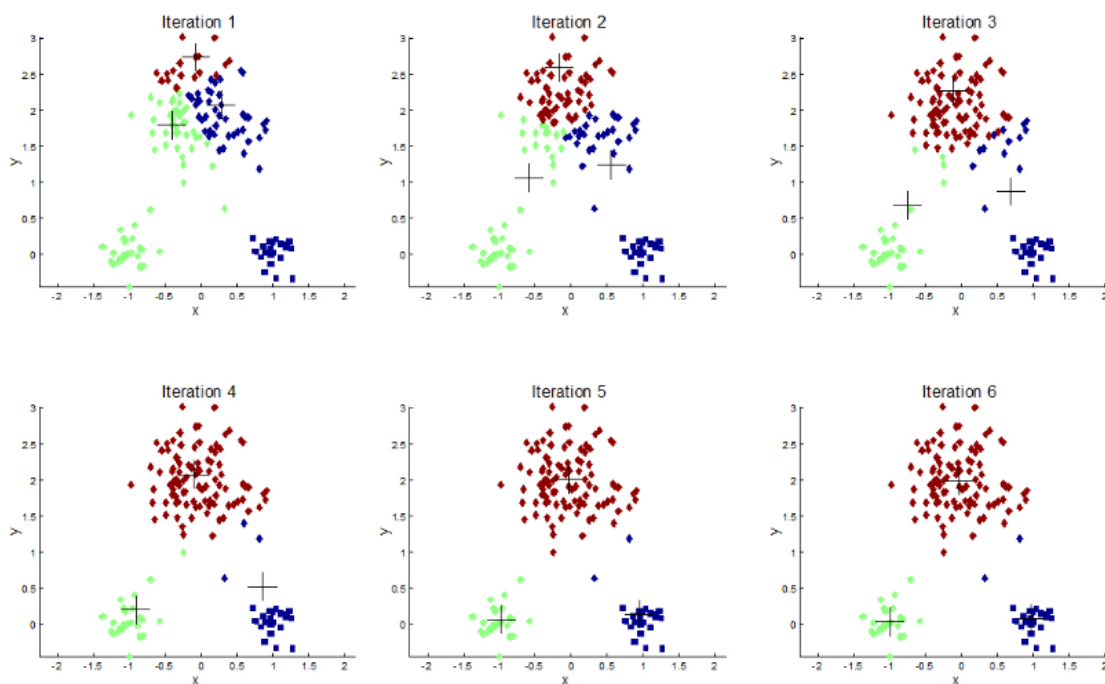


FIGURA 11: RAPPRESENTAZIONE GRAFICA DEL FUNZIONAMENTO DELL'ALGORITMO K – MEANS. FONTE: STANFORD.EDU.

È un algoritmo applicabile in molti scenari come ad esempio per la segmentazione della clientela, la classificazione di documenti, la rilevazione di frodi assicurative o l'identificazione della località dove avvengono, con maggiore frequenza, attività criminali¹¹⁶.

Oltre alla versatilità che lo contraddistingue, esso è efficiente nel gestire quantità molto rilevanti di dati, permette di variare la posizione iniziale dei centroidi per ridurre la dipendenza dalle condizioni iniziali e spesso si risolve con un ottimo

¹¹⁵ Govoni L., (2019), *Algoritmo k – means: cose 'è e come funziona?*, Business e Tecnologia, 02/07/2018.

¹¹⁶ Ibidem.

locale¹¹⁷. Presenta però anche dei limiti, infatti funziona solo su valori numerici, i cluster hanno una forma convessa quindi la sua applicazione per la ricerca di cluster di altra forma risulta difficile ed è necessaria, *ex – ante*, la fissazione di k cluster¹¹⁸.

- gerarchico: in questa tipologia, il clustering può essere agglomerativo se inizia con un numero fisso di cluster, ognuno dei quali rappresenta un singolo oggetto. Successivamente inizia il processo di agglomerazione, infatti ad ogni iterazione vengono accorpati i cluster più simili aumentando di volta in volta la soglia di similarità. L'iterazione termina quando tutti i dati sono presenti in un unico cluster¹¹⁹. Un esempio di algoritmo basato sul cluster agglomerativo è l'algoritmo AGNES che utilizza il metodo del *single - linkage*; in questo algoritmo la distanza tra i cluster corrisponde alla distanza minima fra due punti non appartenenti allo stesso cluster. A ogni iterazione i due cluster più vicini vengono fusi tra di loro¹²⁰.

Una seconda variante di cluster gerarchico è quello divisivo che agisce in maniera inversa rispetto al precedente. Lo stato di partenza in questo caso è un unico cluster contenente tutte le unità statistiche e solo successivamente divide i cluster fino al raggiungimento di una determinata condizione. Un esempio di algoritmo basato su tale tipologia di cluster è l'algoritmo DIANA, introdotto da Kaufmann e Rousseeuw nel 1990.

Per entrambi le tipologie di cluster gerarchico, il risultato finale è un insieme di cluster contenenti uno o più oggetti e il quale, può essere rappresentato da un dendrogramma, un diagramma ad albero, verticale o orizzontale che permette di rappresentare la successione delle partizioni. Le “radici” sono le unità iniziali e a

¹¹⁷Sia $x \in F$. Se esiste $\varepsilon > 0$ tale che $f(x) \leq f(y)$, per ogni $y \in F$ con $||x - y|| \leq \varepsilon$, allora x prende il nome di ottimo locale. Naturalmente un ottimo globale è anche un ottimo locale, ma in generale non è vero il viceversa. L'equivalenza si ha solo se F e f hanno proprietà di convessità.

L'ottimo globale si può rincorrere con tecniche standard come annealing simulato o algoritmi genetici. Fonte: Cotronei M., *Introduzione alla Ricerca Operativa*, Facoltà di Ingegneria dell'Università degli Studi “Mediterranea” di Reggio Calabria.

¹¹⁸ Govoni L., (2019), *Algoritmo k – means: cose ‘è e come funziona?*, Business e Tecnologia, 02/07/2018.

¹¹⁹ Di Pietro S., (2019), *Introduzione alla cluster analysis*, Italian Association for Machine Learning, 05/02/2019.

¹²⁰ Ibidem.

livelli crescenti di distanza di uniscono o si separano i gruppi tra loro, rispettivamente per il cluster agglomerativo e per il cluster divisivo.

A titolo esemplificativo si riportano delle rappresentazioni grafiche:

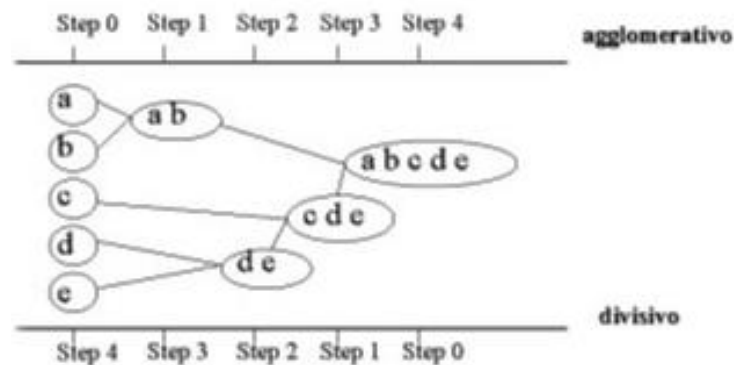


FIGURA 12: RAPPRESENTAZIONE GRAFICA DEL CONFRONTO TRA IL CLUSTER GERARCHICO AGGLOMERATIVO E IL CLUSTER GERARCHICO DIVISIVO. FONTE: [HTTPS://WWW.RESEARCHGATE.NET/PUBLICATION/332291727](https://www.researchgate.net/publication/332291727).

Questa è una metodologia che permette di ottenere una gerarchia organizzata di gruppi anziché un insieme amorfo di cluster, è facile da utilizzare ma purtroppo non si basa su una teoria statistica o dell'informazione, l'algoritmo non può mai annullare ciò che è stato fatto in precedenza, nessuna funzione obiettivo è minimizzata direttamente e presenta una scarsa efficienza se la numerosità campionaria è elevata.

- basato sulla densità: gli algoritmi basati sulla densità sono caratterizzati dalla capacità di identificare cluster di forma arbitraria, definiti come regioni ad alta densità di punti distribuiti arbitrariamente e separati da regioni a bassa densità¹²¹. Questi sono in grado di gestire in maniera efficiente gli errori e gli eventuali punti anomali e non richiedono *ex-ante* la definizione di un numero di cluster. La criticità che presentano, è il fatto di dover determinare la soglia di densità¹²². Un algoritmo utilizzato in questa tipologia è l'algoritmo DBSCAN; esso divide il set di dati in n dimensioni, successivamente, per ogni punto nel set di dati, crea una forma n dimensionale e conta il numero di punti che rientrano in

¹²¹ Research Gate (2018), *Data Clustering Spaziale basati su densità: indagine esplorativa delle tecniche di Data Clustering spaziale e metriche di distanza*, disponibile a <https://www.researchgate.net/publication/332291727>.

¹²² Lutins E., (2017), *DBSCAN: What is it? When to use it? How to use it.*, Medium, 06/09/2017.

quella forma, detta cluster. Poi espande quest'ultimo in modo iterativo esaminando ogni singolo punto al suo interno e tenendo conto degli altri punti nelle vicinanze¹²³. Tale algoritmo svolge in maniera efficiente la separazione dei cluster ad alta densità rispetto a quelli a bassa densità all'interno di un determinato set di dati e gestisce in maniera eccellente i valori anomali, ma per contro non funziona bene con i cluster a densità variabile e con set di dati di dimensioni molto elevate¹²⁴. A titolo di esempio, di seguito si riporta una rappresentazione grafica del funzionamento di tale algoritmo:

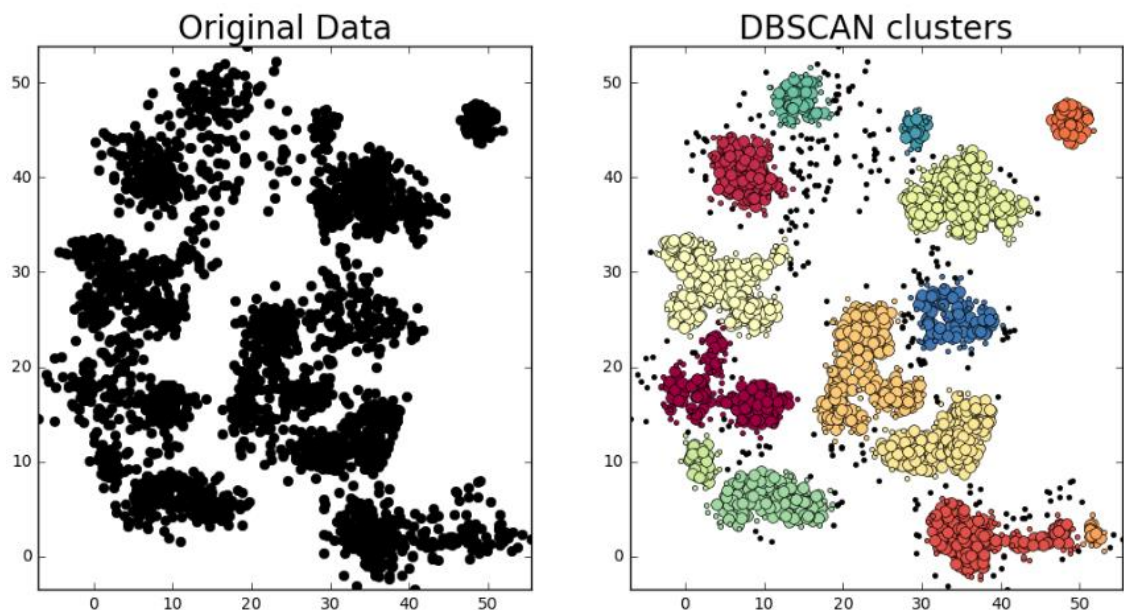


FIGURA 13: RAPPRESENTAZIONE GRAFICA DEL FUNZIONAMENTO DELL'ALGORITMO DBSCAN.
Fonte: LUTINS E., (2017), DBSCAN: WHAT IS IT? WHEN TO USE IT? HOW TO USE IT.,
MEDIUM, 06/09/2017.

- basato sulla distribuzione: questo approccio presuppone che i dati siano composti da distribuzioni, quindi ogni cluster è rappresentato da una distribuzione di probabilità diversa, ognuna delle quali descrive la distribuzione dei valori per i membri di quel cluster¹²⁵. Ogni distribuzione fornisce la probabilità che un determinato elemento del dataset abbia certi valori, supponendo che sia noto a quale cluster appartiene. Un esempio di algoritmo utilizzato in questi casi è l'algoritmo EM (*Expectation – Maximization*) il quale si basa su distribuzioni

¹²³ Lutins E., (2017), DBSCAN: *What is it? When to use it? How to use it.*, Medium, 06/09/2017.

¹²⁴ Ibidem.

¹²⁵ Seif G., (2018), *The 5 Clustering Algorithms Data Scientists Need to Know*, Towards Data Science, 05/02/2018.

gaussiane: i dati vengono utilizzati come modello per risolvere il problema, cioè stimare i vari parametri di ogni distribuzione¹²⁶. Esso calcola la densità di probabilità di appartenenza ai cluster per ogni componente, a partire dai parametri e solo successivamente stima questi ultimi tenendo conto delle rispettive probabilità di appartenenza.

Nella figura che viene riportata di seguito, l'algoritmo raggruppa i dati in tre distribuzioni gaussiane. All'aumentare della distanza dal centro della distribuzione, diminuisce la probabilità che un punto appartenga ad essa. Nel momento in cui però non sono note le distribuzioni, è necessario utilizzare un algoritmo diverso¹²⁷.

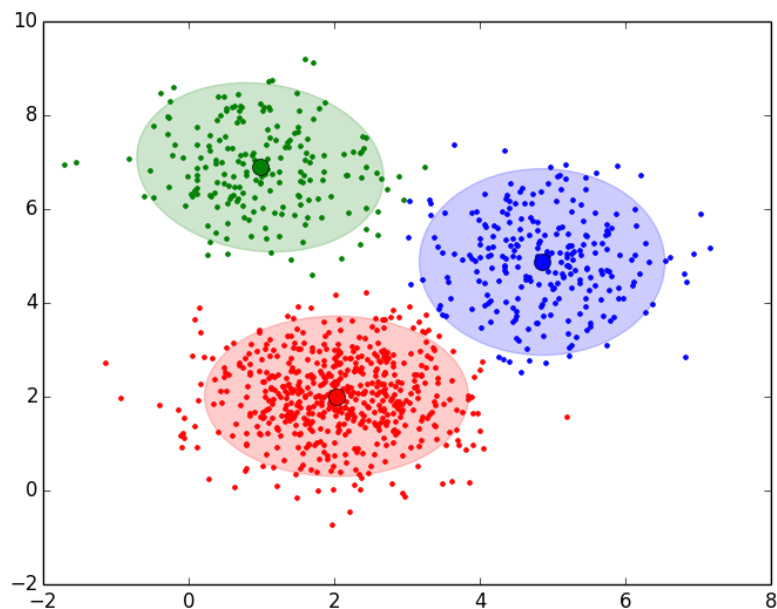


FIGURA 14: RAPPRESENTAZIONE GRAFICA DELL'ALGORITMO EM. FONTE: BROWNLEE J., (2019), A GENTLE INTRODUCTION TO EXPECTATION-MAXIMIZATION (EM ALGORITHM), MACHINE LEARNING MASTERY, 01/11/2019.

Esso rappresenta un algoritmo di raffinamento iterativo che viene utilizzato per il calcolo delle stime dei parametri, ma non viene garantita l'ottimalità del risultato.

¹²⁶ Brownlee J., (2019), *A Gentle Introduction to Expectation-Maximization (EM Algorithm)*, Machine Learning Mastery, 01/11/2019.

¹²⁷ Ibidem.

Una seconda tecnica di apprendimento non supervisionato è una tecnica di *mapping* dei dati e quindi una delle operazioni di pre – elaborazione del dataset nell'apprendimento automatico non supervisionato ovvero:

- la riduzione della dimensionalità: quando il numero n di variabili osservate sulle unità statistiche è elevato, i dati sono dispersi nello spazio n – dimensionale. In circostanze di questo genere, gli eventuali algoritmi che verranno impiegati necessiteranno di più tempo e memoria; con tale tecnica quindi, si riduce la dimensione del dataset, senza perdere le informazioni rilevanti, si eliminano gli eventuali errori e si combinano le informazioni correlate¹²⁸. Esistono diversi metodi di riduzione dimensionale:

- Analisi delle componenti principali (PCA): metodo introdotto da Karl Pearson e detta anche trasformazione discreta di Karhunen – Loeve, è utilizzata nell'ambito della statistica multivariata per la semplificazione dei dati d'origine. Avviene tramite la combinazione lineare delle variabili¹²⁹.

Nel caso di una matrice di dati $n \times p$ si immaginano n punti in uno spazio p -dimensionale. Se p è elevato, l'obiettivo è quello di effettuare una riduzione delle dimensioni pur mantenendo, nel miglior modo possibile la struttura dei punti originari. Le nuove dimensioni sono individuate dalle componenti principali¹³⁰.

Utilizzando tale tecnica quindi si avrà una riduzione del volume dei dati che si andranno ad analizzare, riducendo così la difficoltà computazionale dell'algoritmo di apprendimento, ma per contro, tale diminuzione dimensionale può deteriorare le informazioni predittive del medesimo algoritmo. Di seguito, a titolo di esempio, viene riportata una rappresentazione grafica del funzionamento e del risultato ottenuto grazie all'utilizzo di tale metodologia.

¹²⁸ Raschka S., (2016), *Machine Learning con Python. Costruire algoritmi per generare conoscenza*, Apogeo, Milano.

¹²⁹ James G., Witten D., Hastie T., Tibshinari R., (2013), *An Introduction to Statistical Learning with Applications in R*, Springer, Londra.

¹³⁰ Ibidem.

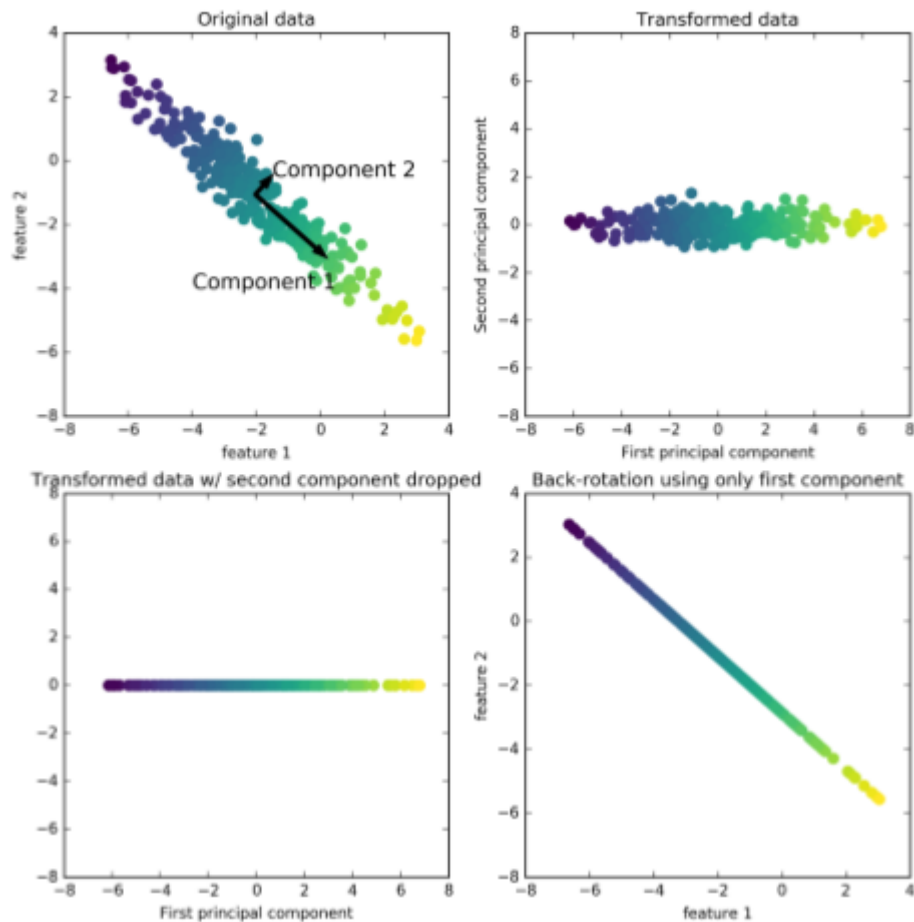


FIGURA 15: RAPPRESENTAZIONE GRAFICA DELLA TRASFORMAZIONE DEI DATI CON L'ALGORITMO PCA.
Fonte: James G., Witten D., Hastie T., Tibshirani R., (2013), An Introduction to Statistical Learning with Applications in R, Springer, Londra.

- Apprendimento tramite le regole di associazione ovvero una tecnica che permette di determinare relazioni e dipendenze all'interno del dataset¹³¹. L'obiettivo di tale metodologia non è quello di prevedere un determinato valore, ma fornire informazioni affini rilevanti per un importante sottoinsieme dei dati disponibili: sostanzialmente si basa sulle relazioni causa/effetto¹³². L'algoritmo più utilizzato per tale metodologia è l'algoritmo A – priori: ideato da R. Agrawal e R. Srikant nel 1994, esso identifica una caratteristica particolare di un set di dati e tenta di determinare la frequenza con cui tale caratteristica viene visualizzata in tutto il set ed è utilizzato principalmente per l'ordinamento di grandi quantità di dati¹³³.

¹³¹ Raschka S., (2016), *Machine Learning con Python. Costruire algoritmi per generare conoscenza*, Apogeo, Milano.

¹³² Ibidem.

¹³³ Wingate J., (2018), *Apriori Machine Learning Algorithm*, R – algo, 01/02/2018.

2.3 Modalità di apprendimento per rinforzo

La seconda metodologia di apprendimento automatico che viene analizzata è l'apprendimento per rinforzo o *Reinforcement learning*. Considerato inizialmente dalla letteratura come un caso particolare di apprendimento supervisionato, oggi esso è utilizzato in diversi ambiti nei quali risulta inefficiente e inefficace un'elaborazione simbolica ovvero la codifica di uno specifico problema e la sua successiva elaborazione tramite un esplicito algoritmo che offre un adeguato output ad ogni possibile input del problema¹³⁴. Tale tecnica ha come obiettivo quello di realizzare algoritmi di apprendimento capaci di adattarsi ai mutamenti dell'ambiente esterno; processo divenuto possibile in quanto è prevista la ricezione di *feedback* esterni generati direttamente dall'ambiente, a seconda delle scelte intraprese dall'algoritmo. Rappresenta essenzialmente l'approccio comportamentale che gli esseri umani adottano nell'ambiente in cui vivono: a una scelta corretta corrisponde un premio, mentre alla scelta inadatta corrisponde una penalità e ciò porta all'ottenimento del miglior risultato ottenibile aumentando la probabilità che le scelte corrette si ripetano e riducendo la probabilità di una ripetizione dell'errore¹³⁵.

L'apprendimento per rinforzo è caratterizzato principalmente da tre filoni di ricerca paralleli che, solo alla fine degli anni 80' si fondono, formando la concezione moderna che lo qualifica. Il approccio si basava sulle possibili applicazioni del *Trial – and – error Learning* applicato inizialmente all'ambito della psicologia dell'apprendimento animale; tale concetto descrive essenzialmente il rafforzamento di uno schema comportamentale come risultato della ricezione di uno stimolo da parte dell'individuo, diretta conseguenza dei cambiamenti dell'ambiente che lo circonda: l'individuo interpreta gli stimoli agendo di conseguenza, ma mantenendo allo stesso tempo, memoria delle azioni passate e ricercando la risposta migliore allo stimolo che l'ambiente invia in quell'istante temporale. Tramite gli errori comportamentali, l'individuo migliorerà la sua capacità adattiva¹³⁶.

¹³⁴ Budek K., Osinski B., (2018), *What is reinforcement learning? The complete guide*, Big Data Science, 05/07/2018.

¹³⁵ Stefenelli M., *Apprendimento automatico nei giochi di strategia*, tesi di laurea, Università di Bologna, a.a. 2013 - 2014, relatore Roli A.

¹³⁶ Barto A., Sutton R., (2016), *Reinforcement Learning: An Introduction.*, The MIT Press, Londra.

Il secondo filone si basa sulla ricerca di una soluzione al problema del controllo ottimo attraverso l'utilizzo di funzioni di valore e di una programmazione dinamica. Richard Bellman ideò l'*optimal return function*¹³⁷, funzione di valore in grado di risolvere problemi stocastici di controllo ottimale.

Il terzo e ultimo filone si sviluppa invece, sull'apprendimento per differenze temporali, basandosi sulla differenza tra due stime temporalmente consecutive della stessa quantità¹³⁸.

Questi tre filoni, come già esplicitato, a partire dagli anni 80', con lo sviluppo del *Q – Learning*, confluirono in un'unica direzione creando così le fondamenta dell'apprendimento per rinforzo moderno. Questa tipologia di apprendimento si basa sulla relazione esistente tra agente e ambiente che può svilupparsi in archi temporali discreti o continui; ad ogni momento, l'agente riceve un input dall'ambiente e sulla base di questo agisce di conseguenza rispetto a tutte le possibili azioni e dato lo stato del mondo. Se l'azione intrapresa dall'agente si rivela adeguata, questo riceverà una ricompensa, in caso contrario, subirà una punizione¹³⁹.

2.3.1 Tipologie di apprendimento per rinforzo

Le più comuni applicazioni dell'apprendimento per rinforzo richiedono:

- Algoritmi ad apprendimento continuo ovvero tecniche in grado di valutare le scelte dell'algoritmo stesso, a seguito dell'input ricevuto dall'ambiente, e premiare o punire lo stesso con l'utilizzo di un segnale numerico di rinforzo a seconda del comportamento istantaneo del sistema¹⁴⁰.

¹³⁷ Gli agenti che sfruttano l'apprendimento per rinforzo imparano a massimizzare la ricompensa futura cumulativa. La parola usata per descrivere la ricompensa futura cumulativa è ricompensa ed è spesso indicata con R . Viene utilizzato anche un pedice t per indicare la ricompensa da un determinato passaggio temporale. In notazione matematica, si presenta così:

$R_t = r_{t+1} + r_{t+2} + r_{t+3} + r_{t+4} + \dots = \sum_{k=0}^{\infty} r_{t+k+1}$. Fonte: Greaves J., *Understanding RL: The Bellman Equations*, disponibile a <https://joshgreaves.com/reinforcement-learning/understanding-rl-the-bellman-equations/>.

¹³⁸ Cesaro E., *Q – Learning: un algoritmo ad apprendimento per rinforzo applicato al trading finanziario*, tesi di laurea, Università di Venezia, a. a. 2016-2017, relatore Corazza M.

¹³⁹ Ibidem.

¹⁴⁰ Stefanelli M., *Apprendimento automatico nei giochi di strategia*, tesi di laurea, Università di Bologna, a.a. 2013 - 2014, relatore Roli A.

- Algoritmi ad addestramento preventivo i quali, diversamente dai precedenti, non hanno la possibilità di valutare continuamente le azioni dell'algoritmo. In una prima fase, queste tecniche istruiscono l'algoritmo e una volta che il sistema è ritenuto affidabile, viene reso immodificabile¹⁴¹.

¹⁴¹ Stefenelli M., *Apprendimento automatico nei giochi di strategia*, tesi di laurea, Università di Bologna, a.a. 2013 - 2014, relatore Roli A.

CAPITOLO III – LE RETI NEURALI ARTIFICIALI: LO STATO DELL’ARTE

3.1 Le reti neurali artificiali

La capacità di prevedere le dinamiche degli strumenti finanziari rappresenta una delle principali sfide nella comunità dell’intelligenza artificiale.

La previsione è quel processo che permette di stimare qualcosa di futuro sulla base di dati passati e presenti, o di analisi delle tendenze, guidando così molte decisioni importanti nei più svariati ambiti tra cui quello economico – finanziario¹⁴².

Per molto tempo, a fine predittivo, sono stati utilizzati modelli lineari ovvero strumenti efficaci ed efficienti per elaborare e risolvere dei problemi che richiedevano la semplice ripetizione di una serie di operazioni ben definite; questi però presentavano dei limiti importanti evidenziati, in particolar modo, agli inizi degli anni Ottanta a seguito di un concorso di previsione su larga scala organizzato da Makridakis, docente della più importante università di Cipro. In tale occasione, la maggior parte dei modelli lineari, comunemente utilizzati, vennero testati su più di mille serie storiche reali: la globalità dei risultati dimostrò che nessun modello lineare era il migliore¹⁴³.

Ma chiaramente, in un contesto, quale il mercato finanziario definito come un ambiente caratterizzato da un’ampia mole di informazioni eterogenee e, influenzato da molteplici fattori, si è presentata la necessità di utilizzare dei modelli in grado di cogliere le diverse correlazioni tra i dati in problemi non lineari: proprio per questo, le tecniche di apprendimento automatico sono divenute nel tempo, molto importanti nella modellizzazione finanziaria¹⁴⁴.

¹⁴² Abbas O., (2015), “Neural Network in Business Forecasting”, *International Journal of Computer*, 1, pp. 114- 128.

¹⁴³ Ibidem.

¹⁴⁴ Lozinskaia M., Zhemchuzhnikov V., (2017), “MICEX INDEX FORECASTING: THE PREDICTIVE POWER OF NEURAL NETWORK MODELING AND SUPPORT VECTOR MACHINE”, *Perm University Herald*, 1, pp. 49 – 60.

La storia delle reti neurali è progredita sia attraverso innovazioni concettuali sia grazie a sviluppi di implementazione; progressi che tuttavia sembrano essere avvenuti “a singhiozzo” piuttosto che attraverso un’evoluzione continuativa¹⁴⁵.

Alcuni dei lavori di base, si sono svolti tra la fine del XIX secolo e l’inizio del XX secolo: si trattava principalmente di lavori interdisciplinari che coinvolgevano la fisica, la psicologia e la neurofisiologia ed enfatizzavano le teorie generali di apprendimento, visione e condizionamento, escludendo quindi modelli matematici specifici.

La visione moderna delle reti neurali si è sviluppata a partire dal 1940 con lo studio di McCulloch W. e Pitts W. che ha mostrato che le reti neurali artificiali potrebbero, in linea generale, essere in grado di calcolare qualsiasi funzione aritmetica o logica. La prima applicazione pratica risale però, alla fine degli anni Cinquanta, con l’invenzione della rete di percezione e della relativa regola di apprendimento da parte di Rosenblatt F., il quale dimostrò la capacità, della rete, di riconoscere dei pattern. Sfortunatamente, in seguito, tale rete dimostrò il suo limite, infatti essa era in grado di risolvere solamente una classe limitata di problemi¹⁴⁶.

All’incirca nello stesso periodo, Windrow B. e Hoff T. introdussero un nuovo algoritmo di apprendimento e lo usarono per addestrare le reti neurali adattive lineari: regola in uso ancora oggi.

Alcuni lavori importanti proseguirono negli anni 70’ quando Kohonen T. e Anderson J. svilupparono indipendentemente delle reti neurali nuove che potevano fungere da memorie, così come l’indagine condotta da Grosseberg S. sulle reti di auto – organizzazione.

Durante gli anni 80’ vennero superati i limiti che avevano caratterizzato i due decenni precedenti, ovvero la mancanza di idee innovative e di computer potenti con i quali sperimentare, così la ricerca nelle reti neurali aumentò esponenzialmente. Nuovi personal computer e postazioni di lavoro divennero disponibili e due concetti vennero introdotti: il primo è stato l’uso della meccanica statistica¹⁴⁷ per spiegare il

¹⁴⁵ Hagan M., Demuth H., Beale M., De Jesus O., (2014), *Neural Network Design*, Martin Hagan, Kansas.

¹⁴⁶ Ibidem.

¹⁴⁷ La meccanica statistica è quella parte della fisica che studia, mediante metodi statistici, il comportamento di insiemi di un grande numero di elementi, allo scopo di prevederne le proprietà macroscopiche. Fonte: DEAGOSTINI Editore, *La meccanica statistica*, 2019.

funzionamento di una certa classe di rete ricorrente e il secondo è stato l'algoritmo di *back-propagation* sviluppato per l'addestramento delle reti neurali multistrato e ideato da Rumelhart D. e McClelland J.¹⁴⁸.

A seguito di questi importanti sviluppi, le reti neurali hanno trovato applicazione in molti campi risultando degli strumenti essenziali per la risoluzione di molti problemi, anche complessi¹⁴⁹.

Le reti neurali artificiali sono modelli quantitativi basati sull'intelligenza artificiale e ispirati alla struttura e al funzionamento del cervello umano. Esse si identificano, più precisamente, in regressori non lineari che esprimono le relazioni funzionali esistenti tra un vettore di input e una o più variabili di output (Soncini S., 2007).

Ogni rete neurale ha la seguente architettura, più o meno complessa:

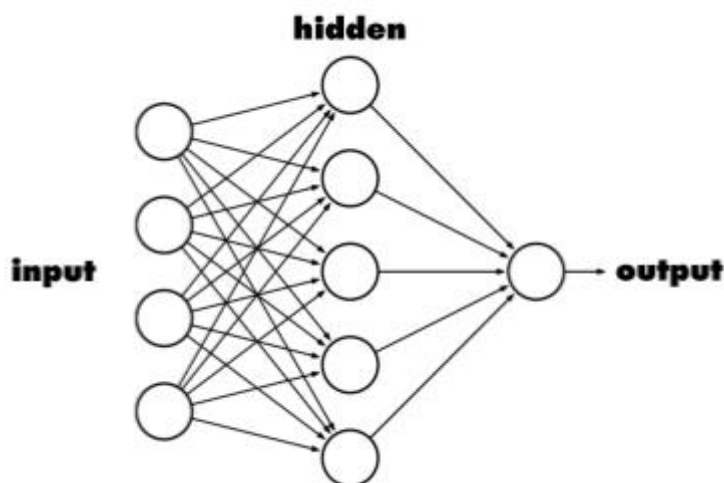


FIGURA 16: RAPPRESENTAZIONE GRAFICA DELL'ARCHITETTURA DI BASE DI UNA RETE NEURALE ARTIFICIALE. FONTE: GIUSEPPE CIABURRO, (2017), NEURAL NETWORKS WITH R, PACKT PUBLISHING, BIRMINGHAM.

Nello specifico, ogni rete neurale artificiale è composta da diversi elementi:

- lo strato degli input composto dai dati che la rete riceve e grazie ai quali si attiva;
- una o più unità intermedie chiamate *hidden layer* o strati nascosti i quali elaborano gli input ricevuti grazie alla capacità di classificazione per la quale la rete è stata addestrata¹⁵⁰;

¹⁴⁸ Hagan M., Demuth H., Beale M., De Jesus O., (2014), *Neural Network Design*, Martin Hagan, Kansas.

¹⁴⁹ Ibidem.

¹⁵⁰ Tecchia M., (2018), *Reti neurali artificiali: cosa sono e come funzionano*, Digitalic, 09/06/2018.

- lo stato di output che raccoglie i risultati e li modella fino ad arrivare alla presentazione della soluzione definitiva del problema al quale la rete è stata sottoposta;
- i pesi nelle reti neurali che sono i fattori più importanti nel processo di conversione di un input per impattare sull'output. Essi, simili alla pendenza nella regressione lineare, sono parametri numerici che determinano quanto fortemente ciascuno dei neuroni influenza gli altri¹⁵¹. Per un tipico neurone, se gli input sono x_1 , x_2 e x_3 , i pesi sinaptici da applicare ad essi sono indicati come w_1 , w_2 e w_3 e l'output che ne risulta sarà:

$$y = f(x) = \sum xi wi$$

dove i identifica l'input. Infine,

- il *bias* ovvero un parametro che viene utilizzato per regolare l'uscita insieme alla somma ponderata degli ingressi al neurone (è simile all'intercetta della regressione lineare).

L'elaborazione effettuata da un neurone è quindi indicata come segue:

$$output = sum(weights * inputs) + bias$$

al quale viene applicata la funzione di attivazione. Ne risulta che l'input del *layer* successivo è l'output dei neuroni presenti nello strato precedente, come mostra la seguente immagine¹⁵²:

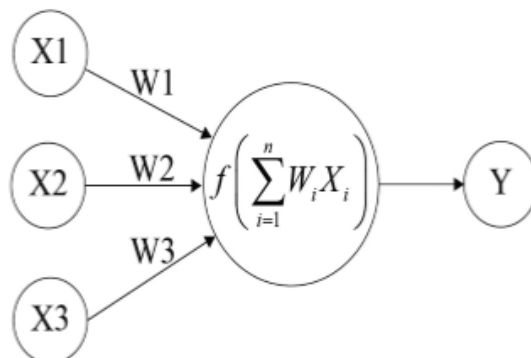


FIGURA 17: RAPPRESENTAZIONE GRAFICA DEL FUNZIONAMENTO DI UNA RETE NEURALE ARTIFICIALE. FONTE: GIUSEPPE CIABURRO, (2017), NEURAL NETWORKS WITH R, PACKT PUBLISHING, BIRMINGHAM.

¹⁵¹ Giuseppe Ciaburro, (2017), Neural Networks with R, Packt Publishing, Birmingham.

¹⁵² Ibidem.

Quindi, maggiore è il numero di strati di cui si compone la rete e maggiore sarà l'efficienza della stessa¹⁵³. Ogni strato è composto da neuroni, detti anche nodi, che intendono simulare il ruolo dei neuroni biologici e che svolgono un'operazione molto semplice che consiste nel diventare attivo se la quantità totale di segnale che riceve supera la propria soglia di attivazione; se ciò accade, questo neurone artificiale emette un segnale che viene trasmesso lungo i canali di comunicazione fino alle altre unità a cui esso è connesso. Ogni nodo di connessione, che simula le sinapsi biologiche, funge da filtro che trasforma il messaggio ricevuto in un segnale inibitorio aumentandone o diminuendone nel contempo l'intensità a seconda delle proprie caratteristiche individuali; inoltre essi, hanno il compito di "pesare" l'intensità dei segnali trasmessi, attraverso appunto i pesi sinaptici¹⁵⁴.

L'astrazione dell'elaborazione delle reti neurali si ottiene principalmente attraverso le funzioni di attivazione ovvero funzioni matematiche che sono in grado di convertire l'input in un output; senza di esse, il funzionamento delle reti neurali sarebbe lo stesso delle funzioni lineari¹⁵⁵. Ci sono diversi tipi di funzioni di attivazione.

La funzione più semplice e comunemente utilizzata è la funzione di attivazione lineare rappresentata dalla seguente formula:

$$y = f(x) = x$$

e nella quale l'output coincide con l'input.

¹⁵³ Tecchia M., (2018), *Reti neurali artificiali: cosa sono e come funzionano*, Digitalic, 09/06/2018.

¹⁵⁴ Floreano D., Mattiussi C., (2002), *Manuale sulle reti neurali*, il Mulino, Bologna.

¹⁵⁵ Ciaburro G., (2017), *Neural Networks with R*, Packt Publishing, Birmingham.

Essa è rappresentata nel seguente grafico:

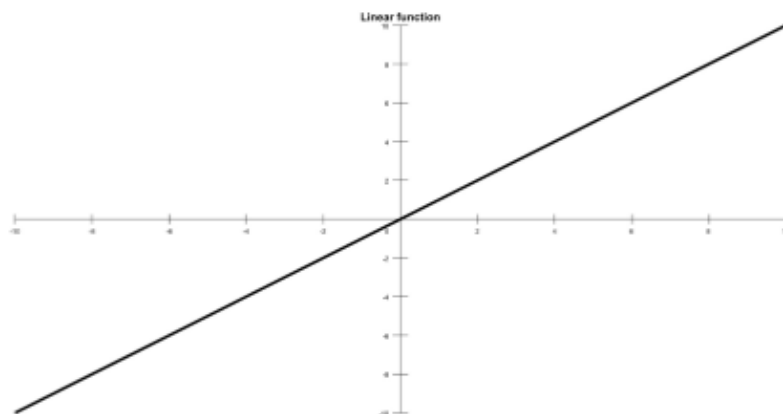


FIGURA 18: RAPPRESENTAZIONE GRAFICA DELLA FUNZIONE DI ATTIVAZIONE LINEARE. FONTE: GIUSEPPE CIABURRO, (2017), NEURAL NETWORKS WITH R, PACKT PUBLISHING, BIRMINGHAM.

Essa viene normalmente utilizzata per lo strato che contiene l'output della rete neurale in quanto, essendo caratterizzata da una maggiore rigidità rispetto alle altre funzioni di attivazione, evita che il risultato tenda verso il minimo o il massimo¹⁵⁶. Meno efficace risulta essere il suo utilizzo negli strati nascosti, soprattutto se costituiti da un elevato numero di neuroni in quanto risulterebbero connessi su una base funzionale che è proprio ciò che si vuole evitare ricorrendo a tale funzione di attivazione. Il limite di tale funzione è quello di non consentire un adattamento adeguato alle serie storiche caratterizzate da trend persistenti¹⁵⁷.

La funzione soglia viene solitamente utilizzata in situazioni in cui l'output assume il valore 0 per l'argomento negativo e 1 per l'argomento positivo: l'uscita è quindi di natura binaria. Modello nel quale il neurone genera un segnale quando la somma degli input ricevuti supera un certo valore di soglia¹⁵⁸.

¹⁵⁶ Crescenzo G., (2007), *Reti Neurali Artificiali: Teoria ed Applicazioni*, Dipartimento di Scienze Economiche, Matematiche e Statistica. Centro di Ricerca Interdipartimentale Bioagromed dell'Università di Foggia, Quaderno di ricerca No. 28.

¹⁵⁷ Ibidem.

¹⁵⁸ Ibidem.

La relativa equazione è:

$$f(x) = 0 \text{ when } x < 0, \\ 1 \text{ when } x \geq 0$$

e il grafico che ne risulta è il seguente:

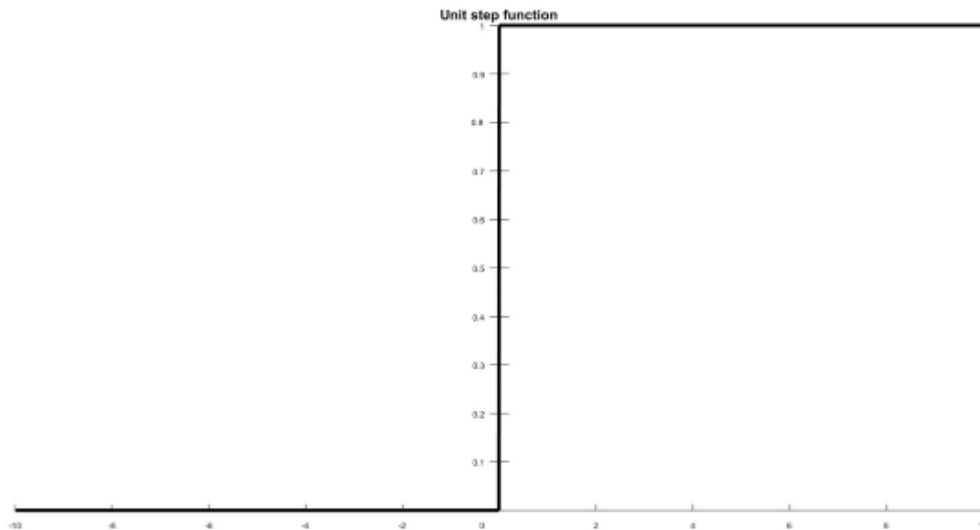


FIGURA 19: RAPPRESENTAZIONE GRAFICA DI UNA FUNZIONE SOGLIA E LA RELATIVA EQUAZIONE. FONTE: GIUSEPPE CIABURRO, (2017), NEURAL NETWORKS WITH R, PACKT PUBLISHING, BIRMINGHAM.

Sebbene essa risulti una funzione di attivazione interessante, a causa dei suoi parallelismi con la biologia, difficilmente viene utilizzata nelle reti neurali artificiali¹⁵⁹.

L'alternativa più comunemente utilizzata è la funzione logistica ovvero una funzione matematica che produce una curva sigmoide caratteristica per la sua forma ad S. Funzione rappresentata dalla formula:

$$f(x) = 1 / (1 + e^{-x})$$

Il segnale di output non risulta più binario, in quanto può assumere valori compresi tra 0 e 1. Tale funzione è fondamentale per la creazione di algoritmi di ottimizzazione efficienti¹⁶⁰ e quindi, il suo utilizzo risulta particolarmente utile negli strati nascosti delle reti applicate alle serie storiche finanziarie.

¹⁵⁹ Lantz B., (2013), *Machine Learning with R*, Packt Publishing, Birmingham.

¹⁶⁰ Ibidem.

A livello grafico, essa viene rappresentata come segue:

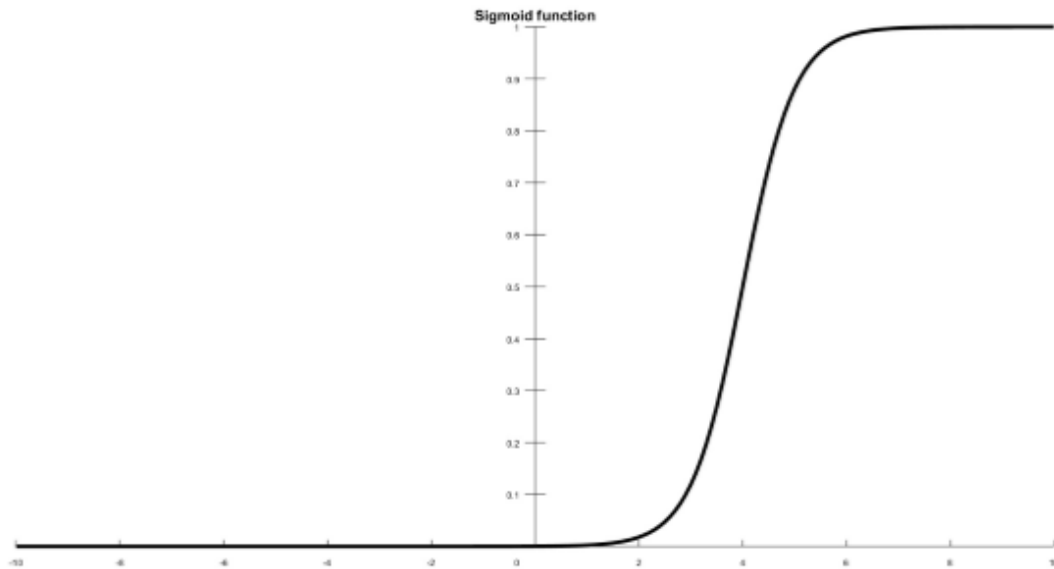


FIGURA 20: RAPPRESENTAZIONE DI UNA FUNZIONE SIGMOIDEA. FONTE: GIUSEPPE CIABURRO, (2017), NEURAL NETWORKS WITH R, PACKT PUBLISHING, BIRMINGHAM.

Un'altra funzione di attivazione molto popolare e ampiamente utilizzata è la funzione iperbolica tangente; si tratta di una funzione non lineare definita nell'intervallo di valori (-1;1) che consente di adattare la rete in modo affidabile negli strati nascosti, soprattutto nel caso in cui l'analista abbia scelto una funzione logistica o lineare per l'output¹⁶¹.

La relativa equazione è la seguente:

$$f(x) = \tanh(x)$$

¹⁶¹ Crescenzo G., (2007), *Reti Neurali Artificiali: Teoria ed Applicazioni*, Dipartimento di Scienze Economiche, Matematiche e Statistica. Centro di Ricerca Interdipartimentale Bioagromed dell'Univeristà di Foggia, Quaderno di ricerca No. 28.

e il relativo grafico, molto simile alla funzione di attivazione logistica è:

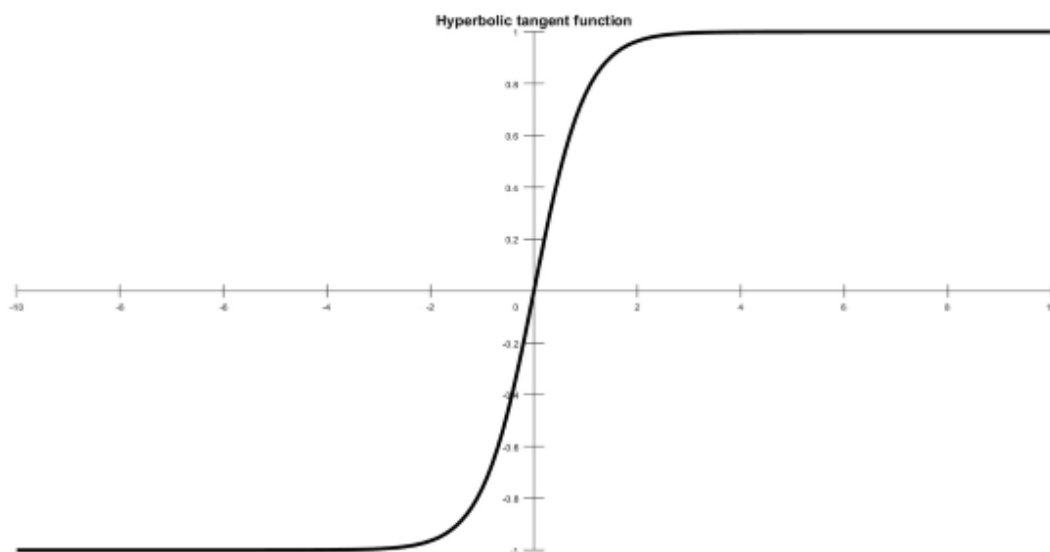


FIGURA 21: RAPPRESENTAZIONE GRAFICA DELLA FUNZIONE DI ATTIVAZIONE IPERBOLICA TANGENTE. FONTE: GIUSEPPE CIABURRO, (2017), NEURAL NETWORKS WITH R, PACKT PUBLISHING, BIRMINGHAM.

In ogni caso, qualsiasi funzione di attivazione venga scelta, questa rende la rete neurale in grado di adattarsi a determinati tipi di dati nel modo più appropriato, consentendo così la costruzione di reti neurali specializzate¹⁶²; inoltre deve essere sufficientemente robusta ed essere contraddistinta dalla differenziabilità, semplicità e velocità computazionale¹⁶³.

La capacità di apprendimento di una rete neurale è intrinseca nella sua topologia ovvero la struttura grazie alla quale i neuroni artificiali, di cui è composta, sono interconnessi¹⁶⁴. Esistono innumerevoli strutture di rete, le quali si differenziano per tre caratteristiche fondamentali:

- il numero di strati o *layer*: le reti a singolo strato (più comunemente chiamate *perceptron*) possono essere utilizzate per problemi di classificazione e in particolar modo per i modelli che sono separabili linearmente. I nodi di input ricevono segnali non elaborati direttamente dal dataset considerato, il valore viene trasformato dalla

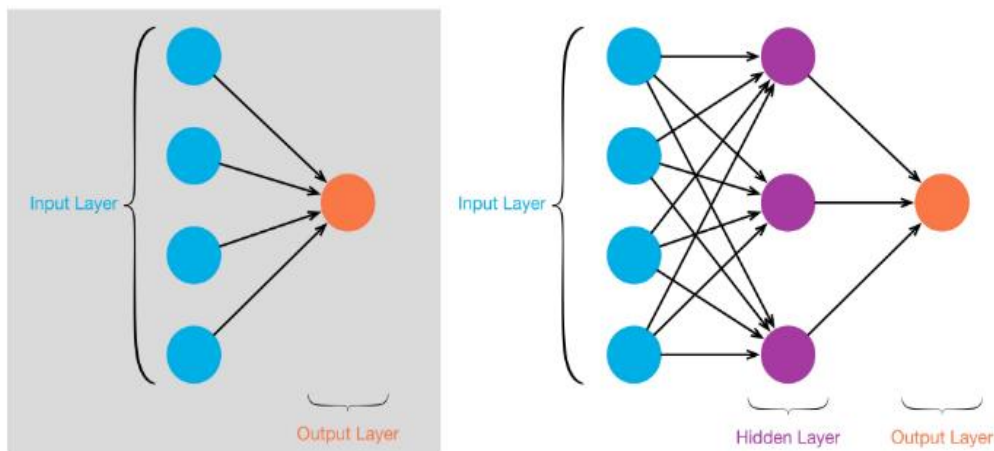
¹⁶² Lantz B., (2013), *Machine Learning with R*, Packt Publishing, Birmingham.

¹⁶³ Ciaburro G., (2017), *Neural Networks with R*, Packt Publishing, Birmingham.

¹⁶⁴ Ibidem.

funzione di attivazione del nodo e infine, i segnali risultanti dai nodi di ingresso, sono trasmessi ai nodi di output che utilizzano la loro funzione di attivazione per generare la previsione finale¹⁶⁵.

Diversamente, le reti multistrato (*multi-layer perceptron*), aggiungono uno o più strati nascosti risultando completamente connesse e proprio per questo, utilizzate nella risoluzione di problemi più complessi¹⁶⁶. A titolo esemplificativo, di seguito si riporta un confronto grafico di due reti neurali caratterizzate da un diverso numero di strati:



**FIGURA 22: RAPPRESENTAZIONE GRAFICA A STRATO SINGOLO (A SINISTRA) E A STRATO MULTIPLO (A DESTRA).
FONTE: [HTTPS://TOWARDSDATASCIENCE.COM/MULTI-LAYER-NEURAL-NETWORKS-WITH-SIGMOID-FUNCTION-DEEP-LEARNING-FOR-ROOKIES-2-BF464F09EB7F](https://towardsdatascience.com/multi-layer-neural-networks-with-sigmoid-function-deep-learning-for-rookies-2-bf464f09eb7f).**

Il numero di nodi presenti in ogni strato è predeterminato dal numero di funzionalità intrinseco nel set di dati di input, così come il numero di nodi di output è determinato a priori in base al numero di risultati da modellare.

Solamente il numero di nodi nascosti è arbitrario e scelto dall'utente prima di addestrare il modello; non esiste una regola generale per determinare tale numero, ma sicuramente esso dipende dal numero di nodi di input, la quantità dei dati di addestramento, la quantità degli errori, nonché, tra molti altri fattori, anche la complessità della risoluzione del problema considerato¹⁶⁷.

¹⁶⁵ Ciaburro G., (2017), *Neural Networks with R*, Packt Publishing, Birmingham.

¹⁶⁶ Ibidem.

¹⁶⁷ Lantz B., (2013), *Machine Learning with R*, Packt Publishing, Birmingham.

- la direzione di propagazione delle informazioni: le reti neurali di tipo *feed – forward* permettono solo collegamenti neuronali tra due diversi strati, mentre le reti di tipo *feed – back* consentono anche connessioni tra i neuroni dello stesso strato.

A livello grafico si può osservare la differenza esistente tra queste due diverse tipologie di rete neurale:

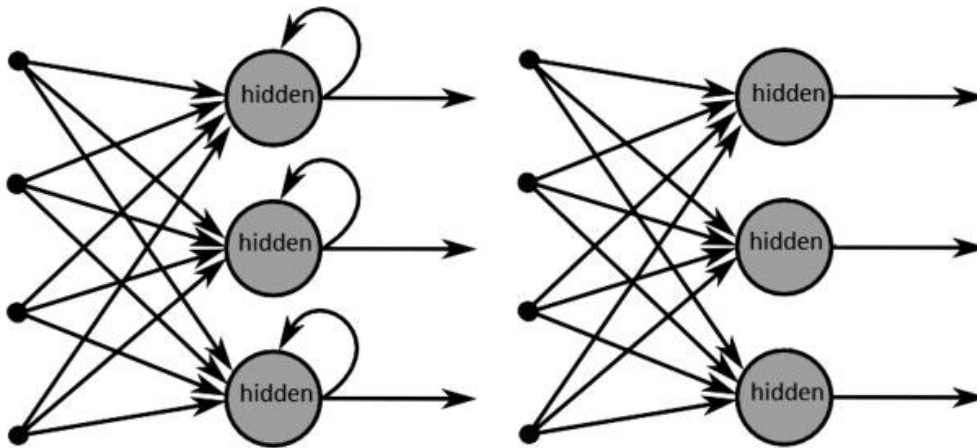


FIGURA 23: RAPPRESENTAZIONE GRAFICA DI UNA RETE NEURALE FEED - BACK (A SINISTRA) E FEED - FORWARD (A DESTRA). FONTE: [HTTPS://DATASCIENCE.STACKEXCHANGE.COM/](https://datascience.stackexchange.com/).

- il processo di apprendimento: l'elaborazione dal livello di input a quello/i nascosto/i e quindi al livello di output è chiamata *forward propagation*: gli input ponderati per i relativi pesi e l'aggiunta dei *bias*, viene applicata ad ogni livello e quindi il valore della funzione di attivazione viene propagato al livello successivo il quale può essere un ulteriore livello nascosto o il livello di output. Una volta che l'output ha raggiunto l'ultimo strato, viene calcolato l'errore (differenza tra output previsto e output originale) necessario per correggere i pesi e i *bias* utilizzati nella *forward propagation*¹⁶⁸.

Differentemente, il processo di *back - propagation* utilizza la derivata parziale della funzione di attivazione per identificare la pendenza (o gradiente) della direzione di ciascun peso in entrata, da cui l'importanza di avere una funzione di attivazione differenziabile.¹⁶⁹ Il gradiente suggerisce quanto rapidamente l'errore sarà ridotto o aumentato per una variazione del peso. L'algoritmo tenterà di modificare i pesi che portano alla massima riduzione dell'errore di una quantità nota come tasso di apprendimento (parametro scalare utilizzato per impostare il tasso di aggiustamento

¹⁶⁸ Ciaburro G., (2017), Neural Networks with R, Packt Publishing, Birmingham.

¹⁶⁹ Ibidem.

per ridurre più velocemente gli errori e utilizzato nella *back - propagation* durante la regolazione dei pesi e degli errori)¹⁷⁰. Maggiore è il tasso di apprendimento, più velocemente l'algoritmo cercherà di scendere lungo le pendenze, il che potrebbe ridurre il tempo di allenamento¹⁷¹.

Grazie a tale processo, il modello si adatta bene ai dati e le previsioni risultano più accurate¹⁷².

3.2 La relativa letteratura: lo stato dell'arte

La previsione dei mercati finanziari e, in particolar modo, degli indici di borsa che si identificano come indicatori altamente stocastici, svolge un ruolo cruciale nell'individuazione di percorsi di investimento redditizi da parte della generalità degli investitori finanziari. Essi comprendono investitori al dettaglio, istituti finanziari nazionali e internazionali nonché banche di investimento alla ricerca di efficaci investimenti sottoforma di apprezzamento del capitale e compravendita borsistica¹⁷³.

La previsione del prezzo o del valore delle azioni, e degli indici azionari, risulta un processo complesso a causa dell'elevata volatilità dei mercati finanziari influenzati da una molteplicità di fattori. Esistono due tipi di analisi che gli investitori effettuano prima di investire:

- l'analisi fondamentale grazie alla quale viene valutato il valore intrinseco delle azioni, la performance di settore e l'economia per decidere se investire o meno in un determinato titolo o indice azionario;
- e l'eventuale analisi tecnica ovvero la valutazione delle azioni o indici azionari basandosi sui valori passati generati dalle attività di mercato¹⁷⁴.

L'ipotesi di mercato efficiente, proposta da Malkiel e Fama nel 1970, dichiara che il prezzo delle azioni è informativamente efficiente, il che significa che è possibile

¹⁷⁰ Ciaburro G., (2017), *Neural Networks with R*, Packt Publishing, Birmingham.

¹⁷¹ Ibidem.

¹⁷² Ibidem.

¹⁷³ Marxiaoli S., Selvam M., Lingaraja K., Vasanth V., (2017), "Forecasting the Stock Index Movements of India: Application of Neural Networks", *International Journal of Soft Computing*, 2, pp. 120-131.

¹⁷⁴ Ibidem.

prevedere il prezzo delle azioni basandosi solo sui dati di trading giornalieri. Pertanto, se le informazioni ottenute dai mercati azionari sono rielaborate in modo adeguato *ex-ante*, può essere previsto il trend di un titolo o indice azionario per un periodo futuro¹⁷⁵.

Proprio a tal fine, la letteratura dimostra che le reti neurali costituiscono uno dei migliori strumenti per l'analisi previsiva in campo finanziario.

Soprattutto negli ultimi decenni, con l'avvento di nuove tecnologie e l'evoluzione dei sistemi informatici, che hanno permesso la formazione di enormi moli di dati, esse sono state oggetto di ricerche e applicazioni da parte di molti studiosi e ricercatori appartenenti ai più svariati ambiti, evidenziando la loro considerevole capacità computazionale che ha portato a risultati più che soddisfacenti.

Chiaramente l'ambito di interesse di questo lavoro risulta essere quello finanziario e proprio per tale ragione, e per successivi fini applicativi, sono stati analizzati i più importanti paper, degli ultimi vent'anni, relativi all'utilizzo finanziario delle reti neurali a fini previsivi.

Il primo articolo che è stato analizzato è la ricerca e l'applicazione delle reti neurali per l'analisi tecnica e conseguente previsione di un indice azionario (Yao, Tan e Poh, 1999). Gli autori di tale ricerca applicano le reti neurali per la previsione dell'indice composito di Kuala Lumpur calcolato sulla base degli 86 maggiori titoli della Malesia. Il periodo considerato va dal 3 gennaio 1984 al 16 ottobre del 1991 e i dati utilizzati sono costituiti da 1911 osservazioni giornaliere. Gli input che vengono considerati sono il valore passato dell'indice, il valore attuale, la media mobile, il momentum rappresentativo della velocità del trend di un titolo azionario¹⁷⁶, l'indice di forza relativa che segnala situazioni di ipercomprato e ipervenduto¹⁷⁷, l'oscillatore stocastico che confronta il prezzo di chiusura dello strumento finanziario con i suoi movimenti nell'ambito di un lasso temporale utilizzato come punto di riferimento e la sua media mobile. Tutti indicatori che possono essere derivati dall'indice composito stesso.

¹⁷⁵ Cochrane J., (2014), *Eugene F. Fama efficient markets and the Nobel Prize*, Chicago Booth Review, 20/05/2014.

¹⁷⁶ Glossario Finanziario – Borsa Italiana, <https://www.borsaitaliana.it>.

¹⁷⁷ Ibidem.

L'output che si vuole prevedere è l'indice malesiano futuro a un giorno¹⁷⁸. I dati storici vengono suddivisi in tre parti, addestramento, validazione e l'insieme di test: il set di addestramento contiene due terzi dei dati raccolti, mentre il validation set e il test - set contengono rispettivamente i due quindicesimi e tre quindicesimi dei dati raccolti. Essi, tramite un processo iterativo che inizia da un singolo strato nascosto costituito da un altrettanto singolo neurone nascosto, procedono fino ad ottenere una minimizzazione dell'errore nel test set. Adottando una semplice procedura per decidere il numero di nodi nascosti, determinato anche dal numero di nodi di input o dello strato precedente.

Tale ricerca ha portato a risultati soddisfacenti e dimostrato che è possibile effettuare delle previsioni utili per l'indice analizzato senza l'uso di molti dati o conoscenze di mercato, ma oltre a ciò, ha evidenziato anche alcuni problemi legati alle reti neurali, infatti il tasso di successo è funzione dell'intervallo scelto per il test - set e la generalizzazione del modello nel tempo è scarsa¹⁷⁹.

Gli autori suggeriscono, che pur avendo ottenuto dei risultati soddisfacenti, per migliorare le capacità delle reti neurali nelle previsioni, dovrebbe essere considerata una mistura di tecniche e fattori fondamentali come input nei diversi periodi di tempo¹⁸⁰.

Nel 2000 Fernández-Rodríguez, González – Martel e Sosvilla – Riviero attuarono delle prove sulla redditività del trading tecnico basato sulle reti neurali artificiali¹⁸¹. Nella loro ricerca considerano l'indice generale della Borsa di Madrid, suggerendo che in assenza di costi di negoziazione, il trading tecnico risulta molto più efficiente ed efficace di una strategia *buy – and – hold* in un contesto di mercato ribassista o stabile. Quest'ultima strategia genera rendimenti più elevati rispetto all'utilizzo delle reti neurali artificiali solo nei periodi caratterizzati da un mercato

¹⁷⁸ Yao J., Tan C. L., Poh H-L. (1999), *Neural Networks for Technical Analysis: a study on KLCI*, International Journal of Theoretical and Applied Finance.

¹⁷⁹ Ibidem.

¹⁸⁰ Ibidem.

¹⁸¹ Fernández-Rodríguez F., González – Martel C., Sosvilla – Riviero S., (2000), On the profitability of technical trading rules based on artificial neural networks: Evidence from the Madrid stock market, Universidad de Las Palmas de Gran Canaria, FEDEA e Universidad Complutense de Madrid, Working Paper No. 7.

rialzista¹⁸². I dati utilizzati comprendono 6931 osservazioni, rappresentanti del periodo che va dal 2 gennaio 1966 al 12 ottobre 1997. Essi utilizzano una rete neurale *feed - back* caratterizzata da nove input, corrispondenti ai rendimenti dei nove giorni precedenti, un livello nascosto costituito da quattro neuroni nascosti, e un livello di output composto da un singolo neurone: ogni unità del livello nascosto riceve la somma ponderata di tutti gli input e un termine di distorsione producendo così, un segnale di output attraverso la funzione di attivazione. Allo stesso modo, l'unità di output riceve la somma ponderata dei segnali di uscita dello strato nascosto e produce un risultato attraverso la funzione di trasferimento di uscita. L'output finale sarà un valore compreso tra -1 e 1: un valore maggiore a 0 rappresenta un segnale di acquisto, mentre un valore inferiore a 0 verrà utilizzato come segnale di vendita. I valori dei pesi utilizzati sono determinati tramite un processo di apprendimento iterativo e la loro trasformazione a ogni livello successivo è determinata da uno specifico tipo di funzione di trasferimento. Il processo di formazione utilizzato è il *back - propagation error*, un metodo progressivo di discesa del gradiente che minimizza la somma degli errori al quadrato del sistema. La funzione di trasformazione è una funzione logaritmica, mentre la funzione di trasferimento è una funzione tangente iperbolica¹⁸³.

Nel 2001, Walczak S., un docente dell'Università di Colorado, conduce una ricerca focalizzata sulle nuove applicazioni delle reti neurali, sistemi esperti e tecnologie intelligenti, in particolar modo nei settori delle serie temporali finanziarie e della medicina¹⁸⁴. Egli ha condotto una ricerca per analizzare le capacità previsive delle reti neurali applicate ai tassi *spot* del dollaro americano rispetto alla sterlina britannica, la moneta tedesca e lo yen. I valori storici, per ciascun tasso di cambio, comprendono il periodo dal 1° marzo 1973 fino al 30 giugno 1995; il set di dati di addestramento termina il 31 dicembre 1994, quello di prova o valutazione comprende gli ultimi due trimestri, quindi dal 2 gennaio 1995 al 30 giugno 1995. Egli utilizza una rete neurale *back -*

¹⁸² Fernández-Rodríguez F., González – Martel C., Sosvilla – Riviero S., (2000), On the profitability of technical trading rules based on artificial neural networks: Evidence from the Madrid stock market, Universidad de Las Palmas de Gran Canaria, FEDEA e Universidad Complutense de Madrid, Working Paper No. 7.

¹⁸³ Ibidem.

¹⁸⁴ Walczak S., (2001), "An empirical Analysis of Data Requirements for Financial Forecasting With Neural Networks", *Journal of Management Information Systems*, 4, pp. 203 - 222.

propagation a un singolo strato nascosto e la confronta con una caratterizzata da due strati nascosti, giungendo alla conclusione che la prima tende a sovraperformare le reti a più strati nascosti¹⁸⁵.

Nel 2003 Chen A., Lueng M. e Daouk H. hanno condotto uno studio grazie al quale hanno provato a modellare e prevedere la direzione del rendimento del tasso di cambio di Taiwan nel mercato azionario. Decisero di condurre tale ricerca in quanto, in quegli anni, molte banche di investimento internazionali e società di intermediazione detenevano quote importanti nei mercati esteri: utilizzarono una rete neurale dopo averla addestrata con i relativi dati storici e ne confrontarono le prestazioni con quelle della classica strategia *buy – and – hold*¹⁸⁶. I dati utilizzati sono stati ottenuti dal database EPS gestito dal dipartimento di educazione di Taiwan e coprono l'orizzonte temporale da gennaio 1982 ad agosto 1992 diviso in due periodi: il primo, che va da gennaio del 1982 ad agosto del 1987, viene utilizzato per la selezione e la convalida del modello, mentre il secondo, che va da settembre 1987 ad agosto 1992, viene utilizzato per confrontare le previsioni e le performance commerciali dei vari modelli considerati. A seconda della durata dell'orizzonte temporale dell'investimento, la variabile da prevedere è il rendimento a tre, sei o dodici mesi dell'indice considerato, mentre le variabili indipendenti sono tutte osservabili entro l'ultimo giorno del mese precedente al primo giorno del mese corrispondente al periodo di previsione¹⁸⁷. Dopo aver confrontato le prestazioni delle reti neurali artificiali con altri metodi previsivi, attraverso una simulazione di trading, i risultati hanno dimostrato che le reti neurali sono utili nella previsione; caratterizzate da una superiorità che viene attribuita alla loro capacità di identificare i valori anomali o errati permettendo così di ottenere profitti superiori rispetto ad altre strategie di investimento¹⁸⁸.

¹⁸⁵ Walczak S., (2001), "An empirical Analysis of Data Requirements for Financial Forecasting With Neural Networks", *Journal of Management Information Systems*, 4, pp. 203 - 222.

¹⁸⁶ Chen A., Lueng M., Daouk H., (2003), *Application of Neural Networks to an emerging financial market: forecasting and trading the Taiwan Stock Index*, Pergamon, Working Paper No. 30.

¹⁸⁷ Ibidem.

¹⁸⁸ Ibidem.

Nel 2004 Jasic T. e Wood D. condussero uno studio utilizzando i valori di chiusura giornalieri dell'indice S&P 500, l'indice tedesco DAX, l'indice giapponese TOPIX e l'indice inglese FTSE All Share. Per l'indice americano, tedesco e inglese, il periodo considerato inizia il 1° gennaio 1965 fino all'11 novembre 1999, mentre il campione per l'indice giapponese, considera i dati storici dal 1° gennaio 1969 all'11 novembre 1999. Tali set di dati vengono suddivisi in modo che circa i due terzi vengano utilizzati per l'addestramento e un terzo, il quale comprende i dati più recenti, per la valutazione delle prestazioni della rete neurale¹⁸⁹.

Per la modellizzazione dei rendimenti degli indici azionari, essi utilizzano una rete *feed-forward* costituita da un singolo strato nascosto; le unità di ingresso corrispondono a vettori comprendenti le osservazioni che costituiscono i diversi ritardi temporali (o *lag*). Essi nel loro studio inoltre definiscono il numero di nodi nascosti pari a circa la metà del numero dei nodi di input.

Ancora una volta si dimostra la superiorità previsiva delle reti neurali artificiali, le quali conducono a una previsione del 90% per ogni set di dati¹⁹⁰.

Un contributo interessante è quello di Pan H., Tilakaratne C. e Yearwood J. (2005), i quali proposero un approccio computazionale per la previsione dell'indice azionario australiano AORD utilizzando reti neurali *feed-forward* multistrato al fine di individuare una rete neurale efficace, o un insieme di reti neurali adattive che sfruttino o modellino le varie oscillazioni dinamiche e possibili influenze tra mercati, ovviamente a fini previsivi¹⁹¹. Essi considerano tre aspetti per l'efficacia della selezione dei dati: gli input effettivi derivanti dal mercato di riferimento, un insieme sufficiente di mercati interconnessi e gli input effettivi derivanti dalle interrelazioni tra i mercati considerati. Inoltre tengono conto del numero ottimale di strati nascosti e, per ognuno di questi, il

¹⁸⁹ Jasic T., Douglas W., (2004), "The profitability of daily stock market indices trades based on neural network predictions: case study for the S&P 500, the DAX, the TOPIX and the FTSE in the period 1965–1999", *Applied Financial Economics*, 14, pp. 285 – 297.

¹⁹⁰ Ibidem.

¹⁹¹ Pan H., Tilakaratne C., Yearwood J., (2005), "Predicting Australian Stock Market Index Using Neural Networks Exploring Dynamical Swings and Intermarket Influences", *Journal of Research and Practice in Information Technology*, 1, pp. 43 – 55.

numero di neuroni nascosti che contengono¹⁹². Essi utilizzando dati giornalieri per un periodo di tredici anni e mezzo, dal gennaio del 1990 all'agosto del 2003, calcolano i rendimenti giornalieri che costituiscono l'input della rete neurale con l'obiettivo di prevedere il trend dell'indice australiano del giorno successivo. Durante il loro lavoro, individuano il comportamento ciclico dell'indice per un periodo di sei giorni; ciò ha portato gli autori a proporre di utilizzare i sei rendimenti precedenti al giorno di previsione come variabili di input. I risultati che hanno ottenuto sono buoni, raggiungendo una correttezza previsiva della rete neurale dell'80% e evidenziando quindi, come le reti neurali siano uno degli strumenti maggiormente promettenti per lo svolgimento di un'attività previsiva sull'andamento nel mercato azionario¹⁹³.

O'Connor N. e Madden M. (2006) condussero uno studio al fine di valutare l'efficacia dell'uso di indicatori esterni come il prezzo delle materie prime e il tasso di cambio, per prevedere i movimenti in media dell'indice industriale Dow Jones attraverso le reti neurali artificiali. Il dataset considerato è costituito dai valori di apertura e di chiusura dell'indice e dai valori corrispondenti agli indicatori esterni dal 2 gennaio 1986 al 4 febbraio 2005. Dopo aver condotto diversi esperimenti, utilizzano una rete neurale *feed – forward* allenata sui dati di 4000 giorni - (dal 11 febbraio 1987 al 17 dicembre 2012) e testata sui 500 giorni successivi – (dal 18 dicembre 2002 al 13 dicembre 2004). I risultati evidenziano che l'aggiunta di indicatori esterni alle variabili di input migliorano in modo significativo le prestazioni in termini di redditività e successo direzionale del modello¹⁹⁴. Le reti neurali dimostrano una percentuale di previsioni direzionali corrette del 55,1%.

Nel 2007 uno studio presentato da Preminger A. e Franck R. suggerisce un solido approccio di regressione al fine di costruire modelli di previsione meno sensibili alla contaminazione dovuta a valori anomali rispetto al metodo di stima dei minimi

¹⁹² Pan H., Tilakaratne C., Yearwood J., (2005), "Predicting Australian Stock Market Index Using Neural Networks Exploring Dynamical Swings and Intermarket Influences", *Journal of Research and Practice in Information Technology*, 1, pp. 43 – 55.

¹⁹³ Ibidem.

¹⁹⁴ O'Connor N., Madden M., (2006), "A neural network approach to predict stock exchange movements using external factors", *Knowledge – Based Systems*, 19, pp. 371 – 378.

quadrati e altri metodi di stima ordinaria. In tale studio viene esplorata la prevedibilità dei prezzi in valuta dello yen giapponese e della sterlina britannica rispetto al dollaro americano con orizzonti temporali di 1, 3, 6 mesi utilizzando dati mensili ottenuti dal database del Federal Reserve Board. Il campione considerato inizia a gennaio 1971 e termina ad ottobre 2004 con 406 osservazioni¹⁹⁵.

Chaigusin S., Chirathamjaree C. e Clayden J. (2008) condussero uno studio sulla previsione dell'indice della borsa thailandese (SET) utilizzando reti neurali *feed – forward* stimate con l'algoritmo di *back - propagation*. Tra i fattori utilizzati come input, globali e locali che influenzano il mercato thailandese, si ritrova l'indice Dow Jones, l'indice Nikkei, l'indice Hang Seng, il prezzo dell'oro, il tasso minimo di prestito e i tassi di cambio tra il Baht thailandese e il dollaro americano. Come set di dati storici è stato utilizzato un periodo di due anni per l'allenamento e per la validazione, dati giornalieri per gli anni 2003-2004. Utilizzarono una rete neurale artificiale a tre strati con rispettivamente 7, 3 e 1 neuroni nascosti e come algoritmo di addestramento quello di discesa del gradiente. I risultati ottenuti supportano l'ipotesi che i movimenti dell'indice SET sono sensibili a tutti i fattori utilizzati come input¹⁹⁶.

Nel 2009, Faria E. L., Albuquerque M., Gonzalez J. e Cavalcante J. condussero uno studio predittivo sul principale indice azionario nel mercato brasiliano attraverso le reti neurali e il metodo *adaptive exponential smoothing* al fine di valutare l'accuratezza di entrambi i metodi nel prevedere l'andamento dei rendimenti del mercato.

Il training - set utilizzato considera i valori di chiusura dell'indice da settembre 1998 ad aprile 2007 e i risultati sono stati testati predicendo il valore assunto dall'indice stesso per periodo che va da aprile 2007 e marzo 2008¹⁹⁷.

¹⁹⁵ Preminger A., Franck R., (2007), "Forecasting exchange rates: A robust regression approach", *International Journal of Forecasting*, 23, pp. 71 – 84.

¹⁹⁶ Chaigusin S., Chirathamjaree C., Clayden J., (2008), "The Use of Neural Networks in the Prediction of the Stock Exchange of Thailand (SET) index", *IEEE computer society*, 10, pp. 670 – 673.

¹⁹⁷ Faria E. L., Albuquerque M., Gonzalez J. e Cavalcante J., (2009), "Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods", *International Journal Expert Systems with Applications*.

I risultati ottenuti dimostrarono che entrambi i modelli forniscono valori simili per la previsione del valore dell'indice, corretti per il 60% dei casi, ma le reti neurali sovraperformano nel prevedere l'andamento di mercato con tassi di successo relativi simili a quelli di altri mercati sviluppati¹⁹⁸.

Nel 2011 Yahuza B. e Haruna C. condussero uno studio, utilizzando le reti neurali artificiali, al fine di prevedere l'indice azionario nel mercato nigeriano. I dati tecnici sono stati raccolti per un periodo di quattro anni, dal 2005 al 2008 e utilizzati per il 70% per l'addestramento, per il 10% per la convalida del modello e il restante 20% come test - set. La rete neurale artificiale utilizzata si configura come una rete *feed – forward* costituita da quattro nodi di input, uno strato nascosto formato da 18 neuroni e un livello di output. I dati inoltre comprendono il prezzo di chiusura del giorno precedente, il prezzo di chiusura del giorno corrente e lo stesso dicasi per i prezzi di apertura, il volume delle azioni scambiate, il prezzo massimo e minimo raggiunti e la data di negoziazione. I risultati ottenuti non contraddicono gli studi precedenti, infatti l'accuratezza delle previsioni raggiunta con l'utilizzo delle reti neurali artificiali è pari all'80%¹⁹⁹.

Nel 2012 Dzikevicius A. e Stabuzyte N. hanno focalizzato la loro ricerca sulla previsione dell'indice azionario OMXV Vilnius. Il periodo considerato va dal 1° gennaio 2000 al 30 aprile 2012 e i dati sono giornalieri suddivisi in due periodicità: la maggior parte dei dati (circa 90%) viene utilizzata per l'addestramento, mentre i dati restanti vengono utilizzati per valutare l'accuratezza della previsione. Essi, nel loro studio, confrontano le abilità previsive della media mobile semplice, regressione multipla e diverse strutture di reti neurali artificiali multistrato²⁰⁰.

¹⁹⁸ Faria E. L., Albuquerque M., Gonzalez J. e Cavalcante J., (2009), "Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods", *International Journal Expert Systems with Applications*.

¹⁹⁹ Yahuza B., Haruna C., (2011), "Utilizing artificial neural network for prediction the nigerian stock market price index", *Computer Science and Telecommunications*, 30, pp. 68 – 77.

²⁰⁰ Dzikevicius A., Stabuzyte N., (2012), "Forecasting OMX Vilnius stock market – A neural network approach", *Business: theory and practice*, 4, pp. 321-332.

Come rivelano i risultati, la previsione maggiormente accurata, si ottiene attraverso le reti neurali multistrato, con un singolo strato nascosto, per il 53,06%. Sebbene tutte le tecniche messe a confronto in questa ricerca offrano approssimativamente il 50% di correttezza, le reti neurali artificiale comunque risultano gli strumenti migliori²⁰¹.

Nel 2013 Niaki S. e Hoseinzade S. condussero una ricerca avente come obiettivo principale la previsione della direzione giornaliera dell'indice S&P 500 utilizzando una rete neurale artificiale *feed - forward*. Il dataset utilizzato nella loro ricerca consisteva nel ritorno giornaliero dell'indice S&P 500 come variabile dipendente, ovvero la variabile output e 27 indicatori finanziari ed economici come variabili indipendenti già note. Il lasso temporale considerato inizia il 1° marzo 1994 e si conclude il 30 giugno 2008: l'80% dei dati considerati viene utilizzato per addestrare la rete neurale artificiale, il 10% per la sua validazione e il restante 10% per testare le prestazioni della rete neurale²⁰². Niaki e Hoseinzade analizzarono la performance finanziaria della rete utilizzata simulando una strategia di investimento che consisteva in un capitale totale di 100,000\$ utilizzato per acquistare un indice quando la previsione mostrava che esso sarebbe aumentato, mentre in caso contrario, sarebbe stato venduto. Confrontando tale strategia di investimento con la classica strategia *buy – and – hold*, essi riscontrarono che la rete neurale artificiale ha aumentato in modo significativo il capitale finale, risultando uno strumento efficace per la previsione del movimento degli indici azionari²⁰³.

Un'ulteriore ricerca volta a dimostrare la capacità delle reti neurali di prevedere i valori di mercato degli indici azionari, fu presentata nel 2014 da Magaji A. e Adeboye K., i quali focalizzarono il loro studio sull'indice NSE del mercato azionario

²⁰¹ Dzikevicius A., Stabuzyte N., (2012), "Forecasting OMX Vilnius stock market – A neural network approach", *Business: theory and practice*, 4, pp. 321-332.

²⁰² Niaki S., Hoseinzade S., (2013), "Forecasting S&P 500 index using artificial neural networks and design of experiments", *Journal of Industrial Engineering International*, 9, pp. 1 - 9.

²⁰³ Ibidem.

nigeriano²⁰⁴. Utilizzarono una rete neurale artificiale multistrato con una funzione di attivazione logistica e allenata tramite l'algoritmo di apprendimento *back – propagation*. I dati sperimentali raccolti sono del periodo che va dal 4 gennaio 2010 al 30 aprile 2012 esclusi i fine settimana e i giorni festivi. Sulla base dei risultati ottenuti, anch'essi dedussero la superiorità delle reti neurali artificiali, le quali avevano effettivamente previsto l'indice NSE²⁰⁵.

Nel 2015 Grigoryan H. presentò un modello di previsione combinato, basato sulle reti neurali artificiali con analisi delle componenti principali. Come caso di studio è stato utilizzato il TAL1T della borsa valori Nasdaq OMX Baltic con periodo campionario dal 12 marzo 2012 al 30 dicembre 2014, per un totale di 700 osservazioni²⁰⁶. I dati storici considerati sono il prezzo di chiusura giornaliero, il prezzo di apertura, il prezzo minimo e quello massimo, il volume scambiato, il fatturato delle azioni del gruppo Tallink e 30 indicatori di analisi tecnica del mercato azionario: 500 coppie di osservazioni vengono utilizzate come training set, test set e validazione, mentre le rimanenti 200 coppie di osservazioni vengono riservate per la previsione. Come variabile di previsione si considera il prezzo di chiusura del titolo Tallink. I risultati ottenuti hanno consentito di concludere che il modello combinato di previsione fornisce un promettente strumento per la previsione delle serie temporali finanziarie²⁰⁷.

Nel 2016, Qiu M. e Song Y. hanno condotto uno studio avente come obiettivo la previsione della direzione del prezzo dell'indice del mercato azionario giapponese²⁰⁸ utilizzando le reti neurali artificiali. Essi utilizzano l'algoritmo di apprendimento *back – propagation* e un solo strato nascosto. Il campione utilizzato è costituito da 1707 osservazioni giornaliere (da gennaio 2007 a dicembre 2013). I dati vengono divisi in due parti: il 78,6% di essi (dal 23 gennaio 2007 al 18 ottobre 2012) viene utilizzato per

²⁰⁴ Magaji A., Adeboye K., (2014), "An intense Nigerian Stock Exchange market prediction using logistic with back – propagation ANN model", *Science World Journal*, 2, pp. 8 – 13.

²⁰⁵ Ibidem.

²⁰⁶ Grigoryan H., (2015), "Stock Market Prediction using Artificial Neural Networks. Case study of TAL1T, Nasdaq OMX Baltic Stock", *Database System Journal*, 2, pp. 14 – 23.

²⁰⁷ Ibidem.

²⁰⁸ L'indice Nikkei 225 è l'indice di mercato più utilizzato nella Borsa di Tokyo e include 225 titoli equamente ponderati. Fonte: Investing.com.

l'addestramento della rete neurale, mentre il restante 21,4% (dal 19 ottobre 2012 al 30 dicembre 2013) viene utilizzato per la valutazione del modello. Il modello ibrido GA – ANN utilizzato risulta un metodo più efficace, rispetto a quelli utilizzati in altri studi simili, con una maggiore precisione di precisione²⁰⁹.

Nel 2017 Lozinskaia A. e Zhemchuzhnikov V. presentarono uno studio attinente alla previsione dell'indice di Borsa russo²¹⁰ utilizzando metodi di apprendimento automatico come la modellizzazione di reti neurali. I dati considerati sono giornalieri per il periodo che va dal 15 gennaio 2002 al 25 aprile 2016 utilizzati in una rete neurale *feed – forward* addestrata utilizzando l'algoritmo di apprendimento *back – propagation* e la tangente iperbolica come funzione di attivazione. Inoltre, per ridurre il rischio di *overfitting*, il 60% dei dati viene utilizzato per la fase di addestramento, il 20% per la fase di validazione e il restante 20% per il set dei test. Dai risultati che essi ottengono, affermano che le reti neurali, utilizzate sia come strumento per prevedere il prezzo degli indici azionari sia per prevedere il loro trend, si possono considerare particolarmente efficaci con un margine di errore estremamente ridotto²¹¹.

Nel 2019, uno dei maggiori studi sulle capacità di previsione degli indici azionari da parte delle reti neurali è stato condotto da Raso H. e Demirci M. Essi partono dal presupposto che la non linearità e gli alti tassi di variazione degli indici del mercato azionario, rendono la previsione un problema difficile da affrontare²¹². Dal momento che negli ultimi decenni, come anche dimostra la letteratura, le reti neurali hanno dimostrato la loro elevata capacità predittiva, essi affrontano il problema di prevedere non solo i prezzi, ma anche i movimenti dell'indice BIST 30 del mercato turco

²⁰⁹ Qiu M., Song Y., (2016), *Predicting the Direction of Stock Market Index Movement Using an Optimized Artificial Neural Network Model*, PLOS ONE, 19/05/2016.

²¹⁰ L'indice MICEX è il principale indicatore del comportamento della Borsa russa e viene calcolato in base alle cinquanta azioni più liquide delle società russe. Fonte: Teleborsa, 07/02/2020.

²¹¹ Lozinskaia A., Zhemchuzhnikov V., (2017), "Micex Index forecasting: the predictive power of neural network modelling and support vector machine", *Perm University Herald*, 1, pp. 49 – 60.

²¹² Raso H., Demirci M., (2019), "Predicting the Turkish Stock Market BIST 30 Index using Deep Learning", *International Journal of Engineering Research and Development*, 1, pp. 253 – 265.

proponendo un modello di apprendimento profondo avente come dati di input diversi indicatori tecnici e oscillatori calcolati sui prezzi storici dell'indice stesso. Il set di dati considerato comprende un periodo di 27 mesi, ovvero da 1° gennaio 2016 all'11 aprile 2018: ogni osservazione contiene la data di negoziazione, i valori di chiusura e di apertura, il prezzo più basso e quello più alto, nonché il volume e il cambiamento percentuale rispetto al giorno di negoziazione precedente²¹³. L'output risultante avrebbe dovuto essere il prezzo dell'Indice BIST 30 nei cinque giorni di negoziazioni successivi. In tale contesto, e in seguito ai risultati ottenuti, l'apprendimento profondo si è dimostrato uno strumento promettente a fini previsivi²¹⁴.

²¹³ Raso H., Demirci M., (2019), "Predicting the Turkish Stock Market BIST 30 Index using Deep Learning", *International Journal of Engineering Research and Development*, 1, pp. 253 – 265.

²¹⁴ Ibidem.

CAPITOLO IV – GLI AMBITI APPLICATIVI DELL’INTELLIGENZA ARTIFICIALE

Oggi l’intelligenza artificiale rappresenta uno dei principali ambiti di interesse della comunità scientifica informatica con temi di ricerca come il Machine Learning, l’elaborazione del linguaggio naturale e la robotica, proprio per questo le aziende informatiche stanno investendo sempre di più in questo settore.

L’intelligenza artificiale, nata inizialmente per semplificare e apportare maggiore efficienza ed efficacia alle più svariate attività umane, sta diventando una parte importante, se non essenziale, della quotidianità di tutti gli individui.

Soprattutto negli ultimi anni si è assistito a una crescente eterogeneità di strumenti, piattaforme e applicazioni basate sull’intelligenza artificiale e sull’apprendimento automatico; proprio questi strumenti tecnologici hanno trasformato, non solo l’industria di Internet e del Software, ma hanno trovato applicazione, e quindi impattato, su una vasta gamma di settori.

Nel settore agricolo i progressi, nonché l’applicazione dell’intelligenza artificiale, hanno condotto a miglioramenti nell’ottenimento della resa e nell’aumento della ricerca e dello sviluppo delle colture in crescita. Con le “tecnologie intelligenti” risulta attualmente possibile prevedere il tempo necessario affinché una specifica coltura sia matura e pronta per la raccolta, ancora è possibile monitorare le colture e il terreno al fine di renderli più facilmente sostenibili per gli agricoltori e aumentandone così l’efficienza. Altre specializzazioni in tale settore riguardano l’automazione delle serre, la simulazione, la modellizzazione e le tecniche di ottimizzazione. Un esempio di applicazione dell’intelligenza artificiale in ambito agricolo è il robot *See & Spray*, sviluppato dalla Blue River Technology, una start – up californiana, che utilizza tecnologie di visione artificiale, come il rilevamento di oggetti, per monitorare e irrorare con precisione un antiparassitario sulle piante di cotone²¹⁵.

²¹⁵ Lateef Z., (2019), *AI Applications: Top 10 Real World Artificial Intelligence Applications*, Edureka, 22/05/2019.

Oltre a questo, la PEAT, start – up tecnologica con sede a Berlino, ha sviluppato un’applicazione chiamata Plantix che è in grado di identificare i potenziali difetti e carenze di nutrienti nel suolo attraverso le immagini²¹⁶.

Come affermato recentemente dal presidente di Coldiretti Puglia, solamente il 2% delle attività agricole italiane utilizza robot e sensori²¹⁷. La diffusione dell’intelligenza artificiale applicata in tale settore, in Italia, risulta irregolare con regioni che la utilizzano su meno dell’1% del terreno, fino ad un massimo del 4-5%; tuttavia è probabile che queste percentuali aumentino dato che, secondo le linee guida del Ministero per lo sviluppo dell’agricoltura in Italia, dovrebbero raggiungere il 10% nel 2021, mentre la loro diffusione è già ben consolidata in Cina, Israele e Stati Uniti le cui percentuali si aggirano intorno al 40 – 70% ; un rapido aumento si è verificato recentemente anche in Brasile, Argentina, India e Malesia²¹⁸.

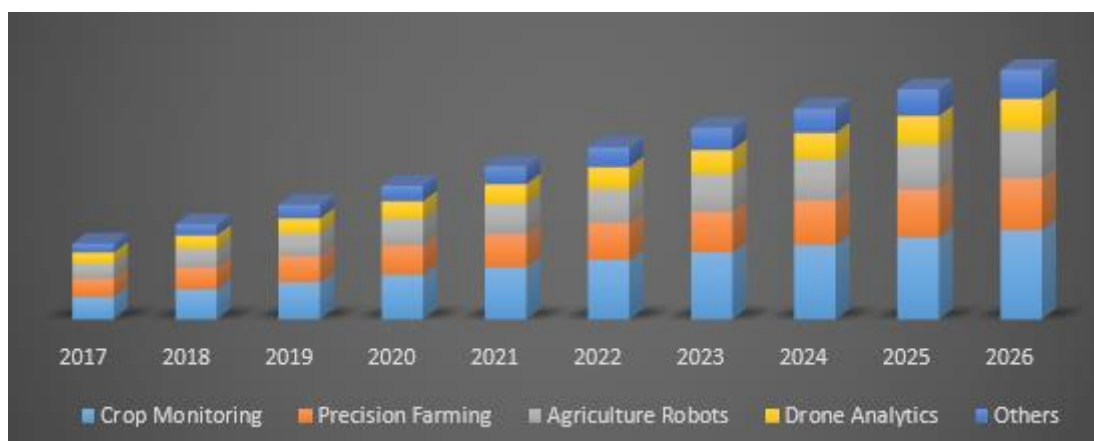


FIGURA 24: GLOBAL ARTIFICIAL INTELLIGENCE IN AGRICULTURE MARKET BY APPLICATION. FONTE: MAXIMIZE MARKET REASEARCH PVT.LDT., 2019.

In ambito sanitario, conseguentemente l’aumento della complessità e dei dati, molte organizzazioni e centri di assistenza medica fanno affidamento sull’intelligenza artificiale per quel che riguarda accertamenti diagnostici e trattamenti, coinvolgimento e aderenza dei pazienti, nonché per lo svolgimento di attività amministrative²¹⁹.

²¹⁶ Lateef Z., (2019), *AI Applications: Top 10 Real World Artificial Intelligence Applications*, Edureka, 22/05/2019.

²¹⁷ Farsagli S., (2018), *Precision Farming: do sensors, data and AI improve production?*, Ingenium: the web Magazine of Engineering and Tech Economy 2030 on the Digital Transformation, 28/11/2018.

²¹⁸ Ibidem.

²¹⁹ Davenport T., Kalakota R., (2019), “The potential for artificial intelligence in healthcare”, *Future Healthcare Journal*, 6, pp. 94 – 98.

La tecnica maggiormente utilizzata in tale settore è l'apprendimento automatico attraverso il quale è possibile applicare la medicina di precisione, la quale è in grado di prevedere quali protocolli di trattamento possono avere successo su un paziente in base a diversi attributi dello stesso considerando anche il contesto di trattamento. Altre tecniche utilizzate sono, senza dubbio, le reti neurali le quali, sfruttano la loro capacità di categorizzazione permettendo di determinare se uno specifico paziente è predisposto o meno a una particolare malattia.

Da un paio di decenni, inoltre, negli Stati Uniti sono stati approvati dei robot chirurgici, vale a dire robot fisici che affiancano i chirurghi, migliorando la loro capacità di vedere, creare incisioni precise e minimamente invasive²²⁰.

A titolo di esempio, un'organizzazione chiamata *Cambio Health Care* ha sviluppato un sistema di supporto alle decisioni cliniche per la prevenzione dell'ictus; tale sistema è in grado di dare al medico un avvertimento quando il paziente è soggetto a un particolare rischio. Un altro esempio è l'azienda *Coala Life* che ha ideato un dispositivo digitalizzato in grado di identificare le malattie cardiache²²¹.

Durante il Milan MIT Critical Care Datathon and ESICM's Big Datatalk avvenuto nel 2019, evento europeo dedicato allo studio delle potenziali applicazioni dell'intelligenza artificiale in ambito sanitario, è stata identificata l'AI come la prossima frontiera digitale nel settore sanitario; inoltre durante il corso del Google Cloud Summit di Milano avvenuto nel maggio dello scorso anno, la Noovle e l'Arsena. IT hanno presentato un progetto che utilizza algoritmi di Machine Learning per classificare automaticamente i documenti clinici digitali ed estrarre da quelli non strutturati, il maggior numero possibile di informazioni rilevanti²²².

Nonostante i numerosi benefici dell'intelligenza artificiale apportati nel settore dell'assistenza sanitaria, secondo un'indagine condotta da *Himss Analytics* solo il 16% delle strutture sanitarie in Europa la utilizza già: il 25% ha un piano specifico mentre il 59% non usa tali strumentazioni né ha in programma di farlo. Al momento, gli operatori

²²⁰ Davenport T., Kalakota R., (2019), "The potential for artificial intelligence in healthcare", *Future Healthcare Journal*, 6, pp. 94 – 98.

²²¹ Lateef Z., (2019), *AI Applications: Top 10 Real World Artificial Intelligence Applications*, Edureka, 22/05/2019.

²²² Anonimo, (2019), *Sanità: telemedicina, big data, intelligenza artificiale e competenze digitali saranno i temi centrali del 2019: gli ultimi eventi sul tema*, Crescita digitale, 14/02/2019.

sanitari europei sfruttano questi strumenti più frequentemente nella gestione dei flussi di lavoro (14%), nella ricerca (13%), nell'amministrazione dei farmaci (12%) e nella radiologia (11%). Queste aree, insieme all'oncologia, sono anche quelle in cui gli operatori sanitari investono maggiormente.²²³.

Adottare l'intelligenza artificiale richiede infatti, la capacità di affrontare alcune sfide e di gestire determinati rischi. Secondo il sondaggio, la mancanza di maturità del prodotto e la fiducia da parte del personale medico, sono considerate le maggiori sfide che ostacolano un uso più diffuso di questi meccanismi nell'assistenza sanitaria in Europa²²⁴. Seguono poi la privacy e i problemi legati all'interoperabilità. In Italia, invece, le preoccupazioni legate all'approvazione legale e la mancanza di fiducia da parte del personale medico, sono percepiti come i più grandi ostacoli per l'uso, maggiormente ricorrente, di queste soluzioni²²⁵. Si prevede un tasso di crescita superiore al 50% da oggi al 2027 in quanto il mercato sanitario sembra essere un settore molto promettente e con alto potenziale per l'applicazione delle tecnologie intelligenti²²⁶, come si evince anche dal grafico sottostante:



FIGURA 25: PREVISIONE DEL TASSO ANNUO DI CRESCITA GLOBALE DEL MERCATO DELL'INTELLIGENZA ARTIFICIALE NEL SETTORE SANITARIO. FONTE: WAINHOUSE RESEARCH, 2017.

²²³ Della Porta M., (2019), *Intelligenza artificiale, l'inarrestabile corsa nel settore sanitario*, Istituto per la competitività, 12/11/2019.

²²⁴ Ibidem.

²²⁵ Ibidem.

²²⁶ Anonimo, (2019), *Sanità: telemedicina, big data, intelligenza artificiale e competenze digitali saranno i temi centrali del 2019: gli ultimi eventi sul tema*, Crescita digitale, 14/02/2019.

Nell'ambito dell'istruzione, la ricerca sull'intelligenza artificiale e l'apprendimento automatico risale alla fine degli anni '70. I metodi di A. I. erano generalmente impiegati per progettare ambienti di apprendimento interattivi al fine di supportare e implementare sistemi di tutoraggio. I rapidi progressi della tecnologia negli ultimi anni però, hanno comportato cambiamenti più importanti, quali nuove opportunità per insegnare e apprendere ovunque e in qualsiasi momento fornendo strumenti e apportando così risultati migliori²²⁷. Kai – Fu Lee, uno dei maggiori esperti mondiali di intelligenza artificiale, paragona l'istruzione moderna a una catena di montaggio, in cui gli studenti avanzano di classe in classe senza prestare attenzione ed evidenziare le differenze in termini di apprendimento, predilezioni e capacità personali²²⁸. Egli ha dimostrato che, proprio grazie l'intelligenza artificiale, si potrebbe creare un sistema di apprendimento su misura per ogni singolo studente; tale tipologia di "educazione intelligente" si concentrerà principalmente sull'insegnamento, i compiti a casa, quelli in classe e il sostegno. Gli algoritmi di *Deep Learning* sarebbero infatti, in grado di analizzare tutte le performance e i risultati ottenuti dagli studenti in ciascuno degli scenari sopracitati, permettendo conseguentemente la creazione di un database complessivo rappresentante il profilo di ogni studente²²⁹.

Oltre a questo, l'utilizzo dell'A. I. nel settore educativo, può portare a una drastica riduzione del tempo che gli educatori impiegano in attività di back office e attività correlate non didattiche²³⁰: è attualmente possibile ricorrere a software per la correzione di test a scelta multipla o a risposta chiusa.

Nell'ambito scolastico italiano, nonché di istruzione e di formazione, l'intelligenza artificiale si fa strada con molta cautela; Giorgio Robino, ricercatore dell'Istituto per le tecnologie didattiche del Consiglio Nazionale delle Ricerche ha dichiarato che i sistemi intelligenti << *sono sistemi software che comunicano con gli esseri umani (o altre macchine) in linguaggio naturale. È necessaria una task force che metta insieme insegnanti (ed esperti del "contenuto"). Esperti di chatbots design, di linguaggio (non*

²²⁷ Kulkarni A., (2019), *AI in Education: Where is It Now and What is the Future?*, Lexalytics, 06/09/2019.

²²⁸ Daniele A., (2019), *La scuola del futuro: come l'intelligenza artificiale rivoluzionerà l'istruzione scolastica*, Le Macchine volanti, 24/06/2019.

²²⁹ Ibidem.

²³⁰ Schmelzer R., (2019), *AI Applications in Education*, Forbes, 12/11/2019.

solo in termini di linguistica computazionale ma proprio in termini di educazione all'uso raffinato della lingua italiana). Ultimi, ma non ultimi, servono esperti di comunicazione, esperti di psicologia, scrittura creativa, umanisti >>²³¹.

In tale settore si prevede una crescita costante per i prossimi tre anni, come dimostra il grafico che segue:

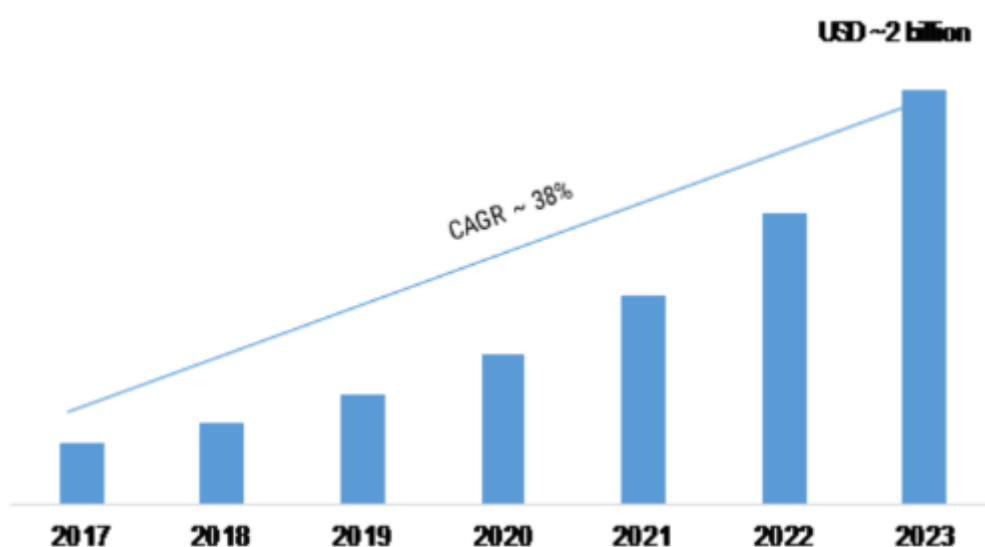


FIGURA 26: PREVISIONE DEL TASSO ANNUO DI CRESCITA GLOBALE DEL MERCATO DELL'INTELLIGENZA ARTIFICIALE NEL SETTORE DELL'ISTRUZIONE. FONTE: ANALISI MARKET RESEARCH FUTURE, 2020.

Nel marketing, definito come il processo di gestione attraverso il quale i beni e i servizi passano dal concetto ai clienti²³², prevale la presenza di assistenti vocali e/o virtuali che, sfruttando l'intelligenza artificiale sia per il riconoscimento del linguaggio naturale sia per l'apprendimento e l'analisi delle abitudini e dei comportamenti degli utenti, sono in grado di perfezionare meccanismi di ingaggio con attività predittive dei comportamenti di acquisto da cui è possibile derivare strategie di comunicazione o proposta di servizi.

Attualmente, a causa della complessità del mercato e una concorrenza più affollata e agguerrita, quasi non esistono prodotti o servizi unici per cui, la maggior parte delle persone che lavorano in tale settore, sono sottoposte a difficoltà non indifferenti

²³¹ Nicita G., (2019), *Intelligenza Artificiale e apprendimento: alte percentuali negli USA*, DIGITAL EDUCATION LAB, 30/08/2018.

²³² Business Dictionary (2018), Marketing, disponibile a <http://www.businessdictionary.com/definition/marketing.html>. 2018

nell'identificare i segmenti di clientela adeguati e nel pubblicizzare al pubblico nel modo più efficace ed efficiente possibile. Conseguentemente, l'obiettivo è competere per ottenere il vantaggio nel mercato.

Individuare il mercato più redditizio per un prodotto e/o un servizio è una delle funzioni attualmente disponibili grazie all'applicazione dell'intelligenza artificiale. Quest'ultima è in grado di anticipare le esigenze dei clienti, fornire assistenza nella creazione di campagne altamente personalizzate, identificare i modelli di acquisto dei clienti, nonché analizzare considerevoli quantità di dati al fine di identificare le caratteristiche di alto valore che contraddistinguono i clienti²³³.

Sostanzialmente, attraverso l'utilizzo e l'implementazione di algoritmi di intelligenza artificiale e Machine Learning, si cerca di persuadere i potenziali clienti a compiere una determinata azione, acquistare uno specifico prodotto o accedere a un particolare servizio²³⁴.

Nel 2019 alcune delle aziende digitali di maggior successo hanno costruito le loro offerte di prodotti e/o servizi attorno alla capacità di fornire consigli su prodotti o contenuti altamente pertinenti e personalizzati, tra cui Amazon, Netflix e Spotify: come ha scritto Lori Goldberg, autrice di Econsultancy, << tutto questo deriva dal clustering basato sull'intelligenza artificiale e dall'interpretazione dei dati dei consumatori associati a informazioni di profilo e dati demografici. Questi sistemi basati sull'intelligenza artificiale si adattano continuamente ai tuoi gusti e alle tue antipatie e reagiscono con nuovi consigli personalizzati in tempo reale>>²³⁵.

In ambito commerciale, oramai da tempo, l'intelligenza artificiale ha portato innumerevoli vantaggi. Essendo che molte aziende si trovano ancora nella fase iniziale dell'applicazione delle tecnologie intelligenti, l'Osservatorio di Intelligenza Artificiale ha ideato un modello in grado di guidare le aziende nel cosiddetto *Artificial Intelligence Journey*, ovvero uno strumento in grado di mettere in evidenza i principali elementi

²³³ Geisel A., (2018), "Current and Future Impact of Artificial Intelligence On Business", *Journal of Scientific & Technology Research*, 7, pp. 116 – 121.

²³⁴ Consoft Sistemi, (2019), *L'intelligenza artificiale al servizio dell'uomo*, Digital Transformation per lo sviluppo sostenibile, 23/05/2019.

²³⁵ Sentance R., (2019), *15 examples of artificial intelligence in marketing*, Econsultancy, 28/02/2019.

organizzativi e tecnologici da prendere in considerazione nel momento in cui si adottano soluzioni intelligenti, in modo da svilupparle, renderle sostenibili e ottenerne una gestione efficace ed efficiente.

Tale modello, al fine di valutare lo stato in cui si trova l'azienda, considera diverse dimensioni: i dati e il patrimonio informatico, i quali forniscono informazioni sul livello qualitativo e quantitativo dei dati necessari per lo sviluppo di algoritmi intelligenti, le metodologie e i relativi algoritmi che indicano le capacità di comprensione dell'A. I., la misura della capacità di possedere o acquisire le competenze necessarie per lo sviluppo delle tecnologie dell'intelligenza artificiale, la relazione con il cliente ovvero la capacità di comprendere le esigenze degli utenti finali e infine la cultura aziendale per evidenziare un'eventuale propensione ai cambiamenti apportati dallo sviluppo e dall'adozione di sistemi di A. I.²³⁶

Grazie a tale strumento, l'Osservatorio ha rilevato sei diversi livelli di maturità nelle aziende italiane; le aziende "in ritardo", ovvero nella fase iniziale del loro percorso, le quali, sebbene dispongano di infrastrutture per la raccolta di dati, questi ultimi sono insufficienti in termini quantitativi e qualitativi; diversamente, le aziende che sfruttano i dati a loro disposizione applicando soluzioni standard offerte dal mercato stesso sono dette "entusiaste". Rispetto a quest'ultima tipologia, le aziende "in cammino" hanno consolidato le metodologie di A. I. e rafforzato le dimensioni di cultura aziendale e organizzazione delle competenze; poi si ritrovano gli "apprendisti", ovvero imprese che hanno investito maggiormente sulla qualità e quantità dei dati in quanto, grazie ad un intervento migliorativo in ambito organizzativo e culturale, sono state in grado di incentivare lo sviluppo tecnologico²³⁷. Successivamente, le aziende caratterizzate da un profilo intermedio, le cosiddette "organizzate", che hanno sviluppato l'ambito organizzativo – culturale in misura maggiore rispetto a quello tecnologico ritenendo che sia necessario intervenire innanzitutto sul miglioramento delle competenze interne e sulla preparazione del cliente che utilizzerà il prodotto e/o servizio.

²³⁶ Redazione Osservatori, (2019), *AI Journey: un modello per fare Artificial Intelligence nelle imprese*, Osservatori. net digital innovation, 26/03/2019.

²³⁷ Ibidem.

Infine le aziende “avanguardiste” che presentano il profilo maggiormente evoluto e includono le aziende che hanno investito per prime nell’A.I.²³⁸

In ogni caso, un numero sempre maggiore di aziende adotta sistemi produttivi in grado di gestire in maniera autonoma eventuali anomalie sulla base dell’esperienza acquisita grazie ad analisi precedentemente effettuate e creando così la rappresentazione del processo produttivo secondo variabili non programmate; tale sistema consente quindi anche un’attività di manutenzione predittiva²³⁹. Tutto ciò, incide positivamente sui benefici che le aziende possono ottenere; il 50% delle aziende, che hanno già investito nelle tecnologie intelligenti, dichiara di essersi posta l’obiettivo di ridurre i costi tramite un miglioramento in termini di efficacia ed efficienza dei processi; il 37% dichiara invece di perseguire l’obiettivo di aumentare i ricavi mentre il restante 13% ha voluto sviluppare soluzioni di supporto decisionale²⁴⁰. Tra le aziende che hanno avviato progetti di A. I., soltanto il 4% non ha raggiunto gli obiettivi prefissati, mentre ben il 68% dichiara di avere in corso progetti che si stanno rivelando di grande successo. Il restante 28%, invece, non è ancora in grado di esprimere un giudizio²⁴¹.

Il 29% delle aziende energetiche di tutto il mondo inoltre, afferma di aver implementato soluzioni basate sull’intelligenza artificiale e aver conseguito così risultati soddisfacenti. Le soluzioni che hanno adottato tali aziende riguardano essenzialmente alcune aree chiave tra cui la catena di approvvigionamento, nella quale, le reti intelligenti sono in grado di prevedere la domanda e l’offerta di energia, effettuare analisi intelligenti in tempo reale o rilevare errori o frodi. Altre soluzioni includono manutenzione predittiva, come già accennato pocanzi, e l’efficienza energetica²⁴².

In particolar modo si mette in evidenza il settore manifatturiero e dell’automazione industriale, nel quale l’Italia, in termini di applicazione di sistemi intelligenti, si trovano ai vertici europei. L’automazione industriale e l’A. I. stanno apportando considerevoli cambiamenti all’operatività grazie all’avvento della robotica di nuova generazione con

²³⁸ Redazione Osservatori, (2019), *AI Journey: un modello per fare Artificial Intelligence nelle imprese*, Osservatori.net digital innovation, 26/03/2019.

²³⁹ Confindustria Digitale (2019), *Intelligenza Artificiale per le PMI*, disponibile a <http://preparatialfuturo.confindustria.it/wp-content/uploads/2019/02/IA-per-PMI.pdf>.

²⁴⁰ Redazione Osservatori, (2019), *Le applicazioni dell’Intelligenza Artificiale e la loro diffusione*, Blog Osservatori Digital Innovation, 26/04/2019.

²⁴¹ Ibidem.

²⁴² Soluciones D., (2018), *How different sectors are using AI*, Medium, 29/10/2018.

importanti margini di miglioramento²⁴³. I robot attualmente utilizzabili, sono in grado di fornire livelli superiore di garanzia e qualità, minimizzare i costi operativi e rendere la produzione più efficace ed efficiente. Negli ultimi cinque anni, la robotica italiana è aumentata del 12% con un picco del 19% nel 2018, ma è prevista una notevole crescita anche della robotica di servizio in un mercato mondiale²⁴⁴.

Anche il settore bancario, storicamente resistente ai cambiamenti, sta subendo notevoli trasformazioni grazie alla tecnologia e il particolar modo all'intelligenza artificiale.

Le tecnologie hanno consentito la diffusione di meccanismi innovativi per creare rapporti con la clientela e l'offerta si è ampliata grazie a un elevato numero di servizi completamente nuovi. Un'indagine condotta da Abi Lab, il centro di ricerca per l'innovazione e la banca, promosso dall'Associazione bancaria, ha evidenziato infatti che per il 65% degli istituti di credito italiani, l'intelligenza artificiale rientra tra le maggiori priorità, mentre il 94% delle banche nazionali dichiara che il trend di investimento nell'A. I. è in crescita o costante rispetto all'anno precedente²⁴⁵.

I principali ambiti di utilizzo dell'intelligenza artificiale in tale settore riguardano tutti i servizi rivolti al cliente, in particolar modo la gestione dei *contact center*, credito, sicurezza e finanza²⁴⁶; oramai è possibile fornire servizi personalizzati e più efficienti per ogni singolo cliente, aumentando così le entrate, prendendo decisioni più rapide e instaurando un buon rapporto con gli stessi clienti. Un esempio tipico di A. I. adottata dagli istituti che compongono il settore bancario sono i *chatbot*, che consistono in un programma di chat automatizzato che viene eseguito automaticamente o segue un percorso predeterminato; tale strumento permette, agli individui aventi a disposizione poco tempo, di richiedere il loro denaro in qualsiasi momento. I *chatbot* sono un modo intelligente di fornire un servizio efficiente ai clienti²⁴⁷.

²⁴³ Ministero dello Sviluppo Economico (2019), *Proposte per una Strategia Italiana per l'Intelligenza Artificiale*, Roma, Maggio 2019.

²⁴⁴ Ibidem.

²⁴⁵ Anonimo, (2019), *Banche, ABI: Intelligenza artificiale per migliorare il servizio al cliente*, QuiFinanza, 25/11/2019.

²⁴⁶ Ibidem.

²⁴⁷ North R., (2019), *Artificial Intelligence – A Boon To The Banking Industry*, Enterprise Edges.

Un particolare tipo di assistente virtuale è il *Robo – Advisor* grazie al quale è possibile ricevere consigli personalizzati di gestione patrimoniale in funzione ai propri obiettivi e, eventualmente, acquistare o vendere prodotti finanziari. Esso è in grado anche di seguire l'andamento degli investimenti offrendo al cliente report periodici e proponendo misure correttive, quando necessario²⁴⁸.

Uno dei maggiori pilastri dell'attività delle banche commerciali è rappresentato dalla valutazione del merito di credito, processo anch'esso affinato, nel corso degli ultimi anni, con metodologie intelligenti e l'accesso a dataset sempre più vasti. Nel *Credit Scoring*, l'A. I. può fare la differenza grazie al suo ruolo di estensione delle procedure in essere: essa, oltre a considerare i parametri consolidati che stanno alla base del *credit scoring*, è in grado di completare la valutazione esaminando dati estremamente eterogenei come ad esempio il comportamento assunto sui *social network*²⁴⁹.

Sempre nel medesimo settore, la prevenzione delle frodi ricopre un'importanza fondamentale; anche in tale ambito, l'implementazione dell'intelligenza artificiale, all'interno di algoritmi consolidati, è divenuta un fattore di successo: ogni qualvolta viene rilevata una transazione fraudolenta in un estratto conto, viene portata all'attenzione delle banche e, le autorità bancarie competenti, adottano immediatamente le misure necessarie per gestire la situazione proteggendo le violazioni di sicurezza²⁵⁰.

²⁴⁸ Anonimo, (2019), *Intelligenza artificiale in banca: cinque scenari di successo*, Magazine Teorema, 21/03/2019.

²⁴⁹ Ibidem.

²⁵⁰ North R., (2019), *Artificial Intelligence – A Boon To The Banking Industry*, Enterprise Edges.

Per un ulteriore approfondimento sulle iniziative già attivate nel settore bancario e i loro rispettivi sviluppi, si veda il grafico sottostante:

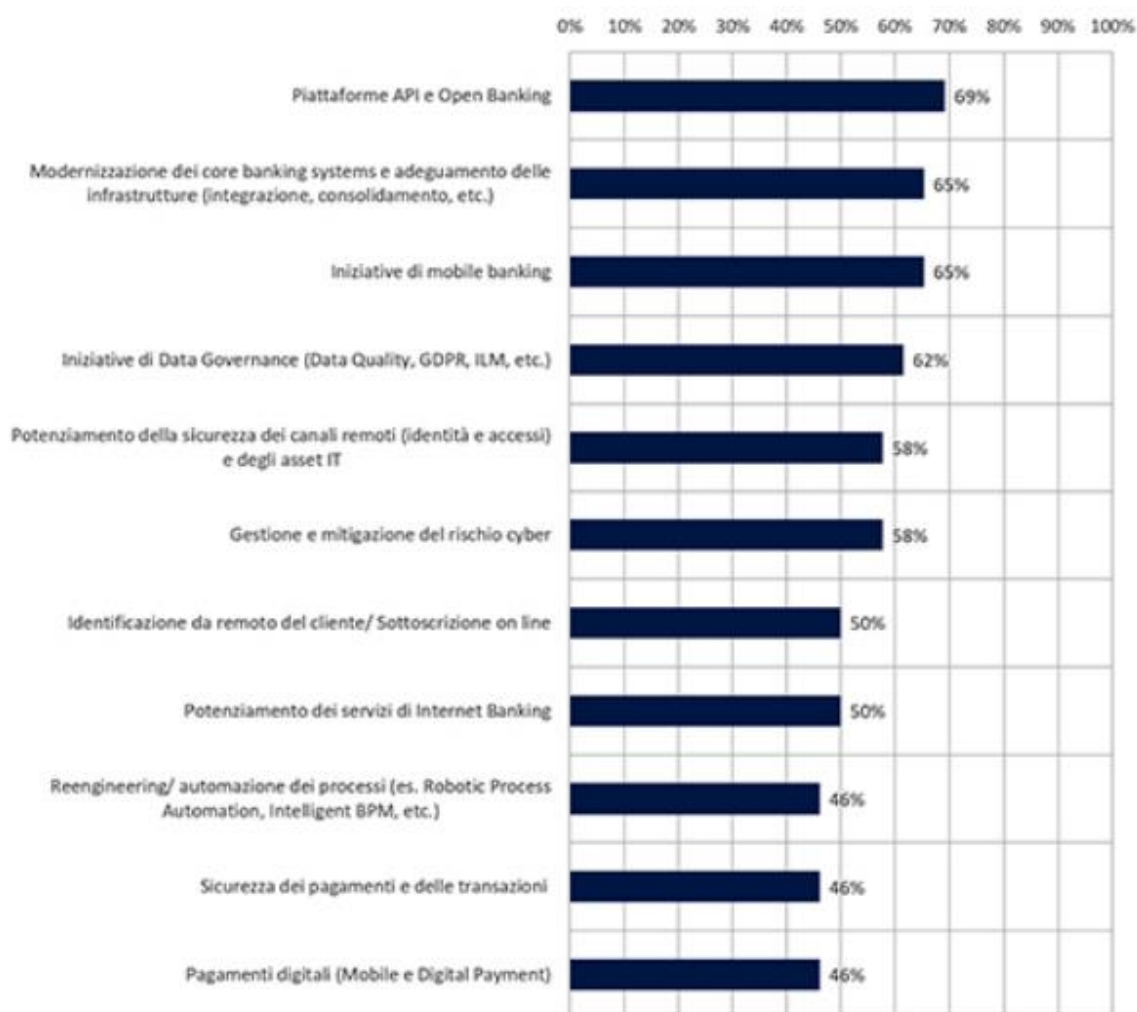


FIGURA 27: RILEVAZIONE SULLE PRIORITÀ ICT DELLE BANCHE ITALIANE, ABI LAB, GENNAIO 2020.

È importante tenere presente però, che con l'aumento costante della concorrenza in tale settore e la quasi oramai inesistente unicità di prodotti e/o servizi finanziari, tale rivoluzione tecnologica ha permesso l'entrata nell'arena competitiva, oltre che degli operatori bancari, anche dei *digital champions* (come Google, Amazon, Facebook, Apple, PayPal)²⁵¹ e delle Fintech, ovvero innovazioni finanziarie rese possibili con l'evoluzione tecnologica che possono concretizzarsi in nuovi modelli di business, processi o prodotti,

²⁵¹ KPMG (2018), *Digital Banking. L'evoluzione delle aspettative dei clienti tra rivoluzione digitale, sfide regolamentari e nuovi competitor*, disponibile a www.kpmg.com/it.

producendo un effetto considerevole sui mercati finanziari, sulle istituzioni o sull'offerta di servizi, nonché sull'economia nazionale²⁵².

Le innovazioni considerate comprendono sia servizi finanziari che tecnologie informatiche e influenzano tutti i settori dell'intermediazione bancaria e finanziaria: dal credito ai servizi di pagamento, dalle valute virtuali ai servizi di consulenza, dalle tecnologie di validazione decentralata delle transazioni all'identificazione biometrica, fino al supporto dell'erogazione di servizi: l'A. I. investe quindi qualsiasi segmento dei mercati bancari e finanziari modificandone la struttura con l'ingresso di start – up tecnologiche, giganti della tecnologia e social media.

Attualmente si parla di *digital banking*, ovvero la digitalizzazione di ogni livello, dal *front – end* al *back – end* che comprende compiti amministrativi ed elaborazione dei dati alleviando così, la pressione esercitata sui dipendenti. Il sistema bancario digitale, nato a seguito delle crescenti esigenze dei consumatori di trovare modi più efficienti per accedere ai registri bancari e completare le transazioni finanziarie al di fuori delle filiali locali, si è sviluppato inizialmente nei servizi bancari online di base come il pagamento di bollette e giroconti²⁵³. L'*online banking* ha poi ispirato e, dato vita, al *mobile banking* il quale offre essenzialmente gli stessi servizi ma dalla comodità del proprio dispositivo mobile²⁵⁴.

In ogni caso, a seguito di un'indagine condotta dalla KPMG risulta che il nostro Paese è caratterizzato da un indice di penetrazione dell'A. I. nel settore bancario inferiore alla media globale, per cui risulta ancora abbastanza arretrato. Secondo dati Eurostat, solo il 31% della popolazione italiana utilizza le tecnologie disponibili per tale settore, diversamente dalla percentuale europea che è superiore al 50%.²⁵⁵ Emerge, in modo chiaro, come il canale Internet sorpassi la filiale nelle preferenze dei clienti, rappresentando il principale canale per l'acquisto e la gestione di prodotti e servizi

²⁵² Financial Stability Board (2017), *Fintech credit Market structure: business models and financial stability implications*, disponibile a <http://www.fsb.org/2017/05/fintech-credit-market-structure-business-models-and-financial-stability-implications/>.

²⁵⁴ Phaneuf A., (2019), *The disruptive digital transformation of banking services in 2019*, Business Insider, 12/09/2019.

²⁵⁵ KPMG (2018), *Digital Banking. L'evoluzione delle aspettative dei clienti tra rivoluzione digitale, sfide regolamentari e nuovi competitor*, disponibile a www.kpmg.com/it.

finanziari. Diversamente, per effettuare operazioni più complesse come la sottoscrizione di un mutuo e la risoluzione di eventuali problematiche, la filiale risulta ancora essere fondamentale.

Dato il ritmo incredibile con il quale le istituzioni bancarie stanno adottando soluzioni di intelligenza artificiale, e dal momento che gli esseri umani sono più inclini a commettere errori, si prevede che arriverà il giorno in cui queste tecnologie diventeranno estremamente avanzate e saranno in grado di compiere qualsivoglia operazione²⁵⁶. Si denota, dal grafico sottostante, un ritmo di crescita globale piuttosto importante infatti, già nel prossimo decennio si prevede che il valore del mercato dell'intelligenza artificiale triplichi rispetto ad oggi:

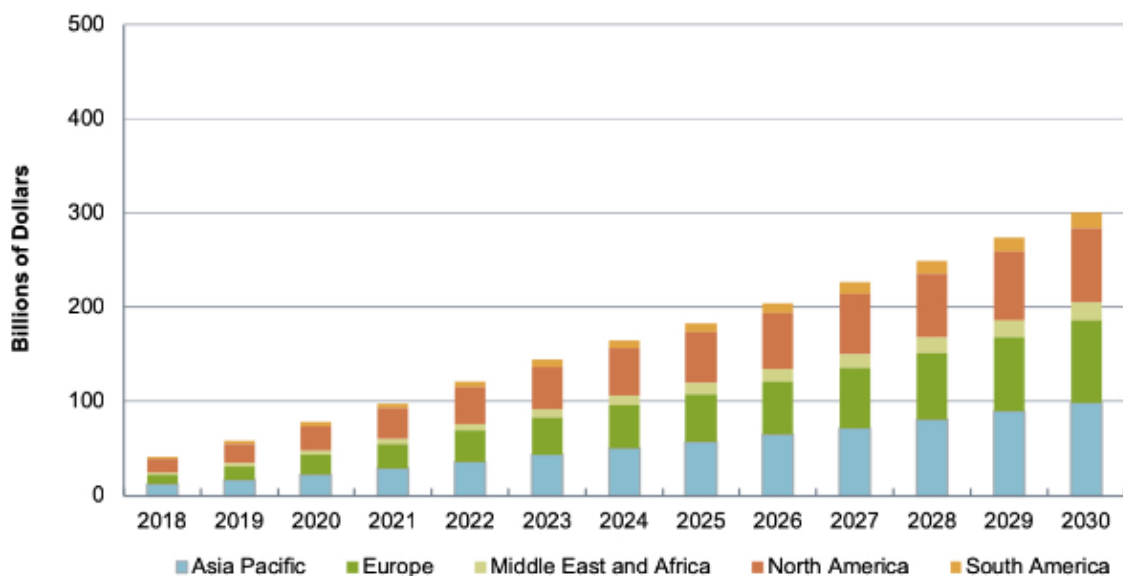


FIGURA 28: THE BUSINESS VALUE FOR THE WORLD MARKET FOR AI IN BAKING BY REGION. FONTE: IHS MARKIT, 2019.

Per il mondo della finanza, l'intelligenza artificiale è il volto dell'innovazione.

L'informatica ha già rivoluzionato il settore finanziario una volta, facilitando un numero enorme di calcoli in una frazione di secondo e ha seguito i mercati che si muovono alla velocità della luce. Ora, i sistemi di intelligenza artificiale sono pronti a favorire una

²⁵⁶ Geisel A., (2018), "Current and Future Impact of Artificial Intelligence On Business", *Journal of Scientific & Technology Research*, 7, pp. 116 – 121.

seconda ondata di innovazione, che rappresenterà l'evoluzione più significativa nella storia della finanza²⁵⁷.

Le società finanziarie hanno congiuntamente investito in modo significativo nell'A. I. in passato e, molte altre, stanno iniziando a studiare e implementare le applicazioni finanziarie dell'apprendimento automatico e del *Deep Learning* nelle loro operazioni²⁵⁸. Nel 2010, il commercio di algoritmi ad alta frequenza rappresentava circa il 70% delle negoziazioni solo negli Stati Uniti; percentuale che, entro il 2014 è salita al 75% ed entro il 2017 rappresentava il circa 90%. In questo stesso anno, Wall Street ha avuto il suo primo fondo di trading azionario basato esclusivamente e totalmente sull'intelligenza artificiale grazie all'IBM Watson ovvero un computer in grado di elaborare ed analizzare le notizie e i rapporti relativi a 6000 società americane; nella prima settimana di attività è cresciuto dell'1% e dopo un anno, nell'agosto 2018, le sue azioni erano aumentate del 20%²⁵⁹.

Sicuramente il caposaldo di tale settore, e anche ai fini di tale studio, è il trading; le società di investimento si affidano a computer e analisti per determinare i modelli futuri sul mercato indispensabili per effettuare trading ed investimenti profittevoli²⁶⁰.

Gli ultimi dati di Aite Group e successivamente ripresi dal Sole 24 Ore, mettono in evidenza come attualmente l'intelligenza artificiale domina il trading di Borsa. Nel 2018 i robot hanno gestito circa il 53% degli scambi delle azioni globali *cash*, percentuale che negli USA arriva al 66% e in Europa si aggira intorno al 47%²⁶¹.

Sono presenti innumerevoli metodologie applicabili nel settore finanziario; nel Capitolo III sono state presentate le principali, con le relative applicazioni.

Sempre in tale ambito, l'A. I. è in grado di fornire un'attività consulenziale suggerendo soluzioni maggiormente adeguate in base alle specifiche esigenze e alla propensione al rischio di ogni investitore oltre che all'attività di gestione delle finanze personali svolta

²⁵⁷ Zamagna R., (2018), *The Future of trading belong to Artificial Intelligence*, Medium, 15/11/2018.

²⁵⁸ Zamagna R., (2018), *The Future of trading belong to Artificial Intelligence*, Medium, 15/11/2018.

²⁵⁹ Ibidem.

²⁶⁰ Ciotola M., (2018), *Cinque modi in cui l'intelligenza artificiale sta rivoluzionando la finanza*, Money, 07/04/2018.

²⁶¹ Tessa M., (2019), *Intelligenza artificiale regina dei listini, gestisce metà trading*, Wall Street Italia, 03/06/2019.

ricorrendo a degli algoritmi che aiutano i consumatori a prendere decisioni intelligenti sull'investire i propri risparmi e gestire il patrimonio personale²⁶².

Successo recentemente raggiunto, ma di rilevante importanza, è il lancio dell'Amundi Stoxx Global Artificial Intelligence UCITS ETF da parte dell'Amundi, una società di gestione, nonché uno dei grandi protagonisti mondiali dell'*asset management* operante nelle principali piazze finanziarie del mondo²⁶³.

Tale ETF innovativo, replicando l'indice Stoxx Global Artificial Intelligence ADTV5 che utilizza un approccio equi - ponderato offrendo una diversificazione in termini di paesi, settori e capitalizzazione di mercato, consente di investire in aziende fortemente esposte alle opportunità di crescita potenziali legate all'intelligenza artificiale; si basa su un processo di selezione sistematico che analizza le azioni in base all'esposizione all'A. I. (misurata come la proporzione tra i brevetti di A. I. concessi ad un'azienda specifica nell'arco di tre anni e il numero totale di brevetti concessi ad essa nello stesso periodo) e al contributo apportato all'A. I. (misurato come il rapporto tra il numero di brevetti di intelligenza artificiale concessi ad un'azienda nell'arco di tre anni e quelli concessi a tutte le aziende nello stesso periodo)²⁶⁴.

Come si può notare dal grafico sottostante, l'andamento di tale indice risulta costantemente crescente:



FIGURA 29: RAPPRESENTAZIONE GRAFICA DELL'AMUNDI STOXX GLOBAL ARTIFICIAL INTELLIGENCE UCITS ETF.

²⁶² Redazione, (2018), *Intelligenza artificiale, come ha modificato e sta modificando il mondo della finanza?*, Conto Facto, 28/12/2018.

²⁶³ Anonimo, (2018), *Ecco il nuovo ETF tematico sull'intelligenza artificiale di Amundi*, Funds people, 20/11/2018.

²⁶⁴ Ibidem.

Sebbene sia ancora nella sua fase nascente, dalla velocità con cui sta progredendo per far evolvere il settore finanziario, ci si può aspettare che le prospettive porteranno a perdite minori, scambi più intelligenti e, naturalmente, un'esperienza di alto livello per i clienti²⁶⁵.

²⁶⁵ Anonimo, (2018), *Ecco il nuovo ETF tematico sull'intelligenza artificiale di Amundi*, Funds people, 20/11/2018.

CAPITOLO V – IMPATTI DELL’INTELLIGENZA ARTIFICIALE ED EVOLUZIONE FUTURA

Come si può intuire dai capitoli precedenti, l’Intelligenza artificiale, per come si conosce attualmente, è caratterizzata da una storia recente e ha raggiunto significativi successi in diversi ambiti essendo in possesso di un enorme potenziale in grado di incidere sull’intera economia nazionale.

L’Intelligenza artificiale rappresenta un mercato enorme, di circa 20 miliardi di dollari nel 2018 e destinato a crescere esponenzialmente fino a raggiungere e superare i 75 miliardi di dollari nel periodo 2019 – 2023 e, catturando infine 94 miliardi di dollari entro il 2027. L’aumento dei finanziamenti e degli investimenti nella tecnologia IA, un numero crescente di start – up in tale mercato, una maggiore adozione delle tecnologie intelligenti per il miglioramento dell’esperienza del servizio clienti nonché una mole sempre più importante di dati, sono i principali fattori “responsabili” dell’avanzamento di questo mercato²⁶⁶.

Come esposto nel capitolo precedente, l’IA è stata molto utile in numerosi settori come l’assistenza sanitaria, l’istruzione, ecc., ma ci sono alcuni vantaggi generali che si possono ottenere in qualsiasi campo grazie al suo utilizzo²⁶⁷.

I metodi di calcolo per il ragionamento, l’apprendimento e la percezione automatizzati sono divenuti un fenomeno molto utilizzato e diffuso nella vita quotidiana di ogni individuo; lo smartphone è un esempio appropriato di come venga utilizzata l’intelligenza artificiale: quando si scatta una foto, un algoritmo identifica e rileva il volto della persona e la associa quando viene pubblicata sui siti di social media. Quello stesso algoritmo è proprio uno strumento di intelligenza artificiale²⁶⁸.

²⁶⁶INKWOOD RESEARCH (2019), Global Artificial Intelligence market forecast 2019 – 2027, Market Research Report, disponibile a <https://www.inkwoodresearch.com/reports/global-artificial-intelligence-market/#report-summary>.

²⁶⁷ Rathi A., (2019), *The Impact of Artificial Intelligence. A critical review of opportunities and risks of AI adoption*, Towards Data Science, 23/09/2019.

²⁶⁸ Ibidem.

Quest'ultima aiuta anche a ridurre gli errori e la possibilità di raggiungere l'accuratezza con un maggior grado di precisione, un esempio è l'esplorazione dello spazio: i robot intelligenti vengono alimentati con dati e inviati a esplorare ed essendo più resistenti e caratterizzati da una maggiore capacità di sopportare dello spazio e dell'atmosfera ostile, non possono essere danneggiati, alterati o malfunzionanti²⁶⁹.

Ulteriore vantaggio è l'applicazione dell'IA nello svolgimento di compiti di natura monotona, e quindi ripetitivi: le macchine sono in grado di pensare più velocemente e possono essere impiegate in compiti pericolosi come l'esplorazione delle miniere o il fondo dell'oceano superando così, i limiti umani. Grazie alla programmazione e alla possibilità di regolare i loro parametri, i robot possono eseguire lavori più laboriosi e pesanti con maggiore responsabilità e senza essere esposti a una rapida usura.

L'intelligenza artificiale, e più precisamente l'utilizzo di assistenti digitali, possono comportare anche un notevole risparmio di risorse umane: tali strumenti, solitamente impiegati dalle organizzazioni altamente avanzate, possono effettivamente interagire con gli utenti e, grazie alla completa assenza del lato emotivo, assumere le giuste decisioni senza essere ostacolate dai diversi stati d'animo che possono offuscare il giudizio e influire sull'efficienza umana.

Infine, contrariamente agli umani, le macchine sono disponibili 24 ore su 24 e 7 giorni su 7; esse non richiedono pause o momenti di riposo. Sono programmati per eseguire il loro compito per molte ore funzionando costantemente senza annoiarsi, distarsi o stancarsi²⁷⁰.

Come ogni cosa, anche l'intelligenza artificiale presenta il rovescio della medaglia caratterizzato da svantaggi di breve e/o medio – lungo termine.

Innanzitutto, l'IA necessita di enormi quantità di dati per essere adeguatamente addestrata ed eseguire correttamente e, nel modo più efficace ed efficiente possibile il compito affidatogli; in molti casi però, i set di dati esistenti non sono abbastanza grandi e non contengono informazioni sufficienti affinché la metodologia intelligente utilizzata riesca ad imparare e funzionare correttamente e, quando ciò accade, gli individui umani

²⁶⁹Rathi A., (2019), *The Impact of Artificial Intelligence. A critical review of opportunities and risks of AI adoption*, Towards Data Science, 23/09/2019.

²⁷⁰ Ibidem.

devono dedicare migliaia di ore alla rielaborazione degli oggetti che poi vengono utilizzati al fine di creare una buona base di conoscenza per gli strumenti intelligenti²⁷¹. Un ulteriore ostacolo chiave da considerare è il costo che risulta essere piuttosto considerevole. Inoltre sono da considerare i costi di installazione, riparazione e manutenzione per strumenti intelligenti fisici, sono da considerare anche i costi per i frequenti aggiornamenti dei software che devono soddisfare le esigenze dell'ambiente in evoluzione²⁷².

Le macchine, nonché l'intelligenza artificiale, non possiedono la caratteristica della creatività. Possono fare solo ciò che viene loro insegnato e sebbene aiutino a progettare e creare, non possono eguagliare il potere del cervello umano il quale, è in grado di pensare fuori dagli schemi²⁷³. In altre parole gli individui umani sono in grado di creare e implementare strumenti intelligenti per eseguire attività creative, ma questi non sono in grado di programmare a loro volta sistemi originali²⁷⁴.

Anche il Presidente dell'Associazione Italiana di Intelligenza artificiale, Piero Pociati, ha evidenziato i limiti attualmente presenti nelle tecnologie intelligenti²⁷⁵. Egli sostiene che i risultati eccezionali ottenuti finora, grazie alla rivoluzione tecnologica in atto, possono trarre in inganno: <<Siamo infatti portati generalmente a credere, da una visione puramente fantascientifica, che l'IA possa essere capace di prendere decisioni da sola come un essere senziente. La verità è che siamo ben lontani dal (voler) raggiungere tale risultato>>²⁷⁶. Egli mette in evidenza la carenza, che la caratterizza, dal punto di vista del ragionamento: le macchine sono in grado di rilevare modelli e fornire le giuste soluzioni, ma sono incapaci nella spiegazione e nell'esternalizzazione del motivo che ha condotto a specifiche decisioni e giudizi, nonché il verificarsi di un'azione specifica. Le macchine intelligenti presentano anche fattori limitanti nell'inferenza e nel ragionamento, infatti risultano ancora essere ben lontane dall'essere considerate sistemi autonomi a cui

²⁷¹ Kaur K., (2019), *The Real World Potential and Limitations of Artificial Intelligence*, Hackernoon, 16/07/2019.

²⁷² Team Dataflair, (2019), *Pros and Cons of Artificial Intelligence – A Threat or a Blessing?*, Data Flair, 25/09/2019.

²⁷³ Ibidem.

²⁷⁴ Pichai S., (2019), *Advantages and Disadvantages of Artificial Intelligence*, World Economic Forum, 21/12/2019.

²⁷⁵ Pociati P., (2019), *Luci e ombre dell'intelligenza artificiale nella diagnostica*, Associazione Italiana per l'Intelligenza artificiale, 21/10/2019.

²⁷⁶ Ibidem.

delegare decisioni: non sono in grado di far fronte all'ambiente dinamico e quindi di modificare le loro risposte e le loro decisioni in base ai cambiamenti degli ambienti²⁷⁷.

Diversi studi di origine accademica e aziendale, hanno inoltre evidenziato, il difficile rapporto tra l'intelligenza artificiale e l'occupazione; si teme sostanzialmente che l'improvviso scoppio delle applicazioni dell'intelligenza artificiale possa perturbare il mercato del lavoro con una portata ancora impreveduta²⁷⁸. Molti analisti prevedono che i progressi nella robotica e nell'intelligenza artificiale, nei prossimi decenni, potrebbero portare a perdite significative di posti di lavoro o alla loro polarizzazione, aumentando così la disparità di reddito e di ricchezza. Un recente report di Bank of America Merrill Lynch ha sottolineato il potenziale di un aumento delle disuguaglianze causato da una maggiore automazione citando poi una ricerca dell'Università di Oxford, che ha rilevato che fino al 35% di tutti i lavoratori del Regno Unito e il 47% di quelli degli Stati Uniti rischiano di rimanere disoccupati a causa della tecnologia nei prossimi due decenni²⁷⁹.

Sebbene altri ricercatori raggiungano conclusioni meno drammatiche, ciò che accomuna questi studi è la loro concentrazione sulla potenziale distruzione di posti di lavoro²⁸⁰.

Risultando ancora non stimabile l'impatto che avrà l'intelligenza artificiale in futuro, potrebbe risultare significativo analizzare gli effetti derivati dalle precedenti ondate di innovazione tecnologica.

A differenza dei timori correnti, l'ondata di innovazione avvenuta con la prima e la seconda rivoluzione industriale durante il XIX e l'inizio del XX secolo ha portato un aumento della domanda di manodopera non qualificata e a basso costo, sollevando preoccupazioni sulla natura degradante del cambiamento tecnologico che vennero smentite in quanto i lavoratori non qualificati trovarono nuove opportunità di lavoro nell'industria manifatturiera o in altri settori che prosperavano grazie appunto, all'automazione. Con l'avanzare della divisione del lavoro, i lavoratori dovevano

²⁷⁷ Team Dataflair, (2019), *Pros and Cons of Artificial Intelligence – A Threat or a Blessing?*, Data Flair, 25/09/2019.

²⁷⁸ International Labour Organization (2018), *The economics of artificial intelligence: Implications for the future of work*, disponibile a https://www.ilo.org/wcmsp5/groups/public/---dgreports/--cabinet/documents/publication/wcms_647306.pdf.

²⁷⁹ Ibidem.

²⁸⁰ Ibidem.

concentrarsi su competenze sempre più ristrette e ripetitive identificandosi così, sempre meno, con il risultato finale del loro lavoro. Negli anni 60' i movimenti sociali hanno cominciato e richiedere migliori condizioni di lavoro e una crescita più rapida dei salari e allo stesso tempo si è verificata una crescita importante della produttività nelle economie avanzate sollevando parte della popolazione dalla povertà e creando così una classe media in rapida espansione.

Con l'aumento del reddito i giovani lavoratori divennero sempre più istruiti. Si stava creando la base per la terza rivoluzione industriale basata sull'introduzione dei computer²⁸¹. Nei decenni successivi, il cambiamento tecnologico è divenuto un fatto di competenze, aumentando gradualmente la domanda di lavoratori mediamente e altamente qualificati a scapito di coloro in possesso di basse qualifiche; i processi di lavoro hanno cominciato a cambiare e ciò ha portato il settore dei servizi ad assumere il ruolo di motore dell'occupazione.

Nel complesso, gli studi esistenti suggeriscono che gli effetti sull'occupazione sono rimasti piuttosto limitati o – a seconda della metodologia impiegata – sono stati addirittura positivi.

Per quel che riguarda l'attuale ondata tecnologica, nella letteratura accademica emerge il consenso del fatto che l'IA avrà un notevole effetto dirompente sul lavoro, con la perdita di alcune occupazioni, la creazione di altre, e altre ancora che saranno caratterizzate da un cambiamento. Risultano presenti diverse prospettive, con un ampio consenso sul fatto che le attuali tecnologie di IA sono maggiormente adatte alle attività di routine, mentre gli esseri umani hanno più probabilità di rimanere dominanti in ambienti imprevedibili o comunque che richiedono una significativa intelligenza sociale²⁸².

Gli studi sulla storia del cambiamento tecnologico, esposti precedentemente, dimostrano che nel lungo periodo le tecnologie contribuiscono all'aumento della

²⁸¹ International Labour Organization (2018), *The economics of artificial intelligence: Implications for the future of work*, disponibile a https://www.ilo.org/wcmsp5/groups/public/---dgreports/--cabinet/documents/publication/wcms_647306.pdf.

²⁸¹ Ibidem.

²⁸² British Academy for the humanities and social sciences (2018), *The impact of artificial intelligence on work An evidence synthesis on implications for individuals, communities, and societies*, disponibile a <https://royalsociety.org/-/media/policy/projects/ai-and-work/summary-the-impact-of-AI-on-work.PDF>.

produttività a livello di popolazione, occupazione e ricchezza economica; tuttavia, tali benefici, richiedono del tempo per emergere portando così alla possibilità della presenza di periodi significativi nei quali parte della popolazione subisce disagi non indifferenti²⁸³.

Con il presente elaborato viene sostenuta quest'ultima tesi, infatti si vuole mettere in evidenza il pensiero che l'intelligenza artificiale è pronta a sostituire le attività, non i lavori: – come già esposto – automatizzando le attività che si basano su analisi, giudizi e risoluzione dei problemi, l'Intelligenza artificiale può rappresentare una minaccia per lavori di bassa competenza, prevedibili e routinari ma, il suo esponenziale utilizzo e il conseguente aumento della produttività potrebbero stimolare molte nuove opportunità, per gli individui, di sviluppare e migliorare maggiormente le proprie competenze e concentrarsi su impieghi che richiedono un alto livello di personalizzazione e creatività.

Sebbene sia difficile prevedere l'impatto esatto, si suppone che dai 50 ai 75 milioni di posti di lavoro a livello globale potrebbero essere potenzialmente coinvolti dalle conseguenze derivanti dall'avvento dell'IA, un numero significativo, ma che risulta quasi insignificante rispetto alle opportunità che l'intelligenza artificiale creerà²⁸⁴.

²⁸³ British Academy for the humanities and social sciences (2018), *The impact of artificial intelligence on work An evidence synthesis on implications for individuals, communities, and societies*, disponibile a <https://royalsociety.org/-/media/policy/projects/ai-and-work/summary-the-impact-of-AI-on-work.PDF>.

²⁸⁴ Stiehler A., (2018), *AI's coming of age. The progress into the AGI phase and the beginning of the true autonomy*, UBS, 02/03/2018.

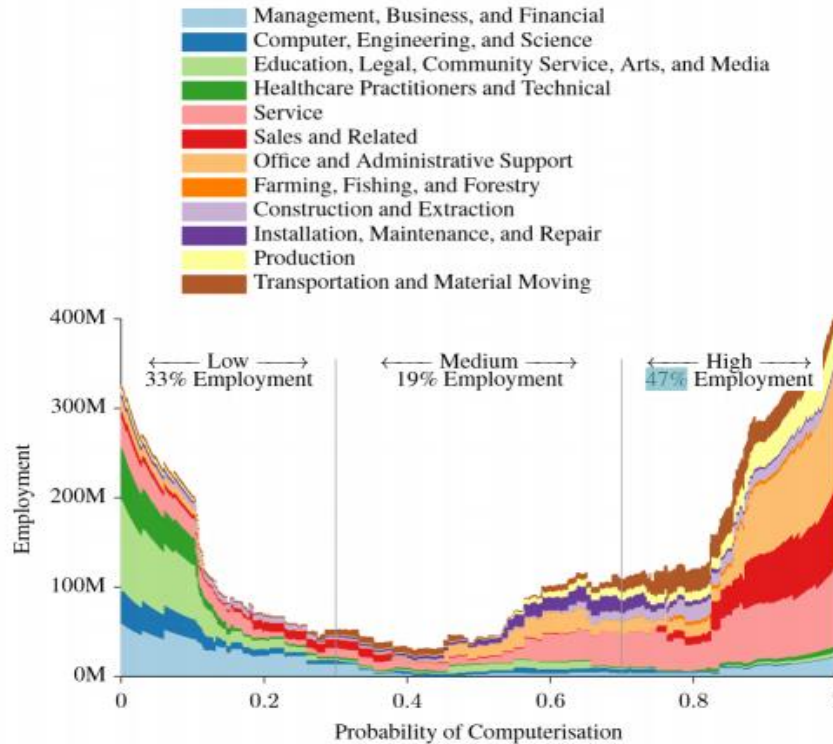


FIGURA 30: GRAFICO SULLE PROBABILITÀ DI COMPUTERIZZAZIONE DEL LAVORO, FONTE: CARL BENEDIKT FREY, MICHAEL A. OSBORNE; THE FUTURE OF EMPLOYMENT.

Per quel che concerne l'evoluzione dell'intelligenza artificiale, la maggior parte degli studi analizzati, sottolineano che essa avrà un impatto economico significativo. Una ricerca condotta dalla società di consulenza Accenture, analizza dodici economie sviluppate che, assieme, coprono più dello 0,5% della produzione economica mondiale e prevede, che a partire dal 2035 l'intelligenza artificiale potrebbe raddoppiare i tassi di crescita economica globale annuale²⁸⁵.

L'economia beneficerà anche della diffusione dell'innovazione che inciderà negativamente su settori diversi, ma creerà nuovi flussi di entrate²⁸⁶.

L'IA avrà effetti positivi diretti e indiretti sul mercato del lavoro, sulla produttività e sul Pil. Gli effetti diretti saranno generati dall'aumento delle entrate e dell'occupazione nelle imprese e nei settori che sviluppano e/o producono tecnologie intelligenti, mentre quelli indiretti saranno generati conseguentemente a un aumento più ampio della

²⁸⁵ Szczepa'ski M., (2019), Economic impacts of artificial intelligence (AI), disponibile su [http://www.europarl.europa.eu/RegData/etudes/BRIE/2019/637967/EPRS_BRI\(2019\)637967_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/BRIE/2019/637967/EPRS_BRI(2019)637967_EN.pdf).

²⁸⁶ Ibidem.

produttività nei settori che utilizzano l'IA per ottimizzare i processi decisionali e quelli aziendali²⁸⁷. Complessivamente si prevedono notevoli guadagni nel prossimo decennio come dimostra la seguente figura:

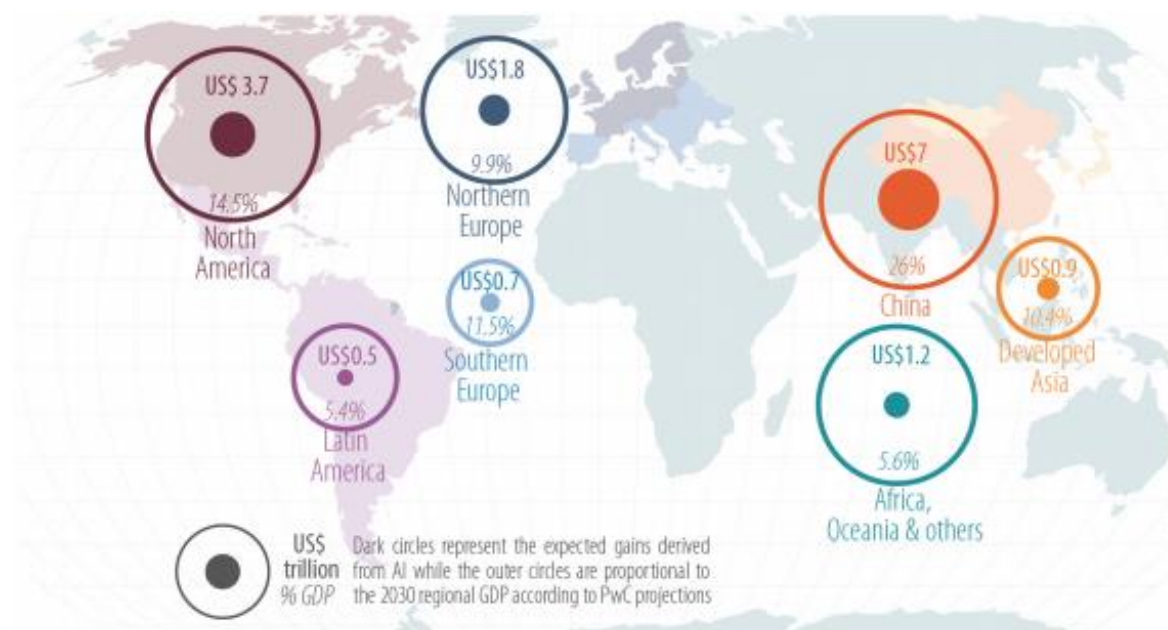


FIGURA 31: GUADAGNI ATTESI DELL'INTELLIGENZA ARTIFICIALE NELLE DIVERSE REGIONI DEL MONDO ENTRO IL 2030. FONTE: THE MACROECONOMIC IMPACT OF ARTIFICIAL INTELLIGENCE, PWC, 2018.

Da uno studio condotto dal McKinsey Global Institute, è emerso inoltre, che l'impatto economico può emergere gradualmente ed essere effettivamente visibile solo con il passare del tempo: l'impatto dell'IA può non essere lineare, ma può accumularsi a un ritmo accelerato nel tempo conferendo così, al modello di adozione un inizio lento a causa dei costi e degli investimenti sostanziali associati all'apprendimento e all'implementazione delle nuove tecnologie con una successiva accelerazione guidata dall'effetto cumulativo della concorrenza e del miglioramento delle capacità complementari infatti viene stimato che il suo contributo può essere tre o più volte superiore entro il 2030 rispetto a quello dei prossimi cinque anni.²⁸⁸

Ulteriore, nonché ultima considerazione emersa dallo studio citato pocanzi, riguarda il fatto che l'adozione dell'IA potrebbe allargare i divari tra paesi, aziende e lavoratori;

²⁸⁷ Szczepański M., (2019), Economic impacts of artificial intelligence (AI), disponibile su [http://www.europarl.europa.eu/RegData/etudes/BRIE/2019/637967/EPRS_BRI\(2019\)637967_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/BRIE/2019/637967/EPRS_BRI(2019)637967_EN.pdf).

²⁸⁸ Bughin J., Seong J., Manyika J., Chui M., Joshi R., (2018), *Notes from the AI frontier. Modeling the impact of AI on the world economy*, McKinsey Global Institute, 09/2018.

dare impulso all'attività economica ma con una distribuzione dei benefici decisamente disomogenea²⁸⁹.

L'IA potrebbe aumentare innanzitutto il divario tra i paesi, infatti questi ultimi potrebbero aver bisogno di strategie e risposte diversificate in base al livello di tecnologie intelligenti adottato: i leader nel settore dell'intelligenza artificiale, ovvero i paesi sviluppati, potrebbero aumentare il loro vantaggio di un ulteriore 20-25%, mentre i paesi in via di sviluppo potrebbero ottenere solamente il 5-15% circa²⁹⁰.

Ulteriore divario che può derivare, riguarda le prestazioni tra i precursori da un lato e gli utenti lenti e i non adottanti dall'altro. Per precursori si intendono le aziende che assorbono completamente gli strumenti di IA nei prossimi 5-7 anni. Queste, secondo le stime, potrebbero ottenere un raddoppio del loro flusso di cassa (al netto degli investimenti associati e i relativi costi di transazione) entro il 2030, mentre all'estremità opposta sono presenti i ritardatari e i non adottanti che non adottano affatto tecnologie di IA o che non le hanno completamente assorbite entro il 2030. Questi ultimi potrebbero subire un calo di circa il 20% del proprio flusso di cassa rispetto agli attuali livelli²⁹¹.

²⁸⁹ Bughin J., Seong J., Manyika J., Chui M., Joshi R., (2018), *Notes from the AI frontier. Modeling the impact of AI on the world economy*, McKinsey Global Institute, 09/2018.

²⁹⁰ Ibidem.

²⁹¹ Ibidem.

CAPITOLO VI – IMPLEMENTAZIONE DI UNA RETE NEURALE ARTIFICIALE PER LA PREVISIONE FINANZIARIA

Lo scopo del presente capitolo risulta essere la previsione dell'Indice Italiano FTSE MIB attraverso l'implementazione di una rete neurale artificiale e la sua successiva correlazione con l'indice asiatico Hang Seng.

6.1 L'indice FTSE MIB. Cenni storici

Con lo scopo di mettere in evidenza le prestazioni delle società italiane quotate nei mercati MTA della Borsa Italiana, dal 1° Giugno 2009 è stata introdotta una nuova serie di indici di borsa, chiamati FTSE, e utilizzata al fine di fornire agli investitori un insieme completo e complementare di indici che misurino le prestazioni dei principali segmenti nel settore dei capitali e nel settore industriale del mercato italiano²⁹².

L'indice di riferimento italiano è nato il 31 dicembre 1992 con il nome di COMIT 30 avente un valore base di 100 punti. Il 17 ottobre 1994 la Borsa Italiana ha rilevato i diritti su COMIT 30 della Banca Commerciale Italiana modificando il nome in Mib30, il quale, fino al 1° giugno 2003, comprendeva i 30 titoli a maggiore capitalizzazione di Piazza Affari.

In seguito ad una partnership tra Borsa Italiana e la società di rating Standard & Poor's, lo storico Mib30 venne sostituito dalla S&P Mib quotato dal 2 giugno 2003 al 29 maggio 2009 per venire successivamente sostituito dall'Indice conosciuto oggi, ovvero il FTSE MIB, nato in seguito alla fusione di Borsa Italiana e la London Stock Exchange²⁹³.

L'evoluzione dell'Indice odierno, dopo un decennio dalla sua nascita, rappresenta un valido indicatore di come si è evoluta l'economia italiana.

Di seguito viene riportato il grafico dell'andamento dell'indice dal 1° giugno 2009 al giorno più recente possibile.

²⁹² Borsa Italiana (2019), *Per la gestione dell'indice FTSE MIB*, Luglio 2019, disponibile a https://research.ftserussell.com/products/downloads/FTSE_MIB_Index_Italian_.pdf.

²⁹³ Autori di Wikipedia, (2020), *FTSE MIB*, Wikipedia. L'enciclopedia libera, 04/01/2020.



FIGURA 32: GRAFICO DELL'ANDAMENTO DELL'INDICE FTSE MIB. FONTE: BLOOMBERG.

Come si evince dal grafico, l'indice considerato presenta un andamento piuttosto regolare, tranne per il periodo 2012 – 2014: il 24 luglio 2012 è stato registrato un minimo di lungo periodo spinto al ribasso dalla forte speculazione che ha colpito i titoli di stato, e di conseguenza, l'intero settore bancario. Poi è iniziato un lento recupero caratterizzato da un nuovo massimo, raggiunto il 10 giugno 2014, grazie alle politiche monetarie accomodanti da parte della BCE; tuttavia l'indice non è riuscito a mantenersi nell'area 22.000 in quanto spinto al ribasso dalla mancanza di crescita nell'Eurozona e in Italia e dal timore di un contagio della nuova crisi del debito alla periferia dell'Eurozona²⁹⁴.

L'anno seguente, il 2015, invece ha registrato un forte rialzo dovuto alle mosse espansive da parte della BCE con il lancio del *Quantitative Easing* (massiccio acquisto di titoli di stato sovrani in asta per la temporanea risoluzione della crisi greca). Il 15 aprile dello stesso anno l'indice arriva a toccare un nuovo record a quota 24.031, ma a causa delle tensioni europee sul caso greco subisce dei contraccolpi, arrivando comunque a superare, per tre chiusure consecutive, il 2% di rialzo nei giorni dell'accordo che scongiura l'uscita della Grecia dalla zona euro. L'anno si chiude con una perdita di quasi mille punti²⁹⁵.

²⁹⁴ Autori di Wikipedia, (2020), *FTSE MIB*, Wikipedia. L'enciclopedia libera, 04/01/2020.

²⁹⁵ Ibidem.

Il primo semestre del 2016 si è contraddistinto da un percorso negativo per l'indice, che ha perso complessivamente il 24,37% a causa del rallentamento della crescita cinese e dell'economia globale, nel crollo dei prezzi del greggio, nelle tensioni che hanno colpito alcune banche italiane tra cui MPS e nell'immediato post – Brexit. L'anno si chiude con un parziale recupero con l'indice a 19.234 punti, con una perdita del 10.2% dovuta al disastroso inizio d'anno²⁹⁶.

Dal 2 gennaio dell'anno seguente, il 2017, in Borsa e nell'indice FTSE MIB debutta Banco BPM a seguito della sua fusione con Banco Popolare. Banca MPS, dopo il fallito aumento di capitale a 5 miliardi di euro, viene sostituita da Banca Generali e Pirelli & C. sostituisce Banca Mediolanum. L'anno si chiude complessivamente in modo positivo, ovvero con un guadagno annuo del 13.59%: il miglior titolo del FTSE MIB è stato Fiat Chrysler Automobiles (FCA), seguita da Stmicroelectronics (+70,1% dopo il +71% del 2016), FinecoBank (+60,1%), Ferrari (+59%) e Moncler (+58.8%). La performance peggiore del 2017 è quella di Saipem (-28,5%), seguita da Leonardo (-25,1%), Tenaris (-22,4%), Mediaset (-21,3%) e BPER Banca (-16,76%).

Enel con oltre 52 miliardi di euro è la prima società italiana per valore di Borsa, dal giugno 2017, quando ha superato Eni, staccata di 2 miliardi di euro, che da anni deteneva il primato. Seguono Intesa San Paolo, Unicredit e Luxottica. Invece la società più piccola, tra le big, è con una capitalizzazione di 2 miliardi di euro BPER Banca, preceduta da Azimut Holding, Yoox NAP, Unipol e Banca Generali.

Il nuovo anno, sebbene caratterizzato inizialmente da un record di 9 sedute consecutive positive con l'indice che raggiunge i 23.544 punti, risulta essere un anno piuttosto burrascoso. Grazie all'aumento del petrolio di inizio anno, Eni torna a essere la società più capitalizzata superando Enel la quale però, pochi giorni dopo ritorna in testa con una capitalizzazione di quasi 55 miliardi di euro. Dal 22 gennaio, dopo lo stacco dell'acconto del dividendo di Enel per un importo complessivo di un miliardo di euro, Eni ritorna ad essere la società maggiormente capitalizzata²⁹⁷.

Il 4 dicembre Luxottica, in seguito al successo dell'OPS di EssilorLuxottica e alla conseguente diminuzione del flottante, sceso sotto il 10%, viene sostituita da DiaSorin. Con la revisione trimestrale, avvenuta il 27 dicembre inoltre, Amplifon e Juventus

²⁹⁶ Autori di Wikipedia, (2020), *FTSE MIB*, Wikipedia. L'enciclopedia libera, 04/01/2020.

²⁹⁷ Ibidem.

Football Club sostituiscono Banca Mediolanum e Mediaset. L'anno si chiude decisamente in modo diverso, con una perdita annua del 16.15% e una quotazione del FTSE MIB di 18.324²⁹⁸.

Il 2019 non è stato un anno facile, l'indice ha vissuto infatti diversi scossoni a causa delle tensioni internazionali e quelle legate alle incertezze della politica italiana. Dopo una prima parte dell'anno positiva che ha portato il FTSE MIB da meno di 18 mila punti a circa 22 mila punti, si sono verificate le prime ricadute che l'anno riportato sotto i 20 mila punti, per poi riprendere un percorso a rialzo continuato fino alla crisi politica che ha portato alla fine del governo Conte I. Da agosto in poi, l'indice è stato caratterizzato da un rialzo che lo ha portato vicino ai massimi dell'anno alla fine di ottobre per arrivare poi a una chiusura di 23.506 punti²⁹⁹. Con la seconda revisione trimestrale dell'anno menzionato inoltre, Nexi sostituisce Banca Generali la quale, ritorna sei mesi dopo, sostituendo UnipolSai³⁰⁰.

Per quel che riguarda l'anno corrente, ovvero il 2020, sembra sarà un anno tutt'altro che tranquillo per i mercati finanziari come lo dimostra la dichiarazione rilasciata da Ann – Katrin Peterse, Investment Strategisti, Global Economics & Strategy di Allianz Global Investore: << Attualmente le borse sono sostenute dalle speranze di un'inversione di tendenza dell'economia globale. Tuttavia, l'attuale situazione complessa sul piano politico e dei fondamentali limita il potenziale di rialzo dei corsi>>³⁰¹.

6.1.1 Composizione e metodologia di calcolo dell'indice FTSE MIB

Come si può notare dalla tabella che segue, la quale rappresenta il paniere di società che costituiscono il FTSE MIB, rispetto a quello originario, si è verificato un ridimensionamento del peso della componente bancaria: gli istituti di credito rappresentavano storicamente la componente principale con una capitalizzazione di listino superiore ai 93 miliardi di euro (con un peso percentuale nel complessivo paniere

²⁹⁸ Autori di Wikipedia, (2020), *FTSE MIB*, Wikipedia. L'enciclopedia libera, 04/01/2020.

²⁹⁹ Buffolo M., (2019), *Com'è stato il 2019 per la Borsa di Milano*, AGI, 28/12/2019.

³⁰⁰ Autori di Wikipedia, (2020), *FTSE MIB*, Wikipedia. L'enciclopedia libera, 04/01/2020.

³⁰¹ Redazione Soldionline, (2020), *Il 2020 sarà un anno tutt'altro che tranquillo per i mercati finanziari*, Soldi Online, 10/01/2020.

pari a 26.4%), mentre oggi, a seguito della crisi che il settore ha dovuto affrontare, rappresentano circa 65 miliardi di euro (con un peso percentuale pari a 16.6%).

Tale lacuna, lasciata non solo dal settore bancario ma anche dal settore delle comunicazioni, è stata colmata soprattutto dall'industria (che attualmente vale l'8.6%) e dal settore automobilistico (che oggi rappresenta il 10.9% dell'intero paniere) grazie all'esponentiale crescita del gruppo FCA e alla scelta di scorporare CNH Industrial e Ferrari³⁰².

Attualmente il FTSE MIB (acronimo di **F**inancial **T**imes **S**tock **E**xchange **M**ilano **I**ndice di **B**orsa) è il principale indice di benchmark dei mercati azionari italiani, che misurando le performance di 40 titoli italiani di società di primaria importanza e a liquidità elevata, riesce a cogliere circa l'80% della capitalizzazione di mercato interna. Le società vengono selezionate dopo una rigorosa analisi dell'universo italiano allargato in modo tale che l'indice possa rappresentare al meglio, per gli investitori, i mercati azionari italiani³⁰³.

Società			
A2A S.p.A.	DiaSorin S.p.A.	Italgas S.p.A.	Saipem S.p.A.
Amplifon S.p.A.	Enel S.p.A.	Juventus Football Club S.p.A.	Salvatore Ferragamo S.p.A.
Atlantia S.p.A.	Eni S.p.A.	Leonardo S.p.A.	Snam S.p.A.
Azimut Holding S.p.A.	Exor N.V.	Mediobanca S.p.A.	STMicroelectronics N.V.
Banca Generali S.p.A.	Ferrari N.V.	Moncler S.p.A.	Telecom Italia S.p.A.
Banco BPM S.p.A.	Fiat Chrysler Automobiles N.V.	Nexi S.p.A.	Tenaris S.A.
BPER Banca S.p.A.	FinecoBank S.p.A.	Pirelli & C. S.p.A.	Terna - Rete Elettrica Nazionale S.p.A.
Buzzi Unicem S.p.A.	Generali S.p.A.	Poste italiane S.p.A.	UBI Banca S.p.A.

³⁰² Autori di Wikipedia, (2020), *FTSE MIB*, Wikipedia. L'enciclopedia libera, 04/01/2020.

³⁰³ Borsa Italiana disponibile a <https://www.borsaitaliana.it/borsa/indici/indici-in-continua/dettaglio.html?indexCode=FTSEMIB&lang=it>.

Campari - Milano S.p.A.	Hera S.p.A.	Prysmian S.p.A.	UniCredit S.p.A.
CNH Industrial N.V.	Intesa Sanpaolo S.p.A.	Recordati S.p.A.	Unipol S.p.A.

TABELLA 1: SOCIETÀ CHE COMPONGONO L'INDICE FTSE MIB. FONTE: BORSA ITALIANA.

L'indice FTSE MIB utilizza l'ultimo prezzo delle attuali contrattazioni in MTA e nella piattaforma elettronica di scambi azionari di Mercato Expandi, della Borsa Italiana durante le ore ufficiali di mercato³⁰⁴. Il valore dell'indice, al tempo t , è calcolato utilizzando la seguente formula:

$$I_t = M_t * D_t$$

dove M_t indica la capitalizzazione totale di mercato e D_t il valore del *Divisor*, entrambe considerate al tempo t , uguale a:

$$M_t = \sum_{i=1}^n p_{i,t} * q_{i,t} * IWF$$

con $p_{i,t}$ che rappresenta il prezzo dell'ultimo contratto concluso alla fine di un intervallo predeterminato e configurabile (attualmente uguale a 30 secondi). Se le azioni di uno dei componenti sono sospese, l'Indice verrà calcolato utilizzando il prezzo dell'ultimo contratto concluso prima della sospensione.

$q_{i,t}$ è il numero di azioni in emissione utilizzate da FTSE per il titolo azionario e IWF è il fattore di ponderazione di investibilità che equivale al 100% al netto della somma delle percentuali delle partecipazioni azionarie possedute dai principali azionisti³⁰⁵.

D_t viene calcolato dividendo il valore della nuova capitalizzazione di mercato corretta per la capitalizzazione di mercato corretta prima delle operazioni sul capitale e moltiplicando poi per il vecchio *Divisor*, valore che soddisfa la relazione $I_t = M_t / D_t$.

Nel caso di una modifica dell'indice o di una operazione sul capitale che altera la capitalizzazione di mercato dell'Indice, viene effettuata una correzione del *divisore*, D_t ,

³⁰⁴ Borsa Italiana (2018), Per la gestione dell'indice FTSE MIB, disponibile a https://www.borsaitaliana.it/azioni/tutti-gli-indici/metodologiadicalcolo/ftse_pdf.htm.

³⁰⁵ Ibidem.

in modo che l'andamento dell'indice rifletta esclusivamente i cambiamenti dovuti ai prezzi dei componenti dell'indice. La correzione del divisore si ottiene dividendo il valore della nuova capitalizzazione di mercato corretta per la capitalizzazione di mercato prima delle operazioni sul capitale e moltiplicando poi per il vecchio divisore³⁰⁶.

6.1.2 Le reti neurali artificiali in R

Tutte le analisi svolte all'interno di questa tesi verranno effettuate con il supporto del software statistico R, un software per l'analisi statistica; tale software è freeware e opera su diversi sistemi operativi quali UNIX, Window e MacOS³⁰⁷.

Questo software è modulare, nel senso che esistono numerosi pacchetti che implementano i vari algoritmi e per quel che riguarda le Reti Neurali Artificiali, R offre principalmente due alternative:

- il pacchetto *neuralnet*, sviluppato nel settembre 2012 da Stefan Fritsch e Frauke Guenther, che permette la formazione di reti neurali che utilizzano, per la stima dei parametri, l'algoritmo di *back - propagation* (sia la *back - propagation resiliente* con o senza peso *backtracking* che la versione globalmente convergente modificata da Anastasiadis *et al* nel 2005). Il pacchetto permette di scegliere la funzione d'errore e della funzione di attivazione, inoltre, il calcolo dei pesi generalizzati è implementato di default³⁰⁸;
- il pacchetto *nnet* che permette la costruzione di reti neurali *feed - forward* con un unico strato nascosto e utilizzato anche per i modelli log - lineari multinomiali³⁰⁹.

Il dataset è costituito dalle quotazioni dell'indice FTSE MIB rilevati con frequenza giornaliera e il periodo campionario totale inizia il 1° giugno 2009 e si conclude il 9 dicembre 2019.

³⁰⁶ Borsa Italiana (2018), Per la gestione dell'indice FTSE MIB, disponibile a https://www.borsaitaliana.it/azioni/tutti-gli-indici/metodologiadicalcolo/ftse_pdf.htm.

³⁰⁷ Hatwell J., (2016), *Artificial Neural Networks in R*, RPub, 10/2016.

³⁰⁸ Stefan F., Frauke G., Marvin N., Sulig M e Mueller S., (2019), *Package "neuralnet"*, CRAN, 07/02/2019.

³⁰⁹ Ripley B., e Venables W., (2016), *Package "nnet"*, CRAN, 02/02/2016.

Il dataset, in base alla letteratura analizzata nel capitolo precedente, è stato suddiviso nel seguente modo:

- L'80% dei dati (ovvero i dati riferiti al periodo campionario che va dal 1° giugno 2009 e termina il 24 ottobre 2017) è stato utilizzato la fase di addestramento (*Training*) ovvero per la stima dei parametri delle reti neurali;
- Il 20% dei dati (più precisamente i dati riferiti al periodo campionario che va dal 25 ottobre 2017 e si conclude il 9 dicembre 2019) è stato utilizzato per la fase di validazione.

Il pacchetto di *R* scelto per l'analisi empirica è stato "*neuralnet*" che è un pacchetto molto flessibile poiché permette all'utente, che lo utilizza, di fissare arbitrariamente diversi parametri nella costruzione delle reti, come la scelta della funzione di attivazione e il metodo di stima.

È possibile scegliere tra molteplici funzioni di attivazione, tutte spiegate accuratamente nel Capitolo 3, e diversi approcci di stima quali il metodo *Backpropagation*, il metodo *Backpropagation resiliente*, con o senza tecnica *Backtracking* sui pesi, e infine, il metodo di *Convergenza Globale Modificata*.

I dati, i quali comprendono il prezzo di chiusura, di apertura, il massimo e il minimo giornaliero per tutto il periodo campionario di riferimento, sono stati normalizzati al fine di ridurre o eliminare la ridondanza dei dati. Sono presenti diverse tipologie di normalizzazione, ma ai fini di tale applicazione, si è scelto il ridimensionamento delle funzionalità (detto anche *Min – Max Scaling*) per ottenere valori compresi tra 0 e 1 mediante la seguente formula:

$$z = \frac{x - \min(x)}{\max(x) - \min(x)}$$

Essa, applicata sia agli input che agli output ha permesso di rendere i dati più adatti ad una convergenza e una comparazione.

Il dataset a disposizione è stato suddiviso in due parti: il training set e il validation set. Come precisato prima, si è deciso di utilizzare l'80% dei dati come esempi, grazie ai quali

la rete neurale può allenarsi (fase di Training) e il restante 20% dei dati come set di validazione. In questo elaborato è stato effettuato il *Resistance test* in quanto i valori utilizzati al fine della comparazione, e quindi nella fase di validazione, appartengono a un periodo diverso, non utilizzato nella fase di *Training*. A livello computazionale, la comparazione consiste nel calcolo della previsione statistica ad un passo dei primi 538 dati, utilizzando i pesi assegnati casualmente dalla rete neurale, e successivamente nel calcolo dell'Errore Quadratico Medio per ogni blocco di dati, ovvero confrontando quelli effettivi e quelli previsti dalla stessa rete. Effettuando la successiva media degli MSE, si è optato per il numero di nodi nascosti che minimizzava tale stimatore di bontà previsiva la cui formula risulta essere la seguente:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2.$$

Nella quale Y_i indica il vettore dei valori osservati della variabile prevista, mentre \hat{Y}_i indica il vettore dei valori previsti della medesima variabile.

In statistica, l'errore quadratico medio. È una funzione di rischio, corrispondente al valore atteso dell'errore al quadrato. Minore risulta l'errore quadratico medio, maggiore è la bontà previsiva³¹⁰.

È risultato ottimale l'utilizzo di una rete neurale *feed – forward* così strutturata:

- 12 nodi di input comprendenti i prezzi di chiusura, i prezzi di apertura e i prezzi massimi e minimi giornalieri degli ultimi tre giorni del periodo campionario di riferimento, vale a dire il 9 dicembre 2019, il 6 dicembre 2019 e i 5 dicembre 2019;
- un unico strato nascosto caratterizzato da 14 nodi;
- la funzione di attivazione logistica, già descritta nel Capitolo 3;
- il metodo di stima Back - propagation resiliente, algoritmo creato da Martin Riedmiller e Heinrich Braun nel 1992, nonché euristica dell'apprendimento supervisionato nelle reti artificiali *feed - forward*. Esso considera solamente il segno della derivata parziale su tutti i modelli e agisce in modo indipendente su

³¹⁰ Binieli M., (2018), *Machine learning: an introduction to mean squared error and regression lines*, freedCodeCamp, 16/10/2018.

ciascun peso.³¹¹ Per ogni peso, se si verifica una variazione di segno della derivata parziale della funzione di errore totale rispetto all'ultima iterazione, il valore di aggiornamento, per quel peso, viene moltiplicato per un fattore η^- , dove $\eta^- < 1$. Se l'ultima iterazione produce lo stesso segno, il valore di aggiornamento viene moltiplicato per un fattore di η^+ , dove $\eta^+ > 1$. I valori di aggiornamento vengono calcolati per ciascun peso nel modo sopra indicato e infine ogni peso viene modificato dal proprio valore di aggiornamento, nella direzione opposta alla derivata parziale di quel peso, in modo da ridurre al minimo la funzione di errore totale. η^+ è empiricamente impostato su 1,2 e η^- su 0,5³¹².

Simile alla Back Propagation comune, quella resiliente presenta degli ulteriori vantaggi, infatti essa è spesso più veloce nella fase di training e inoltre, non richiede di specificare alcun valore dei parametri³¹³;

- un solo nodo di output che indica la previsione del prezzo di chiusura a un passo. Nel caso di studio considerato, il nodo di output restituirà il prezzo di chiusura dell'indice italiano FTSE MIB a un giorno.

Si ottiene così un prezzo di chiusura normalizzato, per cui si procede alla sua de-normalizzazione, ottenendo infine un prezzo di chiusura stimato pari a 23.140,82 euro contro un prezzo di chiusura effettivo 23.122,82 euro riferito sempre allo stesso giorno ovvero il 10 dicembre 2019.

³¹¹ McCaffrey J., (2015), *How to use Resilient Back Propagation to train Neural Networks*, Visual Studio Magazine, 03/09/2015.

³¹² Autori di Wikipedia, (2013), *Rprop*, Wikipedia. L'enciclopedia libera, 26/03/2019.

³¹³ McCaffrey J., (2015), *How to use Resilient Back Propagation to train Neural Networks*, Visual Studio Magazine, 03/09/2015.

Il grafico della rete neurale risultante è il seguente:

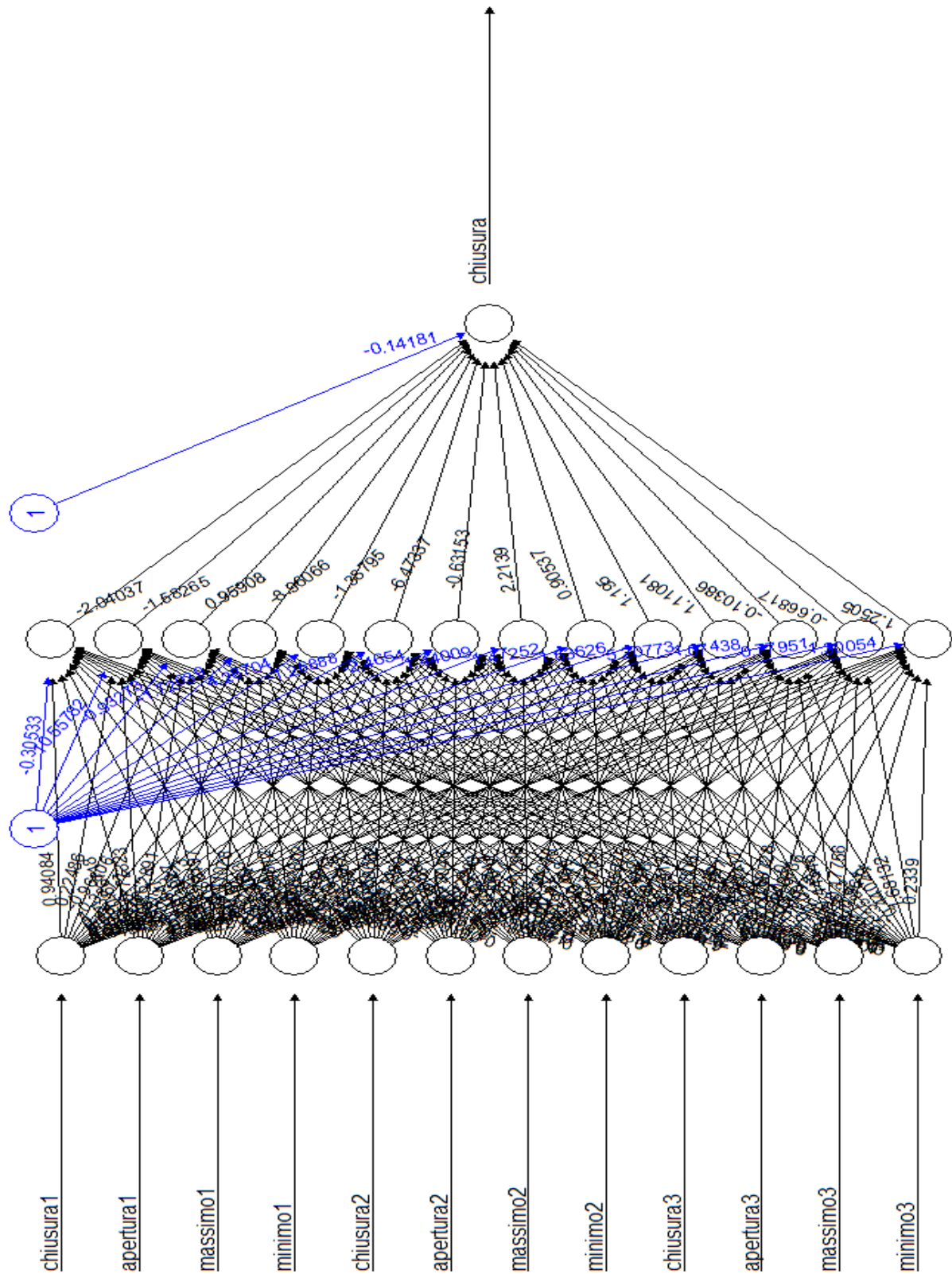


FIGURA 33: GRAFICO RETE NEURALE FEED - FORWARD PER LA PREVISIONE A UN PASSO DELL'INDICE ITALIANO FTSE MIB.

Un limite delle reti neurali riguarda il fatto che non è possibile estrapolare delle informazioni chiare sulla significatività delle variabili in input, proprio per tale ragione, le reti neurali artificiali vengono definite come delle “scatole nere” o *black boxes*.

Per cercare comunque di dare un’interpretazione ai valori dei pesi e per evidenziare quali variabili esplicative incidano maggiormente sull’output previsto, si è deciso di creare una tabella contenente i pesi presenti negli archi che collegano lo strato di input con lo strato intermedio e una che mettesse in evidenza i pesi che collegano lo strato intermedio con lo strato di output, ovviamente ognuna di esse contenente anche i relativi *bias*. Di seguito sono riportate entrambe:

	Nodo1	Nodo2	Nodo3	Nodo4	Nodo5	Nodo6	Nodo7
Bias	-0.305334983	10.55782090	-0.93272764	-1.72039145	4.2870423	-1.268879876	-2.4654028
chiusura1	0.940841415	0.22485641	-0.96806042	-0.38113615	1.5214005	0.462178438	0.6840140
apertura1	-0.667181492	-0.77323032	0.62890650	-0.12475483	-0.2748675	3.749527341	0.4021063
massimo1	-1.443701452	0.33926624	-1.73197190	0.41077867	0.3963953	0.005624018	1.7184141
minimo1	0.835289244	0.33166364	-1.02013241	-0.89812378	-0.6280208	0.469122224	0.6197326
chiusura2	0.272403830	-0.50203396	-1.49265634	0.81032496	0.5345181	-0.586619563	1.9128885
apertura2	-0.589545146	-1.02397220	0.87668729	-0.04010626	0.4811695	0.696618245	0.7304569
massimo2	0.001043932	-1.53620727	-0.11842390	-0.91346025	2.2128265	-1.280388854	-0.7904342
minimo2	0.441778046	-1.05592668	0.29248449	0.55618638	0.4060657	0.002664273	-0.5942776
chiusura3	-0.796380688	-0.89632090	0.94735258	0.13340229	0.2823106	5.751895630	2.4820699
apertura3	-1.227975220	-6.53394614	1.74773017	-0.93328908	-5.3796635	-14.387186866	-0.7551080
massimo3	0.464806639	0.03944751	3.06374690	0.58239146	-1.4657929	-6.264586593	1.0722917
minimo3	-1.335362241	0.03687018	-0.05203902	-0.47459717	-2.8353826	-0.555520221	1.2092811

Nodo8	Nodo9	Nodo10	Nodo11	Nodo12	Nodo13	Nodo14
-1.4400945	1.47252208	-1.6362631	5.207734525	1.07438126	0.7795061	1.30054491
-1.1528151	0.83978053	0.3575105	1.315829873	0.09312772	-0.3755282	0.44941508
1.0496977	-0.50979622	-0.8377152	-0.647741283	-0.66653433	2.2454855	1.15522154
0.6909253	1.26085020	1.7018526	0.240280584	-1.19058681	-1.0403498	-0.04863272
0.1232287	0.41512671	-0.4940730	-0.966474961	-1.99735854	-1.3227205	-0.63136975
1.2265451	1.09750273	1.0767526	-1.181554733	0.34876578	0.9358014	0.02433775
0.7229789	-0.06087422	1.1418469	-2.720690247	0.92906863	0.9192820	0.77392022
-0.2129613	-1.14647700	-0.8704069	-1.119709249	-0.24310346	-1.2370056	-0.10370060
0.4046770	1.10282915	-0.6567524	1.074958297	-0.48109847	0.2482643	-1.17478339
0.3566524	-1.82358575	-0.5913013	-1.730602489	0.36756627	-0.5156665	0.20208022
0.6632098	0.80202334	0.8229518	-0.411169544	-0.69727503	-3.0873589	0.38153700
-0.1722666	-1.06073259	-0.2283611	-0.009919757	0.44345785	-0.1775584	0.96139687
-0.9953338	-0.16818875	1.0173834	-1.358829955	0.44024587	1.6813167	0.23389678

TABELLA 2: PESI CHE COLLEGANO I NODI DELLO STRATO DI INPUT CON I NODI DELLO STRATO NASCOSTO.

	Pesi
Bias	-0.1418080
Nodo1	-2.0403734
Nodo2	-1.5826500
Nodo3	0.9590774
Nodo4	-8.9606614
Nodo5	-1.3879518
Nodo6	-6.4733741
Nodo7	-0.6315269
Nodo8	2.2139023
Nodo9	0.9053742
Nodo10	1.1950002
Nodo11	1.1108135
Nodo12	-0.1038590
Nodo13	-0.6681678
Nodo14	1.2504986

TABELLA 3: PESI CHE COLLEGANO I NODI DELLO STRATO INTERMEDIO CON IL NODO DELLO STRATO DI OUTPUT.

A riprova dei risultati ottenuti e riportati nelle tabelle precedenti, osservando il grafico sottostante, chiamato *Neural Identification Diagram*, si può affermare che il prezzo di chiusura a un giorno dell'indice italiano FTSE MIB è maggiormente condizionato dal prezzo di apertura al tempo t-1, dal prezzo di chiusura al tempo t-3 e dal prezzo massimo avente medesimo ritardo.

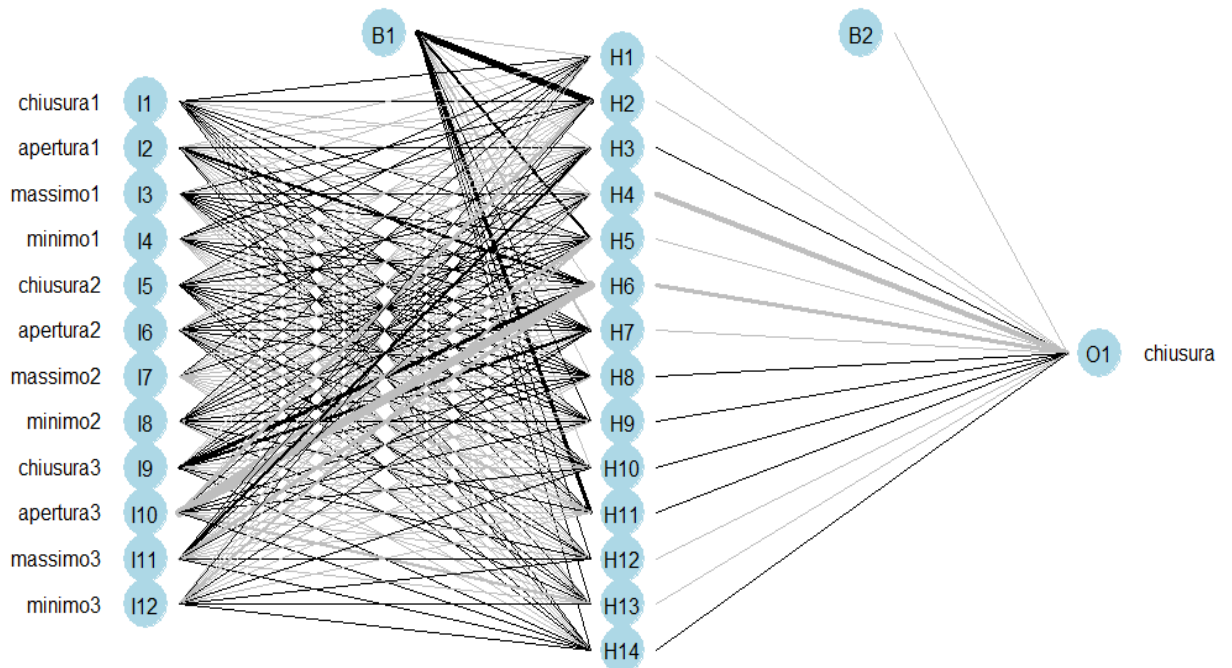


FIGURA 34: GRAFICO NID CHE EVIDENZA LE VARIABILI MAGGIORMENTE SIGNIFICATIVE PER LA RETE NEURALE IMPLEMENTATA.

Solo successivamente, si è proceduto al confronto tra un maggior numero di prezzi di chiusura previsti grazie alla rete neurale e quelli effettivi, ottenendo il seguente risultato:

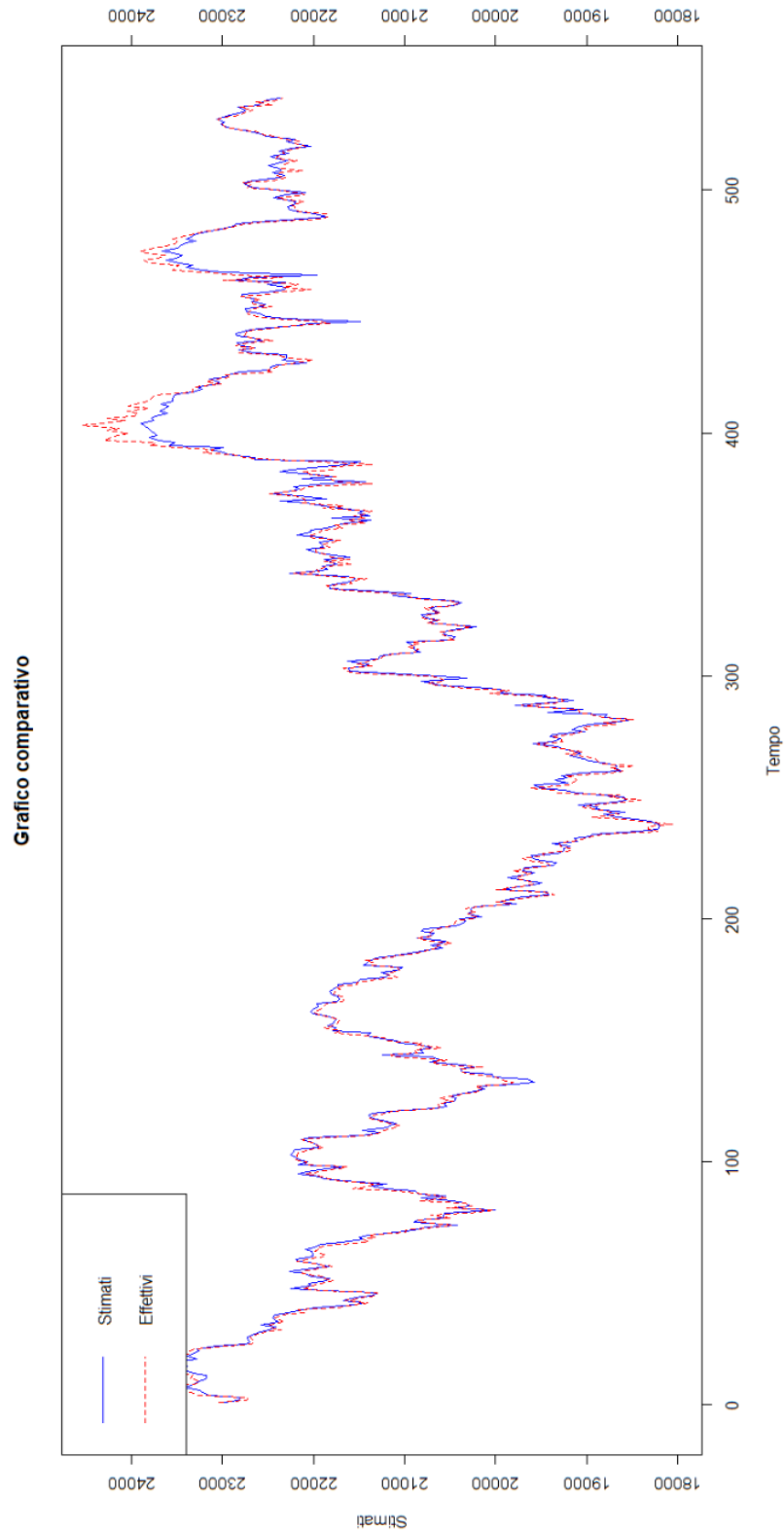


FIGURA 35: GRAFICO COMPARATIVO TRA I PREZZI DI CHIUSURA STIMATI E QUELLI EFFETTIVI.

Come si può notare dal precedente grafico, il quale confronta i primi 538 prezzi di chiusura stimati dell'indice FTSE MIB con quelli effettivi, la rete neurale, nella maggior parte delle osservazioni, risulta essere molto precisa restituendo risultati molto buoni, con un andamento che segue quello dello stesso indice. Si nota solamente una notevole differenza previsiva, nella quale la rete neurale sottoperforma il reale andamento del prezzo di chiusura dell'indice, nel periodo compreso tra la 400esima e la 450esima osservazione, ovvero il periodo che comprende il secondo trimestre del 2018.

Avendo previsto solamente il prezzo di chiusura a un giorno dell'indice italiano FTSE MIB e quindi non potendo ancora dimostrare con certezza la validità e l'efficacia previsiva della rete neurale implementata, si è proceduto a effettuare un test che potesse dimostrarlo. Si sono presi in considerazione, in questo caso un minor numero di dati (e più precisamente dall'11 dicembre 2019 al 31 gennaio 2020) e si è proceduto alla loro previsione, seguendo lo stesso procedimento già spiegato pocanzi.

Dopo averli confrontati con quelli effettivi, si può affermare che la rete neurale implementata presenta una bontà previsiva del 98,23% con un errore dello 0,007145. Tale efficacia la si può osservare anche dal grafico sottostante, il quale mette a confronto i dati effettivi del test set considerato con relativi valori previsti.

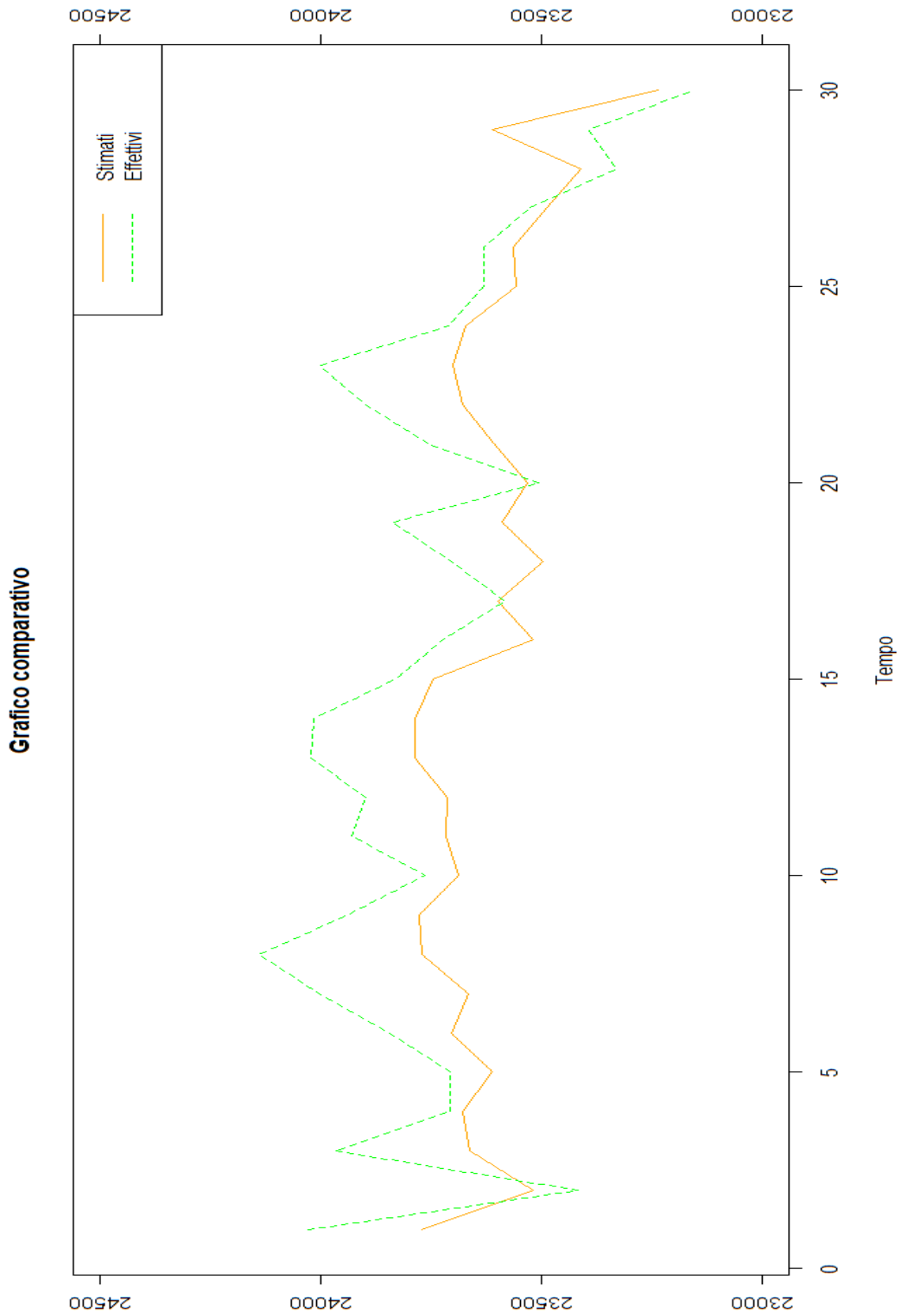


FIGURA 36: GRAFICO COMPARATIVO TRA I DATI EFFETTIVI E LE RELATIVE PREVISIONI DEL TEST SET CONSIDERATO.

Si può notare che l'andamento delle previsioni, nella maggioranza dei casi, segue l'andamento dei dati effettivi sovraperformandoli.

Successivamente, dopo aver implementato una rete neurale artificiale *feed – forward* che permettesse di prevedere il prezzo di chiusura a un giorno dell'indice italiano FTSE MIB, si è deciso di implementare un'ulteriore rete neurale al fine di osservare e verificare se esiste, e in che misura, una correlazione tra l'indice italiano e l'indice Hang Seng di Hong Kong.

6.2 L'indice Hang Seng. Cenni storici.

L'indice Hang Seng è un indice ponderato per la capitalizzazione del flottante di una selezione di società della borsa di Hong Kong. I componenti dell'indice sono divisi in quattro sottoindici: Commercio e Industria, Finanza, Utilità e Proprietà³¹⁴.

Utilizzato al fine di registrare e monitorare i quotidiani cambiamenti delle maggiori società del mercato azionario di Hong Kong, è il principale indicatore delle prestazioni complessive del mercato stesso. Esso è costituito da 50 società che rappresentano circa il 58% della capitalizzazione della borsa asiatica³¹⁵. Queste società, a titolo informativo, sono indicate nella tabella sottostante:

Società				
Tencent	CKH Holdings	China Overseas	Geely Auto	China Shenhua
HSBC Holdings	Link Reit	Sinopec Corp	Petrochina	Bankcomm
AIA	China Life	China Res Land	Power Assets	China Unicom
CCB	CLP Holdings	Boc Hong Kong	CITIC	AAC Tech
Ping AN	HK & China Gas	Sunny Optical	Sino Biopharm	CKI Holdings

³¹⁴ Bloomberg (2020), Informazioni sull'indice Hang Seng di Hong Kong, disponibile a <https://www.bloomberg.com/quote/HSI:IND>.

³¹⁵ Autori di Wikipedia, (2020), *Hang Seng Index*, Wikipedia. L'enciclopedia libera, 08/01/2020.

ICBC	CK Asset	Country Garden	MTR Corporation	Want Want China
China Mobile	SHK PPT	Shenzhou Intl	WH Group	Hangan Int'l
HKEX	Galaxy Ent	Techtronic Ind	New World Dev	Swire Pacific A
Bank of China	Hang Seng Bank	CSPC Pharma	Wharf Reic	Sino Land
CNOOC	Sands China LTD	Mengniu Dairy	Henderson Land	Hang Lung PPT

TABELLA 4: SOCIETÀ CHE COMPONGONO L'INDICE HANG SENG. FONTE: HANG SENG BANK.

Esso è debuttato in Borsa il 24 novembre 1969 ed è attualmente compilato e gestito da Hang Seng Indexes Company Limited, una consociata interamente di proprietà di Hang Seng Bank ovvero una delle maggiori banche registrate e quotate a Hong Kong in termini di capitalizzazione di mercato. Quando l'indice Hang Seng fu pubblicato per la prima volta, la sua base di 100 punti fu stabilita equivalente al valore totale delle azioni alla chiusura del mercato il 31 luglio 1964. Il suo minimo storico è di 58.61 punti, raggiunto retroattivamente il 31 agosto 1967, dopo che è stato stabilito il valore di base, ma prima della pubblicazione stessa. Il 10 dicembre 1993 esso ha superato per la prima volta nella sua storia il traguardo dei 10.000 punti e tredici anni dopo, il 28 dicembre 2006, i 20.000 punti³¹⁶.

Nei dieci mesi successivi ha superato i 30.000 punti, raggiunti il 18 ottobre 2007, dal 30 ottobre 2007 al 9 marzo 2008, l'indice ha perso 9.426 punti, circa il 30%. Il 5 settembre, è passato da oltre i 20.000 punti, a un minimo di 19.708,39, chiudendo in seguito a 19.933,28.

L'8 ottobre 2008, l'indice ha chiuso a 15.431,73, oltre il 50% in meno rispetto al massimo storico e con un valore di chiusura più basso in due anni. Il 27 ottobre 2008, l'indice è ulteriormente sceso a 10.676,29 punti, dopo essere diminuito di quasi due terzi dal suo massimo storico, ma superando nuovamente il traguardo di 20.000 punti a 20.063,93 il 24 luglio 2009. L'indice ha raggiunto 25.000,00 il 19 agosto 2014, e chiudendo

³¹⁶ Autori di Wikipedia, (2020), *Hang Seng Index*, Wikipedia. L'enciclopedia libera, 08/01/2020.

successivamente a 24.909,26 punti. Ha continuato a salire a 26.000 l'8 aprile 2015 con una chiusura di 26.236,86. Il giorno seguente, è salito a 27.922,67 prima di chiudere a 26.944,39. Ancora una volta, un'altra pietra miliare è stata raggiunta il 13 aprile 2015 salendo a oltre 28.000 punti e chiudendo a 28.016.34, il più alto dal dicembre 2007. L'8 luglio 2015 l'indice è sceso fino a 2.139 punti. Il 21 agosto l'indice a rimase in un mercato orso, stabilizzandosi nella fascia dei 18.000-19.000 punti, fino a quando non si concluse nel febbraio 2016. Nel 24 giugno 2016, il mercato ha perso 1.000 punti in risposta ai risultati del referendum nel Regno Unito nell'UE secondo cui "Leave" ha vinto il voto contro "Remain" . Quindi, i mercati sono costantemente diminuiti il 27 e 28 giugno 2016, per poi riprendersi leggermente il 29 giugno 2016. Il suo massimo storico è fissato a 33.223,58 punti, raggiunti il 26 gennaio 2018³¹⁷.

Tutto ciò è evidente nel grafico sottostante:



FIGURA 37: ANDAMENTO DELL'INDICE HANG SENG. FONTE: YAHOO FINANCE

Come per l'indice FTSE MIB, anche per l'indice asiatico Hang Seng, dopo aver ottenuto i dati e dopo averli organizzati e dovutamente sistemati, si è proceduto alla loro normalizzazione con la medesima tecnica, già descritta nel paragrafo precedente, applicata anche in questo caso, sia alle variabili di input, che alla variabile di output.

Successivamente, come nell'applicazione precedente, si è effettuata la *Cross Validation*, sempre suddividendo il data set (comprendente lo stesso periodo campionario 01/06/2009 – 09/12/2019) in due parti:

³¹⁷ Autori di Wikipedia, (2020), *Hang Seng Index*, Wikipedia. L'enciclopedia libera, 08/01/2020.

- L'80% dei dati (ovvero i dati riferiti al periodo campionario che va dal 1° giugno 2009 e termina il 24 ottobre 2017) è stato utilizzato la fase di addestramento (*Training*) ovvero per la stima dei parametri delle reti neurali;
- Il 20% dei dati (più precisamente i dati riferiti al periodo campionario che va dal 25 ottobre 2017 e si conclude il 9 dicembre 2019) è stato utilizzato per la fase di validazione.

Dal momento che l'obiettivo di tale elaborato, oltre che la valutazione della bontà previsiva delle reti neurali per le serie storiche finanziarie, risulta essere lo studio dell'eventuale correlazione esistente tra l'indice italiano e l'indice asiatico, si è proceduto all'implementazione di una rete neurale *feed – forward* caratterizzata da:

- 13 nodi di input i quali comprendono il prezzo di chiusura, il prezzo di apertura, il massimo e minimo giornalieri dell'indice italiano relativi agli ultimi tre giorni disponibili nel data set e il prezzo di apertura dell'indice asiatico a un passo;
- un unico strato nascosto caratterizzato dallo stesso numero nodi della rete neurale precedente cosicché sia possibile un significativo confronto;
- la funzione di attivazione logistica;
- il metodo di stima *Back - propagation resiliente* e, infine,
- un solo nodo di output che indica la previsione del prezzo di chiusura a un passo. Anche in questo caso di studio, il nodo di output restituirà il prezzo di chiusura dell'indice italiano FTSE MIB a un giorno, quindi quello relativo al 10 dicembre 2019.

Si ottiene così un prezzo di chiusura stimato pari a 23.200,53 euro contro un prezzo di chiusura effettivo 23.122,82 euro riferito sempre a un giorno.

Anche in questo caso, come nell'applicazione precedente, si sono create due tabelle contenenti i pesi e i relativi *bias* che collegano rispettivamente lo strato di input con lo stato intermedio, e quest'ultimo, con lo strato di output:

	Nodo1	Nodo2	Nodo3	Nodo4	Nodo5	Nodo6	Nodo7	Nodo8
Bias	1.02096400	-0.04454833	1.01337562	-0.050000909	-0.61798726	2.32311543	-0.5875558	0.1046399654
chiusura1	-0.10302778	4.63497440	3.07103523	2.208335467	-1.07175283	0.23433381	-0.5281010	0.8979816577
apertura1	-2.34071056	-0.80552458	-0.54455640	0.589486683	0.22343026	-0.13210533	1.0858970	0.6624021363
massimo1	0.20725823	-3.03458816	-1.91690967	-0.279408910	1.26897503	0.64241018	2.8778224	-0.1854723304
minimo1	1.40057701	0.73280604	-1.74640219	-0.095004044	0.57614191	0.37190186	-0.4072063	-0.0005994538
chiusura2	-0.30337020	2.92567124	-0.02854049	0.088308414	0.42692563	-0.07556453	0.1454685	0.1070810447
apertura2	1.96337406	-1.09868972	0.82807471	-0.183042866	0.09899274	-1.17653873	-1.3639060	-1.2858216352
massimo2	-0.76304335	-3.41155668	0.01923900	0.790557844	0.88003550	-1.11034148	0.7455302	0.9471857807
minimo2	0.58808279	-0.41174862	1.05124844	-0.776916017	0.69083457	0.84185783	-0.8023957	-1.8958562664
chiusura3	0.10617045	6.82826274	2.30307226	-0.609827263	-3.11237369	-0.33910892	0.3456663	-0.0341734517
apertura3	-0.15020506	-14.39417035	-2.62021258	0.077685145	-0.12616915	-1.22441542	1.4251440	0.1230106166
massimo3	-1.07021906	-1.72260184	-1.06714410	1.139601257	-0.17177733	-2.07584930	-0.1761798	0.4335532321
minimo3	0.04863109	-4.34960267	-0.24612229	-0.004495345	1.12995515	0.25988714	-1.6550165	-2.1065628907
aperturaA	-1.83251553	-0.69368842	-1.13974574	-0.896036513	0.99804912	1.92401113	-0.4768584	0.7826446567

Nodo9	Nodo10	Nodo11	Nodo12	Nodo13	Nodo14
13.8632412	-0.25227882	-0.4315839	-1.26854493	0.6728436	-0.50978975
0.8250855	-0.56335852	0.3229277	-0.31042970	0.4012399	0.71851292
-1.6371281	-1.28340880	-0.1106321	-1.05148422	1.4305633	0.14653555
0.1163034	-1.10075075	-0.2170411	0.07456949	-0.1133328	-1.47193212
-0.2329728	0.47104773	1.1643033	1.62110712	0.5244073	-0.97592909
0.3226387	0.69583402	-1.1136956	-1.30394794	-0.6011762	-0.20577181
-0.4581339	-0.47768203	-0.5160908	0.29120511	1.2148962	2.65967161
-3.0806858	-0.08585338	1.2544929	-0.72541518	0.2261563	-0.65125705
0.1060515	1.10389302	-1.2668775	-1.97682919	2.2630286	0.64388721
-1.2596147	-0.32277410	-0.7758127	0.55192163	1.0756629	-0.06074904
-8.8360866	-1.20522146	1.5828865	5.29665421	-3.4722423	0.86888237
-1.2962955	0.07833132	-1.3865781	-0.09849008	-2.9475914	-0.61089223
0.6150603	0.78669886	1.7810270	-0.47421564	0.6075475	1.12291348
-0.3401522	-0.95479460	2.1725378	-0.28947671	-2.3589640	-0.27077185

TABELLA 5: PESI CHE COLLEGANO I NODI DELLO STRATO DI INPUT CON I NODI DELLO STRATO NASCOSTO.

	Pesi
Bias	-0.1455978
Nodo1	-0.3649324
Nodo2	-3.6403710
Nodo3	-0.3971641
Nodo4	1.6104317
Nodo5	0.8879197
Nodo6	-1.2944288
Nodo7	0.3557323
Nodo8	-0.3983877
Nodo9	-1.5925761
Nodo10	-0.5211282
Nodo11	0.3716888
Nodo12	1.1194362
Nodo13	-0.6115075
Nodo14	1.7860206

TABELLA 6: PESI CHE COLLEGANO I NODI DELLO STRATO INTERMEDIO CON LO STRATO DI OUTPUT.

Come ulteriore verifica, si è proceduto alla creazione di una matrice di correlazione³¹⁸ tra le variabili di input e output del modello.

Come si può notare dalla matrice di correlazione sottostante, la variabile di output, ovvero il prezzo di chiusura a un giorno dell'indice italiano FTSE MIB è correlato da tutte le variabili di input in maniera più o meno rilevante; la variabile che sembra essere meno significativa e quindi con un coefficiente minore rispetto agli altri, è il prezzo di apertura dell'indice asiatico Hang Seng.

³¹⁸ Una tabella contenente dei numeri, compresi tra 0 e 1, detti coefficienti di correlazione. Essi esprimono il grado di dipendenza tra le diverse variabili: un coefficiente pari a 1 significa che la correlazione risulta positiva e che quindi le due variabili presentano un legame di dipendenza, un coefficiente pari a -1 contrariamente indica l'assenza di dipendenza tra le variabili prese in considerazione e infine, un coefficiente pari a 0 indica la presenza di indipendenza stocastica, ovvero che le variabili variano indipendentemente le une dalle altre.

	chiusura1	apertura1	massimo1	minimo1	chiusura2	apertura2	massimo2	minimo2
chiusura1	1.0000000	0.9957508	0.9979042	0.9983873	0.9944248	0.9903746	0.9925000	0.9925000
apertura1	0.9957508	1.0000000	0.9985647	0.9980245	0.9987944	0.9943941	0.9966502	0.9966502
massimo1	0.9979042	0.9985647	1.0000000	0.9979389	0.9975730	0.9937553	0.9962099	0.9962099
minimo1	0.9983873	0.9980245	0.9979389	1.0000000	0.9967486	0.9920332	0.9939625	0.9939625
chiusura2	0.9944248	0.9987944	0.9975730	0.9967486	1.0000000	0.9957459	0.9979009	0.9979009
apertura2	0.9903746	0.9943941	0.9937553	0.9920332	0.9957459	1.0000000	0.9985641	0.9985641
massimo2	0.9925000	0.9966502	0.9962099	0.9939625	0.9979009	0.9985641	1.0000000	1.0000000
minimo2	0.9925000	0.9966502	0.9962099	0.9939625	0.9979009	0.9985641	1.0000000	1.0000000
chiusura3	0.9891883	0.9932772	0.9927707	0.9907696	0.9944187	0.9987937	0.9975717	0.9975717
apertura3	0.9851748	0.9891299	0.9890770	0.9862699	0.9903657	0.9943915	0.9937518	0.9937518
massimo3	0.9874536	0.9913509	0.9915621	0.9882747	0.9924937	0.9966485	0.9962078	0.9962078
minimo3	0.9874536	0.9913509	0.9915621	0.9882747	0.9924937	0.9966485	0.9962078	0.9962078
aperturaA	0.5848348	0.5849518	0.5785671	0.5917354	0.5864726	0.5865897	0.5799502	0.5799502
chiusura	0.9841225	0.9880474	0.9881321	0.9851406	0.9891784	0.9932749	0.9927669	0.9927669

chiusura3	apertura3	massimo3	minimo3	aperturaA	chiusura
0.9891883	0.9851748	0.9874536	0.9874536	0.5848348	0.9841225
0.9932772	0.9891299	0.9913509	0.9913509	0.5849518	0.9880474
0.9927707	0.9890770	0.9915621	0.9915621	0.5785671	0.9881321
0.9907696	0.9862699	0.9882747	0.9882747	0.5917354	0.9851406
0.9944187	0.9903657	0.9924937	0.9924937	0.5864726	0.9891784
0.9987937	0.9943915	0.9966485	0.9966485	0.5865897	0.9932749
0.9975717	0.9937518	0.9962078	0.9962078	0.5799502	0.9927669
0.9975717	0.9937518	0.9962078	0.9962078	0.5799502	0.9927669
1.0000000	0.9957439	0.9978998	0.9978998	0.5883309	0.9944167
0.9957439	1.0000000	0.9985631	0.9985631	0.5884294	0.9987932
0.9978998	0.9985631	1.0000000	1.0000000	0.5814657	0.9975704
0.9978998	0.9985631	1.0000000	1.0000000	0.5814657	0.9975704
0.5883309	0.5884294	0.5814657	0.5814657	1.0000000	0.5902736
0.9944167	0.9987932	0.9975704	0.9975704	0.5902736	1.0000000

FIGURA 39: MATRICE DI CORRELAZIONE.

Ma per verificare, oltre che l'efficacia previsiva della rete neurale, anche la possibile ed eventuale correlazione esistente tra l'indice italiano e l'indice asiatico, si è considerata un'ulteriore serie di dati, chiamata *test set*, comprendenti il periodo che va dall'11 dicembre 2019 al 31 gennaio 2020 al fine di confrontare quest'ultimi dati, che risultano essere quelli effettivi, con la loro relativa previsione, ottenuta grazie alla medesima rete neurale.

Si ottiene un'accuratezza della previsione pari al 98,22% con un errore che risulta essere dello 0,01134.

A fini esemplificativi, il grafico dato da tale confronto risulta essere il seguente:

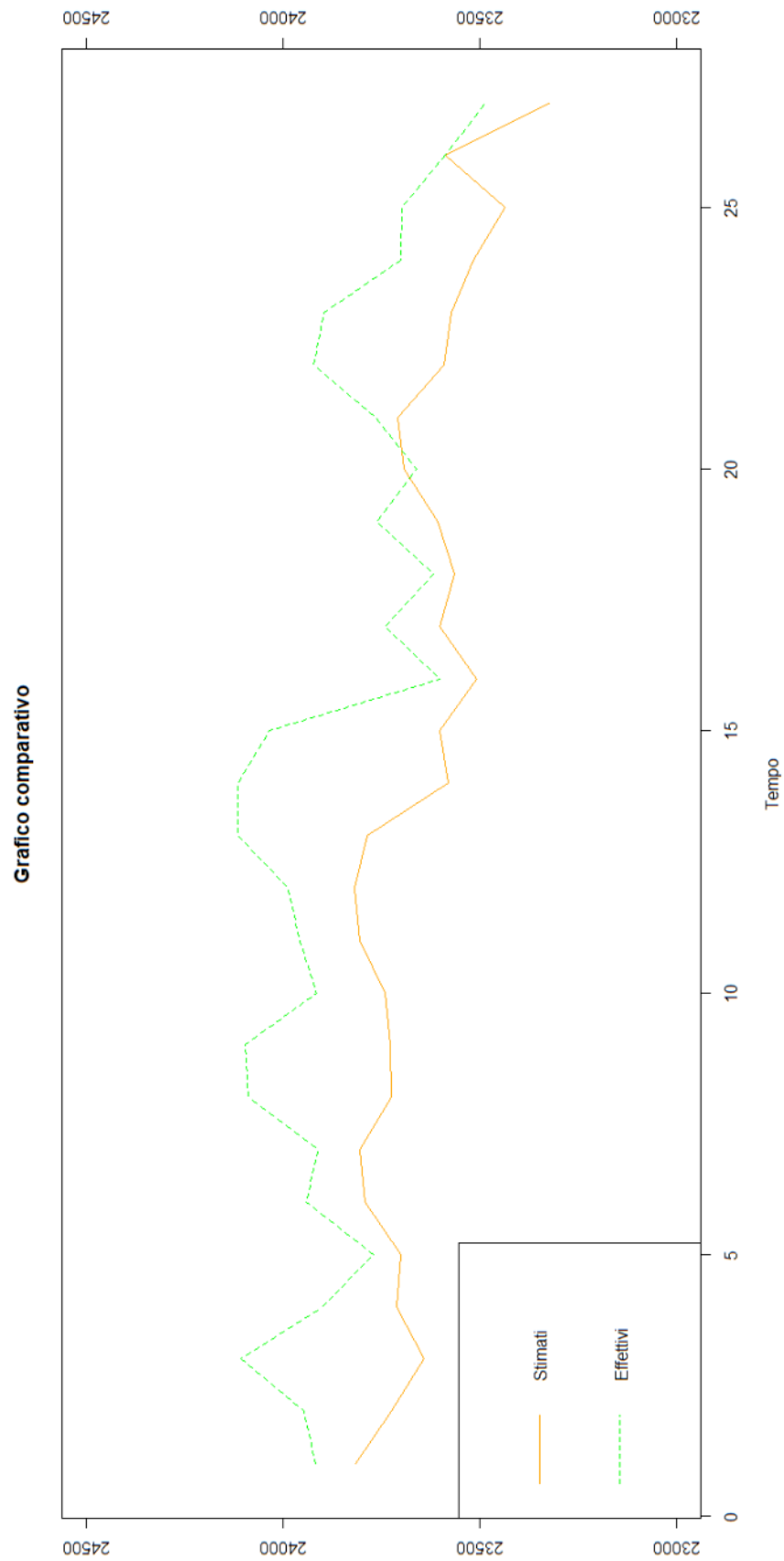


FIGURA 40: GRAFICO COMPARATIVO TRA I DATI EFFETTIVI E LE RELATIVE PREVISIONI DEL TEST SET CONSIDERATO.

Come si può notare, anche in questo caso, nella maggioranza dei casi, l'andamento della previsione segue quello dei dati effettivi sovraperformandoli.

Un'ulteriore considerazione che si può trarre, risulta essere il fatto che pur considerando il prezzo di apertura dell'indice asiatico Hang Seng, la bontà previsiva della rete neurale, caratterizzata dalla stessa struttura e dagli stessi parametri della precedente, risulta pressoché invariata; il che significa che l'indice asiatico non risulta essere un dato significativo al fine di migliorare la previsione dell'indice italiano a un passo.

Conclusioni

L'intelligenza artificiale, come la maggior parte delle discipline, può essere implementata in diversi modi e applicata per raggiungere una molteplicità di obiettivi.

Dopo l'analisi della relativa letteratura, nonché quella di applicazioni simili, l'obiettivo iniziale di tale elaborato è risultato lo studio e la confutabilità della capacità previsiva delle reti neurali artificiali, e in particolar modo delle reti *feed – forward*. A tale scopo si è deciso di testare tale abilità nelle serie storiche finanziarie.

Considerando la serie storica dell'indice italiano FTSE MIB si è implementata una rete neurale caratterizzata da 12 nodi di input comprendenti il prezzo di chiusura, il prezzo di apertura, il prezzo massimo e minimo (tutti caratterizzati da tre ritardi) e ulteriori parametri attentamente ponderati e selezionati al fine di ottenere come output il prezzo di chiusura a un passo dell'indice stesso. Dopo aver ottenuto un risultato più che soddisfacente, si è ritenuto opportuno effettuare un'ulteriore verifica della bontà previsiva della rete; il fatto di aver ottenuto un solo risultato buono non sembrava efficace e confutabile. Si è proceduto quindi, ad effettuare un maggior numero di previsioni a un passo, ottenendo comunque una percentuale di affidabilità molto elevata, pari a 98,23%.

Ulteriore questione alla quale tale elaborato è teso a risolvere però, è quella della possibile ed eventuale esistenza di correlazione, e quindi di dipendenza, tra l'indice Hang Seng di Hong Kong e l'indice italiano FTSE MIB. Proprio per questo, si è proceduto a implementare un'ulteriore rete neurale caratterizzata dagli stessi nodi di input della rete precedente, ma con l'aggiunta del prezzo di apertura dell'indice asiatico non ritardato. Si è optato per la stessa struttura della rete precedente per ottenere un confronto più significativo e attinente con la precedente.

Anche in tal caso, anziché tranne conclusioni poco affidabili basate su un singolo output ottenuto, si sono effettuate un maggior numero di previsioni a un passo, ottenendo una bontà previsiva, anche in questo caso molto soddisfacente, pari al 98,22%.

Ora è possibile affermare che il prezzo di apertura dell'indice asiatico, considerato in quanto la Borsa di Hong Kong apre prima per il fuso orario, non risulta una variabile

significativa; essa infatti non apporta né migliorie, né peggioramenti alla capacità previsiva della rete neurale implementata.

Per contro, come ultima prova effettuata, ma non inserita in tale elaborato, si è deciso di provare a diminuire il numero delle variabili di input della rete neurale, non considerando più i tre ritardi che caratterizzavano il prezzo di chiusura, apertura, massimo e minimo dell'indice italiano, ma andandone a considerare solamente uno e ovviamente con l'aggiunta della variabile del prezzo di apertura dell'indice asiatico non ritardato.

In quest'ultimo caso, però la percentuale di bontà previsiva della rete neurale, ottenuta ovviamente considerando un maggior numero di previsioni, subisce un drastico peggioramento arrivando a 64,22%.

Essendo che però, anche in quest'ultimo caso, il prezzo di apertura dell'indice asiatico era presente tra le variabili di input e avendo già testato in precedenza la sua indipendenza e non correlazione con l'indice italiano, si può azzardare la conclusione che diminuendo il numero di ritardi delle variabili di input del FTSE MIB, la capacità previsiva della rete neurale peggiora, restituendo un output notevolmente dissimile al prezzo di chiusura effettivo.

Le reti neurali, grazie anche a tale elaborato, riescono a riconfermare la loro capacità previsiva nell'ambito delle serie storiche finanziarie, permettendo così, per coloro che le utilizzano, di poter investire con una maggiore consapevolezza e accuratezza.

APPENDICE A. SCRIPT RETE NEURALE FEED – FORWARD FTSE MIB

```
##Caricamento dati, installazione pacchetti e librerie
getwd()
setwd("C:/Users/Alic/Desktop")
getwd()
library(tseries)
dati=read.csv2("Dati.csv", sep=',', dec=',', header = TRUE)
View(dati)
library(quantmod)
library(neuralnet)

##Fase di normalizzazione dei dati
matrice2 = as.matrix(dati)
matrice = matrix(nrow=nrow(matrice2),ncol=ncol(matrice2))

for(row in 1:nrow(matrice)) {
  for(col in 2:5) {
    matrice[row,col]=as.numeric(sub(",", ".", sub(".", "", matrice2[row,col] , fixed = TRUE) , fixed
= TRUE))
  }
}
minnorm = matrix(nrow=1,ncol=4)
maxnorm = matrix(nrow=1,ncol=4)

for(i in 1:4) {
  minnorm[1,i] = min(matrice[,i+1])
  maxnorm[1,i] = max(matrice[,i+1])
}
input=matrix(nrow = 2688, ncol = 12)
for(row in 1:nrow(input)) {
  for(col in 1:ncol(input)) {
    if(col<=4) {
      input[row,col] = matrice[row,col+1]
    } else if (col>4 & col<=8){
      if(col%%4 == 0) {
        input[row,col] = matrice[row+1, 4 ]
      } else {
        input[row,col] = matrice[row+1,(col%%4)+1]
      }
    } else{
      if(col%%4 == 0) {
        input[row,col] = matrice[row+2, 4 ]
      } else {
        input[row,col] = matrice[row+2,(col%%4)+1]
      }
    }
  }
}
}
```

```

inputnorm=matrix(nrow = 2688, ncol = 12)
for(row in 1:nrow(inputnorm)) {
  for(col in 1:ncol(inputnorm)) {
    if(col%%4==0){
      inputnorm[row,col]=((input[row,col]-minnorm[1,4])/(maxnorm[1,4]-minnorm[1,4]))
    } else {
      inputnorm[row,col]=((input[row,col]-minnorm[1,col%%4])/(maxnorm[1,col%%4]-
minnorm[1,col%%4]))
    }
  }
}
output=matrix(nrow = 2688, ncol = 1)
for(row in 1:nrow(input)) {
  output[row,1]=matrice[row+3,2]
}
outputmax = max(output)
outputmin = min(output)

outputnorm=matrix(nrow = 2688, ncol = 1)
for(row in 1:nrow(output)) {
  outputnorm[row,1]=((output[row,1]-outputmin)/(outputmax-outputmin))
}
df=data.frame(inputnorm,outputnorm)

df <- setNames(df,
c("chiusura1","apertura1","massimo1","minimo1","chiusura2","apertura2","massimo2","mini
mo2","chiusura3","apertura3","massimo3","minimo3","chiusura" ))

##CROSS VALIDATION (80% dei dati per il trainig e il 20% per la validazione)

#evaluationrange =1:538
#evaluationoutput = df[evaluationrange,13]
#evaluationinput = df[evaluationrange,1:12]
#evaluationtraining = df[-evaluationrange,]

##Implementazione rete neurale feed - forward

nn=neuralnet(chiusura ~ chiusura1 + apertura1 + massimo1 + minimo1 + chiusura2 +
apertura2 + massimo2 + minimo2 + chiusura3 + apertura3 + massimo3 + minimo3, data=df,
hidden=14, act.fct = "logistic", linear.output = FALSE, algorithm = "rprop+")
plot(nn)

test=data.frame("chiusura1"=0.8882312, "apertura1"= 0.8818474
,
"massimo1"=0.8875129
, "minimo1"= 0.9119267,"chiusura2"=
0.8707197, "apertura2"=0.8813775, "massimo2"=0.8850915,
"minimo2"=0.9095208,"chiusura3"=0.8696935, "apertura3"=0.8949998,
"massimo3"=0.8886413, "minimo3"=0.88156749)
Predict=compute(nn,test)
chiusura=Predict$net.result
chiusuradenorm= (chiusura*(outputmax-outputmin)+outputmin)

```

```

##Visualizzazione pesi rete neurale
pesi2=nn$weights
inw=pesi2[[1]]
inw2=inw[[1]]
pesi=as.matrix(inw2)
matricepesil=matrix(inw2,nrow = nrow(inw2),ncol=ncol(inw2))
colnames(matricepesil)=c("Nodo1","Nodo2","Nodo3","Nodo4","Nodo5","Nodo6","Nodo7","N
odo8","Nodo9","Nodo10","Nodo11","Nodo12","Nodo13","Nodo14")
rownames(matricepesil)=c("Bias","chiusura1","apertura1","massimo1","minimo1","chiusura2"
,"apertura2","massimo2","minimo2","chiusura3","apertura3","massimo3","minimo3")
View(matricepesil)

ouw=pesi2[[1]]
ouw2=ouw[[2]]
pesi=as.matrix(ouw2)
matricepesiO=matrix(ouw2,nrow = nrow(ouw2),ncol=ncol(ouw2))
colnames(matricepesiO)=c("Pesi")
rownames(matricepesiO)=c("Bias","Nodo1","Nodo2","Nodo3","Nodo4","Nodo5","Nodo6","No
do7","Nodo8","Nodo9","Nodo10","Nodo11","Nodo12","Nodo13","Nodo14")
View(matricepesiO)

#library(NeuralNetTools)
#plotnet(nn)

##Validazione
#nn=neuralnet(chiusura ~ chiusura1 + apertura1 + massimo1 + minimo1 + chiusura2 +
apertura2 + massimo2 + minimo2 + chiusura3 + apertura3 + massimo3 + minimo3,
data=evaluationtraining, hidden=14, act.fct = "logistic", linear.output = FALSE)

#test=evaluationinput
#Predict=compute(nn,test)
#chiusura=Predict$net.result
#chiusuradenorm= (chiusura*(outputmax-outputmin)+outputmin)

#plot(chiusuradenorm[1:538,], type = "l", ylab = "Stimati",
#main = "Grafico comparativo", xlab = "Tempo",
#col = "blue",ylim=c(18000,24500))
#par(new = TRUE)
#plot(output[1:538], type = "l", xaxt = "n", yaxt = "n",
#ylab = "", xlab = "", col = "red", lty = 2,ylim=c(18000,24500))
#axis(side = 4)
#mtext("Effettivi", side = 4, line = 3)
#legend("topleft", c("Stimati", "Effettivi"),
#col = c("blue", "red"), lty = c(1, 2))

##Fase di test della bontà previsiva della rete
getwd()
setwd("C:/Users/Alic/Desktop")
getwd()
library(tseries)

```

```

datitest=read.csv2("Datitest.csv", sep=',',dec=',', header = TRUE)
View(datitest)
library(quantmod)
library(neuralnet)

##Fase di normalizzazione dei dati
matrice1 = as.matrix(datitest)
matricetest = matrix(nrow=nrow(matrice1),ncol=ncol(matrice1))

for(row in 1:nrow(matricetest)) {
  for(col in 2:5) {
    matricetest[row,col]=as.numeric(sub(" ", ".", sub(".", "", matrice1[row,col] , fixed = TRUE) ,
fixed = TRUE))
  }
}
for(i in 1:4) {
  minnorm[1,i] = min(matrice[,i+1])
  maxnorm[1,i] = max(matrice[,i+1])
}
inputest=matrix(nrow = 30, ncol = 12)
for(row in 1:nrow(inputest)) {
  for(col in 1:ncol(inputest)) {
    if(col<=4) {
      inputest[row,col] = matricetest[row,col+1]
    } else if (col>4 & col<=8){
      if(col%%4 == 0) {
        inputest[row,col] = matricetest[row+1, 4 ]
      } else {
        inputest[row,col] = matricetest[row+1,(col%%4)+1]
      }
    } else{
      if(col%%4 == 0) {
        inputest[row,col] = matricetest[row+2, 4 ]
      } else {
        inputest[row,col] = matricetest[row+2,(col%%4)+1]
      }
    }
  }
}
inputestnorm=matrix(nrow = 30, ncol = 12)
for(row in 1:nrow(inputestnorm)) {
  for(col in 1:ncol(inputestnorm)) {
    if(col%%4==0){
      inputestnorm[row,col]=((inputest[row,col]-minnorm[1,4])/(maxnorm[1,4]-minnorm[1,4]))
    } else {
      inputestnorm[row,col]=((inputest[row,col]-minnorm[1,col%%4])/(maxnorm[1,col%%4]-
minnorm[1,col%%4]))
    }
  }
}
}

```

```

evaluationoutput=matrix(nrow = 30, ncol = 1)
for(row in 1:nrow(inputest)) {
  evaluationoutput[row,1]=matricetest[row+3,2]
}
evaluationoutputnorm=matrix(nrow = 30, ncol = 1)
for(row in 1:nrow(evaluationoutput)) {
  evaluationoutputnorm[row,1]=((evaluationoutput[row,1]-outputmin)/(outputmax-
outputmin))
}
dfest=data.frame(inputestnorm)

dfest <- setNames(dfest,
c("chiusura1t", "apertura1t", "massimo1t", "minimo1t", "chiusura2t", "apertura2t", "massimo2t",
"minimo2t", "chiusura3t", "apertura3t", "massimo3t", "minimo3t"))

Predictest=compute(nn,dfest)
chiusuratest=Predictest$net.result
chiusuratestdenorm= (chiusuratest*(outputmax-outputmin)+outputmin)

errortest=sum((chiusuratest-evaluationoutputnorm)^2)

comparison=data.frame(chiusuratest, evaluationoutputnorm)
deviation=((evaluationoutputnorm-chiusuratest)/evaluationoutputnorm)
comparison1=data.frame(chiusuratest,evaluationoutputnorm,deviation)
accuracy=1-abs(mean(deviation))

plot(chiusuratestdenorm, type = "l", ylab = "Stimati",
main = "Grafico comparativo", xlab = "Tempo",
col = "orange",ylim=c(23000,24500))
par(new = TRUE)
plot(evaluationoutput, type = "l", xaxt = "n", yaxt = "n",
ylab = "", xlab = "", col = "green", lty = 2,ylim=c(23000,24500))
axis(side = 4)
mtext("Effettivi", side = 4, line = 3)
legend("topright", c("Stimati", "Effettivi"),
col = c("orange", "green"), lty = c(1, 2))

```


APPENDICE B. SCRIPT RETE NEURALE FEED – FORWARD FTSE MIB –

HANG SENG

```
## Caricamento dati, installazione pacchetti e librerie
getwd()
setwd("C:/Users/Alic/Desktop")
getwd()
library(tseries)
dati=read.csv2("Dati.csv", sep=',', dec=',', header = TRUE)
library(readxl)
datiA <- read_excel("DatiA.xlsx")
library(quantmod)
library(neuralnet)

##Caricamento e sistemazione dei dati
matrice2 = as.matrix(dati)
matrice2A = as.matrix(datiA)

todelete <- c()
for(row in 1:nrow(matrice2)){
  rowA <- 1
  trovato <- 0
  while (rowA <= nrow(matrice2A) && trovato==0) {
    if(matrice2[row,1]==matrice2A[rowA,1]){
      trovato <-1
    }
    rowA <- rowA+1
  }
  if(trovato==0){
    print(paste("cancello data",matrice2[row,1]))
    todelete <- append(todelete,row)
  }
}
matrice2 <- matrice2[-todelete,]

todelete <- c()
for(rowA in 1:nrow(matrice2A)){
  row <- 1
  trovato <- 0
  while (row <= nrow(matrice2) && trovato==0) {
    if(matrice2A[rowA,1]==matrice2[row,1]){
      trovato <-1
    }
    row <- row+1
  }
  if(trovato==0){
    print(paste("cancello data",matrice2A[rowA,1]))
    todelete <- append(todelete,rowA)
  }
}
```

```

matrice2A <- matrice2A[-todelete,]

matrice = matrix(nrow=nrow(matrice2),ncol=ncol(matrice2))
matriceA = matrix(nrow=nrow(matrice2A),ncol=ncol(matrice2A))

for(row in 1:nrow(matrice)) {
  for(col in 2:5) {
    matrice[row,col]=as.numeric(sub(",", ". ", sub(".", "", matrice2[row,col] , fixed = TRUE) , fixed
= TRUE))
  }
}
minnorm = matrix(nrow=1,ncol=4)
maxnorm = matrix(nrow=1,ncol=4)

for(i in 1:4) {
  minnorm[1,i] = min(matrice[,i+1])
  maxnorm[1,i] = max(matrice[,i+1])
}
inputl=matrix(nrow = 2578, ncol = 12)
for(row in 1:nrow(inputl)) {
  for(col in 1:ncol(inputl)) {
    if(col<=4) {
      inputl[row,col] = matrice[row,col+1]
    } else if (col>4 & col<=8){
      if(col%%4 == 0) {
        inputl[row,col] = matrice[row+1, 4 ]
      } else {
        inputl[row,col] = matrice[row+1,(col%%4)+1]
      }
    } else{
      if(col%%4 == 0) {
        inputl[row,col] = matrice[row+2, 4 ]
      } else {
        inputl[row,col] = matrice[row+2,(col%%4)+1]
      }
    }
  }
}
inputlnorm=matrix(nrow = 2578, ncol = 12)
for(row in 1:nrow(inputlnorm)) {
  for(col in 1:ncol(inputlnorm)) {
    if(col%%4==0){
      inputlnorm[row,col]=((inputl[row,col]-minnorm[1,4])/(maxnorm[1,4]-minnorm[1,4]))
    } else {
      inputlnorm[row,col]=((inputl[row,col]-minnorm[1,col%%4])/(maxnorm[1,col%%4]-
minnorm[1,col%%4]))
    }
  }
}
}

```



```

matriceI=as.matrix(inputInorm)

for(row in 1:nrow(matriceA)) {
  for(col in 2:5) {
    matriceA[row,col]=as.numeric(matrice2A[row,col])
  }
}
matriceA1=as.matrix(matriceA[4:nrow(matriceA),-c(1,2,4,5,6,7)])

minnormA = min(matriceA1)
maxnormA = max(matriceA1)

inputA=as.matrix(matriceA1)

inputAnorm=matrix(nrow = 2578, ncol = 1)
for(row in 1:nrow(inputAnorm)) {
  inputAnorm[row,1]=((inputA[row,1]-minnormA)/(maxnormA-minnormA))
}

matriceasian= as.matrix(inputAnorm)
DATI=data.frame(matriceI, matriceasian)

matriceDATI=as.matrix(DATI)

output=matrix(nrow = 2578, ncol = 1)
for(row in 1:nrow(matriceDATI)) {
  output[row,1]=matrice[row+3,2]
}

outputmax = max(output)
outputmin = min(output)

outputnorm=matrix(nrow = 2578, ncol = 1)
for(row in 1:nrow(output)) {
  outputnorm[row,1]=((output[row,1]-outputmin)/(outputmax-outputmin))
}

df=data.frame(matriceDATI,outputnorm)

df <- setNames(df,
c("chiusura1","apertura1","massimo1","minimo1","chiusura2","apertura2","massimo2","mini
mo2","chiusura3","apertura3","massimo3","minimo3","aperturaA", "chiusura" ))

##Implementazione rete neurale feed - forward
nn=neuralnet(chiusura ~ chiusura1 + apertura1 + massimo1 + minimo1 + chiusura2 +
apertura2 + massimo2 + minimo2 + chiusura3 + apertura3 + massimo3 + minimo3 + aperturaA,
data=df, hidden=14, act.fct = "logistic", linear.output = FALSE, algorithm = "rprop+")
plot(nn)

test=data.frame("chiusura1"=0.8882312, "apertura1"= 0.8818474 ,
"massimo1"=0.8875129 , "minimo1"= 0.9119267,"chiusura2"=
0.8707197, "apertura2"=0.8813775, "massimo2"=0.8850915,

```

```

"minimo2"=0.9095208,"chiusura3"=0.8696935, "apertura3"=0.8949998,
"massimo3"=0.8886413, "minimo3"=0.88156749, "aperturaA"=0.5625070)
Predict=compute(nn,test)
chiusura=Predict$net.result
chiusuradenorm= (chiusura*(outputmax-outputmin)+outputmin)

library(ggplot2)
library(corrplot)
data.frame.cor=as.matrix(df)
correlation.matrix=cor(data.frame.cor)

##Visualizzazione pesi rete neurale
pesi2=nn$weights
inw=pesi2[[1]]
inw2=inw[[1]]
pesi=as.matrix(inw2)
matricepesil=matrix(inw2,nrow = nrow(inw2),ncol=ncol(inw2))
colnames(matricepesil)=c("Nodo1","Nodo2","Nodo3","Nodo4","Nodo5","Nodo6","Nodo7","N
odo8","Nodo9","Nodo10","Nodo11","Nodo12","Nodo13","Nodo14")
rownames(matricepesil)=c("Bias","chiusura1","apertura1","massimo1","minimo1","chiusura2"
,"apertura2","massimo2","minimo2","chiusura3","apertura3","massimo3","minimo3","apertur
aA")
View(matricepesil)

ouw=pesi2[[1]]
ouw2=ouw[[2]]
pesi=as.matrix(ouw2)
matricepesiO=matrix(ouw2,nrow = nrow(ouw2),ncol=ncol(ouw2))
colnames(matricepesiO)=c("Pesi")
rownames(matricepesiO)=c("Bias","Nodo1","Nodo2","Nodo3","Nodo4","Nodo5","Nodo6","No
do7","Nodo8","Nodo9","Nodo10","Nodo11","Nodo12","Nodo13","Nodo14")
View(matricepesiO)

##Fase di test della bontà previsiva della rete
getwd()
setwd("C:/Users/Alic/Desktop")
getwd()
library(tseries)
datitest=read.csv2("Datitest.csv", sep=',',dec=',', header = TRUE)
library(readxl)
datitestA=read.csv2("DatitestA.csv", sep=',',dec=',', header = TRUE)
library(quantmod)
library(neuralnet)

matrice1 = as.matrix(datitest)
matrice1A = as.matrix(datitestA)

todelete <- c()
for(row in 1:nrow(matrice1)){
  rowA <- 1
  trovato <- 0

```

```

while (rowA <= nrow(matrice1A) && trovato==0) {
  if(matrice1[row,1]==matrice1A[rowA,1]){
    trovato <-1
  }
  rowA <- rowA+1
}
if(trovato==0){
  print(paste("cancello data",matrice1[row,1]))
  todelete <- append(todelete,row)
}
}
matrice1 <- matrice1[-todelete,]

todelete <- c()
for(rowA in 1:nrow(matrice1A)){
  row <- 1
  trovato <- 0
  while (row <= nrow(matrice1) && trovato==0) {
    if(matrice1A[rowA,1]==matrice1[row,1]){
      trovato <-1
    }
    row <- row+1
  }
  if(trovato==0){
    print(paste("cancello data",matrice1A[rowA,1]))
    todelete <- append(todelete,rowA)
  }
}
matrice1A <- matrice1A[-todelete,]

matricetest = matrix(nrow=nrow(matrice1),ncol=ncol(matrice1))
matricetestA = matrix(nrow=nrow(matrice1A),ncol=ncol(matrice1A))

for(row in 1:nrow(matricetest)) {
  for(col in 2:5) {
    matricetest[row,col]=as.numeric(sub(",", ". ", sub(".", "", matrice1[row,col] , fixed = TRUE) ,
fixed = TRUE))
  }
}
inputestI=matrix(nrow = 27, ncol = 12)
for(row in 1:nrow(inputestI)) {
  for(col in 1:ncol(inputestI)) {
    if(col<=4) {
      inputestI[row,col] = matricetest[row,col+1]
    } else if (col>4 & col<=8){
      if(col%%4 == 0) {
        inputestI[row,col] = matricetest[row+1, 4 ]
      } else {
        inputestI[row,col] = matricetest[row+1,(col%%4)+1]
      }
    } else{
      if(col%%4 == 0) {

```

```

    inputestl[row,col] = matricetest[row+2, 4 ]
  } else {
    inputestl[row,col] = matricetest[row+2,(col%%4)+1]
  }
}
}
}
inputestnorml=matrix(nrow = 27, ncol = 12)
for(row in 1:nrow(inputestnorml)) {
  for(col in 1:ncol(inputestnorml)) {
    if(col%%4==0){
      inputestnorml[row,col]=((inputestl[row,col]-minnorm[1,4])/(maxnorm[1,4]-minnorm[1,4]))
    } else {
      inputestnorml[row,col]=((inputestl[row,col]-minnorm[1,col%%4])/(maxnorm[1,col%%4]-
minnorm[1,col%%4]))
    }
  }
}
matriceltest=as.matrix(inputestnorml)

for(row in 1:nrow(matricetestA)) {
  for(col in 2:5) {
    matricetestA[row,col]=as.numeric(sub(" ", ". ", sub(".", "" , matrice1A[row,col] , fixed = TRUE)
, fixed = TRUE))
  }
}
inputestA=as.matrix(matricetestA)
inputestAnorm=matrix(nrow = 27, ncol = 1)
for(row in 1:nrow(inputestAnorm)) {
  inputestAnorm[row,1]=((inputestA[row,3]-minnormA)/(maxnormA-minnormA))
}

matriceasiantest= as.matrix(inputestAnorm)

DATI=data.frame(matriceltest, matriceasiantest)

matriceDATItest=as.matrix(DATI)

outputest=matrix(nrow = 27, ncol = 1)
for(row in 1:nrow(outputest)) {
  outputest[row,1]=matriceDATItest[row,2]
}
outputestdenorm= (outputest*(outputmax-outputmin)+outputmin)
dfctest=data.frame(matriceDATItest)

dfctest <- setNames(dfctest,
c("chiusura1","apertura1","massimo1","minimo1","chiusura2","apertura2","massimo2","mini
mo2","chiusura3","apertura3","massimo3","minimo3","aperturaA"))

Predictest=compute(nn,dfctest)
chiusuratest=Predictest$net.result
chiusuratestdenorm= (chiusuratest*(outputmax-outputmin)+outputmin)

```

```
errortest=sum((chiusuratest-outputest)^2)
```

```
comparison=data.frame(chiusuratest, outputest)
```

```
deviation=((outputest-chiusuratest)/outputest)
```

```
comparison1=data.frame(chiusuratest,outputest,deviation)
```

```
accuracy=1-abs(mean(deviation))
```

```
plot(chiusuratestdenorm, type = "l", ylab = "Stimati", main = "Grafico comparativo", xlab =  
"Tempo", col = "orange", ylim=c(23000,24500))
```

```
par(new = TRUE)
```

```
plot(outputestdenorm, type = "l", xaxt = "n", yaxt = "n", ylab = "", xlab = "", col = "green", lty =  
2,ylim=c(23000,24500))
```

```
axis(side = 4)
```

```
mtext("Effettivi", side = 4, line = 3)
```

```
legend("bottomleft", c("Stimati", "Effettivi"), col = c("orange", "green"), lty = c(1, 2))
```


APPENDICE C. SCRIPT RETE NEURALE FEED – FORWARD FTSE MIB – HANG SENG CON RIDUZIONE DELLE VARIABILI DI INPUT

```
## Caricamento dati, installazione pacchetti e librerie
getwd()
setwd("C:/Users/Alic/Desktop")
getwd()
library(tseries)
dati=read.csv2("Dati.csv", sep=',', dec=',', header = TRUE)
library(readxl)
datiA <- read_excel("DatiA.xlsx")
library(quantmod)
library(neuralnet)

##Caricamento e sistemazione dei dati
matrice2 = as.matrix(dati)
matrice2A = as.matrix(datiA)

todelete <- c()
for(row in 1:nrow(matrice2)){
  rowA <- 1
  trovato <- 0
  while (rowA <= nrow(matrice2A) && trovato==0) {
    if(matrice2[row,1]==matrice2A[rowA,1]){
      trovato <-1
    }
    rowA <- rowA+1
  }
  if(trovato==0){
    print(paste("cancello data",matrice2[row,1]))
    todelete <- append(todelete,row)
  }
}
matrice2 <- matrice2[-todelete,]

todelete <- c()
for(rowA in 1:nrow(matrice2A)){
  row <- 1
  trovato <- 0
  while (row <= nrow(matrice2) && trovato==0) {
    if(matrice2A[rowA,1]==matrice2[row,1]){
      trovato <-1
    }
    row <- row+1
  }
  if(trovato==0){
    print(paste("cancello data",matrice2A[rowA,1]))
    todelete <- append(todelete,rowA)
  }
}
```

```

}
matrice2A <- matrice2A[-todelete,]

matrice = matrix(nrow=nrow(matrice2),ncol=ncol(matrice2))
matriceA = matrix(nrow=nrow(matrice2A),ncol=ncol(matrice2A))

for(row in 1:nrow(matrice)) {
  for(col in 2:5) {
    matrice[row,col]=as.numeric(sub(",", ". ", sub(".", "", matrice2[row,col] , fixed = TRUE) , fixed
= TRUE))
  }
}
for(row in 1:nrow(matriceA)) {
  for(col in 2:5) {
    matriceA[row,col]=as.numeric(matrice2A[row,col])
  }
}

##Normalizzazione dei dati
DATI=data.frame(matrice[-c(1,6,7)], matriceA[-c(1,2,4,5,6,7)])

minnorm = matrix(nrow=1,ncol=5)
maxnorm = matrix(nrow=1,ncol=5)

for(i in 2581:2){
  DATI[i,5]=DATI[i-1,5]
}
DATI <- setNames(DATI, c("chiusural","apertural","massimoI","minimoI","aperturaA" ))

matriceDATI=as.matrix(DATI)

for(i in 1:5) {
  minnorm[1,i] = min(matriceDATI[,i])
  maxnorm[1,i] = max(matriceDATI[,i])
}
output=matrix(nrow = 2580, ncol = 1)
for(row in 1:nrow(output)) {
  output[row,1]=matriceDATI[row,2]
}
outputmax = max(output)
outputmin = min(output)

for(i in 1:5) {
  minnorm[1,i] = min(matriceDATI[,i])
  maxnorm[1,i] = max(matriceDATI[,i])
}
matriceDATI=matriceDATI[-1,]

input=matriceDATI

```



```

inputnorm=matrix(nrow = 2580, ncol = 5)
for(row in 1:nrow(inputnorm)) {
  for(col in 1:ncol(inputnorm)) {
    inputnorm[row,col]=((input[row,col]-minnorm[1,col])/(maxnorm[1,col]-minnorm[1,col]))
  }
}
outputnorm=matrix(nrow = 2580, ncol = 1)
for(row in 1:nrow(output)) {
  outputnorm[row,1]=((output[row,1]-outputmin)/(outputmax-outputmin))
}
df=data.frame(inputnorm,outputnorm)
df <- setNames(df, c("chiusural","apertural","massimol","minimol","aperturaA","chiusura"))

##CROSS VALIDATION (80% dei dati per il trainig e il 20% per la validazione)

#evaluationrange =2068:2580
#evaluationoutput = df[evaluationrange,6]
#evaluationinput = df[evaluationrange,1:5]
#evaluationtraining = df[-evaluationrange,]
#nn=neuralnet(chiusura ~ chiusural + apertural + massimol + minimol + aperturaA,
data=evaluationtraining, hidden=10, act.fct = "logistic", linear.output = FALSE, algorithm =
"rprop+")

#test = evaluationinput
#Predict=compute(nn,test)
#Predict$net.result
#mean(abs(Predict$net.result-evaluationoutput)^2)

##Implementazione rete neurale feed-forward
nn=neuralnet(chiusura ~ chiusural + apertural + massimol + minimol + aperturaA, data=df,
hidden=7, act.fct = "logistic", linear.output = FALSE, algorithm = "rprop+")
plot(nn)

test=data.frame("chiusural"=0.8696936, "apertural"=0.8949999, "massimol"=0.8886455,
"minimol"=0.8815675,"chiusuraA"=0.5810327)
Predict=compute(nn,test)
chiusura=Predict$net.result
chiusuradenorm= (chiusura*(outputmax-outputmin)+outputmin)

##Visualizzazione pesi rete neurale
pesi2=nn$weights
inw=pesi2[[1]]
inw2=inw[[1]]
pesi=as.matrix(inw2)
matricepesil=matrix(inw2,nrow = nrow(inw2),ncol=ncol(inw2))
colnames(matricepesil)=c("Nodo1","Nodo2","Nodo3","Nodo4","Nodo5","Nodo6","Nodo7")
rownames(matricepesil)=c("Bias","chiusural","apertural","massimol","minimol","aperturaA")
View(matricepesil)

```

```

ouw=pesi2[[1]]
ouw2=ouw[[2]]
pesi=as.matrix(ouw2)
matricepesiO=matrix(ouw2,nrow = nrow(ouw2),ncol=ncol(ouw2))
colnames(matricepesiO)=c("Pesi")
rownames(matricepesiO)=c("Bias","Nodo1","Nodo2","Nodo3","Nodo4","Nodo5","Nodo6","Nodo7")
View(matricepesiO)

library(NeuralNetTools)
plotnet(nn)

##Validazione
#nn=neuralnet(chiusura ~ chiusural + apertural + massimol + minimol + aperturaA,
data=evaluationtraining, hidden=7, act.fct = "logistic", linear.output = FALSE)

#test=evaluationinput
#Predict=compute(nn,test)
#chiusura=Predict$net.result
#chiusuradenorm= (chiusura*(outputmax-outputmin)+outputmin)

#plot(chiusuradenorm[1:516,], type = "l", ylab = "Stimati",
#main = "Grafico comparativo", xlab = "Tempo",
#col = "blue",ylim=c(18000,24500))
#par(new = TRUE)
#plot(output[1:516], type = "l", xaxt = "n", yaxt = "n",
#ylab = "", xlab = "", col = "red", lty = 2,ylim=c(18000,24500))
#axis(side = 4)
#mtext("Effettivi", side = 4, line = 3)
#legend("topleft", c("Stimati", "Effettivi"),
#col = c("blue", "red"), lty = c(1, 2))

##Fase di test
getwd()
setwd("C:/Users/Alic/Desktop")
getwd()
library(tseries)
datitest=read.csv2("Datitest.csv", sep=',',dec=',', header = TRUE)
library(readxl)
datitestA=read.csv2("DatitestA.csv", sep=',',dec=',', header = TRUE)
library(quantmod)
library(neuralnet)

##Caricamento e sistemazione dei dati
matrice1 = as.matrix(datitest)
matrice1A = as.matrix(datitestA)

todelete <- c()
for(row in 1:nrow(matrice1)){
  rowA <- 1

```

```

trovato <- 0
while (rowA <= nrow(matrice1A) && trovato==0) {
  if(matrice1[row,1]==matrice1A[rowA,1]){
    trovato <-1
  }
  rowA <- rowA+1
}
if(trovato==0){
  print(paste("cancello data",matrice1[row,1]))
  todelete <- append(todelete,row)
}
}
matrice1 <- matrice1[-todelete,]

todelete <- c()
for(rowA in 1:nrow(matrice1A)){
  row <- 1
  trovato <- 0
  while (row <= nrow(matrice1) && trovato==0) {
    if(matrice1A[rowA,1]==matrice1[row,1]){
      trovato <-1
    }
    row <- row+1
  }
  if(trovato==0){
    print(paste("cancello data",matrice1A[rowA,1]))
    todelete <- append(todelete,rowA)
  }
}
matrice1A <- matrice1A[-todelete,]

matricetest = matrix(nrow=nrow(matrice1),ncol=ncol(matrice1))
matricetestA = matrix(nrow=nrow(matrice1A),ncol=ncol(matrice1A))

for(row in 1:nrow(matricetest)) {
  for(col in 2:5) {
    matricetest[row,col]=as.numeric(sub(",", ". ", sub(".", "", matrice1[row,col] , fixed = TRUE) ,
fixed = TRUE))
  }
}
for(row in 1:nrow(matricetestA)) {
  for(col in 2:5) {
    matricetestA[row,col]=as.numeric(sub(",", ". ", sub(".", "", matrice1A[row,col] , fixed = TRUE)
, fixed = TRUE))
  }
}

##Normalizzazione dei dati
DATItest1=data.frame(matricetest[,-c(1,6,7)], matricetestA[,-c(1,2,4,5,6,7)])

for(i in 30:2){

```

```

  DATI[i,5]=DATI[i-1,5]
}
DATItest1 <- setNames(DATItest1,
c("chiusuralt","aperturalt","massimolt","minimolt","aperturaAt" ))

matriceDATIt=as.matrix(DATItest1)

outputest=matrix(nrow = 29, ncol = 1)
for(row in 1:nrow(outputest)) {
  outputest[row,1]=matriceDATIt[row,2]
}
for(i in 1:5) {
  minnorm[1,i] = min(matriceDATIt[,i])
  maxnorm[1,i] = max(matriceDATIt[,i])
}
matriceDATIt=matriceDATIt[-1,]

inputest=matriceDATIt

inputestnorm=matrix(nrow = 29, ncol = 5)
for(row in 1:nrow(inputestnorm)) {
  for(col in 1:ncol(inputestnorm)) {
    inputestnorm[row,col]=((inputest[row,col]-minnorm[1,col])/(maxnorm[1,col]-
minnorm[1,col]))
  }
}
outputestnorm=matrix(nrow = 29, ncol = 1)
for(row in 1:nrow(outputest)) {
  outputestnorm[row,1]=((outputest[row,1]-outputmin)/(outputmax-outputmin))
}

dftest=data.frame(inputestnorm)

dftest <- setNames(dfest, c("chiusuralt","aperturalt","massimolt","minimolt","aperturaAt"))

Predictest=compute(nn,dfest)
chiusuratest=Predictest$net.result
chiusuratestdenorm= (chiusuratest*(outputmax-outputmin)+outputmin)

errortest=sum((chiusuratest-outputestnorm)^2)

comparison=data.frame(chiusuratest, outputestnorm)
deviation=((outputestnorm-chiusuratest)/outputestnorm)
comparison1=data.frame(chiusuratest,outputestnorm,deviation)
accuracy=1-abs(mean(deviation))

plot(chiusuratestdenorm, type = "l", ylab = "Stimati",
      main = "Grafico comparativo", xlab = "Tempo",
      col = "orange",ylim=c(12000,24500))
par(new = TRUE)
plot(outputest, type = "l", xaxt = "n", yaxt = "n",
      ylab = "", xlab = "", col = "green", lty = 2,ylim=c(12000,24500))

```

```
axis(side = 4)
```

```
mtext("Effettivi", side = 4, line = 3)
```

```
legend("bottomleft", c("Stimati", "Effettivi"),
```

```
col = c("orange", "green"), lty = c(1, 2))
```


Bibliografia e sitografia

1. A. Turing, *Computing Machinery and Intelligence*, Mind 1950, pp. 433 - 460.
2. Abbas O., (2015), "Neural Network in Business Forecasting", *International Journal of Computer*, 1, pp. 114- 128.
3. Alkhatib K., Najadat H., Hmeidi I. e Shatnawi M., (2013), "Stock Price Prediction Using K – Nearest Neighbor (Knn) Algorithm", *International Journal of Business, Humanities and Technology*, 3, pp. 32 - 33.
4. Anonimo, (2018), *Ecco il nuovo ETF tematico sull'intelligenza artificiale di Amundi*, Funds people, 20/11/2018.
5. Anonimo, (2019), *Banche, ABI: Intelligenza artificiale per migliorare il servizio al cliente*, QuiFinanza, 25/11/2019.
6. Anonimo, (2019), *Intelligenza artificiale in banca: cinque scenari di successo*, Magazine Teorema, 21/03/2019.
7. Anonimo, (2019), *Sanità: telemedicina, big data, intelligenza artificiale e competenze digitali saranno i temi centrali del 2019: gli ultimi eventi sul tema*, Crescita digitale, 14/02/2019.
8. Aristotele, (2016), "*Analitici Primi*", Giorgio Colli (a cura di), *Organon*, Giulio Einaudi Editore, Torino, pp. 105 – 110.
9. Aristotele, *Analitici Primi in Organon*, Maurizio Migliori (a cura di), Bompiani, Milano 2016, pp. 375 - 377.
10. Autori di Wikipedia, (2020), *FTSE MIB*, Wikipedia. L'enciclopedia libera, 04/01/2020.
11. Autori di Wikipedia, (2020), *Hang Seng Index*, Wikipedia. L'enciclopedia libera, 08/01/2020.
12. Autori di Wikipedia, "Go (gioco)", Wikipedia, l'Enciclopedia libera, Tratto il 12 dicembre 2019 da [https://it.wikipedia.org/wiki/Go_\(gioco\)](https://it.wikipedia.org/wiki/Go_(gioco)).
13. B. Lenzi, Passato, (2012), "Presente e futuro dell'intelligenza artificiale", *Divenire*, 5, p. 235.
14. Barto A., Sutton R., (2016), *Reinforcement Learning: An Introduction.*, The MIT Press, Londra.

15. Bloomberg (2020), Informazioni sull'indice Hang Seng di Hong Kong, disponibile a <https://www.bloomberg.com/quote/HSI:IND>.
16. Boldrini N., (2019), *Cos'è l'intelligenza artificiale, perché tutti ne parlano e quali sono gli ambienti applicativi*, AI4 BUSINESS, 01/03/2019.
17. Borsa Italiana (2018), Per la gestione dell'indice FTSE MIB, disponibile a https://www.borsaitaliana.it/azioni/tutti-gli-indici/metodologiadicalcolo/ftse_pdf.htm.
18. Borsa Italiana (2019), *Per la gestione dell'indice FTSE MIB*, Luglio 2019, disponibile a https://research.ftserussell.com/products/downloads/FTSE_MIB_Index_Italian_.pdf.
19. Borsa Italiana disponibile a <https://www.borsaitaliana.it/borsa/indici/indici-in-continua/dettaglio.html?indexCode=FTSEMIB&lang=it>.
20. British Academy for the humanities and social sciences (2018), *The impact of artificial intelligence on work An evidence synthesis on implications for individuals, communities, and societies*, disponibile a <https://royalsociety.org/-/media/policy/projects/ai-and-work/summary-the-impact-of-AI-on-work.PDF>.
21. Brownlee J., (2019), *A Gentle Introduction to Expectation-Maximization (EM Algorithm)*, Machine Learning Mastery, 01/11/2019.
22. Buchanan B. G., Lederberg J., (1971), *Il programma Dendral per spiegare i dati empirici*, Dipartimento di scienze del computer, Università di Stanford, no. CS203.
23. Budek K., Osinski B., (2018), *What is reinforcement learning? The complete guide*, Big Data Science, 05/07/2018.
24. Buffolo M., (2019), *Com'è stato il 2019 per la Borsa di Milano*, AGI, 28/12/2019.
25. Bughin J., Seong J., Manyika J., Chui M., Joshi R., (2018), *Notes from the AI frontier. Modeling the impact of AI on the world economy*, McKinsey Global Institute, 09/2018.
26. Business Dictionary (2018), Marketing, disponibile a <http://www.businessdictionary.com/definition/marketing.html>. 2018.
27. Campbell M., Hoane J. A., Hsu F. H., (2002), "Deep Blue", *Elsevier*, 134, pp. 57 - 83.
28. Cangini N., *Quantum Supervised Learning: algoritmi e implementazione*, tesi di laurea, Università di Bologna, a. a. 2017-2018, relatore Sartori C.

29. Cesaro E., *Q – Learning: un algoritmo ad apprendimento per rinforzo applicato al trading finanziario*, tesi di laurea, Università di Venezia, a. a. 2016-2017, relatore Corazza M.
30. Chaigusin S., Chirathamjaree C., Clayden J., (2008), “The Use of Neural Networks in the Prediction of the Stock Exchange of Thailand (SET) index”, *IEEE computer society*, 10, pp. 670 – 673.
31. Chen A., Leung M., Daouk H., (2003), *Application of Neural Networks to an emerging financial market: forecasting and trading the Taiwan Stock Index*, Pergamon, Working Paper No. 30.
32. Ciaburro G., (2017), *Neural Networks with R*, Packt Publishing, Birmingham.
33. Ciotola M., (2018), Cinque modi in cui l’intelligenza artificiale sta rivoluzionando la finanza, *Money*, 07/04/2018.
34. Cochrane J., (2014), *Eugene F. Fama efficient markets and the Nobel Prize*, Chicago Booth Review, 20/05/2014.
35. Confindustria Digitale (2019), *Intelligenza Artificiale per le PMI*, disponibile a <http://preparatialfuturo.confindustria.it/wp-content/uploads/2019/02/IA-per-PMI.pdf>.
36. Consoft Sistemi, (2019), *L’intelligenza artificiale al servizio dell’uomo*, Digital Transformation per lo sviluppo sostenibile, 23/05/2019.
37. Cordeschi R., Tamburrini G., (2001), *L’intelligenza artificiale: la storia e le idee*, Carocci, Roma.
38. Cotronei M., *Introduzione alla Ricerca Operativa*, Facoltà di Ingegneria dell’Università degli Studi “Mediterranea” di Reggio Calabria.
39. Crescenzo G., (2007), *Reti Neurali Artificiali: Teoria ed Applicazioni*, Dipartimento di Scienze Economiche, Matematiche e Statistica. Centro di Ricerca Interdipartimentale Bioagromed dell’Università di Foggia, Quaderno di ricerca No. 28.
40. Daniele A., (2019), *La scuola del futuro: come l’intelligenza artificiale rivoluzionerà l’istruzione scolastica*, *Le Macchine volanti*, 24/06/2019.
41. Davenport T., Kalakota R., (2019), “The potential for artificial intelligence in healthcare”, *Future Healthcare Journal*, 6, pp. 94 – 98.

42. Davis M., (2000), *The Universal Computer. The Road from Leibniz to Turing*, Adelphi Edizioni, Milano.
43. DEAGOSTINI Editore, *La meccanica statistica*, 2019.
44. Della Porta M., (2019), *Intelligenza artificiale, l'inarrestabile corsa nel settore sanitario*, Istituto per la competitività, 12/11/2019.
45. Deng H., (2018), *An Introduction to Random Forest*, Towards Data Science, 07/12/2018.
46. Di Pietro S., (2019), *Introduzione alla cluster analysis*, Italian Association for Machine Learning, 05/02/2019.
47. Dopico J. R. R., De la Calle J. D. e Sierra A. P., (2009), *Encyclopedia of Artificial Intelligence*. Information sci, New York.
48. Dreyfus H. L., (2002), *Phenomenology and the Cognitive Sciences*, Dipartimento di Filosofia edell'Università della California a Berkeley, No. 1.
49. Dzikevicius A., Stabuzyte N., (2012), "Forecasting OMX Vilnius stock market – A neural network approach", *Business: theory and practice*, 4, pp. 321-332.
50. Elliot, Robert J.; Mamon Rogemar S., (2007), *Hidden Markov models in finance.*, Springer, New York.
51. Faria E. L., Albuquerque M., Gonzalez J. e Cavalcante J., (2009), "Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods", *International Journal Expert Systems with Applications*.
52. Farsagli S., (2018), *Precision Farming: do sensors, data and AI improve production?*, Ingenium: the web Magazine of Engineering and Tech Economy 2030 on the Digital Transformation, 28/11/2018.
53. Fernández-Rodríguez F., González – Martel C., Sosvilla – Riviero S., (2000), On the profitability of technical trading rules based on artificial neural networks: Evidence from the Madrd stock market, Universidad de Las Palmas de Gran Canaria, FEDEA e Universidad Complutense de Madrid, Working Paper No. 7.
54. Financial Stability Board (2017), *Fintech credit Market structure: business models and financial stability implications*, disponibile a <http://www.fsb.org/2017/05/fintech-credit-market-structure-business-models-and-financial-stability-implications/>.
55. Floreano D., Mattiussu C., (2002), *Manuale sulle reti neurali*, il Mulino, Bologna.

56. Gandhi R., (2018), *Support Vector Machine – Introduction to Machine Learning Algorithms*, Towards Data Science, 07/06/2018.
57. Geisel A., (2018), “Current and Future Impact of Artificial Intelligence On Business”, *Journal of Scientific & Technology Research*, 7, pp. 116 – 121.
58. Glossario Finanziario – Borsa Italiana, <https://www.borsaitaliana.it>.
59. Govoni L., (2019), *Algoritmo k – means: cose ‘è e come funziona?*, Business e Tecnologia, 02/07/2018.
60. Goy A., Torre I., (2009), “Agenti artificiali e agenti intelligenti: paradigmi, applicazioni e prospettive”, *LEXIA*, 03/04, pp. 297 - 313.
61. Greaves J., Understanding RL: The Bellman Equations, disponibile a <https://joshgreaves.com/reinforcement-learning/understanding-rl-the-bellman-equations/>.
62. Grigoryan H., (2015), “Stock Market Prediction using Artificial Neural Networks. Case study of TAL1T, Nasdaq OMX Baltic Stock”, *Database System Journal*, 2, pp. 14 – 23.
63. Guerini M., (2017), *Intelligenza Artificiale, forte e debole, tra mito e realtà*, Corriere della sera Blog, 10/11/2017.
64. Haenlein M., Kaplan A., *A brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence*, University of California.
65. Hagan M., Demuth H., Beale M., De Jesus O., (2014), *Neural Network Design*, Martin Hagan, Kansas.
66. Harrison O., (2018), *Machine Learning Basics with the K – Nearest Neighbors Algorithm*, Towards Data Science, 10/09/2018.
67. Hatwell J., (2016), *Artificial Neural Networks in R*, RPubS, 10/2016.
68. Imandoust S. B., Bolandraftar M., (2014), “Forecasting the direction of stock market index movement using three data mining techniques: the case of Tehran Stock Exchange”, *International Journal of Engineering and Applications*, 4, p. 106-117.
69. INKWOOD RESEARCH (2019), Global Artificial Intelligence market forecast 2019 – 2027, Market Research Report, disponibile a <https://www.inkwoodresearch.com/reports/global-artificial-intelligence-market/#report-summary>.

70. International Labour Organization (2018), *The economics of artificial intelligence: Implications for the future of work*, disponibile a https://www.ilo.org/wcmsp5/groups/public/---dgreports/cabinet/documents/publication/wcms_647306.pdf.
71. J.J., (2016), *MAE and RMSE – Which Metric is Better?*, Medium, 23/03/2016.
72. Jain S., Kain M., (2018), “Prediction for Stock Marketing Using Machine Learning”, *International Journal on Recent and Innovation Trends in Computing and Communication*, 6, pp. 131-135.
73. James G., Witten D., Hastie T., Tibshinari R., (2013), *An Introduction to Statistical Learning with Applications in R*, Springer, Londra.
74. Jasic T., Douglas W., (2004), “The profitability of daily stock market indices trades based on neural network predictions: case study for the S&P 500, the DAX, the TOPIX and the FTSE in the period 1965–1999”, *Applied Financial Economics*, 14, pp. 285 – 297.
75. Kaur K., (2019), *The Real World Potential and Limitations of Artificial Intelligence*, Hackernoon, 16/07/2019.
76. KPMG (2018), *Digital Banking. L’evoluzione delle aspettative dei clienti tra rivoluzione digitale, sfide regolamentari e nuovi competitor*, disponibile a www.kpmg.com/it
77. Kulkarni A., (2019), *AI in Education: Where is It Now and What is the Future?*, Lexalytics, 06/09/2019.
78. L. Carlucci e M. Aiello, (2004), “Intelligenza Artificiale: i primi 50 anni”, *Mondo Digitale*, 2, p. 10.
79. Lantz B., (2013), *Machine Learning with R*, Packt Publishing, Birmingham.
80. Lateef Z., (2019), *AI Applications: Top 10 Real World Artificial Intelligence Applications*, Edureka, 22/05/2019.
81. Le J., (2018), *Decision Trees in R*, DataCamp, 19/06/2018.
82. Lozinskaia A., Zhemchuzhnikov V., (2017), “Micex Index forecasting: the predictive power of neural network modelling and support vector machine”, *Perm University Herald*, 1, pp. 49 – 60.
83. Lutins E., (2017), *DBSCAN: What is it? When to use it? How to use it.*, Medium, 06/09/2017.

84. Magaji A., Adeboye K., (2014), "An intense Nigerian Stock Exchange market prediction using logistic with back – propagation ANN model", *Science World Journal*, 2, pp. 8 – 13.
85. Mahajan S., Deshmukh K., Thite P., Bhavana S., (2016), *Stock Market Prediction and Analysis Using Naïve Bayes*, International Journal on Recent and Innovation Trends in Computing and Communication, Working Paper No.4.
86. Marxiaoli S., Selvam M., Lingaraja K., Vasanth V., (2017), "Forecasting the Stock Index Movements of India: Application of Neural Networks", *International Journal of Soft Computing*, 2, pp. 120-131.
87. Matsumoto D., *The Cambridge Dictionary of Psicology*, Cambridge 2009, p. 555.
88. Mignucci M., *La teoria aristotelica della scienza*, Sansoni, Firenze 1965, pp. 110 - 111.
89. Ministero dello Sviluppo Economico (2019), *Proposte per una Strategia Italiana per l'Intelligenza Artificiale*, Roma, Maggio 2019.
90. Mitchell T., (2006), *The Discipline of Machine Learnig*, School of Computer Science Carnegie Mellon University, 07/2006.
91. Mohammed M., Khan B. M., e Bashier M.B.E., (2017), *Machine Learning. Algorithms and Applications*, CRC Press, Boca Raton.
92. Nasteski V., (2018), *An Overview of the supervised machine learning methods*, Research Gate, 11/12/2018.
93. Niaki S., Hoseinzade S., (2013), "Forecasting S&P 500 index using artificial neural networks and design of experiments", *Journal of Industrial Engineering International*, 9, pp. 1 - 9.
94. Nicita G., (2019), *Intelligenza Artificiale e apprendimento: alte percentuali negli USA*, DIGITAL EDUCATION LAB, 30/08/2018.
95. Nilsson N. J., (1998), *Introduction to Machine Learning. An early draft of a proposed textbook*, Standford University, Standford.
96. North R., (2019), *Artificial Intelligence – A Boon To The Banking Industry*, Enterprise Edges.
97. O'Connor N., Madden M., (2006), "A neural network approach to predict stock exchange movements using external factors", *Knowledge – Based Systems*, 19, pp. 371 – 378.

98. Pan H., Tilakaratne C., Yearwood J., (2005), "Predicting Australian Stock Market Index Using Neural Networks Exploring Dynamical Swings and Intermarket Influences", *Journal of Research and Practice in Information Technology*, 1, pp. 43 – 55.
99. Patel S., (2017), *Support Vector Machine – Theory*, Medium, 03/05/2017.
100. Phaneuf A., (2019), *The disruptive digital transformation of banking services in 2019*, Business Insider, 12/09/2019.
101. Pichai S., (2019), *Advantages and Disadvantages of Artificial Intelligence*, World Economic Forum, 21/12/2019.
102. Pociati P., (2019), *Luci e ombre dell'intelligenza artificiale nella diagnostica*, Associazione Italiana per l'Intelligenza artificiale, 21/10/2019.
103. Possati L., (2017), *Osservazioni sul numero e sulla computabilità*, Hal, 2017.
104. Preminger A., Franck R., (2007), "Forecasting exchange rates: A robust regression approach", *International Journal of Forecasting*, 23, pp. 71 – 84.
105. Proverbio D., (2018), *Intelligenza artificiale e robotica: le sfide del nuovo millennio*, Rivista di affari europei, 11/01/2018.
106. Qiu M., Song Y., (2016), *Predicting the Direction of Stock Market Index Movement Using an Optimized Artificial Neural Network Model*, PLOS ONE, 19/05/2016.
107. Rabiner L. R. e Juang B. H., (1986), "An Introduction to Hidden Markov Models", *ASSP MAGAZINE*, 86, pp. 5 - 6.
108. Raschka S., (2016), *Machine Learning con Python. Costruire algoritmi per generare conoscenza*, Apogeo, Milano.
109. Raso H., Demirci M., (2019), "Predicting the Turkish Stock Market BIST 30 Index using Deep Learning", *International Journal of Engineering Research and Development*, 1, pp. 253 – 265.
110. Rathi A., (2019), *The Impact of Artificial Intelligence. A critical review of opportunities and risks of AI adoption*, Towards Data Science, 23/09/2019.
111. Redazione Osservatori, (2019), *AI Journey: un modello per fare Artificial Intelligence nelle imprese*, Osservatori. net digital innovation, 26/03/2019.
112. Redazione Osservatori, (2019), *Le applicazioni dell'Intelligenza Artificiale e la loro diffusione*, Blog Osservatori Digital Innovation, 26/04/2019.

113. Redazione Soldionline, (2020), *Il 2020 sarà un anno tutt'altro che tranquillo per i mercati finanziari*, Soldi Online, 10/01/2020.
114. Redazione, (2018), *Intelligenza artificiale, come ha modificato e sta modificando il mondo della finanza?*, Conto Facto, 28/12/2018.
115. Redazione, (2020), *Robotica. Storia, evoluzioni e ultime novità*, Impactscool. The future is opensource.
116. Research Gate (2018), *Data Clustering Spaziale basati su densità: indagine esplorativa delle tecniche di Data Clustering spaziale e metriche di distanza*, disponibile a <https://www.researchgate.net/publication/332291727>.
117. Ripley B., e Venables W., (2016), *Package "nnet"*, CRAN, 02/02/2016.
118. Russell S. e Norving P., (2010), *Intelligenza artificiale. Un approccio moderno*, Pearson, Milano.
119. Schmelzer R., (2019), *AI Applications in Education*, Forbes, 12/11/2019.
120. Schrage M., (2020), *Data Mining. What it is and why it matters*, SAS Insights.
121. Seif G., (2018), *The 5 Clustering Algorithms Data Scientists Need to Know*, Towards Data Science, 05/02/2018.
122. Seif G., (2019), *Una facile introduzione all'apprendimento senza supervisione con 4 tecniche di base*, Towards Data Science, 19/08/2019.
123. Sentance R., (2019), *15 examples of artificial intelligence in marketing*, Econsultancy, 28/02/2019.
124. Soluciones D., (2018), *How different sectors are using AI*, Medium, 29/10/2018.
125. Stefan F., Frauke G., Marvin N., Sulig M e Mueller S., (2019), *Package "neuralnet"*, CRAN, 07/02/2019.
126. Stefanelli M., *Apprendimento automatico nei giochi di strategia*, tesi di laurea, Università di Bologna, a. a. 2013-2014, relatore Roli A.
127. Stiehler A., (2018), *AI's coming of age. The progress into the AGI phase and the beginning of the true autonomy*, UBS, 02/03/2018.
128. Subha M. V., Nambi S.T., (2012), *Classification of Stock Index movement using k-Nearest Neighbours (k-NN) algorithm*, Wseas Transaction on Information Science and applications, 09/09/2012.
129. Suman F., (2019), *Dove sta andando oggi l'intelligenza artificiale?*, Il Bo live Università di Padova, 11/03/2019.

130. Szczepański M., (2019), Economic impacts of artificial intelligence (AI), disponibile su [http://www.europarl.europa.eu/RegData/etudes/BRIE/2019/637967/EPRS_BRI\(2019\)637967_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/BRIE/2019/637967/EPRS_BRI(2019)637967_EN.pdf).
131. Team Dataflair, (2019), *Pros and Cons of Artificial Intelligence – A Threat or a Blessing?*, Data Flair, 25/09/2019.
132. Tecchia M., (2018), *Reti neurali artificiali: cosa sono e come funzionano*, Digitalic, 09/06/2018.
133. Tessa M., (2019), *Intelligenza artificiale regina dei listini, gestisce metà trading*, Wall Street Italia, 03/06/2019.
134. Walczak S., (2001), "An empirical Analysis of Data Requirements for Financial Forecasting With Neural Networks", *Journal of Management Information Systems*, 4, pp. 203 - 222.
135. Wang J., (2011), *The Turing Test, Searle's chinese room, and E.T.'s chinese room*, School of Business, Richard Stockton College of New Jersey, No. 4628.
136. Weizenbaum, (1966), *Joseph Weizenbaum's ELIZA: Communications of the ACM*, Cambridge.
137. Wingate J., (2018), *Apriori Machine Learning Algorithm*, R – algo, 01/02/2018.
138. Wright R., (1996), "Can machines think?", *Time Magazine*, Volume. 147, p. 50.
139. Yahuza B., Haruna C., (2011), "Utilizing artificial neural network for prediction the nigerian stock market price index", *Computer Science and Telecommunications*, 30, pp. 68 – 77.
140. Yao J., Tan C. L., Poh H-L. (1999), *Neural Networks for Technical Analysis: a study on KLCI*, International Journal of Theoretical and Applied Finance.
141. Zaidi M., Amirat A., (2016), "Forecasting stock market trends by logistic regression and neural networks", *International Journal of Economics, Commerce and Management*, 4, pp. 221-222.
142. Zamagna R., (2018), *The Future of trading belong to Artificial Intelligence*, Medium, 15/11/2018.

Ringraziamenti

Ed eccomi, al termine della mia carriera universitaria. Questa tesi rappresenta la fine di un percorso molto importante e impegnativo della mia vita che mi ha portato ad affrontare molte sfide con conseguenti gioie e delusioni che mi hanno spinto ad impegnarmi maggiormente e cercare di superare i miei limiti.

Nel 2013 ho iniziato il mio percorso universitario. Decisi di intraprenderlo per mettermi alla prova ed affrontare la materia che, fino a quel momento, aveva costituito una mia grande difficoltà. Oggi posso dire di essere riuscita ad affrontare le mie difficoltà e vincere le sfide che mi ero posta e questo con molto impegno e determinazione: nulla è facile, ma niente è impossibile.

Ci tengo innanzitutto a ringraziare il Professor Pizzi Claudio per tutto l'aiuto e l'assistenza fornita. La pazienza e la dedizione, che lo contraddistinguono, nel rispondere ad ogni mio dubbio e perplessità, sono state di fondamentale importanza nella stesura di tale elaborato.

Un ringraziamento importante va ai miei genitori, i quali hanno vissuto con me, giorno dopo giorno, l'ansia, le preoccupazioni e le gioie che hanno caratterizzato il mio percorso di studi. Oltre ad avermi sostenuto economicamente, mi avete sostenuto moralmente e spero che questo traguardo possa, almeno in parte, ricompensare i vostri sacrifici rendendovi fieri di me.

Grazie a mia sorella Alice, uno dei pilastri della mia vita. Io e te abbiamo condiviso tutto, fin dalla nascita, gioie, preoccupazioni, ansie e dolori. Tu con la tua intelligenza e tranquillità sei sempre riuscita a darmi degli ottimi consigli evitando che commettessi degli errori nei momenti più difficili.

Grazie a te Marco, l'altro pilastro della mia vita. Tu che mi hai sempre sostenuto in modo incondizionato, spronandomi sempre e in ogni modo a cercare di superare le mie paure e i miei limiti. Senza di te probabilmente non avrei mai raggiunto questo importante

traguardo. Sei entrato nella mia vita quasi per caso, e ora sei una certezza, una costante di fondamentale importanza. Ogni giorno mi ritengo fortunata ad averti nella mia vita, non avrei mai potuto desiderare qualcosa di meglio.

Grazie ai miei nonni, materni e paterni. Anche voi mi avete sempre sostenuto e creduto in me spronandomi a fare sempre del mio meglio, soprattutto quando non ottenevo i risultati che desideravo. Grazie per le preghiere e le candele che mi dedicate.

Grazie ai miei zii e cugini che hanno sempre sopportato le mie confessioni serali disperate. Tra qualche risata e crisi di nervi, siete sempre riusciti ad incoraggiarmi nel percorso universitario e nelle mie scelte.

E infine, ma non per questo meno importante, un ringraziamento a te Matteo. Abbiamo condiviso l'intero percorso universitario, anche se la vera amicizia è nata solamente durante la Magistrale. Ho scoperto in te un vero amico e volevo ringraziarti per avermi sempre ascoltato e condiviso le mie preoccupazioni e relative crisi di nervi, grazie di avermi sostenuto, incoraggiato e aver sempre accettato un confronto con me quando si presentavano dubbi e incertezze.