



UNIVERSITÀ DEGLI STUDI DI PALERMO
SCUOLA POLITECNICA

Ingegneria informatica e delle telecomunicazioni
Dipartimento dell'Innovazione Industriale e Digitale

STRUMENTI E METODI PER IL TRADING ALGORITMICO

TESI DI LAUREA DI
GAETANO DI GRAZIA
0646355

RELATORE
PROF. ROBERTO PIRRONE

ANNO ACCADEMICO 2020-2021

TRIENNALE



Guardo ogni giorno il mare, dalla mia finestra, blu.

Onde, ora piccole, ora grandi, che si infrangono sugli scogli; mai stanche di riprovare.

Mai stanche di venire alla luce per generare altre onde e schiantarsi, alla fine, per lasciar posto a chi verrà.

Questa è stata la mia vita, ora che ci penso.

Un provare e riprovare di continuo e, ora, che son vecchio, mi preparo all'ultimo tentativo, l'ultimo infrangersi prima d'essere solo un ricordo; fervido, vivo, tangibile, ma ricordo.

Ho riposto, per tutta la durata di questa esistenza, pezzi di me nei miei oggetti.

Ho riposto speranze e rabbia, gioie e dolori.

Profumi.

Ho cercato, con vera tenacia, di imprimere un messaggio in ogni oggetto che toccavo.

Ho immaginato, come fosse un puzzle, che un giorno qualcuno avrebbe rimesso insieme tutti questi piccoli pezzi, permettendomi di rivelarmi: fervido, vivo, tangibile, ma ricordo.

Non ho pensieri, non ho ricordi. Sai perché?

Ognuno di essi è accuratamente conservato. Una foto, riporta alla luce degli eventi, il profumo di un fiore altri eventi, una carezza, altri ancora. I miei ricordi non son miei ma di chi, visibile o no, li ha condivisi con me.

*Alla mia famiglia che ha creduto in me,
a Clara che ha fatto sì che io credessi in me,
ai miei amici più cari che mi hanno supportato,
a me stesso, che non ho mai mollato.*

Indice

Sommario	3
1 Cenni introduttivi	4
1.1 Trading	4
1.2 Analisi quantitativa	5
1.2.1 Analisi fondamentale vs analisi tecnica	5
1.3 Finanza quantitativa	6
1.3.1 Momentum	6
1.3.2 Efficienza	7
1.3.3 Volatilità	7
1.3.4 Utilità totale e marginale	8
1.3.5 Mercati e strumenti prevedibili	9
2 Strumenti esistenti	10
2.1 Getting data	10
2.1.1 Python: Pandas for finance	11
2.2 Notizie	12
2.2.1 Feed RSS e telegram	12
2.3 Piattaforme	13
3 Analisi di serie temporali	15
3.1 Correlazione	16
3.1.1 Autocorrelazione	16
3.2 Pattern	16
3.3 Stazionarietà	17
3.4 Processi di Wiener (Brownian motion): Markov	17
3.4.1 Random walk	18
3.4.2 Black-Scholes	18
3.5 Modelli autoregressivi	19
3.5.1 Modello ARMA(p, q)	19
3.5.2 Modello ARIMA(p, d, q)	20
3.6 Misuratori di errore	21
3.6.1 MAE, MSE, RMSE, SSE e Goodness-of-fit	21
3.6.2 Errori percentuali e scalari: MAPE e MASE	22
3.7 Exponential smoothing	23
3.7.1 Single exponential smoothing	23
3.7.2 Double exponential smoothing: Holt's method	24
3.7.3 Triple exponential smoothing: Holt-Winters	24

4	Tecniche avanzate di forecasting	26
4.1	Intelligenza artificiale	26
4.2	NLP	27
4.2.1	FOMO, FUD e panic selling	27
4.2.2	NLP e NLFF	27
4.2.3	Algoritmi per l’NLP	28
4.3	Deep learning	31
5	Conclusione	32
	Bibliografia	33

Sommario

Questa tesi nasce come incontro tra due discipline diverse eppure molto simili: l'ingegneria e l'economia.

Sebbene possa sembrare più un argomento legato alla disciplina economica, oggi, il modo di fare trading, di investire, di speculare è stato quasi completamente stravolto dalla tecnologia e dai servizi che questa mette a disposizione; se, dunque, gran parte dei concetti che troviamo in questo documento sono di natura micro e/o macro economica, questo è "solo" un trampolino di lancio per un lavoro che, a Dio piacendo, sarà rivisto e approfondito magari anche in modo sperimentale alla fine del percorso di laurea magistrale.

L'elaborato ha come obiettivo dapprima la disamina di alcune tra le più comuni piattaforme di trading online e in secondo luogo lo studio di alcune delle più note tecniche (algoritmi) matematiche usate per ottenere una previsione più o meno precisa dei prezzi di mercato; più nello specifico si indagheranno i metodi attraverso i quali vengono studiate le serie storiche, dette time series.

Nel primo capitolo si introduce il concetto di trading e le principali definizioni in ambito economico riguardanti l'analisi dei mercati; mentre nel secondo capitolo si illustrano alcune delle piattaforme più diffuse e, inoltre, vengono proposti due esempi di acquisizione dati attraverso il linguaggio di programmazione Python. Nel terzo capitolo, invece, si affronta quello che è il vero e proprio argomento cardine dell'elaborato in questione: lo studio delle serie temporali.

In questa sezione vengono proposti, oltre alle caratteristiche che una serie storica può presentare, i principali metodi usati per lo studio delle serie storiche, tra i quali i modelli autoregressivi e quelli di tipo exponential smoothing; altrettanto interessante è il paragrafo in cui si affrontano i principali misuratori di errore di forecasting dei valori. Infine, nel quarto e ultimo capitolo, viene illustrata una tecnica innovativa di forecasting delle serie storiche: l'NLP.

Infatti, oltre al classico studio prettamente matematico delle serie storiche, mediante l'utilizzo del deep learning è possibile analizzare le serie storiche anche dal punto di vista del natural language processing cercando di capire qual è la relazione, se esiste, tra le notizie, di mercato e non, e l'oscillazione del prezzo di un asset.

Capitolo 1

Cenni introduttivi

1.1 Trading

Partendo dal presupposto che nella lingua inglese il verbo "to trade" si traduce con commerciare possiamo definire il trading come segue: "Il trading è l'attività di base di tutti gli investitori o servizi d'investimento e non è altro che l'acquisto, la vendita o la permuta di attività.

Nei mercati finanziari, le persone negoziano titoli quali azioni, valute, materie prime e derivati.

Le azioni sono negoziate in Borsa, mentre le materie prime e i titoli di capitale sono acquistati e venduti nella sala delle contrattazioni.

Lo scopo generale consiste nel cercare di realizzare un utile acquistando a un prezzo più basso e vendendo a un prezzo più alto, solitamente entro un periodo di tempo relativamente breve.

Un trader può essere sia un investitore privato sia un'istituzione globale. Il trading può avvenire direttamente o attraverso un broker e può essere effettuato di persona, per telefono o tramite una piattaforma di negoziazione online."

Fonte: Capital.com

Di conseguenza il trading esiste da che è stata introdotta la moneta mentre per il trading finanziario si è dovuto aspettare il 1531 quando la borsa ebbe origine in Belgio; le cronache dicono che a quel tempo un gruppo di commercianti/banchieri fiamminghi cominciarono a riunirsi periodicamente nel palazzo dei Van der Bourse, che era una famiglia di banchieri, per scambiarsi titoli di credito e stipulare delle compravendite.

Negli ultimi decenni con l'avvento di internet, invece, quello che si è sviluppato sempre di più è il trading *online*, che ha dato modo ad un vasta parte della popolazione di accedere a strumenti che prima erano proibitivi e riservati solo agli addetti ai lavori.

Dunque con questa apertura e con l'avvento di tecnologie come: big-data, python e machine learning sempre più tra appassionati di vari campi si sono posti l'obiettivo di individuare sempre più dei trend, delle forme ricorrenti, per prevedere l'andamento dei mercati, questo è detto: **forecasting**.

Per quanto concerne la vigilanza sono Autorità di vigilanza sui mercati regolamentati di strumenti finanziari la Consob, la Banca d'Italia, il Ministero dell'economia e delle finanze e, inoltre, settorialmente l'ISVAP e la COVIP.

Riguardo alle competenze delle Autorità occorre distinguere tra vigilanza sugli intermediari e vigilanza sui mercati.

La vigilanza sugli intermediari mira a garantire la trasparenza e la correttezza dei comportamenti e la sana e prudente gestione dei soggetti abilitati, avendo riguardo alla tutela degli investitori e alla stabilità, alla competitività e al buon funzionamento del

sistema finanziario.

La competenza della Banca d'Italia in questo campo riguarda il contenimento del rischio e la stabilità patrimoniale degli intermediari.

La Consob invece opera nel senso di garantire la trasparenza e la correttezza dei comportamenti degli intermediari.

Questi obiettivi vengono perseguiti attraverso un'articolazione della funzione di vigilanza da parte della Consob e della Banca d'Italia, nelle forme: a) della vigilanza regolamentare; b) dei poteri di intervento sui soggetti abilitati; c) della vigilanza informativa; d) della vigilanza ispettiva; e) della vigilanza sui gruppi.

La vigilanza sui mercati, riguarda la definizione dell'organizzazione e della gestione di mercati regolamentati di strumenti finanziari, intendendo l'attività che vi si esercita come attività di impresa esercitata da società per azioni (di costituzione obbligatoria nel caso della privatizzazione dei mercati già esistenti all'atto dell'entrata in vigore della riforma).

La privatizzazione dei mercati regolamentati presuppone che l'iniziativa per l'istituzione, l'organizzazione e il funzionamento dei mercati sia rimessa all'iniziativa privata, sia pure soggetta ad autorizzazione della Consob.

Sempre la Consob disciplina con regolamento i profili organizzativi essenziali delle società di gestione, mentre il Ministero dell'economia e delle finanze si occupa di delineare i requisiti di professionalità e di onorabilità dei soggetti che amministrano e controllano queste società, oltre che di coloro che partecipano in misura rilevante al loro capitale.

Le Autorità vigilano sui passaggi di partecipazioni nel capitale delle società di gestione, comunque avvengano.

Con l'intervenuta riforma sono cessate le funzioni del Consiglio di borsa e sono state soppresse le sedi locali delle borse valori.

L'attività di gestione e organizzazione dei mercati è ora di competenza delle società di gestione dei mercati regolamentati secondo quanto previsto dal TUF.

Fonte: Bankpedia

1.2 Analisi quantitativa

1.2.1 Analisi fondamentale vs analisi tecnica

Nel mondo dei mercati azionari si sono sviluppate due diverse discipline contrapposte, analisi fondamentale e tecnica, che sono di supporto ai trader per analizzare e prevedere l'andamento di un determinato titolo.

Nella fattispecie l'analisi fondamentale riguarda *strettamente* l'azienda di cui si sta analizzando l'azione ovvero il suo bilancio: più l'azienda va bene, più aumenta il prezzo del titolo.

Però, questo tipo di analisi si porta dietro dei problemi, ossia: spesso i bilanci arrivano in ritardo rispetto a quando può verificarsi una discesa o un'impennata del prezzo, dunque sono già dati "vecchi" e che questa non tiene conto, per la valutazione, dell'enorme effetto che le speculazioni e rumors vari possano avere sul titolo; per fare un esempio se anche Intel dovesse essere in discesa sul mercato, un'immissione di liquidità di qualche milione di dollari potrebbe sensibilmente arrestare la decrescita e/o innescare un trend crescente.

Dunque per ottenere informazioni di questo tipo si è sviluppata un'apposita disciplina finanziaria, detta analisi tecnica, che va ad integrare l'analisi fondamentale di borsa.

In generale in questa disciplina si va alla ricerca di pattern, ovvero segnali anticipatori, che ci permettano di riconoscere una situazione storico/psicologica già manifestatasi in passato e che stia riproponendosi.

Uno dei principali strumenti utilizzati dall'analisi tecnica tradizionale è l'analisi del grafico del corso di un titolo; per arrivare ad una previsione bisogna analizzare quindi prezzi, tempo e volumi di mercato.

Una nota importante è che i prezzi minimi e massimi storici sono considerati supporti e resistenze psicologiche dell'andamento futuro del prezzo e che possono quindi dirci molto sulle nostre previsioni.

Il problema cruciale di queste analisi però è che, come detto precedentemente, il mercato si basa molto sulla psicologia degli attori (trader esperti, trader improvvisati, aziende, etc...) coinvolti: paura, dubbi, entusiasmo che influenzano in modo significativo l'andamento dei prezzi e che creando una sorta di rumore nel grafico, complicano la rilevazione di pattern.

1.3 Finanza quantitativa

La finanza quantitativa è l'applicazione della matematica, in particolare della teoria della probabilità - ai mercati finanziari.

Viene utilizzata, di solito, per concentrarsi sui contratti negoziati più frequentemente ovvero azioni e obbligazioni piuttosto che immobiliare o polizze sulla vita.

La base della finanza quantitativa è un'osservazione empirica di prezzi, tassi di cambio e tassi di interesse piuttosto che uno studio della teoria economica; il suo scopo è quello di arrivare a rispondere a domande chiave come: "Quanto vale un contratto?".

1.3.1 Momentum

Il momentum è un oscillatore e indica la forza del mercato misurando il tasso di variazione dei prezzi rispetto ai loro livelli effettivi. Questo concetto è fondamentale per tutta l'analisi tecnica basata sugli oscillatori.

Il calcolo viene effettuato sui dati del passato, come per tutti gli oscillatori, registrando le continue variazioni di prezzo su intervalli di tempo prefissati.

Per costruire la linea del momentum a cinque giorni occorre sottrarre all'ultimo prezzo segnato quello di chiusura di cinque sedute precedenti; la formula quindi sarà:

$$M = V - V_x$$

dove V è l'ultimo prezzo di chiusura e V_x è il prezzo di cinque giorni prima; è possibile trovare esempi in cui i giorni presi in considerazione siano 10 o 14.

La formulazione di questo oscillatore non è così facile quando il mercato passa di giorno in giorno o addirittura di ora ad oscillare tra prezzi alti e bassi.

A questo punto se il prezzo è in rialzo, quindi questo numero è positivo, e anche se il prezzo domani dovesse scendere ci si aspetta che sia comunque positivo, questo diventa negativo solo se si verifica un forte calo del prezzo nonché quindi un'inversione di tendenza; allo stesso modo accade viceversa quando il momentum è negativo. Un'altra problematica relativa a questo oscillatore è l'assenza di una banda superiore ed inferiore; infatti, in generale moltissimi oscillatori si muovono in una banda orizzontale ma per molti di essi, come per il momentum, è impossibile stabilire un intervallo ben preciso; questo fa sì che non vi sia una rigorosa regola da rispettare per individuare le aree di eccesso del mercato; nel corso degli anni quindi è stata definita come soluzione più

efficiente per ovviare a questo problema dell'individuazione delle zone di ipercomprato e ipervenduto, ovvero situazioni estreme del mercato, l'analisi dei movimenti passati dell'oscillatore, ovvero un'analisi visiva.

1.3.2 Efficienza

Un pilastro dell'economia finanziaria è rappresentato dall'EMH (Efficient market Hypothesis); i primi sviluppi teorici sono dovuti a Samuelson (1965) e Mandelbrot (1966). Esistono innumerevoli definizioni progressive di Efficient market Hypothesis, una delle principali è quella data da Eugene Fama, nel 1970, ovvero

Definizione 1.3.1 *Un mercato finanziario è efficiente se in ogni istante il prezzo delle attività scambiate riflette pienamente le informazioni rilevanti disponibili per cui non sono possibili ulteriori operazioni di arbitraggio.*

In tal caso, la concorrenza garantisce che i rendimenti delle attività siano ai loro livelli di equilibrio (eguaglianza tra domanda e offerta).

In un mercato finanziario siffatto né l'analisi tecnica (previsione dei prezzi futuri basata sullo studio dei prezzi passati) né l'analisi fondamentale (studiando l'andamento del valore delle imprese attraverso l'analisi della redditività si tenta di capire se esistono nuove prospettive sul valore delle azioni) possono consentire ad un investitore di conseguire profitti maggiori di quelli che un altro investitore otterrebbe detenendo un portafoglio di titoli scelti a caso, con il medesimo grado di rischio.

Poiché l'EMH è formulato in termini di aggiustamento del rischio, effettua previsioni verificabili solo se abbinato a un particolare modello di rischio, di conseguenza, la ricerca in economia finanziaria almeno dagli anni '90 si è concentrata sulle anomalie del mercato, cioè sulle deviazioni da specifici modelli di rischio.

L'idea che i rendimenti dei mercati finanziari siano difficili da prevedere risale a Bachelier (1900), Mandelbrot (1963) e Samuelson (1965) ma è strettamente associata a Eugene Fama, in parte a causa della sua influente revisione del 1970 sulla ricerca teorica ed empirica (Fama 1970).

L'efficienza dei mercati finanziari può essere intesa in varie accezioni: allocativa, valutativa, tecnico-operativa e informativa ed è importante notare che in questo elaborato facciamo maggiormente riferimento alla seconda ed alla quarta.

Il paradosso Grossman-Stiglitz

In contrapposizione alla teoria dei mercati efficienti troviamo, però, le autorevoli voci di Grossman e Stiglitz.

La tesi è che: "se il mercato è informativamente efficiente, vale a dire, cioè tutte le informazioni rilevanti si riflettevano nei prezzi di mercato, allora nessun singolo agente avrebbe sufficiente incentivo per acquisire le informazioni su cui i prezzi si basano".

Grossman e Stiglitz osservarono che nonostante i prezzi erano sempre fissati al valore corretto e riflettevano tutta l'informazione importante, nessun investitore voleva sostenere i costi attinenti la ricerca, l'elaborazione e la produzione dell'informazione, in quanto da essa non potevano trarre nessun tipo di profitto, si conclude dunque che un mercato efficiente dal punto di vista informativo è impossibile.

1.3.3 Volatilità

In finanza, la volatilità è una misura della variazione percentuale del prezzo di uno strumento finanziario nel corso del tempo mentre la volatilità storica deriva dalla

effettiva serie storica dei prezzi misurabile nel passato; per indicarla viene usato il simbolo σ che corrisponde alla deviazione standard.

Per quanto riguarda, invece, l'indice di volatilità (Var) questo è dato dalla varianza, ovvero la differenza quadratica media tra il rendimento e la sua media

$$Var(x) = \frac{1}{N} \sum_{i=1}^N (x_i - m)^2$$

che molto somigliano ai momenti centrali n -esimi di una variabile aleatoria discreta. La volatilità è una quantità importante nella finanza quantitativa, rappresenta un po' l'umore del mercato: lunghi periodi di calma possono essere seguiti da mercati nervosi con elevata volatilità.

Questa è anche strettamente connessa all'idea di rischio, infatti, il rischio è spesso misurato utilizzando la varianza, che è il quadrato della volatilità e per questo motivo, quest'ultima viene spesso utilizzata per calcolare la dimensione di posizioni di trading finanziario in modo da mantenere un livello di rischio costante in un portafoglio.

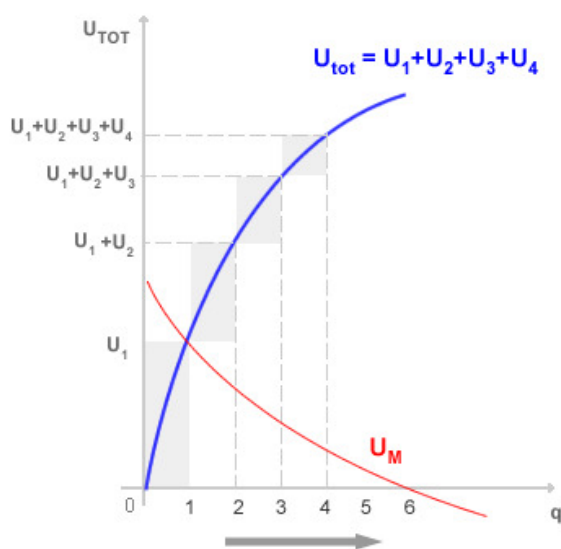
1.3.4 Utilità totale e marginale

L'utilità totale è il grado di soddisfazione complessiva che deriva dal consumo di un certo bene.

L'utilità marginale, invece, è definibile come l'incremento del livello di utilità (ΔU) ovvero della soddisfazione che un individuo trae dal consumo di un bene, ricollegabile ad aumenti marginali nel consumo del bene (Δx_i), dato e costante il consumo di tutti gli altri beni.

In termini matematici, l'aggettivo marginale è ben rappresentato dal concetto di derivata; essa, infatti, è proprio un rapporto tra variazioni (rapporto incrementale) con incrementi infinitesimi.

La legge dell'utilità marginale decrescente afferma che all'aumentare del consumo di un bene, l'utilità marginale di quel bene diminuisce e, una volta, arrivata a zero questa si dice che ha raggiunto il punto di sazietà.



Fonte: Okpedia

L'utilità totale è in funzione crescente con la quantità consumata (q) del bene soltanto se l'utilità marginale dell'ultima unità consumata è positiva.

Quando il bisogno viene completamente soddisfatto, qualsiasi ulteriore consumo del bene genera disutilità ossia utilità negativa, tale è da ridurre l'utilità totale del consumo.

Nell'esempio a fianco il soggetto economico massimizza l'utilità totale alla sesta unità del bene, quando soddisfa completamente il suo bisogno. Ogni ulteriore consumo del bene, dopo la settima unità, genera utilità negativa (fastidio) e riduce l'utilità totale del soggetto.

Dunque ciò che si potrebbe ricercare è il punto di sazietà, dato quello potremmo tentare di prevedere il potenziale andamento al ribasso del soggetto economico.

1.3.5 Mercati e strumenti prevedibili

Il rendimento futuro di qualsiasi attività rischiosa è intrinsecamente difficile da prevedere in generale, anche con le analisi che faremo in questo elaborato, tuttavia, ci sono diversi modi possibili per prevedere i rendimenti futuri a lungo termine attraverso analisi adeguate o, comunque, per cercare di limitare le perdite.

Previsione di un intervallo

Se il valore equo può essere sufficientemente definito e stimato e ci si può aspettare che i prezzi di mercato convergano entro un periodo di tempo ragionevole, è possibile prevedere in una certa misura l'intervallo ragionevole dei rendimenti medi futuri su diversi anni.

Ad esempio, i rendimenti dei titoli di Stato sono distribuiti attorno al loro rendimento carry return (ovvero la cedola sulle obbligazioni meno i costi di interesse del prestito a breve termine), definito come il rendimento atteso delle obbligazioni quando la curva dei rendimenti non cambia.

L'accuratezza di questo tipo di previsione dipende dalla velocità con cui i prezzi delle attività di rischio convergono ai loro valori fondamentali e dalla facilità con cui questi valori fondamentali possono essere stimati.

L'equity, invece, sembra essere l'asset class più difficile su cui prevedere i rendimenti futuri perché il valore fondamentale.

Trovare azioni trascurate

Ci sono troppe azioni nei mercati perché gli investitori prestino la stessa attenzione a tutte. Molte azioni, in particolare quelle emesse da piccole società, non sono molto popolari e talvolta vengono trascurate dagli investitori e non sempre hanno un prezzo adeguato a quello che è il reale valore dell'azienda, dunque, in questo caso, il compito dell'investitore è quello di cercare l'ago nel pagliaio, di trovare una valida azienda che mantenga basso il suo prezzo visto che è trascurata da altri ma che sia pronta ad esplodere sul mercato con dei progetti o dei risultati interessanti.

Diversificare

Uno dei principi più importanti dell'investimento è assicurarsi di avere un portafoglio diversificato, ciò significa assicurarsi di distribuire il proprio capitale tra diversi investimenti in modo da non fare affidamento su un singolo investimento per tutti i rendimenti, cosicché, se qualcuno di questi andasse male potremmo compensare con gli altri assets disponibili.

Un contributo fondamentale alla materia è arrivato da Harry Markowitz, nel 1952, con la nascita della teoria detta "Portfolio selection", oggi tramutatasi nella Modern Portfolio Theory (MPT).

Capitolo 2

Strumenti esistenti

2.1 Getting data

Esistono due modi per estrarre i dati dal web: scansione del web e scraping del web.

In generale sono due facce della stessa medaglia: entrambe analizzano automaticamente e interagiscono, in qualche modo, con i siti web.

Il primo modo, il crawling, era il più conosciuto fino a pochi anni fa, infatti è questo il modo in cui google e altri motori di ricerca analizzano ogni sito del web per indicizzarlo; dopotutto chi non ha mai sentito parlare del famoso "robot.txt"?!

La seconda, scraping, in tal senso è una "new entry".

Questo meccanismo è necessario agli sviluppatori per far fronte alla mancanza di alcune API oppure per evitare di abbonarsi ai fornitori di tali servizi.

In questo metodo viene creato uno script automatico che scarica continuamente i dati da un "singolo" sito web, per analizzarli ed eseguire alcune azioni.

In conclusione, la vera differenza tra loro è l'obiettivo dei robot coinvolti nelle attività e il numero di siti web analizzati.

In tal senso ci viene in aiuto il linguaggio di programmazione Python mediante il quale possiamo acquisire dati in modo diretto attraverso lo scraping, come si vede nell'esempio sotto.

```
import requests
import bs4
from bs4 import BeautifulSoup

def parseAmazon():
    r = requests.get('https://finance.yahoo.com/quote/AMZN?p=AMZN')
    soup = bs4.BeautifulSoup(r.text, "lxml")
    price = soup.find_all('div', {'class': 'My(6px) Pos(r)o
        smartphone_Mt(6px)'})[0].find('span').text
    return price

while True:
    print('Amazon current price: '+str(parseAmazon()))
```

In particolare, questo script ci permette di fare lo scraping del sito Yahoo Finance per ottenere l'Amazon stock price attuale.

Nel codice possiamo notare la libreria requests che permette di fare richieste GET ai server e ottenere delle risposte, questo caso otterremo come risposta la pagina web

richiesta e tramite BeautifulSoup analizzeremo la pagina alla ricerca del *div* in cui è contenuta l'informazione che cerchiamo.

2.1.1 Python: Pandas for finance

Analogamente possiamo, una volta ottenuti i dati, utilizzare la libreria Pandas per manipolarli ed analizzarli

```

from pandas_datareader import data
from pandas_datareader._utils import RemoteDataError
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
from datetime import datetime

START_DATE = '2005-01-01'
END_DATE = str(datetime.now().strftime('%Y-%m-%d'))
USA_STOCK = 'AMZN'

def get_stats(stock_data):
    return {
        'last': np.mean(stock_data.tail(1)),
        'short_mean': np.mean(stock_data.tail(20)),
        'long_mean': np.mean(stock_data.tail(200)),
        'short_rolling': stock_data.rolling(window=20).mean(),
        'long_rolling': stock_data.rolling(window=200).mean(),
    }

def clean_data(stock_data, col):
    weekdays = pd.date_range(start = START_DATE, end = END_DATE)
    clean_data = stock_data[col].reindex(weekdays)
    return clean_data.fillna(method = 'ffill')

def create_plot(stock_data, ticker):
    stats = get_stats(stock_data)
    plt.subplots(figsize=(12, 8))
    plt.plot(stock_data, label = ticker)
    plt.plot(stats['short_rolling'], label = '20 day rolling mean')
    plt.plot(stats['long_rolling'], label = '200 day rolling mean')
    plt.xlabel('Date')
    plt.ylabel('Adj Close (p)')
    plt.legend()
    plt.title('Stock Price over time')
    plt.show()

def get_data(ticker):
    try:
        stock_data = data.DataReader(ticker, 'yahoo', START_DATE, END_DATE )

        adj_close = clean_data(stock_data, "Adj Close")
        create_plot(adj_close, ticker)

    except RemoteDataError:

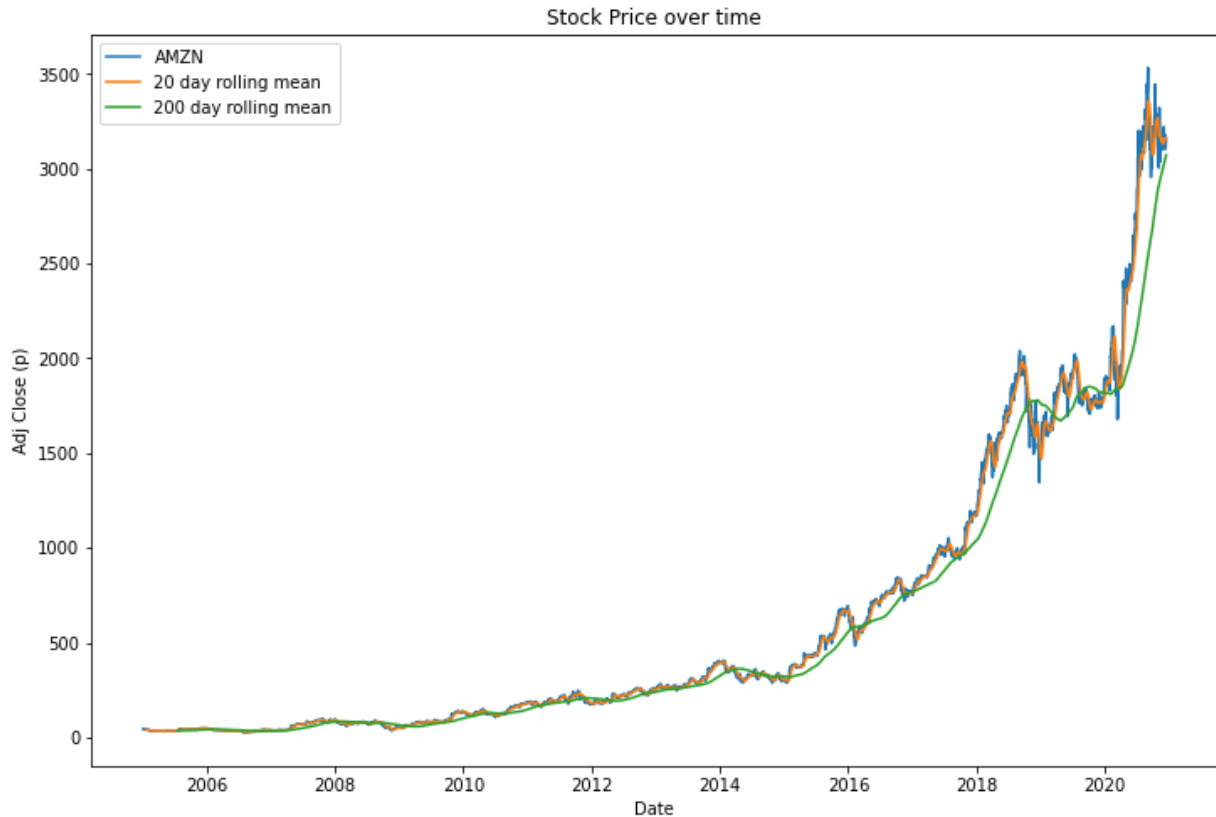
```

```
print('No data found for {t}'.format(t=ticker))
```

```
get_data(USA_STOCK)
```

Fonte: Shane Lee, Youtube

In particolare, il codice sopra produce un plot della quotazione di mercato di Amazon nell'intervallo di date desiderato per le posizioni si short e long; di seguito il grafico prodotto.



2.2 Notizie

Fin dalla nascita della borsa è noto che le notizie di carattere politico e/o economico influenzano positivamente e/o negativamente il mercato e il suo andamento; per questo, infatti, è vietata nel mondo la pratica dell'insider trading, che indica la compravendita di titoli (valori mobiliari: azioni, obbligazioni, derivati) di una determinata società da parte di soggetti che, per la loro posizione all'interno della stessa o per la loro attività professionale, sono venuti in possesso di informazioni riservate non di pubblico dominio. Il metodo più comune per il reperimento delle notizie è quello dei Feed RSS ma, di recente, si sta sempre più affermando l'utilizzo di telegram per altri scopi diversi da quello della messaggistica.

2.2.1 Feed RSS e telegram

Per quanto riguarda i Feed RSS si trova una lista abbastanza completa nel sito blog.Feedspot; qui troviamo fornitori di notizie tra i più noti come Investing, CNBC, Reddit et al.

Diversamente, oltre ai già presenti gruppi di notizie in tempo reale, attraverso degli script, detti anche bot, è possibile fare in modo che arrivino delle notizie di mercato in

tempo reale, come risultato di richieste Json, in un canale creato appositamente per lo scopo.

Sicuramente, comunque, nessuno dei due metodi sembra essere utile allo scopo di analizzare i prezzi veri e propri in quanto esistono un gran numero di microvariazioni di prezzo, soprattutto nelle valute, che risulterebbero ingestibili.

2.3 Piattaforme

Prima di analizzare le piattaforme, è necessario sapere che ciascuna di queste è un cosiddetto "broker finanziario".

Per definizione, un broker finanziario è un operatore che si occupa di vendere o acquistare beni per conto di un cliente.

Se ci focalizziamo nel settore della finanza, un broker è un intermediario che garantisce ai suoi clienti l'esecuzione di operazioni di trading sui principali mercati finanziari.

Questi soggetti, inoltre, fungono da market maker, infatti è impossibile effettuare del trading online senza la loro presenza.

Per quanto riguarda il loro guadagno, alcuni applicano dei costi fissi al momento dell'operazione di trading, nel caso dei broker di contratti per differenza (CFD), invece, c'è una piccola differenza (spread in inglese) tra il prezzo a cui è possibile acquistare e il prezzo a cui è possibile vendere.

Questa piccola differenza è appunto il profitto del broker finanziario; infatti è intuitivo notare che in tal caso il profitto del broker è direttamente proporzionale al profitto degli utenti: ecco perché questo tipo di piattaforma è sviluppata con tutto l'interesse di far guadagnare i propri clienti.

Si spiega così, ad esempio, il fatto che tutti questi broker mettono a disposizione dei corsi di trading completi o dei servizi di segnali di trading veramente affidabili.

Fonte: Osservatorio finanza Ultimamente si è sempre più diffuso il social trading che permette ai trader di copiare le operazioni di altri trader, il tutto in maniera completamente automatica.

Questo è un vantaggio sia per trader esperti che per trader meno esperti in quanto i primi potrebbero avere dei guadagni potenzialmente "facili" ed automatici, i secondi, invece, guadagnano commissioni extra ogni mese a seconda del numero di trader che li sta seguendo e copiando le loro operazioni; questa è l'innovazione che ha reso famoso il broker eToro.

Per quanto riguarda Plus500 e IQOption, sono fondamentalmente delle piattaforme in cui si possono aprire conti, anche demo, e tradare sia opzioni binarie (eToro) sia azioni e valute (Plus500). Un'altra piattaforma cardine del trading online è Metatrader con il suo MQL.

Partiamo col dire che MQL è un linguaggio object oriented sviluppato dall'azienda MetaQuote, molto simile al C++ e che è fortemente orientato all'analisi economica di grafici ed utilizzato nella piattaforma, di proprietà della stessa azienda, MetaTrader. All'interno della piattaforma è possibile aggiungere dei conti e fare del trading su azioni e indici alla stregua delle altre piattaforme precedentemente introdotte anche se ci sono delle notevoli differenze.

MetaTrader è famosa perché non solo ha introdotto il trading automatico, ovvero il trading sia impostato con una gestione ponderata del rischio (stop-loss e take-profit), ma anche perché permette, attraverso i cosiddetti Expert Advisor, di aggiungere degli script (plug-in in sostanza) che analizzano e predicono l'andamento dei titoli/indici azionati. Con MQL i traders possono strutturare strategie automatizzate che comprendono l'utilizzo di 1 o più indicatori di trading. Si possono integrare istogrammi, linee, livelli,

aree, individuazione di pattern e molto altro ancora.

Gli indicatori di trading possono essere usati per fare negoziazioni sia in maniera automatica sia discrezionale, questo è uno strumento grafico pensato per fornire un'immagine sintetica di analisi di mercato, dunque, di base, gli indicatori si prestano al trading discrezionale.

Capitolo 3

Analisi di serie temporali

Partendo dalla considerazione che dicesi serie una collezione, una classificazione, qualitativa di osservazioni di un fenomeno in base ad un criterio stabilito, se prendiamo come criterio il tempo e ordiniamo le osservazioni, ci ritroviamo il concetto di serie storica.

Dunque, in statistica descrittiva, una serie storica (o temporale) si definisce come un insieme di variabili casuali ordinate rispetto al tempo, ed esprime la dinamica di un certo fenomeno nel tempo.

È interessante notare il parallelo tra l'evoluzione del machine learning ed un altro termine diventato oramai di uso *estremamente* comune: big data; quest'ultimo, infatti, è stato coniato da Doug Laney (Douglas, 2001) per definire le capacità di trattare enormi moli di dati da parte di alcuni particolari algoritmi, in poco tempo e, in termini informatici, con poche risorse.

Approfondendone il significato dei due termini si può facilmente capire l'importante connessione, e sovrapposizione, che avviene tra le definizioni di big data e machine learning, i primi, infatti, coinvolgono i campi dell'informatica e della statistica, e richiedono delle tecnologie specifiche non solo nelle fasi di raccolta ed immagazzinamento di grandi volumi di dati, ma soprattutto nell'elaborazione e nella visualizzazione, per sviluppare forme di lettura ed interpretazione complesse, che, se vogliamo, è in molte parti simile al concetto di machine learning dato nel paragrafo 4.1. Per rendere l'idea delle potenzialità legate allo sviluppo del machine learning e dei big data in ambito di analisi delle serie storiche finanziarie, dunque del forecasting, è sufficiente pensare a questo elaborato: nei paragrafi successivi, infatti, non solo sono trattati dei metodi matematici che possono essere automatizzati ma possono anche ottenere informazione da, appunto, big data collezionati in modo opportuno riguardo un particolare asset o un mercato.

A queste informazioni potrebbero essere affiancati ad esempio degli indicatori che derivano da un processo dell'elaborazione dei testi e del linguaggio noto come sentiment analysis, come si vede in 4.2.2. Ciò ci dice che le serie temporali e gli approcci di machine learning non devono esistere isolati l'uno dall'altro, anzi, devono essere combinati insieme per sfruttare i vantaggi delle due tecniche: da un parte le serie temporali fanno un ottimo lavoro nello scomporre i dati in serie con trend e stagionalità, dall'altra il modello ad apprendimento automatico può decidere quando ed in che misura tenere in considerazione queste caratteristiche nel forecasting.

Giusto per dirne uno, l'ARIMA è uno degli algoritmi di apprendimento automatico più semplici ed efficaci per eseguire previsioni di serie temporali.

3.1 Correlazione

La correlazione è una relazione tra due variabili tale che a ciascun valore della prima corrisponda un valore della seconda a patto che si segua una certa regolarità; è comune, nella statistica, calcolare gli *indici di correlazione* che misurano la *forza* di questa relazione, di seguito una delle formule più usate, (ne esistono anche altre come r_s di Spearman o il τ di Kendall), detta coefficiente di correlazione lineare di Pearson

$$r = \frac{\sum (x_t - \hat{x})(y_t - \hat{y})}{\sqrt{\sum (x_t - \hat{x})^2 (y_t - \hat{y})^2}}$$

In particolare si dice lineare la relazione il cui grafico su assi cartesiani si avvicina alla forma di una retta, non lineare quando ha un andamento curvilineo.

È importante notare che intanto $r \in [-1, 1]$ e poi che tale coefficiente misura la forza della relazione **lineare** tra le due variabili ma non quella non lineare; infatti, ad esempio, potrebbe accadere di poter misurare una correlazione lineare, magari anche positiva, tra due variabili ma che se ne misuri una ancora più forte ma *non lineare*.

3.1.1 Autocorrelazione

Proprio come la correlazione misura l'entità di una relazione lineare tra due variabili, l'autocorrelazione misura la relazione lineare tra i valori ritardati di una serie temporale. In generale, possiamo dunque misurare l'indice di autocorrelazione mediante la formula

$$r_k = \frac{\sum_{t=k+1}^T (y_t - \hat{y})(y_{t-k} - \hat{y})}{\sum_{t=1}^T (y_t - \hat{y})^2}$$

dove T è la lunghezza della serie storica considerata.

Dunque esaminando questo risultato possiamo dire che:

- quando i dati presentano un trend, allora le autocorrelazioni per piccoli ritardi tendono ad essere ampie e positive. Questo è chiaro in quanto, salvo casi eccezionali, osservazioni vicine nel tempo hanno anche dimensioni pressoché coerenti;
- quando i dati sono stagionali allora le autocorrelazioni saranno maggiori per i ritardi stagionali e si presenteranno a multipli della frequenza stagionale;
- quando, invece, i dati mostrano sia un trend che una stagionalità allora è possibile osservare una combinazione dei due effetti sopra descritti.

3.2 Pattern

Le serie storiche possono presentare dei comportamenti, o dei pattern, che sono detti: trend, stagionalità, ciclicità, irregolarità e white noise. Ognuno di questi termini caratterizza non solo il comportamento ma anche il modo di approcciare allo studio di una serie storica.

Per trend si intende la direzione complessiva, spesso riconosciuta tramite un metodo grafico, della serie storica; in particolare, nell'ambito economico, un trend può essere rialzista, ribassista o laterale.

Una serie si dice a comportamento stagionale quando è possibile riconoscere un pattern che si ripete nel tempo ad intervalli ben definiti: ad esempio ogni inverno aumenta il consumo del gas nelle case.

A differenza della stagionalità, si ha una ciclicità quando la serie ha un pattern che si ripete ad intervalli non definiti e potenzialmente non definibili; chiaramente, rispetto alla stagionalità, non avendo dei limiti temporali risulta più difficile individuare l'inizio e la fine più o meno precisa di questi intervalli.

Alcune serie, invece, possono presentare delle piccole (o grandi) variazioni rispetto al loro andamento regolare mentre altre presentano delle vere e proprie irregolarità; ad esempio l'andamento del prezzo dei dogecoin era piuttosto regolare prima che Elon Musk si pronunciasse, mentre per irregolarità potremmo considerare la caduta di un aeroplano che avviene sostanzialmente senza una particolare regolarità (per fortuna). Come ultimo caso, si definisce white noise (rumore bianco) una serie storica che non mostra particolari segni di autocorrelazione, infatti se volessimo calcolare gli indici di tali dati ci aspettiamo che questi siano grossomodo vicini a zero.

Si noti però che se stabiliti un limite superiore e uno inferiore a $\pm \frac{2}{\sqrt{T}}$, dove T è la lunghezza della serie, più del 5% dei picchi si trova oltre questi allora la serie con molta probabilità non sarà un rumore bianco.

3.3 Stazionarietà

Una serie di dati si dice stazionaria se le sue proprietà non dipendono dal tempo in cui vengono osservate.

In particolare, quindi, delle serie storiche che presentano trend o stagionalità non sono stazionarie mentre delle serie di tipo white noise generalmente lo sono in quanto non importa quando la si osserva dovrebbe sembrare la stessa più o meno in ogni momento. Parte integrante della stazionarietà è la differenziazione, che è usata per rendere la serie stazionaria, per de-trend e per controllare le autocorrelazioni; tuttavia, alcune analisi di serie temporali non richiedono la differenziazione e le serie eccessivamente differenziate possono produrre stime imprecise.

3.4 Processi di Wiener (Brownian motion): Markov

Definizione 3.4.1 Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità e τ un insieme di tempi: un processo stocastico è una famiglia di variabili aleatorie $(X_t)_{t \in \tau}$ definite su Ω .

A volte è più comodo identificare un processo stocastico come una funzione $\mathbf{X} : \Omega \times \tau \rightarrow \mathbb{R}$ tale che, per ogni $t \in \tau$ fissato, è una variabile aleatoria.

Definizione 3.4.2 Un processo stocastico discreto $(X_n)_{n \geq 0}$ si dice di Markov se

$$P(X_{n+1} = i_{n+1} | X_0 = i_0, \dots, X_n = i_n) = P(X_{n+1} = i_{n+1} | X_n = i_n)$$

L'interpretazione della proprietà di Markov è che la previsione statistica dello stato $X_{n+1} = i_{n+1}$ che si può effettuare conoscendo la storia passata e presente $X_0 = i_0, \dots, X_n = i_n$, è identica alla previsione statistica effettuabile conoscendo il solo stato presente $X_n = i_n$; in sostanza è una forma di perdita di memoria del passato, una volta noto lo stato presente.

I processi di Markov possono essere caratterizzati sia da un insieme di stati continui che discreti, in particolare, nel secondo caso vengono detti *catene di Markov*.

Definizione 3.4.3 Il moto browniano geometrico (a volte detto moto browniano esponenziale) è un processo stocastico in tempo continuo in cui il logaritmo della quantità variabile nel tempo segue un moto browniano, o, più precisamente, un processo di Wiener.

Il processo è ritenuto appropriato per modellizzare alcuni fenomeni dei mercati finanziari; in particolare, è usato nell'ambito dell'option pricing, in quanto una quantità che segue un moto browniano geometrico può assumere soltanto valori maggiori di zero, il che riflette la natura del prezzo di un'attività finanziaria.

3.4.1 Random walk

Nella teoria della probabilità, una random walk è un processo stocastico in cui il cambiamento nella variabile casuale non è correlato ai cambiamenti passati; ne segue che il cambiamento nella variabile casuale non può essere previsto, ovvero non esiste nessun pattern per il cambiamento di valore della variabile, altrimenti questo potrebbe essere determinato, visto che a quel punto non dovremmo far altro che riconoscere il pattern. Secondo questa teoria, sebbene non sia possibile prevedere la variazione del prezzo delle azioni, il cambiamento non è del tutto irrazionale: infatti, notizie di una notevole importanza economica, come vendite, tasse, cicli economici sono in grado di variare i prezzi sul mercato.

3.4.2 Black-Scholes

L'equazione di Black-Scholes è l'equazione che è responsabile di un numero ingente di modelli, tecnicamente un'equazione differenziale parziale per il prezzo di un'opzione; il motivo per cui si ha bisogno di un'equazione così complicata, per modellare il prezzo di un'opzione, è dovuto alla natura randomica dei movimenti di prezzo.

La formula si basa sulle seguenti ipotesi

1. il mercato è perfetto, ovvero:
 - è perfettamente competitivo e ciò comporta che gli operatori non sono in grado di influenzare il prezzo dei titoli con le loro operazioni;
 - è privo di attriti, ciò non ci sono costi di transazioni e di tasse ed è possibile vendere allo scoperto senza nessuna penalità.
Inoltre si può acquistare e/o vendere in quantità arbitrarie ed infinitamente divisibili ad un tasso di interesse costante, che coincide con il tasso di rendimento dei titoli a capitalizzazione integrale (zero coupon bond);
 - c'è assenza di arbitraggi non rischiosi. Si noti che le prime due ipotesi sono tecniche di minimo impatto mentre l'ipotesi di non arbitraggio è fondamentale.
2. il prezzo del titolo sottostante è un moto browniano geometrico con media e varianza noti e costanti nel tempo;
3. il prezzo di esercizio K dell'opzione è noto e costante nel tempo e il titolo sottostante non distribuisce dividendi.

Siano C il prezzo della call option, S il prezzo corrente dello stock, K il prezzo di esercizio della option, r il tasso di interesse costante calcolato in maniera continua (con $r \in [0, 1]$), σ la volatilità dello stock (con $\sigma \in [0, 1]$), t il tempo di scadenza (in anni) ed N la distribuzione normale, allora, a partire dalla seguente equazione

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial C^2} + rS \frac{\partial C}{\partial S} = rC \quad (3.1)$$

possiamo derivarla per ottenere la formula di Black-Scholes per call e put di opzioni europee, ovvero

$$C(S, t) = N(d_1)S - N(d_2)Ke^{-rt} \quad (3.2)$$

Interpretazione finanziaria

L'equazione ha un'interpretazione concreta che viene spesso utilizzata dai professionisti, possiamo dunque esprimere la PDE in questo modo:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = rV - rS \frac{\partial V}{\partial S} \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} = rV - rS \frac{\partial V}{\partial S}$$

Ora, il lato sinistro è costituito da un termine di "decadimento temporale", ovvero la variazione del valore derivativo rispetto al tempo, chiamato Θ , e da un termine che coinvolge la seconda derivata spaziale γ , la convessità del valore derivativo rispetto al valore sottostante.

Il lato destro è il rendimento privo di rischio da una posizione a lungo termine nel derivato e una posizione a breve costituita da $\frac{\partial V}{\partial S}$ azioni del sottostante.

In estrema sintesi, l'intuizione di Black e Scholes è che il portafoglio rappresentato dal lato destro è privo di rischio: quindi l'equazione dice che il rendimento privo di rischio su qualsiasi intervallo di tempo infinitesimale, può essere espresso come la somma di theta e un termine che incorpora gamma.

3.5 Modelli autoregressivi

Per quanto riguarda i modelli statistici questi sono suddivisi in: semplici (con un x e una y), multipli (con più di una x e soltanto una y , multivariati (più x e più y), lineari e non lineari.

Uno dei modelli di cui sentiamo più parlare, ad esempio, è la regressione lineare.

Il termine *regressione* deriva dall'esperimento svolto dal biologo Galton (dobbiamo a lui anche l'introduzione del termine correlazione) che, nel 1886, esaminò le altezze dei figli (Y) in funzione delle altezze dei genitori (X) e notò una relazione funzionale tra le due variabili: più alti erano i genitori, più alti erano i figli e viceversa.

Tuttavia, v'erano dei problemi agli estremi, infatti, a genitori troppo bassi o troppo alti non corrispondevano figli altrettanto bassi o alti, quest'ultimi si avvicinavano, invece, alla media.

Galton quindi concluse che questo costituiva una *regression towards mediocrity* e la relazione funzionale fu, appunto, chiamata "modello di regressione".

I modelli ARMA(p , q), Autoregressive Moving Average, sono usati da tempo nell'analisi delle serie finanziari e nascono dalla fusione di altri due modelli di tipo, appunto, Autoregressive e Moving Average.

3.5.1 Modello ARMA(p , q)

Per arrivare a formulare il modello $ARMA(p, q)$, dobbiamo prima esaminare i due modelli che lo compongono: $AR(p)$ e $MA(q)$.

In particolare, il modello $AR(p)$ assume che il valore di una certa serie storica al tempo t , dipende linearmente dai p valori precedenti e si avrà quindi

$$X_t = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i} + \varepsilon_t$$

dove ε_t si assume essere un processo a componenti incorrelate (già definito, white noise), con media pari a zero e varianza σ_a^2 mentre α_i sono costanti reali diverse da zero.

Un tale modello ci permette di incorporare l'effetto dei valori passati nel nostro modello; intuitivamente, questo sarebbe simile ad affermare che è probabile che domani sia caldo se è stato caldo negli ultimi 3 giorni. Il modello, invece, MA(q), si può definire tramite la seguente formula

$$X_t = c_0 + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t$$

Dunque i processi a media mobile di ordine q risultano essere le somme ponderate dei valori del termine di errore ε_t , presente e passati, fino ad un ordine q già definito. Come accennato, dunque, il modello ARMA(p, q) nasce dalla combinazione dei modelli AR(p) e MA(q), infatti, in questo caso si ha che

$$X_t = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i} + \sum_{j=1}^q \beta_j \varepsilon_{t-j}$$

Si ha, inoltre, una certa invertibilità, nel senso che ponendo a zero una volta uno (q) e una volta l'altro (p) parametro, riusciamo a risalire rispettivamente ai modelli AR(q) e MA(p).

3.5.2 Modello ARIMA(p, d, q)

In statistica per modello ARIMA (acronimo di AutoRegressive Integrated Moving Average) si intende una particolare tipologia di modelli atti ad indagare serie storiche che presentano caratteristiche particolari.

Fa parte della famiglia dei processi lineari non stazionari.

Mentre i modelli di tipo exponential smoothing si basano su una descrizione della tendenza e della stagionalità nei dati, i modelli ARIMA mirano a descrivere le autocorrelazioni nei dati.

In questo caso, la formula risulta essere

$$X_t = (1 - B)^d Y_t$$

$$X_t = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i} + \varepsilon_t - \sum_{j=1}^q \beta_j \varepsilon_{t-j}$$

E si hanno i seguenti casi particolari

- se $d = q = 0$, allora la classe di processi ARIMA(p, d, q) diventa un processo AR(p);
- se $p = d = 0$, allora la classe di processi ARIMA(p, d, q) diventa un processo MA(q);
- se $d = 0$
- se $p = q = 0$ è un processo di tipo random walk;
- se $p = q = d = 0$ è un processo di tipo white noise.

Box-Jenkins

Box e Jenkins hanno proposto una metodologia per l'adattamento di un modello ARIMA alla serie temporale, X_t , che consiste nell'iterare le fasi di Identificazione del Modello, Stima e Controllo Diagnostico.

La procedura per lo studio del modello si distingue in quattro fasi:

- Analisi preliminare: verifica della stazionarietà della serie, analisi grafica, identificazione di eventuali valori anomali, ricerca delle trasformazioni più adeguate a renderla stazionaria;
- Identificazione del modello: individuazione degli ordine p,d,q del modello, dove il parametro d viene scelto nella fase precedente, mentre p e q si possono identificare mediante l'analisi delle funzioni di autocorrelazione parziale e totale;
- Stima dei parametri: stima dei parametri del modello ARIMA con il metodo della massima verosimiglianza o dei minimi quadrati;
- Verifica del modello: controllo sui residui del modello stimato per verificare se sono una realizzazione campionaria di un processo rumore bianco a componenti gaussiane.

3.6 Misuratori di errore

Considerando che per errore di forecasting s'intende lo scarto, o, più propriamente, la differenza tra un valore osservato di una determinata variabile y e la sua previsione \hat{y} , ovvero

$$e_{T+h} = y_{T+h} - \hat{y}_{T+h|T}$$

3.6.1 MAE, MSE, RMSE, SSE e Goodness-of-fit

Il MAE rappresenta l'errore assoluto che è dato da

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$

mentre il MSE è l'errore quadratico medio

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{x})^2$$

dove \hat{x}_i è il valore i -esimo del forecasting.

Di conseguenza, possiamo calcolare il Rooted Mean Square Error, ovvero la radice quadrata del MSE

$$\text{RMSE} = \sqrt{\text{MSE}} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$$

Per quanto riguarda l'SSE, come ben sappiamo, nella maggior parte dei casi abbiamo una raccolta di osservazioni ma non conosciamo i valori dei coefficienti $\beta_0, \beta_1, \dots, \beta_k$ dunque, il principio dei minimi quadrati fornisce un modo per scegliere i coefficienti in

modo efficace riducendo al minimo la somma degli errori al quadrato ovvero scegliamo i valori di $\beta_0, \beta_1, \dots, \beta_k$ che minimizzano

$$\sum_{t=1}^T e_t^2 = \sum_{t=1}^T (y_t - \beta_0 - \beta_1 x_{1,t} - \beta_2 x_{2,t} - \dots - \beta_k x_{k,t})^2$$

Questa è chiamata stima dei minimi quadrati perché fornisce il valore minimo per la somma degli errori al quadrato; trovare le migliori stime dei coefficienti è spesso chiamato "adattamento" del modello ai dati, o talvolta "apprendimento" o "addestramento" del modello.

Prendendo invece in considerazione un modello a regressione lineare, un modo comune per riepilogare quanto bene questo si adatti ai dati è tramite il coefficiente di determinazione o anche detto R^2 conosciuto anche come *Goodness-of-fit*; questo può essere calcolato come il quadrato della correlazione tra le variabili osservate y e i corrispondenti valori predetti \hat{y} o anche

$$R^2 = \frac{\sum (\hat{y}_t - \bar{y})^2}{\sum (y_t - \bar{y})^2}$$

dove le somme sono su tutte le osservazioni.

Pertanto, riflette la proporzione di variazione nella variabile di previsione che è rappresentata (o spiegata) dal modello di regressione.

Nelle regressione lineare, il valore di R^2 è uguale al quadrato della correlazione tra y e x ; per quanto riguarda l'intervallo di valori, $R^2 \in [0, 1]$ dove più R^2 è vicino ad 1, più le previsioni sono simili, o coincidono, con le variabili osservate.

3.6.2 Errori percentuali e scalari: MAPE e MASE

Oltre ai precedenti, è possibile calcolare un errore percentuale, il più comune è il MAPE, che è dato da

$$p_t = 100 \frac{e_t}{y_t}$$

che ha sicuramente il vantaggio di essere adimensionale e dunque la capacità di essere confrontato tra datasets diversi; mentre lo svantaggio è che può essere potenzialmente infinito, qualora y_t fosse molto piccolo, o indefinito nel caso in cui y_t fosse esattamente zero.

Nel 2006, però, Hyndman & Koehler, in alternativa all'utilizzo dell'errore percentuale proposero un errore di tipo scalare, per misurare le performance di serie con differenti unità di misura, in particolare per una serie storica che non presenta stagionalità possiamo calcolare il coefficiente q_i

$$\begin{cases} q_i = \frac{e_i}{\frac{1}{T-1} \sum_{t=2}^T |y_t - y_{t-1}|}, & \text{per una serie senza stagionalità} \\ q_i = \frac{e_i}{\frac{1}{T-m} \sum_{t=m+1}^T |y_t - y_{t-m}|}, & \text{per una serie con stagionalità} \end{cases}$$

dunque a questo punto possiamo calcolare l'errore, anche detto MASE, ovvero

$$\text{MASE} = \frac{1}{N} \sum_{i=1}^N q_i$$

3.7 Exponential smoothing

L'exponential smoothing è un metodo di previsione di serie temporali per dati univariati che può essere esteso per supportare i dati con un trend od una componente stagionale.

3.7.1 Single exponential smoothing

Il single exponential smoothing è un potente metodo di previsione che può essere utilizzato come alternativa alla popolare famiglia di metodi Box-Jenkins ARIMA che possiamo applicare, in particolare, quando non riscontriamo un particolare comportamento della serie storica e ci ritroviamo, "semplicemente", a dover analizzare una serie di valori.

Mentre nei modelli ARIMA le osservazioni passate sono ponderate allo stesso modo, l'exponential smoothing assegna pesi decrescenti esponenzialmente man mano che l'osservazione è meno recente.

Nel caso delle medie mobili, i pesi assegnati alle osservazioni sono gli stessi e sono pari a $\frac{1}{N}$ mentre nello smoothing esponenziale, invece, ci sono uno o più parametri (α, β, γ) di smoothing da determinare (o stimare) e queste scelte determinano i pesi assegnati alle osservazioni.

In linea generale, la formula che si usa per la versione più semplice di questo metodo è:

$$\begin{cases} \hat{y}_{t+h|t} = l_t \\ l_t = \alpha y_t + (1 - \alpha)l_{t-1} \end{cases}$$

di seguito un esempio

Period	Data	Forecast ($\alpha = 0$)	Forecast ($\alpha = 1$)
1	10		
2	18	10	10
3	29	10	18
4	15	10	29
5	30	10	15
6	12	10	30
7	16	10	12

Il metodo in questione viene detto esponenziale in quanto attraverso una serie di sostituzioni algebriche è possibile ricondursi ad una formula al cui interno è presente la serie

$$\alpha \sum_{i=0}^{t-1} (1 - \alpha)^i = 1 - (1 - \alpha)^t$$

In particolare, possiamo notare che $\alpha \in [0, 1]$ e che nei due casi limite si ha il solo valore 10 ripetuto in quanto viene del tutto trascurata la componente data da $\alpha(y_t - \hat{y}_t)$ e, dall'altra parte, viene ripetuto il dato che si ha nel periodo precedente.

In questo metodo, come è possibile notare dalla tabella sopra, la parte fondamentale del lavoro è determinare il coefficiente α che governa il peso dei dati raccolti e delle previsioni precedenti, infatti, se variamo α possiamo ottenere la seguente tabella

Forecast ($\alpha = 0,3$)	Error ($\alpha = 0,3$)	Error squared ($\alpha = 0,3$)	SSE ($\alpha = 0,3$)
3	15	225	
12,4	16,6	275,56	888,46
15,7	-0,70	0,49	MSE($\alpha = 0,3$)
11,5	18,5	342,25	148,0766667
16	-4	16	
10,6	5,4	29,16	
Forecast ($\alpha = 0,6$)	Error ($\alpha = 0,6$)	Error squared ($\alpha = 0,6$)	SSE ($\alpha = 0,6$)
6	12	144	
14,8	14,2	201,64	798,64
21,4	-6,4	40,96	MSE($\alpha = 0,3$)
13	17	289	133,1066667
22	-10	100	
11,2	4,8	23,04	

Valori grandi indicano che il modello presta attenzione principalmente alle osservazioni passate più recenti, mentre valori più piccoli significano che vengono presi valori storici anche meno recenti quando si effettua una previsione.

Come possiamo vedere, dunque, si può calcolare il valore dell'MSE in modo da renderlo il più piccolo possibile, quindi, nel caso di cui sopra preferiamo il forecast con $\alpha = 0,6$. Possiamo fare di meglio? Potremmo applicare un metodo iterativo per tentativi ed errori, iniziando dunque con intervallo di α compreso tra 0,1 e 0,9 e procedendo col determinare la migliore scelta iniziale per α e quindi cerchiamo tra $\alpha - \Delta$ e $\alpha + \Delta$; possiamo arrivare a trovare un α con tre cifre decimali.

Ma ci sono metodi di ricerca migliori, come la procedura Marquardt che è un ottimizzatore non lineare che riduce al minimo la somma dei quadrati dei residui.

3.7.2 Double exponential smoothing: Holt's method

Come abbiamo evidenziato, solitamente, si usa la versione semplice (o singola) dell'exponential smoothing nel caso in cui non ci sia un particolare trend da seguire, diversamente, invece, quando è presente un trend che possa darci una qualche informazione sull'andamento della serie, allora è possibile ricorrere al metodo di Holt che è descritto dalle seguenti equazioni

$$\begin{cases} \hat{y}_{t+h|t} = l_t + h b_t & (1) \\ l_t = \alpha y_t + (1 - \alpha)(l_{t-1} + b_{t-1}) & (2) \\ b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1} & (3) \end{cases}$$

Dove la prima è detta equazione di previsione (forecast equation), la seconda equazione di livello (level equation) e l'ultima equazione di trend (trend equation).

Anche in questo, il parametro $\beta \in [0, 1]$

3.7.3 Triple exponential smoothing: Holt-Winters

A questo punto arriviamo all'ultima versione dei metodi di exponential smoothing, quello che utilizziamo per le serie che presentano sia un trend sia una stagionalità: il metodo Holt-Winters.

Come per il trend, la stagionalità può essere modellata come un processo additivo o moltiplicativo per una variazione lineare o esponenziale della stagionalità; dunque potremo avere: stagionalità additiva e moltiplicativa.

Oltre ai fattori di smoothing, già noti, α e β , viene aggiunto un nuovo parametro chiamato γ che controlla l'influenza sulla componente stagionale.

A questo punto, per il metodo additivo, dovremo chiaramente aggiungere un'ulteriore equazione e dunque l'equazione di forecasting sarà

$$\hat{y}_{t+h|t} = l_t + h b_t + s_{t+h-m(k+1)}$$

dove

$$\begin{cases} l_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} + b_{t-1}) & (1) \\ b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1} & (2) \\ s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m} & (3) \end{cases}$$

e k è la parte intera di $\frac{h-1}{m}$ che garantisce che le stime degli indici stagionali utilizzati per la previsione provengano dall'ultimo anno del campione.

Capitolo 4

Tecniche avanzate di forecasting

4.1 Intelligenza artificiale

In relazione a questi argomenti, affinché possano essere maggiormente chiari, possiamo dare delle definizioni precise, ad esempio

Definizione 4.1.1 *L'intelligenza artificiale è una disciplina appartenente all'informatica che studia i fondamenti teorici, le metodologie e le tecniche che consentono la progettazione di sistemi hardware e sistemi di programmi software capaci di fornire all'elaboratore elettronico prestazioni che, a un osservatore comune, sembrerebbero essere di pertinenza esclusiva dell'intelligenza umana.*

L'intelligenza artificiale fa abbondantemente uso delle reti neurali che vengono utilizzate per risolvere problemi ingegneristici legati a diversi ambiti tecnologici come l'informatica, l'elettronica o altre discipline; più precisamente

Definizione 4.1.2 *Una rete neurale (in inglese neural network) è un modello matematico composto da neuroni artificiali di ispirazione alle reti neurali biologiche (quella umana o animale).*

Le reti neurali fanno sì che i computer siano in grado di risolvere i problemi in modo indipendente e che migliorino le loro capacità.

A sua volta, facenti parte della branca dell'intelligenza artificiale troviamo i concetti di machine learning e deep learning.

Il Machine Learning è la tecnologia storicamente più antica e più semplice; questa utilizza un algoritmo che il sistema adatta, solo dopo aver ricevuto un feedback umano e presuppone l'esistenza di dati strutturati disponibili all'elaborazione.

In tal caso il sistema viene prima alimentato con dati strutturati e categorizzati e quindi "capisce" come classificare i nuovi dati a seconda del tipo e poi, in base alla classificazione, il sistema esegue le attività programmate.

Ad esempio è in grado di riconoscere un cane o un gatto in una foto e di spostare i file nelle cartelle corrispondenti.

Dopo una fase iniziale di applicazione, l'algoritmo è ottimizzato dal feedback umano, che indica al sistema le classificazioni errate e le categorizzazioni corrette.

Nel caso del Deep Learning, invece, i dati strutturati non sono necessari, infatti il sistema funziona nelle reti neurali multistrato, che combinano diversi algoritmi e sono modellate sul cervello umano; ciò consente al sistema di elaborare anche dati non strutturati.

Ciò che è importante da notare è che nel Deep Learning, è il sistema stesso ad identificare nei dati le caratteristiche distintive adeguate, senza la necessità di una

categorizzazione dall'esterno; l'addestramento da parte di uno sviluppatore non è necessario, è il sistema stesso a controllare se le classificazioni cambiano a causa di un nuovo input o se ne vanno introdotte di nuove.

4.2 NLP

4.2.1 FOMO, FUD e panic selling

Quando parliamo di uno tra: FOMO, FUD e panic selling, ci riferiamo a dei comportamenti che i trader meno esperti, o comunque grossomodo ogni essere umano, possono assumere nei confronti del mercato e degli avvenimenti.

In particolare si parla di F.U.D. (Paura - Incertezza - Dubbio) quando si hanno delle immissioni volontarie di notizie false e/o accentuatamente negative per causare uno scoraggiamento degli investitori e una conseguente discesa dei prezzi, con vantaggio per i grandi investitori che acquistano a prezzi bassi.

Si dice F.O.M.O., invece, letteralmente la paura di perdere un investimento, un'opportunità, che genera negli investitori inesperti il desiderio spasmodico di comprare uno strumento finanziario che, per esempio, era al rialzo fino a qualche secondo prima e che potrebbe poi calare inesorabilmente; in sintesi è una presa di posizione non ponderata e dettata solo dall'impeto.

Per panic selling si intende il meccanismo secondo il quale, in seguito ad un particolare evento, magari catastrofico, la quasi totalità degli investitori alle prime armi presi dal panico e convinti di essere in perdita comincia a svendere ciò che possiede facendo crollare in modo smisurato e irrazionale il mercato dello strumento finanziario.

4.2.2 NLP e NLFF

Il Natural Language Processing è una branca dell'intelligenza artificiale che si occupa della progettazione di sistemi in grado di comprendere il linguaggio umano (o linguaggio naturale) e di produrre risposte in linguaggio naturale.

Negli ultimi anni, a partire da questa branca, si è sviluppata quella che adesso viene chiamata Natural Language Financial Forecasting (NLFF) che si occupa di prevedere i prezzi di mercato in base al sentiment analysis di parte del web. Secondo Kelly (2007) con le informazioni rapidamente disponibili i mercati sono diventando sempre più efficienti, poiché gli esseri umani progettano algoritmi complessi per trarne vantaggio continuamente eventuali errori di prezzo percepiti nei mercati e uno dei primi studi si è concentrato specificatamente sulla quantificazione della relazione tra comunicati stampa e movimenti nei mercati azionari è stato condotto non molto tempo fa (Gillam, Ahmad, & Ahmad, 2002).

Caso Dogecoin

Di recente si è osservato uno dei fenomeni che forse poche volte si sono verificati nella storia.

Se, da una parte, da sempre le grandi multinazionali sono riuscite a far muovere il mercato, anche se di poco, in un verso o in un altro, questa volta è stato tutto frutto, tramite dei tweet, di un uomo solo: Elon Musk; la fama lo precede, non è questa la sede per esaminare il suo curriculum, a lui dobbiamo Tesla e SpaceX.

In particolare, in un solo giorno, la criptovaluta Dogecoin ha avuto un incremento del prezzo di mercato dell'800%, in particolare da un prezzo di circa 0.002 USD a 0.08,

dunque chi si fosse trovato ad investire 50\$ avrebbe potuto comprare 25.000 Dogecoin, ritrovandosi con un capitale, in controvalore, di 2.000\$



Fonte: BittsAnalytics

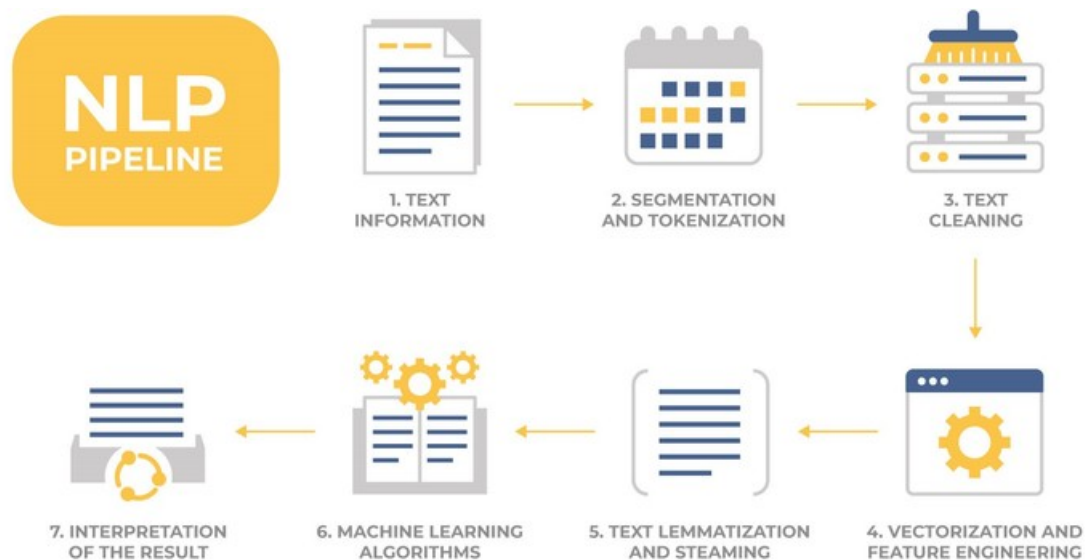
4.2.3 Algoritmi per l’NLP

Tra i compiti più importanti che il Natural Language Processing risolve sono:

- la traduzione automatica è il primo compito classico assegnato agli sviluppatori di tecnologie NLP;
- controllo grammaticale e ortografico - come conclusione del primo compito;
- classificazione del testo - definizioni della semantica del testo per ulteriori elaborazioni (uno dei compiti più popolari fino ad oggi);
- riconoscimento di entità denominate (NER) - definizione e selezione di entità con un significato predefinito (utilizzato per filtrare le informazioni di testo e comprendere la semantica generale);
- riassunto - la generalizzazione del testo in una forma di versione semplificata (reinterpretazione del contenuto dei testi);
- generazione di testo - uno dei compiti che vengono utilizzati per costruire sistemi di intelligenza artificiale;
- modellazione di argomenti: tecnica per estrarre argomenti nascosti da grandi volumi di testo.

È importante notare che tutti questi compiti nell’elaborazione del linguaggio naturale moderno e attuale sono spesso integrati in uno solo nella creazione di sistemi di intelligenza artificiale interattivi: i chatbot.

In generale, il funzionamento dei sistemi che utilizzano la PNL può essere descritto come la prossima pipeline:



Alcune delle tecniche utilizzate sono: Levenshtein Distance, text normalization, Naive Bayes algorithm, Word embedding, cosine similarity, Long short-term memory e la vettorizzazione.

Levenshtein Distance

L'elaborazione del linguaggio naturale di solito indica l'elaborazione di testo o informazioni basate su testo che può essere sia audio che video; un passo importante in questo processo è il trasformare diverse parole e forme verbali in un'unica forma vocale. Inoltre, spesso abbiamo bisogno di misurare quanto sono simili o diverse le stringhe e di solito in questo caso, utilizziamo varie metriche che mostrano la differenza tra le parole. Una delle metriche semplici e allo stesso tempo comunemente utilizzabili è Modifica distanza (a volte è nota come distanza di Levenshtein), un algoritmo per stimare la somiglianza di due valori di stringa (parola, forma delle parole, composizione delle parole), confrontando il numero minimo di operazioni per convertire un valore in un altro.

Cosine similarity

La somiglianza del coseno è una metrica utilizzata per la misurazione della somiglianza del testo in vari documenti.

I calcoli per questa metrica si basano sulle misure della somiglianza del vettore con la formula dei vettori del coseno, ovvero

$$\cos(\Theta) = \frac{\overline{A} \cdot \overline{B}}{\|\overline{A}\| \cdot \|\overline{B}\|}$$

È inoltre possibile utilizzare varie funzioni o caratteristiche del testo come vettori che descrivono questo testo, ad esempio, utilizzando metodi di vettorizzazione del testo.

Vettorizzazione

La vettorizzazione è una procedura per convertire le parole, all'interno di un testo, in cifre, per estrarre caratteristiche del testo e ulteriori informazioni utili per gli algoritmi di apprendimento automatico. In altre parole, il metodo di vettorizzazione del testo è la

trasformazione del testo in vettori numerici; i metodi di vettorializzazione più popolari sono due: "bag of words" e "TF-IDF".

Per quanto riguarda il primo metodo, il modo più intuitivo e semplice per vettorializzare le informazioni di testo prevede quanto segue:

- assegnare ad ogni parola un indice intero univoco per costruire un dizionario di parole con indici interi;
- contare il numero di apparizioni di ogni parola e salvarle il risultato numerico con l'indice relativo.

Di conseguenza, otteniamo un vettore con un valore di indice univoco e le frequenze di ripetizione per ciascuna delle parole nel testo.

Per quanto riguarda il secondo metodo, invece, TF-IDF sta per Term Frequency e Inverse Document Frequency.

Questa tecnica consente di stimare l'importanza del termine per il termine (parole) rispetto a tutti gli altri termini in un testo.

Come possiamo notare dal nome stesso, utilizziamo gli algoritmi TF e IDF, in particolare

- TF: mostra la frequenza del termine nel testo rispetto al numero totale delle parole nel testo;
- IDF - è la frequenza inversa dei termini nel testo e mostra semplicemente l'importanza di ogni termine.
Viene calcolato come un logaritmo del numero di testi diviso per il numero di testi contenenti questo termine.

Come combinazione, quindi, si ottiene l'algoritmo TF-IDF che dunque

- valuta i valori TF per ogni termine (parola);
- estrai i valori IDF per questi termini;
- ottieni i valori TF-IDF per ogni termine: moltiplicando TF per IDF;
- otteniamo un dizionario con TF-IDF calcolato per ogni termine.

Naive Bayes algorithm

Il Naive Bayesian Analysis (NBA) è un algoritmo di classificazione che si basa sul teorema bayesiano, con l'ipotesi sull'indipendenza condizionale tra coppie di caratteristiche ed è dato dalla formula

$$P(c|x) = \frac{P(c|x) \times P(c)}{P(x)}$$

In altre parole, l'NBA presume che l'esistenza di qualsiasi caratteristica nella classe non sia correlata a nessun'altra caratteristica; ecco perché questo approccio è detto naive (o ingenuo).

Il vantaggio di questo classificatore è il piccolo volume di dati necessario per l'addestramento del modello, la stima dei parametri e la classificazione.

4.3 Deep learning

Il deep learning è diventato popolare in molti aspetti del data science e la domanda che ci poniamo è: perché dovremmo voler utilizzare il deep learning piuttosto che i metodi tradizionali, per il forecasting delle serie temporali?

I metodi tradizionali hanno a che fare con problemi multivariati ad alta dimensione, relazioni non lineari e set di dati incompleti; queste sono situazioni che si verificano comunemente nei dati del mondo reale e sono molto più facilmente risolvibili utilizzando reti neurali profonde.

Esistono anche degli svantaggi però, infatti le reti neurali sono progettate come scatole nere che possono essere di difficile interpretabilità statistica.

La modellazione sequence-to-sequence per le serie temporali è stata abbastanza popolare in passato e continua sempre di più ad esserlo oggi; questi modelli si riferiscono specificamente a reti neurali in cui l'input è una sequenza (ad esempio, una serie temporale) e l'output è analogamente una sequenza.

Per consentire la manipolazione della serie, il modello deve:

- tenere traccia dello storico delle dipendenze della sequenza;
- ricordare l'ordinamento dell'indice temporale;
- consentire l'ottimizzazione dei parametri sulla durata della sequenza.

Esistono vari tipi di architetture di rete neurale in base ai diversi tasks da eseguire, le principali sono: Reti Neurali Convoluzionali (CNN) e Reti Neurali Ricorrenti (RNN).

Le CNN sono utilizzate principalmente nella visione artificiale, dove gli input sono spesso matrici con voci corrispondenti ai valori dei pixel.

Queste sono comunemente utilizzate grazie alla relativa semplicità nella progettazione delle funzionalità; infatti non necessitano, per esempio, di imparare da una serie di osservazioni ritardate nel tempo.

Le CNN sono ottime per apprendere rappresentazioni di input di grandi dimensioni.

Le reti neurali ricorrenti sono utilizzate principalmente nell'elaborazione del linguaggio naturale, l'esempio più classico è la traduzione.

Condividono molti aspetti positivi con le CNN quando si tratta di sequenze, con il plus di applicare una ricorrenza in ogni fase temporale durante l'elaborazione delle sequenze.

Gli RNN consentono ai layer di avere stati nascosti quando vengono inseriti in un livello successivo, in modo che i valori storici possano giocare un ruolo nella previsione.

Nella fattispecie, nell'ambito del time series forecasting nella letteratura si trovano due metodi rispettivamente legati alle CNN e alle RNN: Dilated Convolutional Neural Networks (DCNN) e Long Short-Term Memory (LSTM). Le reti neurali convoluzionali dilatate (DCNN) sono un tipo specifico di CNN che consentono di avere una portata temporale molto collegando un esponenziale numero di valori di input per l'output (utilizzando strati convoluzionali dilatati di varie dimensioni): ciò, naturalmente, consente prestazioni migliori a lungo serie temporali con forti dipendenze storiche.

La Long Short-Term Memory (LSTM) è un tipo di RNN che apprende in modo simile una mappatura dall'input all'output nel tempo, ma la mappatura non è più fissa come nell'RNN standard; infatti, nella backpropagation, un problema che sorge spesso è l'annullarsi o il tendere ad infinito dei gradienti (rispettivamente alla moltiplicazione e divisione per zero).

Per risolvere questo problema, gli LSTM utilizzano "porte computazionali" per controllare il flusso di informazioni nel tempo, che consentono al metodo di dimenticare (trascurare) informazioni storiche meno utili; gli LSTM, quindi, danno migliori risultati quando si ha a che fare con dipendenze a lungo termine.

Capitolo 5

Conclusione

Nel presente elaborato si è fatto un generale excursus sulla terminologia utilizzata in ambito economico: dal concetto base di trading alla Modern Portfolio Theory passando per la volatilità; si è fatto inoltre ampio riferimento alle piattaforme oggi esistenti per ridurre la distanza tra i mercati e i, esordienti e non, trader.

È stata svolta un'indagine minuziosa, e mi sentirei di dire ancora comunque superficiale, disamina di alcune tra le dozzine di formule matematiche che supportano le analisi di mercato e, in ultimo, si è trattato dell'importanza delle più avanzate tecniche di forecasting legate all'uso massiccio dell'intelligenza artificiale, con una breve illustrazione degli algoritmi attualmente adoperati, nella previsione dei mercati, esaminando anche il recente caso Dogecoin.

Ciò che potrebbe dunque essere oggetto di studi è la realizzazione di alcuni bot che possano automatizzare sia le previsioni secondo i vari algoritmi, sia un motore di sentiment analysis che riesca a processare le informazioni in modo adeguato restituendo un'analisi dettagliata che possa dare supporto alle previsioni fatte o, in alternativa, possa anche essere usata da sola per lo scalping. Dunque, se in generale si guarda al mercato come un complesso, come un sorta di box della quale è possibile, sotto determinate ipotesi prevedere qualche risultato con l'avvento della NLP (NLFF) possiamo cercare di prevedere l'impatto che una singola notizia, od un gruppo di notizie affini, possono avere su un determinato mercato, tentando di ridurre al minimo l'errore commesso.

Questa disciplina, come detto, permette di interpretare le informazioni generate da fonti come la stampa, il web ed i social media, e di trasformarle in "indicatori di sentimento" che descrivono in maniera numerica quale sia l'opinione in un particolare momento su un determinato argomento, che potrebbe anche essere il valore del titolo azionario stesso o il settore a cui appartiene il titolo o le notizie economiche della regione di appartenenza dell'azienda.

La sentiment analysis in un bot per il trading permetterebbe pertanto di introdurre una proxy della componente emotiva che, a sua volta, potrebbe rappresentare un elemento importante a spiegare una parte della volatilità del titolo, al fine di prevederne il valore in periodi successivi al momento di analisi.

Infatti, se lo scopo risulta essere quello di ricavare un profitto speculando, allora un meccanismo di scalping trading o, al massimo intraday trading, potrebbero essere notevolmente remunerative.

Rimane da capire, però, quanto velocemente si può procedere con l'acquisizione e l'elaborazione di, per esempio, un tweet al fine di aprire una posizione finanziaria, le tempistiche, infatti, sono fondamentali per una buona riuscita dello scalping. In definitiva ci troviamo in un mondo in cui, forse, è possibile un "nuovo" modo di fare trading.

Se, in effetti, l'analisi matematica non sempre ci può essere d'aiuto, o meglio, non sempre può essere precisa al punto tale da evitare del tutto le perdite di capitale possiamo, a mio avviso, trovare una valida alternativa nella NLFF.

Bibliografia

- [1] Bell, S. (2016). Quantitative finance for dummies. John Wiley & Sons Inc.
- [2] Wilmott, P. (2010). Paul Wilmott on quantitative finance. Vol. 1: ... (2. ed., reprinted). Wiley.
- [3] Brownlee, J. (s.d.). Deep learning for time series forecasting (2018o ed.).
- [4] Forecasting: Principles and practice(2nd ed). (s.d.). Recuperato 15 dicembre 2020, da <https://Otexts.com/fpp2/>

Sitografia

- [5] 6.4.3.5. Triple Exponential Smoothing. (s.d.). Recuperato 12 febbraio 2021, da <https://www.itl.nist.gov/div898/handbook/pmc/section4/pmc435.htm>
- [6] Baldo, D. (s.d.). Modellazione Statistica per l'Ottimizzazione di Sistemi Energetici [Ca' Foscari]. Recuperato 18 febbraio 2021, da <http://dspace.unive.it/bitstream/handle/10579/2753/815758-1165874.pdf?sequence=2>
- [7] Bankpedia —. (s.d.). Recuperato 2 febbraio 2021, da https://www.bankpedia.org/termine.php?c_id=19250
- [8] Bhulai, S., & Semanjski, I. (2016). Data analytics 2016: Proceedings of the fifth international conference on data analytics. https://www.researchgate.net/publication/309853606_Data_analytics_2016_proceedings_of_the_fifth_international_conference_on_data_analytics
- [9] Board of Governors of the Federal Reserve System, Heston, S. L., & Sinha, N. R. (2016). News versus sentiment: Predicting stock returns from news stories. Finance and Economics Discussion Series, 2016(048), 1–35. <https://doi.org/10.17016/feds.2016.048>
- [10] Boso, I. (2018). Tecniche di Deep Learning per l'implementazione di un Trading System [Ca' Foscari]. <http://dspace.unive.it/bitstream/handle/10579/16211/809757-14309.pdf?sequence=2>
- [11] Bousquet, O., Luxburg, U. von, & Rätsch, G. (A c. Di). (2004). Advanced lectures on machine learning: ML Summer Schools 2003, Canberra, Australia, February 2-14, 2003 [and] Tübingen, Germany, August 4-16, 2003: revised lectures. Springer.
- [12] Broker finanziario, come scegliere il migliore [2021]. (s.d.). [osservatoriofinanza.it](https://www.osservatoriofinanza.it). Recuperato 2 marzo 2021, da <https://www.osservatoriofinanza.it/broker-finanziari>
- [13] Brownlee, J. (2018, agosto 19). A gentle introduction to exponential smoothing for time series forecasting in python. Machine Learning Mastery. <https://machinelearningmastery.com/exponential-smoothing-for-time-series-forecasting-in-python/>
- [14] Cavallin, S. (2016). Modelli per la previsione di serie storiche finanziarie [Università degli studi di Padova]. http://tesi.cab.unipd.it/59331/1/Cavallin_Steve.pdf
- [15] Chen, C. (2017). Combinazioni previsionive: Focus sull'algoritmo AFTER [Università degli studi di Padova]. http://tesi.cab.unipd.it/59655/1/CHUN_CHEN.pdf

- [16] Chen, J. (s.d.). Modern portfolio theory(Mpt). Investopedia. Recuperato 17 febbraio 2021, da <https://www.investopedia.com/terms/m/modernportfoliotheory.asp>
- [17] Correlazione. (s.d.). <https://www.dsu.univr.it/documenti/Avviso/all/all845198.pdf>
del Vigna, A. (2011). Il moto browniano.
<http://users.dma.unipi.it/~flandoli/Il%20moto%20browniano-1.pdf>
- [18] Dieffenbach, C. B. (2013). Random Walk [Financial Economics].
https://www.albany.edu/bd445/Economics_802_Financial_Economics_Slides_Fall_2013/Random_Walk.pdf
- [19] Dogecoin's price jumped as high as 35% after Elon Musk tweeted «Dogecoin to the Moooonn» in a Twitter poll on the future of crypto. (s.d.). Business Insider.
Recuperato 2 marzo 2021, da <https://www.businessinsider.in/tech/news/dogecoins-price-jumped-as-high-as-35-after-elon-musk-tweeted-dogecoin-to-the-moooonn-in-a-twitter-poll-on-the-future-of-crypto/articleshow/80734720.cms>
- [20] Economia, R. (2021, gennaio 29). Reddit (E musk) ora soffiano sul Bitcoin, Dogecoin vola a +800% in un giorno. Ma cosa succede? Corriere della Sera.
https://www.corriere.it/economia/finanza/21_gennaio_29/reddit-elon-musk-ora-soffiano-criptovalute-dogecoin-800percento-un-giorno-bitcoin-vola-8ed7ec48-6242-11eb-b7b0-378dab96ebfa.shtml
- [21] Emh definitions. (s.d.). Recuperato 2 marzo 2021, da <http://www.e-m-h.org/definition.html>
- [22] Flandoli, F. (2017). Catene di Markov.
<http://users.dma.unipi.it/~flandoli/catene1.pdf>
- [23] Forecasting with machine learning techniques. (2016, giugno 21). Cardinal Path.
<https://www.cardinalpath.com/blog/forecasting-with-machine-learning-techniques>
- [24] Frego, M., Pizzato, M., Tasin, L., & Tubaro, L. (2007). Appunti del corso Processi stocastici. <http://www.science.unitn.it/~tubaro/corso3/ps.pdf>
- [25] How ml can take demand forecasting to the next level in supply chain. (s.d.).
Recuperato 2 marzo 2021, da <https://www.tcs.com/blogs/how-ml-can-take-demand-forecasting-to-next-level-in-supply-chain>
- [26] Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679–688.
<https://doi.org/10.1016/j.ijforecast.2006.03.001>
- [27] info@numxl.com, S. F. (s.d.). Mase—Mean absolute scaled error. Help Center.
Recuperato 2 marzo 2021, da <https://support.numxl.com/hc/en-us/articles/115001223523-MASE-Mean-Absolute-Scaled-Error>
- [28] Intelligenza artificiale. (2021). In Wikipedia.
https://it.wikipedia.org/w/index.php?title=Intelligenza_artificiale&oldid=118899035
- [29] Koenecke, A. (s.d.). Applying Deep Neural Networks to Financial Time Series Forecasting.
https://stanford.edu/~koenecke/files/Deep_Learning_for_Time_Series_Tutorial.pdf
- [30] Loreti, M. (2006). Teoria degli Errori e Fondamenti di Statistica.

- [31] Luglio, M. (2020). Processi di Markov.
<https://didattica-2000.archived.uniroma2.it//SenTra/deposito/Markov.pdf>
- [32] Minnini, R. M. (s.d.). Valutazione delle opzioni col modello di Black e Scholes. Università di Bari Aldo Moro.
http://galileo.dm.uniba.it/mininni/Lezione_Modello%20Black&Scholes.pdf
- [33] Modello ARIMA: Come prevedere le serie temporali. (2020, novembre 9). Lorenzo Govoni. <https://lorenzogovoni.com/modello-arima/>
- [34] Natural language processing—Okpedia. (s.d.). Recuperato 29 gennaio 2021, da https://www.okpedia.it/natural_language_processing
- [35] Neill, S. P., & Hashemi, M. R. (2018). Ocean modelling for resource characterization. In *Fundamentals of Ocean Renewable Energy* (pagg. 193–235). Elsevier. <https://doi.org/10.1016/B978-0-12-810448-4.00008-2>
- [36] Posted by Igor Bobriakov on December 21, 2019 at 6:56am, & Blog, V. (s.d.). Top nlp algorithms & concepts. Recuperato 22 gennaio 2021, da <https://www.datasciencecentral.com/profiles/blogs/top-nlp-algorithms-amp-concepts>
- [37] Puggiotto, L. (2014). La comunicazione finanziaria in tempo di crisi [Ca' Foscari]. <http://dspace.unive.it/bitstream/handle/10579/5374/816986-1168496.pdf?sequence=2>
- [38] Python for finance stock data with pandas and numpy. (s.d.). Recuperato 2 marzo 2021, da <https://www.youtube.com/watch?v=DOHg16zcUCc>
- [39] Sigman, K. (2006). Geometric Brownian motion.
<http://www.columbia.edu/ks20/FE-Notes/4700-07-Notes-GBM.pdf>
- [40] Trevisan, D. (2016). Appunti sulle catene di Markov.
<https://people.dm.unipi.it/trevisan/didattica/2016-2017/markov-chains.pdf>
- [41] Untitled. (s.d.). Recuperato 22 febbraio 2021, da <https://capital.com/it/trading-definizione>
- [42] Utilità totale—Okpedia. (s.d.). Recuperato 26 febbraio 2021, da https://www.okpedia.it/utilita_totale