

# **Dipartimento di Impresa e Management**

Cattedra: Financial Market Analysis

**L'intelligenza artificiale applicata ai  
portafogli finanziari**

---

Prof. Claudio Boido

RELATORE

---

Mario Trerotola Matr. 226201

CANDIDATO



# Sommario

INTRODUZIONE .....	4
L'INTELLIGENZA ARTIFICIALE .....	6
TIPOLOGIE DI MACHINE LEARNING .....	9
APPLICAZIONI IN FINANZA .....	12
ANALISI DELLA LETTERATURA .....	14
SELEZIONE DEGLI ASSET E CREAZIONE DI UN PORTAFOGLIO EFFICIENTE .....	17
ASSET CLUSTERING .....	17
ASSET SELECTION .....	22
ASSET ALLOCATION .....	25
IMPLEMENTAZIONE DI UN TOOL PER LA VALUTAZIONE DI ASSET .....	30
SOLUZIONI ATTUALMENTE IN USO .....	30
ESTRAZIONE DI DATI DA REPORT TRIMESTRALI, BILANCIO E FONTI ESTERNE .....	32
SENTIMENT ANALYSIS CON L'IMPIEGO DI WEB SCRAPING .....	35
ALGORITHMIC TRADING .....	37
IMPLEMENTAZIONE DI UN TOOL BASATO SU INTELLIGENZA ARTIFICIALE .....	40
CRITICITÀ DELLE SOLUZIONI ATTUALMENTE IN USO .....	43
CONCLUSIONI .....	45
BIBLIOGRAFIA .....	47
SITOGRAFIA .....	49

# Introduzione

L'intelligenza artificiale (AI) è una tecnologia che si sta diffondendo in tutti i settori ed è in continuo sviluppo. Grazie alla sua capacità di analizzare e ricavare pattern da enormi quantitativi di dati in ambienti in costante evoluzione, l'intelligenza trova molte applicazioni nel settore della finanza e in particolare in quello della gestione dei portafogli.

Il seguente elaborato analizza le principali applicazioni dell'intelligenza artificiale al settore della finanza fornendo un approccio teorico e pratico nell'analisi di prodotti finanziari, di scenari di crescita, della loro selezione per la creazione di un portafoglio efficiente che possa soddisfare le esigenze in un investitore utilizzando strumenti basati su intelligenza artificiale. Per lo sviluppo di algoritmi utilizzeremo Anaconda, una distribuzione di Python specifica per applicazioni in data science e machine learning, e per l'ottenimento dei dati storici utilizzeremo il servizio Quandl.

Nella trattazione sono state analizzate le caratteristiche principali della maggior parte delle tecnologie attualmente in uso in particolare nel campo della gestione dei portafogli. Sono state mostrate nella seguente tesi applicazioni pratiche solo di algoritmi, che potranno essere testati dai lettori.

Il primo capitolo tratta dell'intelligenza artificiale e degli elementi base della struttura di un algoritmo basato su questa tecnologia. Continua poi analizzando una sottocategoria dell'intelligenza artificiale, nota come machine learning, descrivendone quattro tipologie, denominate supervised, unsupervised, semi-supervised e reinforcement learning. Dopo aver analizzato la struttura di un sistema di questo tipo, si procede con l'analisi dei possibili settori ed applicazioni, in cui può essere utilizzata questa tecnologia, della quale nell'ultima parte descriveremo lo stato attuale di sviluppo e le soluzioni in uso nel settore della finanza.

La seconda parte tratta della creazione di un portafoglio efficiente. Parte dal modo in cui possono essere raggruppati dei prodotti finanziari, attraverso algoritmi di clustering e come ottenere una classificazione in base al profilo di rischio per un investitore. Considera poi il profilo di rischio di un investitore, nel rispetto della direttiva 2004/39/CE (MiFID), della direttiva di attuazione 2006/73/CE e della struttura dei questionari che permettono la creazione di un profilo di rischio di un cliente. Divideremo gli investitori in cinque profili in base alla capacità di sottoporsi ad un certo livello di rischio per avere un determinato livello di crescita. Si procede poi con l'analisi di un algoritmo che permetta, dato un gruppo di strumenti finanziari, di ricavare la composizione di due portafogli che li contengano in modo da minimizzare il rischio e massimizzare il profitto potenziale.

Il terzo capitolo tratta delle soluzioni attualmente in uso nella gestione automatica di portafogli e dei sistemi informativi per gli investitori. Descrive poi la struttura base di un tool che permetta l'ottenimento di dati da bilanci e report trimestrali, individuando quali sono i rapporti fondamentali da considerare quando si analizza un'azienda. Nel considerare tutti i fattori che influenzano l'andamento dei prezzi di mercato, questo

capitolo si sofferma sull'influenza delle news e di Twitter sui mercati finanziari e sul ruolo che hanno informazioni di questo tipo sull'oscillazione dei prezzi nel breve e nel medio periodo. Viene introdotto poi l'algorithmic trading e si studia come questo permetta di ottenere, da dati di mercato, previsioni riguardo l'andamento di un titolo attraverso l'analisi dei principali algoritmi in uso. Si procede poi all'ideazione e alla progettazione di un servizio che permetta analisi di mercato di medio e breve periodo, utilizzando tutte le tecnologie precedentemente descritte, e consenta una gestione automatica ed attiva di un portafoglio, nel rispetto delle direttive MiFID.

L'obiettivo dello strumento, progettato nell'ultimo capitolo, è mostrare come sia possibile combinare i diversi strumenti analizzati nella trattazione per fornire nuovi modi atti a generare valore per clienti, aziende ed istituzioni. L'applicazione descritta (Figura 13) è indirizzata al settore della gestione dei portafogli ed è dedicata principalmente ad investitori non professionisti. Nella prima parte verrà mostrato come il sistema crea un portafoglio da zero per il cliente, partendo dall'analisi del mercato, utilizzando strumenti basati su algorithmic trading, fino alla convalida del portafoglio suggerito all'investitore. Successivamente verrà implementato un meccanismo di ribilanciamento, attraverso sistemi di monitoraggio costante del mercato. Sono state successivamente analizzate anche tecniche di sorveglianza, sviluppate in ottemperanza delle direttive per la regolamentazione di strumenti informatici.

# L'intelligenza artificiale

L'intelligenza artificiale è una branca dell'informatica che si occupa del costruire algoritmi che, per funzionare, fanno affidamento su una collezione di dati di fenomeni. I dati possono essere ottenuti tramite osservazioni o costruiti in base alle esigenze del modello. A differenza degli esseri umani che possono processare dati in un numero limitato di dimensioni, basato sulle nostre percezioni, le macchine possono processare facilmente e velocemente dati multidimensionali. Il Machine Learning è un approccio specifico all'intelligenza artificiale che usa tecniche statistiche, per far in modo tale che un computer abbia la possibilità di imparare e di aumentare le sue performance senza essere esplicitamente programmato. I tool basati su Machine Learning eseguono una task, trovando dei pattern in dataset, facendo inferenza.

L'algoritmo impara sulla base dei risultati ottenuti e si ottimizza per rendere le previsioni più veritieri.

I passaggi essenziali per sviluppare un algoritmo sono:

- Studio dei dati
- Selezione di un modello
- Il modello deve essere allenato sui training data
- Viene applicato il modello per effettuare delle previsioni su nuovi dati, per verificare che l'algoritmo generalizzi bene.

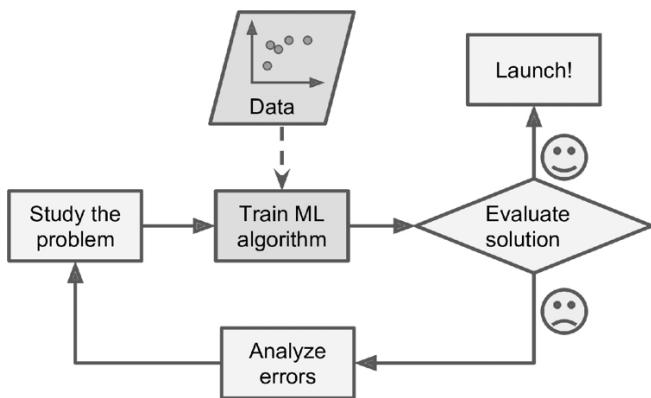


Figura 1. Approccio ad un problema con l'impiego di machine learning.

Fonte: Géron, (2019)

Poiché il compito principale è quello di selezionare un modello e allenarlo su alcuni dati, solo due cose possono portare un algoritmo ad essere inefficiente, ossia la scelta dei dati o del modello.

Per sviluppare un algoritmo efficace i dati devono essere rappresentativi, sufficienti, verificati, e rilevanti. Per generalizzare bene è molto importante che i dati siano rappresentativi dei nuovi casi sui quali si vuole generalizzare, se l'esempio è troppo piccolo, ci può essere del “sampling noise”, al contrario troppi dati possono essere non rappresentativi in caso di molte imperfezioni. Nel caso in cui nei dati siano presenti molte imperfezioni, come errori e outliers può risultare difficile ritrovare dei pattern e il sistema non sarà

efficiente, per questo l'elaborazione dei dati è un passaggio molto importante. È necessario che i training data contengano sufficientemente dati rilevanti. Una scelta sbagliata dei dati può portare ad una generalizzazione dei training data che risulta inefficiente su dati reali (overfitting) o dati non sufficienti possono portare ad un'eccessiva semplificazione del modello (underfitting).

Gli algoritmi basati su machine learning si differenziano anche in base alla possibilità di imparare incrementalmente da un flusso di nuovi dati.

Viene definito Batch learning un algoritmo che non è capace di apprendere in modo incrementale e quindi deve essere allenato utilizzando tutti i dati a disposizione, dato che nel momento in cui viene avviato, non apprende più, viene applicato solo quando ha imparato abbastanza e si dimostra efficiente. Questa tipologia viene definita offline learning.

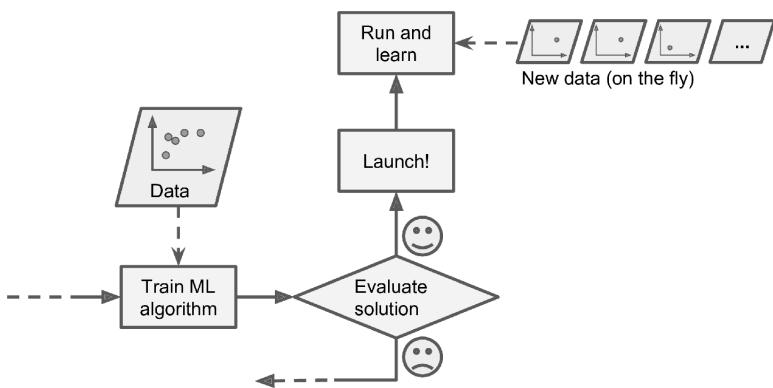


Figura 3. Struttura di un algoritmo offline learning

Fonte: Géron, (2019)

Un sistema basato su online learning impara man mano che gli vengono sottoposti nuovi dati. Questo è l'ideale per soluzioni in cui il sistema riceve dati come un flusso continuo e nel caso in cui si abbia poca capacità sia in termini di calcolo che di spazio.

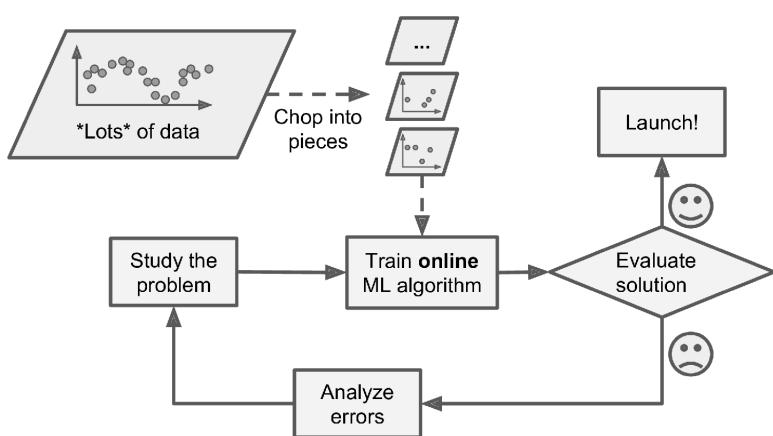


Figura 4. Struttura di un algoritmo online learning

I problemi che si prestano ad essere risolti con algoritmi basati su machine learning sono:

- Problemi le cui attuali soluzioni richiedono molto tempo per aggiustare gli algoritmi, in questa direzione un algoritmo di machine learning può spesso semplificare il codice ed essere più efficiente degli approcci tradizionali
- Problemi molto complessi per i quali, utilizzando approcci tradizionali, non si ottengono buone soluzioni.
- Problemi con un ambiente in costante evoluzione. Il vantaggio degli algoritmi di machine learning, in particolare quelli online, è la possibilità di adattarsi ai nuovi dati e trovare una soluzione dinamica.
- Problemi che richiedono l'analisi di una quantità di dati molto elevata. Tecniche basate su machine learning risultano più efficaci e veloci rispetto a tecniche statistiche.
- Problemi che richiedono l'analisi di dati da fonti e di tipologie diverse. Il machine learning permette l'analisi di tipologie di dati diversi, anche testuali e di effettuare previsioni aumentando il numero di dati rilevanti in modo da ottenere previsioni più efficaci.

Moltissimi sono i vantaggi dell'impiego di queste tecnologie. Il machine learning permette di identificare facilmente dei pattern e trend senza la necessità di un intervento umano. Questo può imparare continuamente, aumentando l'efficienza e l'accuratezza, utilizzando dati multidimensionali e vari, che, nei precedenti modelli statistici, era impossibile integrare. Proprio per queste sue caratteristiche si presta ad applicazioni in tutti i settori. Possono presentare però anche degli inconvenienti dipendenti sia dal campo di applicazione, dai dati e dai soggetti coinvolti.

## Tipologie di machine learning

Esistono molti tipi di algoritmi per il machine learning e possiamo quindi dividerli in categorie, a seconda del loro approccio e utilizzo.

Gli algoritmi per il machine learning possono essere: supervised, unsupervised, semi-supervised o reinforcement.

Nel supervised machine learning, il dataset è composto da un insieme di dati classificati  $\{(x_i, y_i)\}_{i=1}^N$ . Ogni elemento  $x_i$  in  $N$  è chiamato feature vector. Un feature vector è un vettore che in ogni dimensione "j" contiene un valore che descrive in qualche modo l'esempio. Questo valore è chiamato feature e viene indicato come  $x^{(j)}$ .

La classificazione " $y_i$ " può essere un elemento che appartiene ad una classe di elementi finita, un numero reale o una struttura più complessa.

L'obiettivo è usare il dataset per produrre un modello che prende un feature vector come input e ha come output informazioni che lasciano dedurre la classificazione di questo vettore.

Un esempio di applicazione molto importante è l'anomaly detection, che per esempio viene impiegato per eliminare outlier da dataset, prima dell'impiego di altri algoritmi, cercare difetti di manifattura e rilevare transazioni insolite su carta di credito per prevenire frodi

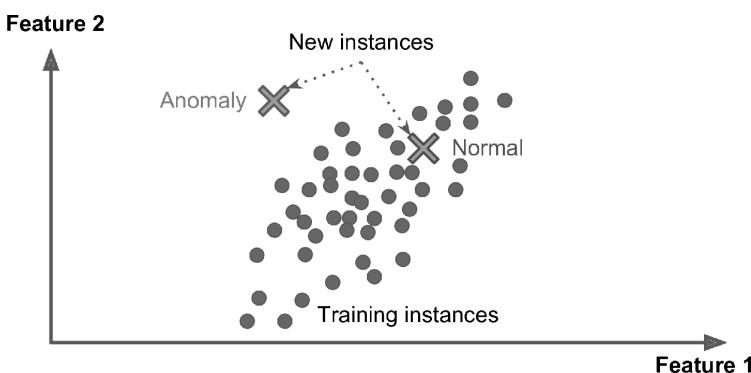


Figura 4. Esempio di anomaly detection

Fonte: Géron, (2019)

Nell'unsupervised Machine Learning il dataset è un insieme di dati non classificati. L'obiettivo è prendere le feature di un vettore  $x$  come input e trasformarle in un altro vettore o un valore che può essere usato per risolvere un problema

Per esempio, questa tipologia di algoritmi è usata nel clustering, nella dimensionality reduction e nell'outlier detection.

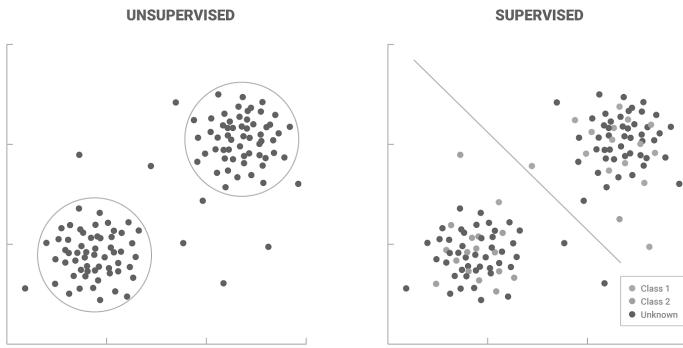


Figura 5. Differenza tra Unsupervised e supervised machine learning. A differenza del supervised machine learning il training set non ha delle vere e proprie label. Il modello identifica dei cluster e a questi vengono associati nuovi dati. I dati che si possono generare sorgono nell’acquisizione di dati, nelle tempistiche, nell’interpretazione degli stessi per la grande suscettibilità agli errori

Fonte: Géron, (2019)

Nel Semi-Supervised Machine Learning dataset contiene sia elementi classificati che non classificati. Di solito la quantità di dati non classificati è più alta di quelli classificati. L’obiettivo è lo stesso del supervised machine learning, partendo dal presupposto che usando più elementi non classificati l’algoritmo può generare un modello migliore.

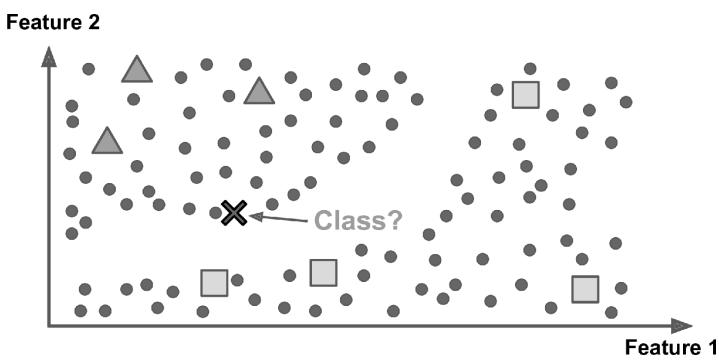


Figura 6. Esempio di semi-supervised learning con due classi. Gli elementi non classificati aiutano a classificare nuovi esempi, anche se distanti da quelli classificati

Fonte: Géron, (2019)

Il reinforcement learning è una branca del machine learning dove la macchina si trova in un ambiente ed è capace di percepire lo stato, attraverso un vettore di feature. La macchina può eseguire azioni in ogni stato e azioni differenti portano ricompense differenti. L’obiettivo è quello di imparare una policy, che è una

funzione che come input ha un feature vector e come output l'azione da eseguire in quello stato. L'azione migliore in questo caso è quella che permette di massimizzare la ricompensa media attesa.

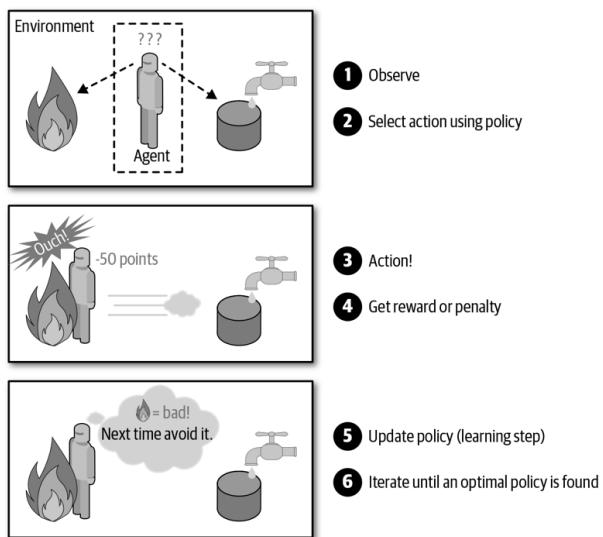


Figura 7. Reinforcement learning

Fonte: Géron, (2019)

## Applicazioni in finanza

Un sistema finanziario efficiente ha bisogno di un'analisi efficace dei dati. Questo aiuta diverse aree come l'investment optimization, comprensione di mercati finanziari, contratti di assicurazione, decisioni di credito e interazione con i clienti. Istituzioni finanziarie stanno ottimizzando la loro allocazione di capitale utilizzando tool basati su intelligenza artificiale. I manager di mutual fund, broker e agenzie governative usano l'intelligenza artificiale per trovare segnali nel mercato al fine di ottenere guadagni più elevati. Agenzie governative e società private utilizzano l'AI per l'analisi della qualità dei dati, fraud detection e per l'analisi del mercato azionario.

I mercati finanziari generano un grande volume di dati che devono essere analizzati con efficienza. Questo richiede molta attenzione ed è molto facile commettere errori. Per questo motivo sono state utilizzate tecniche di machine learning che permettono di superare queste limitazioni ed ottenere efficienza. In questo modo è possibile essere avvisati in caso di cambi di direzione del mercato e intervenire in caso di rischio elevato.

Un'altra applicazione molto importante del Machine Learning e dell'intelligenza artificiale è utilizzare ricerche di mercato e serie storiche e, grazie all'enorme quantità di dati disponibili, segnali di un movimento nei prezzi. Questa tecnica necessita dell'impiego di ingenti quantità di dati per essere efficiente e generare previsioni efficaci che garantiscono rendimenti sugli investimenti più elevati. Questa tecnica viene applicata nel campo della gestione e ottimizzazione dei portafogli.

Le agenzie governative e le istituzioni private hanno bisogno di supervisionare i mercati in modo da evidenziare movimenti fraudolenti e, grazie all'impiego di language processing e machine learning, per monitorare i comportamenti dei trader. I dati sono ottenuti da diverse fonti come e-mail, conversazioni, messaggi, documenti e social. Questo può essere usato per stimare il rischio e rilevare anomalie.

Queste tecnologie possono essere utilizzate anche per risolvere problemi di asimmetria informativa, come selezione avversa ed azzardo morale.

L'intelligenza artificiale e il machine learning sono utili nel predire i comportamenti di un utente. Le aziende assicurative usano queste tecniche per analizzare gli utenti di un servizio di polizza e stimare il premio assicurativo in base ai possibili comportamenti e garanzie di questi. In questo modo possono essere ridotti rischi di commettere errori, permettendo di aumentare i ricavi.

I modelli attuali nel campo della copertura, nel caso di potenziali vulnerabilità di mercato, sono basati su tecniche statistiche e regole matematiche basate su modelli. L'utilizzo di machine learning permette a questi algoritmi di coprire anche scenari inaspettati. L'input necessario per questi algoritmi è, proprio per questo, uno dei problemi principali nell'utilizzo e la mancanza di dati può generare delle previsioni inaccurate. L'impatto dell'intelligenza artificiale consiste nell'ottimizzazione di liquidità, scadenze e leva.

Nel settore finanziario può essere molto utile analizzare i social network, perché possono essere usati come indicatore di sentimento di mercato o lasciare trasparire incertezza da parte del management. Le piattaforme social hanno milioni utenti che pubblicano contenuti di categoria diversa ogni giorno. Twitter è una

piattaforma social e di microblogging su cui gli utenti interagiscono attraverso messaggi chiamati tweet. Questa piattaforma è usata dalle più grandi organizzazioni private ed istituzionali per comunicare con il pubblico e molte volte alcune comunicazioni vengono effettuate direttamente dai board di queste società. Può quindi essere comodo analizzare, attraverso “sentiment indicator”, l’atteggiamento nei confronti di questi messaggi, che permette di rispecchiare l’ottimismo o il pessimismo di chi scrive. Questo è applicabile per ogni piattaforma, in quanto algoritmi di questo genere sono capaci di rintracciare sul web, ogni volta che un’azienda viene citata, l’atteggiamento di mercato.

Gli esperti utilizzano tecniche di ottimizzazione dei portafogli e diversificazione degli investimenti. Il modello di selezione di Markowitz indica che ottimizzare il rischio e il rendimento di un portafoglio consiste nel risolvere un problema di minimizzazione tra il rischio e il possibile rendimento di ogni posizione, tenendo presente la percentuale di portafoglio ad essa dedicata. Attraverso tecniche di machine learning è possibile prevedere il futuro rendimento di una determinata posizione e stimare il rischio. Questi modelli danno risultati più accurati delle applicazioni delle tecniche classiche, in quanto si basano su grandi quantità di dati storici a disposizione.

La nascita e lo sviluppo di algoritmi e tecnologie che permettono di processare un grande numero di dati da diverse fonti. risulta di essenziale nel campo dell’analisi tecnica. Attraverso l’analisi di documenti, serie storiche e social, è possibile prevedere come andrà una determinata organizzazione sul mercato. Ciò permette di ridurre i costi e aumentare la produttività. Bisogna considerare però che i dati devono essere veritieri in quanto, nel caso vengano forniti dati erronei, l’algoritmo potrebbe restituire previsioni non accurate o false.

Proprio per questo motivo è importante verificare che questi siano integri e corretti. I sistemi devono quindi garantire sicurezza, comprensività, chiarezza. Per essere controllati vengono usati algoritmi basati su intelligenza artificiale che permettono anti-money laundering (AMT), countering financial terrorism (CFT) e fraud detection.

## **Analisi della letteratura**

Negli ultimi anni anche il mondo degli investitori dei mercati azionari ha introdotto l'impiego dell'intelligenza artificiale. Anche se gli esseri umani hanno sempre un ruolo principale Greenwich Associate (2019) riporta che il 34.4% degli scambi di investment-grade corporate bond avviene su piattaforme elettroniche. Anche se alcuni gestori di fondi sono riluttanti nell'accettare algoritmi automatizzati, molti di loro utilizzano analisi basate sull'intelligenza artificiale per avere idee di investimento e costruire portafogli. Utilizzando l'intelligenza artificiale, i cosiddetti robo-advisor analizzano milioni di dati provenienti da fonti diverse per effettuare operazioni al prezzo ottimale, gli analisti sviluppano delle previsioni accurate dei prezzi e dei rendimenti con grande accuratezza e grazie a tutto questo le società che si occupano di trading hanno a disposizione soluzioni che offrono ricavi più elevati a parità di rischio. Con l'intelligenza artificiale è possibile prendere e mettere insieme tipologie di dati molto diverse e identificare pattern avanzati, tali da rendere in molte situazioni i modelli statistici meno efficienti. Molte startup e società, anche molto conosciute e leader nel settore della finanza utilizzano l'intelligenza artificiale.

Un esempio molto rilevante è Bloomberg con la sua piattaforma Terminal. Bloomberg Terminal è la piattaforma che consente ai professionisti del settore dei servizi finanziari di monitorare e analizzare i dati del mercato in tempo reale, offrendo notizie, quotazioni e una electronic trading platform. Negli anni questa piattaforma si è evoluta investendo molto nel settore del natural language processing (NLP) per analizzare documenti, raccomandazioni, il sentimento di mercato e lo stesso rapporto con gli utenti.

Al cuore di questo sistema c'è un sistema di relationship extractor, che riesce a recuperare informazione da documenti. I clienti hanno la possibilità di accedere a tutte le informazioni, di cui hanno bisogno, semplicemente scrivendo ciò che desiderano in inglese.

Negli ultimi anni abbiamo assistito anche allo sviluppo di robo advisor, Un robo advisor è un applicativo software che permette di costruire e gestire un investment portfolio.

Questa tipologia di servizi è adatta ad un pubblico di investitori occasionali o di investitori che non vogliono occuparsi della gestione del loro portafoglio.

Questi applicativi sono basati su algoritmi di intelligenza artificiale, che, in base all'attuale situazione del mercato o all'approccio del cliente, reagiscono e adattano il portafoglio. Per stabilire il livello di rischio richiesto al cliente vengono poste delle domande come l'età, le entrate, per quanto tempo si vuole investire e quanto è importante per il cliente la somma investita e, in base a questo, viene stabilita l'allocazione ottimale degli asset nel suo portafoglio, per esempio tra bond, azioni, real estate e liquidità. Man mano che il portafoglio cambia, il robo advisor ribilancerà il portafoglio per mantenere lo stesso criterio di rischio.

Questo compito di solito era svolto dai financial advisor, ma in questo modo è possibile abbattere il TCI (Total cost of investing). Di solito i costi di transazioni e le commissioni sono zero e al robo advisor spetta una percentuale degli asset nel portafoglio gestito.

Negli ultimi anni sono state sviluppate da startup e imprese nel settore fintech soluzioni basate su intelligenza artificiale che aiutano banche e prestatori di credito nel migliorare le decisioni della

sottoscrizione di prestiti. Dopo le crisi finanziarie, i Regolatori hanno posto grande attenzione sulla gestione del rischio e prevedono che le istituzioni finanziarie abbiano misure per la gestione del rischio trasparente. Le banche devono allinearsi con queste nuove soluzioni, per ridurre i tempi decisionali e sviluppare il proprio business. Gli algoritmi basati su machine learning trovano però scetticismo da parte degli esperti del settore finanziario, anche in base alla regolamentazione del settore, a causa della mancanza di trasparenza e alla sua cosiddetta definizione di blackbox, molte volte infatti l'accuratezza del modello aumenta a costo della sua trasparenza. Questo porta ad un problema sia pratico che etico. È necessario quindi che l'Intelligenza Artificiale nel settore bancario ed assicurativo soddisfa requisiti minimi di governance e controllo. In questo settore Deloitte nel 2018 ha sviluppato una piattaforma chiamata Zen Risk che dà la possibilità agli utenti di accedere a tool moderni. Il suo compito è quello di automatizzare parte del lavoro e offrire un modo per integrare trasparentemente la nuova tecnologia nei modelli già esistenti.

Da novembre 2019 il Nasdaq, caratterizzato da alti volumi di transazioni, ha introdotto un sistema di sorveglianza per monitorare tentativi di atti illeciti per vincere il sistema.

Questo sistema, basato sul deep-learning, lavora insieme ad analisti per analizzare una media di 17,5 milioni di transazioni al giorno. I vecchi sistemi segnalavano all'incirca 1000 transazioni fraudolente al giorno e solo una piccola frazione di questi veniva identificata e venivano prese delle misure di sicurezza.

Il vantaggio di questo sistema è la possibilità, oltre che di analizzare un numero maggiore di casi per ovviare al problema precedente, di rilevare dei pattern molto più complessi e molto più difficili da individuare.

Il sistema è stato allenato utilizzando dati basati su esempi storici. Ogni volta che vengono trovate delle attività fraudolente simili, viene avvisato l'analista. Questo dopo aver verificato l'accuratezza dell'outcome, lo reinserisce nel sistema. In questo modo un algoritmo basato sul deep-learning può aumentare anche il numero di casi differenti che riesce a rilevare.

Il sistema si rivela cruciale, per la sua capacità di adattarsi più velocemente a nuovi pattern delle tecniche fraudolente, man mano che queste diventano più complesse e quindi più difficili da tracciare. Nel caso in cui il sistema risulti efficace, l'azienda si propone di applicarlo a livello globale.

Lewis (2010) ha attirato l'attenzione del pubblico sull'Algorithmic Trading, che utilizza complessi algoritmi uniti con modelli matematici per prendere decisioni di comprare o vendere financial security. L'algorithmic trading utilizza anche tecniche di high-frequency trading, che possono aiutare ad eseguire decine di migliaia di transazioni al secondo. Questo sistema può essere usato in innumerevoli situazioni, come l'esecuzione di ordini e l'arbitraggio.

Il vantaggio nell'utilizzo di algoritmi consiste nel competere anche per la velocità con cui questi ordini vengono eseguiti, basti pensare che, nei pressi degli stock exchange, hanno sedi diverse società attive nel trading, in modo tale da abbassare i tempi di connessione con i server degli stock exchange.

L'algorithmic trading permette la rilevazione anche di rapidi e piccoli movimenti di prezzo, a volte invisibili ad occhi umani, e permette, se si verificano determinate condizioni prestabilite, di effettuare ordini in modo automatico. Ciò presenta dei vantaggi e svantaggi in quanto si possono verificare degli errori anche clamorosi, ne è sicuramente esempio il flash crash del 2010

Il 6 Maggio 2010 (in quei giorni c'erano preoccupazioni per la crisi del debito della Grecia) indici azionari come S&P 500, Down Jones Industrial Average e Nasdaq Composite crollarono di 998,5 punti in pochissimi minuti per poi recuperare la maggior parte delle perdite nei venti minuti successivi.

# Selezione degli asset e creazione di un portafoglio efficiente

## Asset clustering

La cluster analysis è una tecnica statistica usata per classificare osservazioni simili in gruppi chiamati cluster, e l'obiettivo è di raggruppare osservazioni che sono simili in subset omogenei del dataset di partenza. Questo tipo di analisi viene definita anche analisi di classificazione. Questa viene utilizzata quando non sono disponibili informazioni riguardo l'appartenenza o meno di un oggetto ad una classe. La procedura operativa per l'applicazione di questa tecnica consiste nel:

- formulare il problema
- scegliere uno o più parametri rappresentativi
- scegliere una misura della distanza
- scegliere un numero di cluster opportuno
- verificare la validità del modello

L'assunto di partenza è di costruire gruppi (o cluster) assicurando che le osservazioni in un gruppo siano il più possibile simili tra loro mentre le osservazioni appartenenti a gruppi diversi siano il più differenti possibili.

Ci sono due tipi di classificazione:

1. K-means clustering
2. Clustering gerarchico.

Il primo metodo di solito viene utilizzato quando il numero di classi è definito in partenza, mentre il secondo quando il numero di classi è sconosciuto in partenza e ci aiuta a trovare il numero di classi ideali. Un algoritmo di tipo k-means divide n osservazioni in k gruppi, nel quale ogni osservazione appartiene al cluster con la media più vicina, servendo da prototipo.

Per sviluppare un algoritmo di questo tipo in Python abbiamo bisogno dei seguenti package:

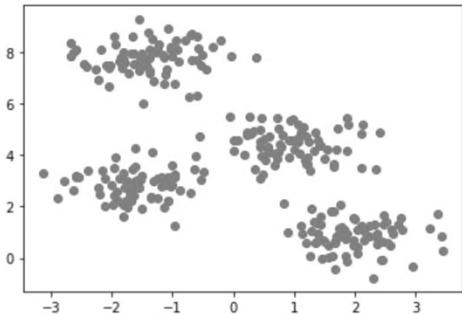
- pandas,
- NumPy,
- scikit-learn,
- Seaborn
- Matplotlib.

Il primo passo è cominciare con l'importare le librerie necessarie:

```
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
from sklearn.datasets.samples_generator import make_blobs
from sklearn.cluster import KMeans
```

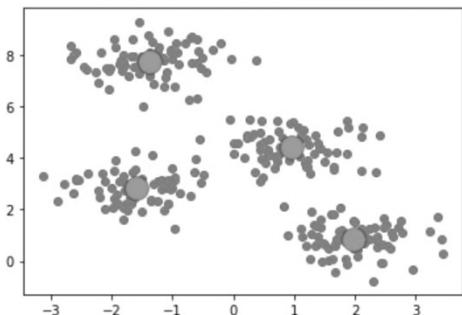
I dati verranno generate utilizzando la funzione disponibile in `sklearn.datasets` e il parametro `centers` rappresenta il numero di cluster, e poi rappresentiamo i dati ottenuti

```
X, y = make_blobs(n_samples=300, centers=4, cluster_std=0.60, random_state=0)
plt.scatter(X[:,0], X[:,1])  
<matplotlib.collections.PathCollection at 0x117a64610>
```



Procediamo successivamente con la categorizzazione dei dati

```
kmeans = KMeans(n_clusters=4, init='k-means++', max_iter=300, n_init=10, random_state=0)
pred_y = kmeans.fit_predict(X)
plt.scatter(X[:,0], X[:,1])
plt.scatter(kmeans.cluster_centers_[:, 0], kmeans.cluster_centers_[:, 1], s=300, c='red')
plt.show()
```



Nella classificazione gerarchica non si ha come input il numero di cluster ma le osservazioni simili devono essere aggregate. Il criterio per rilevare osservazioni simili è quello di calcolare la distanza rispetto al centro di ogni cluster.

In Python per applicare il cluster gerarchico dobbiamo prima importare le librerie

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
```

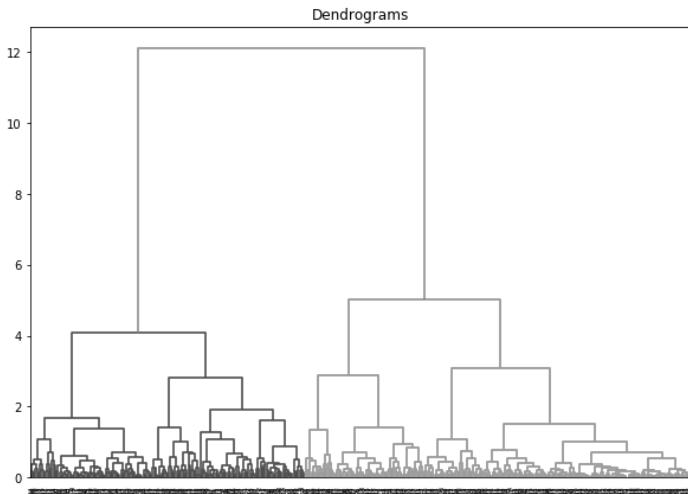
Successivamente devono essere caricati I dati dal file di prova “test.csv”

```
data = pd.read_csv('test.csv')
data.head()
```

Prima di applicare uno strumento di analisi i dati devono essere normalizzati in modo da essere confrontabili tra di loro.

```
from sklearn.preprocessing import normalize
data_scaled = normalize(data)
data_scaled = pd.DataFrame(data_scaled, columns=data.columns)
data_scaled.head()
```

A questo punto bisogna rappresentare i risultati su un dendogramma per decidere il numero di cluster dove sull'asse x sono rappresentate le osservazioni e sull'asse y la distanza



A questo punto possiamo scegliere il numero di cluster che rispecchia meglio i dati e si appresta meglio alle esigenze in questo caso ne sceglieremo 2

Applichiamo ora la tecnica del clustering gerarchico

```
from sklearn.cluster import AgglomerativeClustering
cluster = AgglomerativeClustering(n_clusters=2, affinity='euclidean', linkage='ward')
cluster.fit_predict(data_scaled)
```

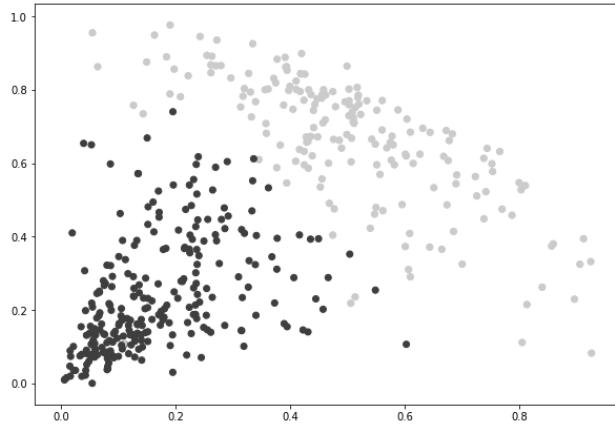
e potremo osservare che i valori di output corrisponderanno all'appartenenza al primo o al secondo cluster

```
array([1, 1, 1, 0, 0, 1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0,
       0, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 0, 1, 1,
       1, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 0, 1, 1, 1, 0, 1,
       1, 0, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0,
       0, 0, 0, 0, 1, 0, 1, 1, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 1, 1, 1, 1,
       0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0,
       0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 1, 0,
       0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1,
       0, 0, 0, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 0, 0, 1, 0, 0, 0, 1,
       0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 1, 0,
       0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 1, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1,
       1, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 0,
       0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 1, 1, 1, 0, 0, 0, 1, 0, 0, 0, 1,
       0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 0, 0,
       0, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 1, 0,
       0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1,
       1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0,
       1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1,
       1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0,
       0, 0, 1, 0, 1, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0,
       1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1,
       1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0])
```

A questo punto occorre rappresentare i dati divisi per cluster

```
plt.figure(figsize=(10, 7))
plt.scatter(data_scaled['dataa'], data_scaled['datab'], c=cluster.labels_)
```

ed otterremo il seguente risultato



Dopo questo esempio pratico bisogna stabilire quali parametri di riferimento scegliere nel caso in cui si vogliano analizzare strumenti finanziari.

La correlazione, come misura della somiglianza di due titoli, non funziona benissimo in periodi di stress, per questo potrebbero essere considerati degli indicatori correlati alle performance passate o alle aspettative di crescita, in modo da non osservare cambiamenti drastici nella struttura dei gruppi, nel caso di periodi di stress del mercato.

Considereremo ora due rapporti finanziari: revenue to assets e net income to assets. Successivamente viene presa una media ponderata di questi rapporti e la differenza dei rapporti tra i vari asset verrà usata come misura della somiglianza.

Le entrate sono il quantitativo di denaro che una società riceve in un periodo specifico, viene indicato anche come vendite (Sales). L'utile netto è il guadagno complessivo o profitto ed è ottenuto sottraendo alle revenue i costi a cui è soggetto la società. Gli asset vengono rilasciati ogni trimestre e sono contenuti nel bilancio. Questa misura di solito è presa di riferimento per indicare le dimensioni di una società.

Ricavi, net income e assets sono presenti nei report trimestrali e quindi facilmente accessibili.

I report aziendali (per grandi aziende) sono disponibili presso:

- SEC, in particolare accedendo al database della SEC chiamato EDGAR
- Sito della società
- Altri regolatori: le banche non devono presentare report alla SEC ma devono presentarli ai regolatori del settore bancario.

## Asset selection

Il primo passo nella realizzazione di un portafoglio per un investitore è la pianificazione dello stesso. In questo stadio vengono definiti le linee da seguire in modo da rispecchiare gli interessi dell'investitore ed in particolare la sua propensione al rischio.

Il risk profiling è un elemento essenziale nella composizione di un portafoglio, senza una conoscenza appropriata degli obiettivi dell'investitore, dell'orizzonte temporale, del bisogno di liquidità e dell'avversione al rischio è impossibile raccomandare investimenti o costruire delle strategie di investimento a lungo termine. Un profilo di rischio identifica il livello di rischio che è disposto ad accettare un investitore. Questo non deve rispecchiare solo la volontà di un individuo di accettare un rischio, ma anche le capacità dello stesso nel prendere decisioni in merito.

La possibilità di assumere rischi è valutata attraverso la quantificazione delle sue attività e passività. Un individuo con diversi asset e poche passività ha la possibilità di assumere livelli di rischio più elevati rispetto ad un investitore che ha diverse attività.

Il tema del rilevamento della tolleranza al rischio è un tema molto importante per i regolatori e per gli intermediari finanziari. Il legislatore della MiFID segnala la preferenza verso il rischio e il profilo di rischio tra le informazioni che gli intermediari sono tenuti a rilevare per indicare raccomandazioni adeguate alle consulenze e gestione di portafogli.

La direttiva 2004/39/CE (MiFID) e la direttiva di attuazione 2006/73/CE prevedono che l'intermediario prima di poter offrire i propri servizi debba ottenere dal cliente:

- Esperienza e conoscenza del tipo di prodotto o servizio offerto
- Situazione finanziaria
- Obiettivi di investimento

La direttiva di attuazione suggerisce indica altre informazioni che sarebbe opportuno ottenere da un cliente.

Riguardo le conoscenze del cliente:

- tipi di servizi, operazioni e strumenti finanziari con i quali il cliente ha dimestichezza
- frequenza e periodo in cui il cliente ha messo in atto operazioni dello stesso tipo
- livello di istruzione e professione o, se rilevante, precedente professione.

Della situazione finanziaria:

- informazioni sul reddito regolare,
- attività, comprese le attività liquide,
- investimenti e beni immobili,
- impegni finanziari regolari.

Riguardo gli obiettivi di investimento:

- orizzonte temporale,

- propensione al rischio,
- profilo di rischio,
- fine dell'investimento.

<i>Items</i>	Variabili	<i>Items</i>	Variabili
		caratteristiche socio-demografiche	genere età stato civile stato di famiglia e previsioni di variazioni del nucleo di familiari a carico
esperienze e conoscenze	professione istruzione natura, volume, frequenza delle operazioni su strumenti finanziari realizzate dal cliente e periodo durante il quale queste operazioni sono state eseguite tipi di servizi, operazioni e strumenti finanziari con i quali il cliente ha dimostrato	esperienze e conoscenze	professione istruzione esperienza precedente in materia di investimento esiti delle scelte d'investimento precedenti (positivo / negativo) conoscenza dei prodotti finanziari conoscenza del funzionamento dei mercati finanziari <i>trade off</i> rischio-rendimento dimensioni del rischio (credito/mercato/cambio) diversificazione di portafoglio percezione delle probabilità <i>overconfidence</i> e ottimismo
situazione finanziaria	investimenti e beni immobili; attività, comprese le attività liquide fonte ed entità del reddito regolare impegni finanziari regolari	situazione finanziaria	investimenti e beni immobili; attività, comprese le attività liquide fonte ed entità del reddito regolare impegni finanziari regolari altri impegni finanziari (previsioni e aspettative di variazioni delle spese regolari)
obiettivi d'investimento	periodo di tempo per il quale il cliente desidera conservare l'investimento  finalità dell'investimento  preferenze in materia di rischio profilo di rischio	obiettivi d'investimento	periodo di tempo per il quale il cliente desidera conservare l'investimento preferenze verso il tempo (tasso di sconto/grado di impazienza) esigenze di liquidità finalità dell'investimento ammontare (in rapporto alla ricchezza/reddito) atteggiamento verso il rischio (rischio oggettivo) capacità emotiva di assumere rischio (rischio soggettivo) avversione alle perdite

Figura 8. Le indicazioni della MiFID e della letteratura a confronto

Fonte: Luciano, Soccorso, (2012)

La profilazione del rischio può essere ottenuta in modi diversi, ma generalmente avviene attraverso un questionario. Questi questionari valutano le risposte di un individuo attraverso delle domande a risposta multipla per lo più e attraverso questo è possibile ricavare un profilo di rischio, che verrà usato successivamente dai financial advisor per formare il portafoglio dell'investitore.

Molte volte la volontà e la possibilità di sottoporsi a rischio non coincidono, per questo motivo è molto utile non solo profilare un investitore, ma anche comprendere la sua propensione al rischio.

È possibile distinguere diverse tipologie di profili, di solito sono invidiati cinque macro-profili in cui è possibile raggruppare gli investitori:

Conservativo	Un investitore conservativo non desidera sottoporsi a nessun rischio di investimento. Il requisito fondamentale è la salvaguardia del capitale ed è preparato a sacrificare alte potenzialità di guadagno per essere più Tranquillo
Moderatamente conservativo	Un investitore moderatamente conservativo è preparato ad accettare piccole quantità di rischio. La sua priorità resta quella di preservare il suo capitale a medio-lungo termine e potrebbe avere competenze nel settore finanziario
Bilanciato	Un investitore bilanciato ha delle conoscenze dei mercati ed è disposto ad accettare del rischio a breve termine sul capitale. È disposto ad assumersi una piccola quantità di rischio per fare un modo che il suo capitale non venga eroso dalle tasse e dall'inflazione
Orientato alla crescita	Un investitore orientato alla crescita comprende i movimenti del mercato . Un investitore di questo tipo desidera massimizzare la crescita del suo capitale ma non desidera prendere scelta non bilanciate. È disposto a sacrificare la sicurezza a breve termine per massimizzare la crescita del capitale
Orientato ad alta crescita	Un investitore che è disposto ad assumersi alti livelli di rischio è preparato a perdite anche molto importanti sul capitale per avere un profitto potenziale più elevato a lungo termine

## Asset allocation

Una volta selezionati gli asset di cui deve far parte il portafoglio di un investitore, bisogna ora trovare le proporzioni in cui un investitore dovrebbe investire in ciascuno strumento finanziario.

Nell'effettuare questa operazione la Teoria Moderna del Portafoglio (MTP) riveste un ruolo fondamentale. Questa è una teoria sviluppata da Harry Markowitz e pubblicata nel 1952 con il titolo "Portfolio Selection" nel "Journal of finance".

L'assunto principale è il risk-return trade-off, cioè ad un alto rischio è associata una più alta probabilità di risultati migliori, mentre a rischi più bassi corrispondono performance minori. La teoria moderna del portafoglio assume che gli investitori sono avversi al rischio, in modo che dati due portafogli che offrono lo stesso rendimento l'investitore preferirà quello con rischiosità più bassa.

Nella MTP gioca un ruolo essenziale la diversificazione, infatti, investendo in più di un'azione, un investitore può avere una riduzione della rischiosità del suo portafoglio.

Con il principio della diversificazione, il rischio di un portafoglio non è uguale alla media ponderata delle singole posizioni in un portafoglio, al contrario, le performance generali sono la media ponderata delle singole posizioni. Il rischio rappresenta la volatilità di una posizione e per calcolare il rischio complessivo bisogna tenere conto anche dei movimenti rispetto ad ogni altra posizione.

Il primo passo per la creazione di un portafoglio di esempio è la scelta del dataset. Ci sono numerose librerie che permettono l'ottenimento di dati, in questo caso prenderemo come riferimento "Quandl", per il suo utilizzo c'è bisogno di una API key che si ottiene attraverso la registrazione sulla loro piattaforma.

Le librerie necessarie per lo sviluppo dell'applicativo sono:

- quandl
- pandas
- numpy
- matplotlib.pyplot

Procediamo ora con l'importazione delle librerie

```
import quandl
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

Ora bisogna definire un intervallo temporale di riferimento da cui prendere i dati e scegliere i titoli, dei quali di vogliono ottenere i dati.

In questo caso come intervallo temporale sceglieremo il periodo che va dal 01/01/2018 al 30/04/2020 e come titoli le azioni di Apple , General Motors, General Electric, Tesla , Allign Technology, Exxon Mobil, Simon Property Group e American International Group (AIG).

```

quandl.ApiConfig.api_key = 'Your API Key'
selected = ['AAPL', 'GM', 'GE', 'TSLA', 'ALGN', 'XOM', 'SPG', 'AIG']
data = quandl.get_table('WIKI/PRICES', ticker = selected,
                        qopts = { 'columns': ['date', 'ticker', 'adj_close'] },
                        date = { 'gte': '2016-1-1', 'lte': '2020-04-30' }, paginate=True)

```

Impostiamo correttamente i dati e calcoliamo il rendimento e la covarianza annuale e giornaliera.

Successivamente creiamo gli array nei quali si troveranno i rendimenti, la volatilità, l'indice di Sharpe e il peso dei portafogli. Successivamente impostiamo il numero di portafogli casuali che andremo a generare.

```

clean = data.set_index('date')
table = clean.pivot(columns='ticker')
returns_daily = table.pct_change()
returns_annual = returns_daily.mean() * 250
cov_daily = returns_daily.cov()
cov_annual = cov_daily * 250
port_returns = []
port_volatility = []
sharpe_ratio = []
stock_weights = []
num_assets = len(selected)
num_portfolios = 70000
np.random.seed(101)

```

Definiamo ora una funzione che simuli pesi casuali per ogni posizione e calcoliamo le performance e la variabilità del portafoglio.

Per calcolare la variabilità del portafoglio dobbiamo prima analizzare la formula della deviazione standard che è la seguente:

$$\sigma_{portafoglio} = \sqrt{w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2w_1 w_2 Cov_{1,2}}$$

Una misura molto efficace per determinare l'efficienza di un portafoglio è l'indice di Sharpe

Questo misura il rendimento tenendo conto del rischio. Questo indice indica quanto rendimento in più ottengo per la maggiore volatilità data dall'impiego di asset più rischiosi. L'indice di Sharpe si calcola con la seguente formula

$$Indice\ di\ Sharpe = \frac{\bar{r}_p - r_f}{\sigma_p}$$

$\bar{r}_p$  = Rendimento atteso del portafoglio

$r_f$  = tasso privo di rischio

$\sigma_p$  = Deviazione standard del portafoglio

```

for single_portfolio in range(num_portfolios):
    weights = np.random.random(num_assets)
    weights /= np.sum(weights)
    returns = np.dot(weights, returns_annual)
    volatility = np.sqrt(np.dot(weights.T, np.dot(cov_annual, weights)))
    sharpe = returns / volatility
    sharpe_ratio.append(sharpe)
    port_returns.append(returns)
    port_volatility.append(volatility)
    stock_weights.append(weights)

portfolio = {'Returns': port_returns,
             'Volatility': port_volatility,
             'Sharpe Ratio': sharpe_ratio}

for counter,symbol in enumerate(selected):
    portfolio[symbol+' Weight'] = [Weight[counter] for Weight in stock_weights]

df = pd.DataFrame(portfolio)

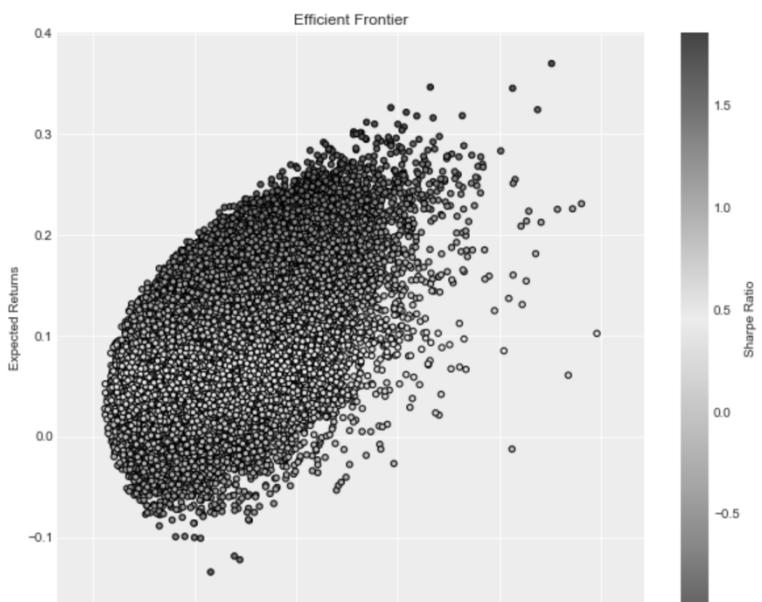
column_order = ['Returns', 'Volatility', 'Sharpe Ratio'] + [stock+' Weight' for stock in selected]

df = df[column_order]

plt.style.use('seaborn-dark')
df.plot.scatter(x='Volatility', y='Returns', c='Sharpe Ratio',
                 cmap='RdYlGn', edgecolors='black', figsize=(10, 8), grid=True)
plt.xlabel('Volatility (Std. Deviation)')
plt.ylabel('Expected Returns')
plt.title('Efficient Frontier')
plt.show()

```

Otteneremo il seguente risultato:



Possiamo osservare ora come si sia formata una linea. Questa è chiamata frontiera efficiente. Questa è chiamata così perché tutti poiché punti su questa linea danno il livello più basso di rischio per ogni livello di performance.

I portafogli che si trovano al di sotto della frontiera efficiente sono definiti inefficienti mentre quelli che si trovano sulla curva sono poiché migliori in termini di rischio-rendimento.

Pertanto, nell'esempio sopra indicato verranno presi di riferimento poiché portafogli in alto a sinistra perché avranno il miglior rapporto tra rischio e rendimento.

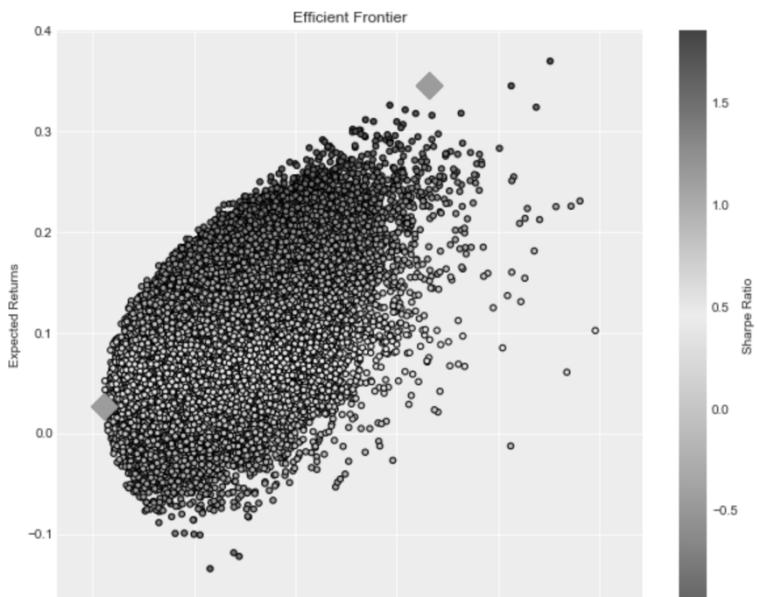
Rappresentiamo ora graficamente poiché portafogli che presentano la minore volatilità e il maggiore indice di Sharpe

```
min_volatility = df['Volatility'].min()
max_sharpe = df['Sharpe Ratio'].max()

sharpe_portfolio = df.loc[df['Sharpe Ratio'] == max_sharpe]
min_variance_port = df.loc[df['Volatility'] == min_volatility]

plt.style.use('seaborn-dark')
df.plot.scatter(x='Volatility', y='Returns', c='Sharpe Ratio',
                 cmap='RdYlGn', edgecolors='black', figsize=(10, 8), grid=True)
plt.scatter(x=sharpe_portfolio['Volatility'], y=sharpe_portfolio['Returns'], c='red', marker='D', s=200)
plt.scatter(x=min_variance_port['Volatility'], y=min_variance_port['Returns'], c='blue', marker='D', s=200 )
plt.xlabel('Volatility (Std. Deviation)')
plt.ylabel('Expected Returns')
plt.title('Efficient Frontier')
plt.show()
```

Ed otterremo poiché seguenti risultati.



Visualizziamo ora poiché due portafogli:

```
print(min_variance_port.T)
print(sharpe_portfolio.T)
```

```
28530
Returns      0.027493
Volatility   0.122322
Sharpe Ratio 0.224760
AAPL Weight  0.102443
GM Weight    0.151027
GE Weight    0.050100
TSLA Weight  0.100252
ALGN Weight  0.031474
XOM Weight   0.222589
SPG Weight   0.037362
AIG Weight   0.304753
19460
Returns      0.346517
Volatility   0.186429
Sharpe Ratio 1.858711
AAPL Weight  0.174135
GM Weight    0.037662
GE Weight    0.461866
TSLA Weight  0.002236
ALGN Weight  0.034526
XOM Weight   0.176810
SPG Weight   0.038077
AIG Weight   0.074688
```

Un investitore avverso al rischio sceglierà il portafoglio con la volatilità più bassa, perché corrisponde ad un livello di rischio potenziale più basso, in questo caso particolare del 12.23%

Un investitore disposto ad assumersi rischi più elevati avrebbe potuto scegliere il portafoglio con la volatilità più alta, in questo caso del 18.64%.

# Implementazione di un tool per la valutazione di asset

## Soluzioni attualmente in uso

Per analizzare i servizi che utilizzano Intelligenza artificiale nell'ambito finanziario bisogna suddividerli in quelli utilizzati dai professionisti del settore e quelli disponibili agli investitori non professionisti.

Tra i sistemi utilizzati dai professionisti il più importante e completo è Bloomberg Terminal.

Questo è un sistema hardware e software che permette ai suoi utenti di accedere al servizio di dati di Bloomberg, per ottenere dati in tempo reale da tutto il mondo le ultime notizie e previsioni di mercato da più di 15000 fonti. Bloomberg terminal è basato sul cloud e questo riduce l'incidenza di errori di sistema e attraverso il remote access è possibile accedere anche al di fuori della propria postazione di lavoro.

Una parte molto importante del terminal è la possibilità di riconoscere di cosa l'utente abbia bisogno senza utilizzare comandi specifici per ogni funzione. Il Terminal offre inoltre la possibilità di una live chat tra investitori e qui possono avvenire anche degli scambi.

È possibile utilizzare il Bloomberg Terminal per effettuare transazioni finanziarie. Questo servizio è molto intuitivo ed è il più completo per poiché professionisti del settore. La barriera di ingresso per usufruirne è il costo che oscilla tra poiché 20000\$ e 25000\$ per postazione. Per questi suoi costi molto alti il Terminal è usato di solito da grandi investitori istituzionali, portfolio manager e analisti finanziari.

Ci sono altre alternative a Bloomberg Terminal come Koyfin, Thomas Reuters Eikon, FactSet, Capital IQ e Symphony, ma, nonostante abbia numerosi competitor, la piattaforma di Bloomberg è diventato lo standard di settore.

Per gli investitori non professionisti, che desiderano reinvestire poiché propri risparmi e che non hanno tempo o hanno poche conoscenze nell'ambito della finanza sono disponibili sul mercato diverse soluzioni di robo advisor. Un robo advisor è un prodotto software che permette di gestire i propri investimenti senza il bisogno di consultare un financial advisor. Il primo passo per un nuovo cliente è registrarsi e compilare un questionario dal quale un algoritmo potrà dedurre la propensione al rischio e suggerire i portafogli corrispondenti a tale parametro, e il cliente dovrà confermarne la sua scelta. Avviato il processo di investimento il software può ribilanciare autonomamente il portafoglio, in modo da riavvicinare il proprio portafoglio agli obiettivi selezionati dal cliente. Di solito queste soluzioni si basano sull'impiego di mutual funds e ETF più che su azioni, infatti, di solito tendono a seguire un indice o modelli ibridi che creano soluzioni più complesse ad esempio con la selezione di azioni, sempre nel rispetto del profilo di rischio target.

Dal 2012 questa categoria di prodotti è presente in Italia grazie al servizio Moneyfarm. Questo offre servizi di gestione patrimoniale e di fondi pensione. È possibile monitorare lo sviluppo del proprio portafoglio sia su piattaforma mobile che desktop. Inoltre, questo servizio offre un consulente dedicato per aggiornamenti sul portafoglio e per aiutare sulle scelte di investimento.

Secondo la CONSOB nel documento “Digitalizzazione della consulenza in materia di investimenti finanziari” la consulenza automatizzata effettuata da un Robo Advisor è un vero e proprio servizio di

consulenza e deve pertanto essere operato da una piattaforma che sia autorizzata dalla CONSOB. Chi eroga questo servizio, deve inoltre rispettare le regole di base affiche l'utente possa accedere senza problemi alla piattaforma. I Robo Advisor sono quindi responsabili, come un intermediario finanziario, della qualità del servizio erogato e di una consulenza non corretta

In Italia oltre a Moneyfarm ci sono altre soluzioni come: eToro, Onlinesim, Tinaba, Yellow Advice e Euclidea.

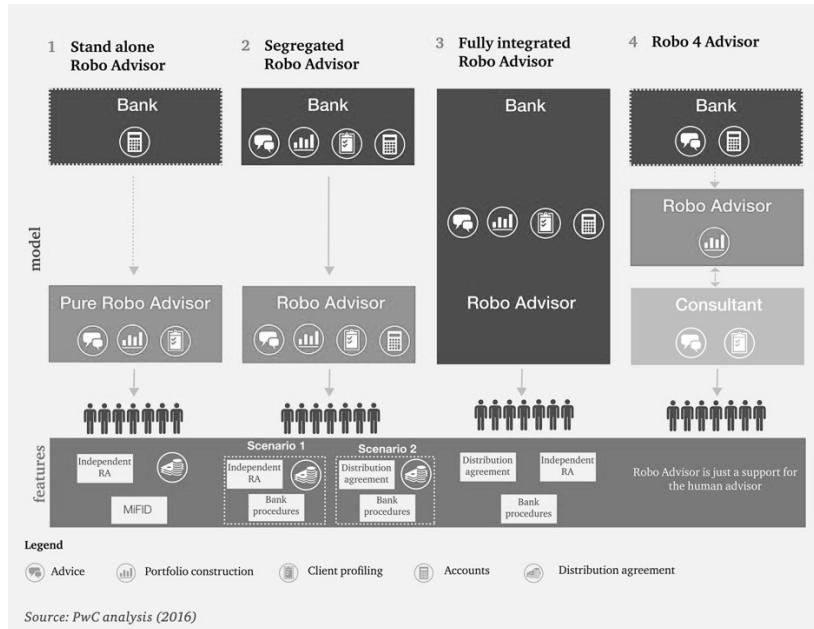


Figura 9. Diverse tipologie di robo advisor

Fonte: Pwc Analysis, (2016)

## Estrazione di dati da report trimestrali, bilancio e fonti esterne

Un report trimestrale è un insieme di rendiconti finanziari pubblicati da una azienda ogni trimestre, da inviare alle autorità di vigilanza competenti. Inoltre, questo può offrire delle comparazioni con le performance del trimestre precedente o dello stesso dell'anno precedente.

Negli USA le public traded company sono obbligate a comunicare i propri report alla Securities and Exchange Commission. (SEC)

Il calendario finanziario di molte aziende termina il 31 Dicembre, ma questo è a discrezione della stessa azienda in quanto per esempio per Apple Inc. (AAPL) l'anno fiscale parte da ottobre a settembre dell'anno successivo. I report trimestrali devono essere inviati entro poche settimane dalla fine del trimestre

	Azienda con l'anno fiscale Gennaio - dicembre	Apple Inc. Con l'anno fiscale ottobre – settembre
Fine Q1	31 Marzo	31 Dicembre
Fine Q2	30 Giugno	31 Marzo
Fine Q3	30 Settembre	30 Giugno
Fine Q4 (fine dell'anno fiscale)	31 Dicembre	30 Settembre

Tabella 1. Esempio di calendario dei trimestri delle aziende, accanto calendario trimestri Apple

Fonte: A. Hayes (2020)

I report trimestrali includono i dati finanziari principali di un'azienda come il reddito lordo, il profitto netto, le spese operative e i cash flow.

La Security and Exchange Commission richiede che i report debbano essere riportati secondo il modello Form 10-K per i report annuali e secondo il Form 10-Q per i report trimestrali entro sessante giorni dalla fine del periodo.

I report trimestrali sono di solito accompagnati da un'introduzione del management dove i Key Performance Indicators (KPIs) sono presentati ad investitori ed analisti. Di solito sono presenti anche delle previsioni di risultati ed obiettivi futuri.

Il Form 10-Q per report trimestrali è composto essenzialmente da due parti. La prima parte contiene le informazioni finanziarie rilevanti del periodo e la seconda contiene tutte le altre informazioni pertinenti

Il report di tipo Form 10-Q dà la possibilità di valutare quali sono i cambiamenti che stanno avvenendo all'interno dell'azienda, ancor prima della pubblicazione dei guadagni trimestrali. Ogni trimestre gli analisti ed investitori attendono l'annuncio dei guadagni delle compagnie. La pubblicazione di un report, specialmente da parte di azioni ad alta capitalizzazione, può portare a delle fluttuazioni del mercato. Per un'azienda non è importante solo dimostrare la sua abilità di crescere rispetto all'anno o al periodo precedente, ma anche l'abilità di rispettare le stime effettuate dagli analisti. Anche in presenza di crescita è possibile che un'azione presenti delle oscillazioni negative, nel caso non riesca a rispettare le previsioni degli analisti.

Per ottenere i dati dai report trimestrali è possibile utilizzare un algoritmo di tipo web scraping. Questo tipo di algoritmi vengono usati per estrarre dati da pagine web e documenti online. Ciò può essere fatto sia manualmente che in modo sistematico, utilizzando un bot o un web crawler. Il riconoscimento dei contenuti da estrarre è ottenuto attraverso la ricerca di tag nella pagina, e, nel momento in cui corrispondano a quelli impostati, il sistema provvederà a memorizzarli in un database. Quando i dati sono stati caricati su un database, altri software, dopo aver stabilito una connessione con il database stesso, potranno utilizzarli per effettuare altre operazioni.

In Python sono presenti numerosi framework che permettono allo sviluppatore di ridurre drasticamente il tempo di sviluppo di strumenti di questo tipo. Uno di Questo è Scrapy.

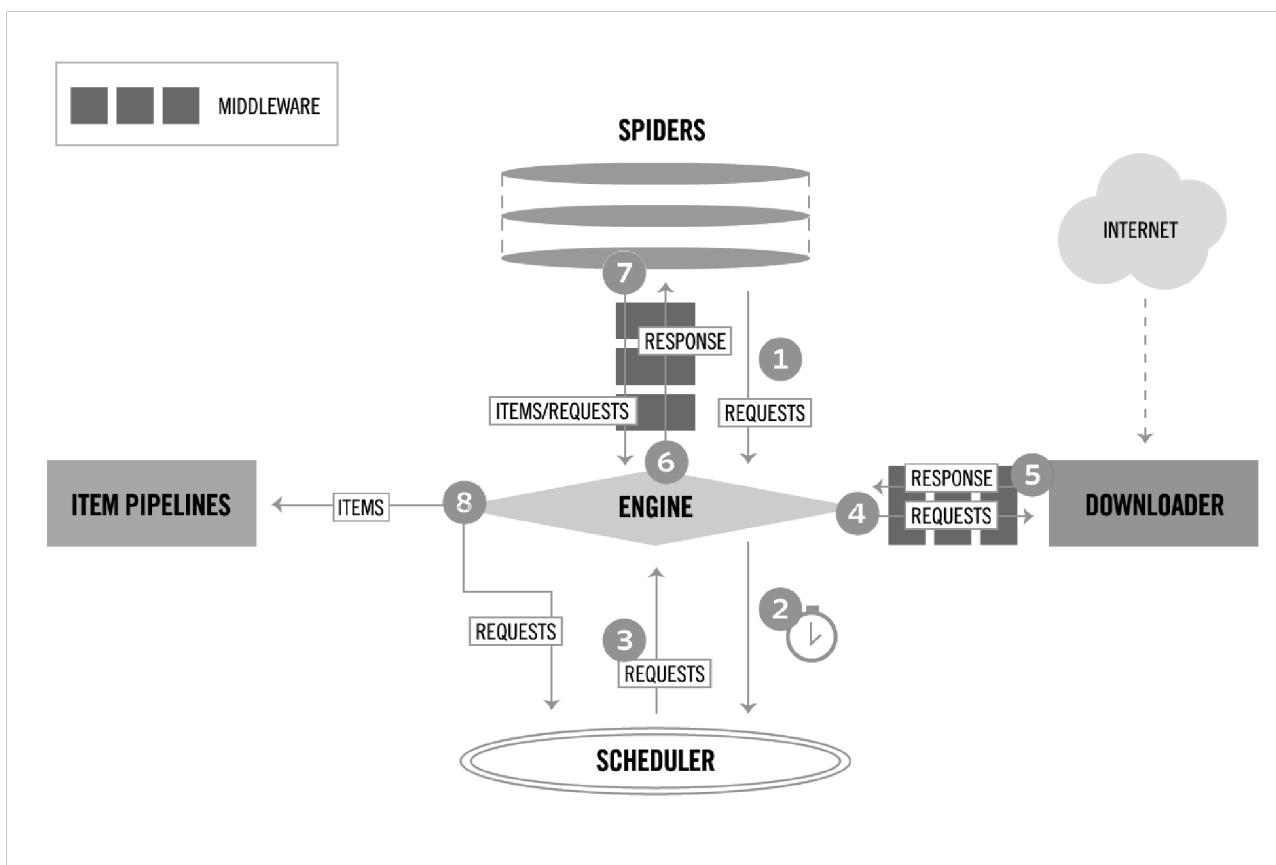


Figura 10. Componenti principali dell'architettura di scrapy

Fonte: [scrapy.org](http://scrapy.org)

Nella creazione del Bot i rapporti finanziari di cui tener conto sono

- Debt-to-Equity( serve a quantificare la leva finanziaria dell'azienda. Questo indica il rapporto tra debito ed equity usato dall'azienda per finanziare i propri asset)

$$\frac{\text{passività totali}}{\text{patrimonio netto}}$$

- Margine di profitto netto (Questo rapporto indica quanto l'azienda è capace di gestire i costi rispetto al profitto)

$$\frac{\text{reddito netto}}{\text{patrimonio netto}}$$

- Return on Equity ( Questo indica l'utile netto rispetto al patrimonio netto)

$$\frac{\text{reddito netto}}{\text{entrate totali}}$$

## **Sentiment Analysis con l'impiego di web scraping.**

Nel mondo degli investimenti finanziari è molto importante la tempistica di un'operazione, infatti pochi decimi di secondi possono corrispondere a migliaia di euro di perdita potenziale. Prima della diffusione di internet, i trader si affidavano a connessioni personali, giornali e media tradizionali per intraprendere le proprie decisioni. Con la nascita dei social media questa situazione è cambiata ed è molto più facile monitorare il sentimento di mercato in modo da poter fare delle previsioni in tempi più ristretti e migliori. Per un singolo investitore può essere difficile analizzare una mole di dati così grande e continua, e quando specialmente possono presentarsi dei problemi di rilevazioni errate o interpretazioni soggettive.

La sentiment analysis o emotion AI permette di identificare, estrarre e quantificare informazioni soggettive. L'esempio più semplice di sentiment analysis è la polarity detection, che si occupa del classificare informazioni come positive, negative o neutrali.

La sentiment analysis si basa sull'Natural Language Processing (NLP) che si occupa del processare testo in qualsiasi forma, trasformandolo in modo tale da essere compreso da un calcolatore. In questo caso è possibile osservare intenti ed entità. Gli intenti corrispondono all'area di interesse generale del testo, mentre le entità tendono a contestualizzare gli intenti.

La sentiment analysis è molto impiegata dai reparti marketing delle grandi e medie compagnie, in quanto è possibile ricavare l'opinione che il pubblico ha dei prodotti di un'azienda. Dal punto di vista finanziario questa tecnica risulta molto più importante sotto due punti di vista.

Da un lato analizzando le opinioni dei clienti sui social è possibile fare una previsione orientativa della crescita delle vendite, sia sul breve che sul medio termine. D'altro canto, è possibile analizzare attraverso l'analisi del management, delle principali fonti di informazione finanziaria e delle opinioni degli investitori rilevanti, l'effettivo grado di fiducia o sfiducia del mercato.

Rispetto ad un essere umano, un prodotto software di questo tipo permette di monitorare in tempo reale la diffusione di notizie da fonti selezionate e affidabili e, attraverso un'analisi del contenuto, di ricavare il grado di positività o negatività di una notizia, nei confronti di un prodotto o di un'azienda, potendo prevedere, in base al numero di notizie con gradi di fiducia concordanti, possibili movimenti del mercato.

Dalla sua diffusione Twitter viene visto ed usato come strumento per esprimere le opinioni in merito al mercato azionario. Su questa piattaforma si può assumere che per una convenzione informale venga usato il prefisso "\$" per parlare di azioni. Questo tag rende più semplice ottenere i dati di un particolare titolo di interesse.

Per realizzare un sistema di questo tipo è possibile suddividere questo algoritmo in quattro parti :

- Tweet Extractor
- Language Detector
- Non duplicate detector
- Sentiment classifier

Il modulo Tweet Extractor utilizza le Twitter Search API per ottenere i tweet delle azioni selezionate.

Il modulo Language Detector si occupa invece del rilevare la lingua utilizzata nel tweet, per semplificare l'algoritmo possono essere presi in considerazione i tweet scritti in lingua inglese. Un metodo per ottenere un classificatore di questo tipo è utilizzare un algoritmo basato sul modello n-gran.

Il modulo Near Duplicate Detector si occupa del trovare tweet che sono duplicati e eliminarli dall'analisi. Questo permette che i dati risultino più veritieri in quanto i duplicati alterano la percezione del mercato.

Il modulo sentiment classifier è la parte principale dell'analisi e può essere sviluppato utilizzando dei servizi esterni come Google Cloud Natural Language API, che possono offrire dei modelli già allenati e nel caso di sviluppo di un nuovo modello, si può configurare come un problema di classificazione di tipo SVM. In questo caso bisognerà allenare il modello a riconoscere il grado di positività o negatività.

Un sistema di questo tipo può essere utilizzato anche per analizzare articoli e informazioni sul web. Può essere implementato un algoritmo di tipo quantitativo in base al numero di news in merito ad un determinato prodotto finanziario nelle ultime 24 ore che implicherebbe una maggiore variabilità nel breve periodo e, attraverso l'analisi del tono degli articoli, è possibile analizzare anche la direzione dello spostamento previsto del prezzo.

Un'analisi di tipo qualitativa implica l'analisi dei maggiori media alla ricerca di informazioni riguardo aziende nel periodo di analisi. In questo modo, raggruppando anche più fonti è possibile analizzare all'uscita delle news quale titolo potrà essere apprezzato.

Un esempio di questo effetto è stato esposto in una ricerca pubblicata nel 2001 dal Journal of Finance. Il New York Times dedicò una pagina intera su un'azienda del settore delle biotecnologie riguardo i progressi effettuati nel campo dei trattamenti per il cancro. Anche se nessuna delle scoperte era avvenuta nel periodo il prezzo delle azioni salì da 12 a 85 dollari . Questo è un esempio dell'effetto dei media sulle dinamiche di comportamento degli investitori.

## Algorithmic trading

L'algorithmic trading è una tecnica che usa degli applicativi per automatizzare il processo della compravendita di strumenti finanziari.

Ci sono molti esempi di strategie di trading. Le strategie più popolari sono:

- Mean reversion
- Momentum
- Sentiment based
- Statistical arbitrage

In una strategia di tipo **mean reversion** una serie stocastica di prezzi si considera che torni a convergere con la sua media mobile. Si dice che l'azione si trova in un trend rialzista se il prezzo risulta superiore rispetto alla media mobile, si dice invece che è in un trend ribassista se i valori assunti dal prezzo sono inferiori alla media mobile.



Figura 11. Algoritmo di tipo mean regression

Fonte: International Journal of Computer Science and Network, 2015

Una strategia di tipo **Momentum** consiste nell'avere una posizione di tipo long per azioni, futures o ETF, nel caso in cui sia possibile vedere un trend rialzista e una posizione di tipo short, nel caso di un trend di tipo ribassista. Di solito questa tipologia di trend può durare molto tempo e può convenire mantenere la posizione fino alla fine del trend. Solitamente in questo caso sono impiegate due medie mobili rispettivamente di 30 e 90 giorni. Nel caso la media mobile dei 30 giorni superi quella dei 90 giorni è consigliabile comprare, nel caso contrario è consigliato vendere.



Figura 12. Esempio dell'applicazione della Momentum Strategy

Fonte: LaptrinhX, 2018

Un'analisi sentiment-based consiste nell'uso dell'NLP e nell'analisi del testo per determinare informazioni soggettive o lo stato emozionale di chi scrive un qualsiasi testo.

Questa tipologia di algoritmi si basa sempre su un sistema di mean regression, però in questo caso al posto di una media mobile del prezzo di uno strumento, utilizza la media mobile dell'indicatore delle opinioni di mercato.

Uno strategia di statistical arbitrage è una tecnica che permette di osservare i trend delle fluttuazioni dei prezzi e le differenze di prezzo tra strumenti finanziari. Ci sono diverse strategie di statistical arbitrage:

- Market Neutral Arbitrage
- Cross Asset Arbitrage
- Pairs trading

Il Market Neutral Arbitrage consiste nell'avere simultaneamente una posizione long e una short su asset sotto e sopra valutati . L'obiettivo di trarre profitto dalle due posizioni antitetiche.

Il Cross Asset Arbitrage analizza le discrepanze tra uno strumento derivato e il sottostante come per esempio un indice azionario future e le azioni di questo indice. Una sottoclasse di questa tecnica è l'ETF Arbitrage che identifica discrepanze tra il valore di un ETF e l'asset sottostante.

Il Pairs trading consiste invece nel raggruppare le azioni in gruppi di due in base a similarità di pricing o dissettore. Quando un'azione ha delle performance migliori dell'altra, quella che ha delle performance peggiori viene acquistata. La posizione è protetta da cambiamenti o movimenti di mercato vendendo l'azione che aveva avuto un trend positivo.

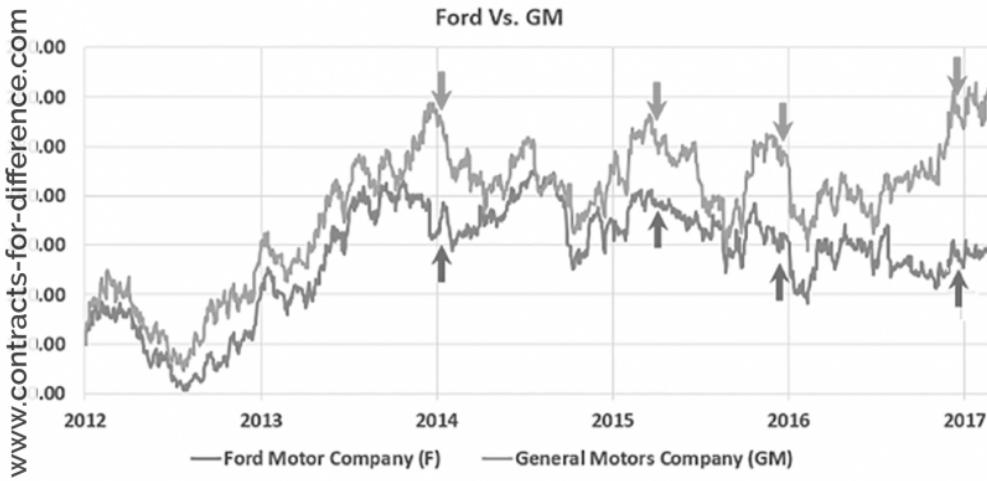


Figura 13. Esempio dell'applicazione di una Pairs trading strategy

Fonte: LaptrinhX, 2018

Secondo la direttiva MiFID II una società di trading basata su algoritmi dovrà disporre di sistemi efficaci di controllo dei rischi per garantire che i suoi sistemi di negoziazione siano resistenti a situazioni di estrema volatilità del mercato, pertanto devono disporre di sistemi efficaci con soglie limite, che impediscano l'invio di ordini errati. Ed è strettamente vietato che questi sistemi vengano usati per scopri contrari alle regole di una sede di negoziazione. Queste società devono far fronte a eventuali guasti del sistema e garantire che i loro sistemi siano testati e monitorati e inoltre devono comunicare l'uso di tali tecnologie all'autorità competente.

I requisiti organizzativi impongono che questi algoritmi debbano essere controllati in un ambiente simulato e le imprese saranno inoltre tenute a monitorare i loro sistemi processi e procedure per identificare e annullare eventuali impatti negativi che possono avere gli algoritmi e per poter annullare tutti gli ordini in sospeso, attraverso un comando di arresto.

## **Implementazione di un tool basato su intelligenza artificiale**

L'algorithmic trading è utilizzato per segnalare trend nei movimenti di prezzo ed effettuare acquisti in modo automatico nel momento in cui si verifichino determinate condizioni con un sistema di tipo IF/ELSE.

Queste tecniche sono utilizzate maggiormente per il trading nel breve periodo e di tipo high frequency perché le anomalie di mercato tendono a scomparire nel medio lungo termine. Attraverso questi algoritmi, in particolare con quelli di mean reversion, sentiment-based, è possibile determinare dei trend a medio termine di mercato. L'algorithmic trading ha bisogno di una grande quantità di dati storici per effettuare delle previsioni affidabili ed ha la necessità inoltre di ottenerli con pochissima latenza.

La gestione dell'estrazione dei dati è un processo che può essere effettuato sia attraverso l'utilizzo di provider di dati esterni che offrono API pubbliche, che con l'impegno di un processo di data mining basato su algoritmi di web scraping. Poiché è un prodotto per investimenti a medio e lungo termine, la latenza nella trasmissione dei dati non è un parametro da considerare. Un primo problema che si incontra è la mole di dati da ricevere e sulla quale effettuare le analisi. Per evitare che l'ottenimento e l'analisi di tutti i dati venga effettuata da un solo calcolatore, i processi verranno divisi in diversi container.

I dati ottenuti vengono inviati ad un modulo che si occuperà dell'analisi dei dati con un algoritmo di mean regression. Un algoritmo di questo tipo permette una previsione dell'andamento di un'azione.

Una volta selezionate le azioni soggette a dei trend rialzisti, in quanto si sta progettando il portafoglio, bisogna procedere alla convalida delle informazioni.

La convalida consiste nel raccogliere informazioni da diverse fonti.

Attraverso un'analisi dei media più importanti si procederà con una sentiment-based analysis. Un'analisi di questo tipo permette di analizzare il livello di confidenza del mercato nei confronti di un titolo.

Per aumentare la qualità delle informazioni bisogna stabilire un peso diverso per ogni fonte. Dopo questa fase si procede con uno scraping dei report delle società, per un'analisi strutturale per ottenere gli indici sintetici che permettono di valutare l'effettiva efficacia dell'investimento.

Selezionate le azioni con delle previsioni di crescita si formeranno dei cluster in base al settore di appartenenza e alla volatilità della stessa.

La volatilità permette di analizzare il rischio di uno strumento tenendo conto della serie storica dei prezzi.

La divisione per settore permette di minimizzare il rischio specifico di settore quando si procederà alla costruzione dei portafogli.

Dopo aver raggruppato i dati si potranno osservare quali settori sono più promettenti, in modo da aggiustarne il peso nel portafoglio.

A questo punto si procede con la simulazione dei portafogli ottenuti generando in modo casuale i pesi dei diversi asset, ma mantenendo posizioni più rilevanti nei settori che presentano alti livelli di crescita.

Successivamente si costruisce la frontiera efficiente e considerando rispettivamente i portafogli con il maggiore indice di Sharpe e con la minore variabilità.

Potrebbe verificarsi il caso che siano selezionabili due portafogli. e sarà selezionato quello conforme al profilo di rischio precedentemente indicato dal cliente che dovrà validare la sua scelta, specificando la frequenza di ribilanciamento del portafoglio.

Nel momento in cui ci sia un aumento improvviso della variabilità del prezzo o delle opinioni di mercato, il sistema deve segnalare che il mercato è instabile, specificando il livello di fiducia del mercato. Nel momento in cui non c'è fiducia e il trend è ribassista, bisogna bilanciare il portafoglio. Contemporaneamente un operatore riceverà una segnalazione da parte del sistema e sarà tenuto a verificare se questo comportamento sia dovuto a dinamiche reali di mercato o ad un errore di sistema.

Nel caso sia dovuto ad un errore di sistema l'operatore sarà tenuto a bloccare la possibilità di fare ordini e ribilanciare il portafoglio e il servizio non sarà più utilizzabile fino alla verifica del corretto funzionamento del sistema. In tal caso il supporto sarà tenuto a contattare tutti gli utenti in modo da segnalare il problema, che nello stesso tempo saranno affiancati da un consulente da remoto che possa aiutarli nella gestione del portafoglio fino al momento in cui l'intera infrastruttura non torni ad un funzionamento ottimale.

Tutti i flussi informativi generati al sistema, per rispettare le condizioni espresse dalla direttiva MiFID II, devono essere controllati da operatori qualificati. Questi verificheranno il funzionamento ottimale degli algoritmi nelle condizioni del mercato. Inoltre, il sistema selezionerà, in modo del tutto casuale, i portafogli dei clienti con i rispettivi dati storici per osservare che tutte le operazioni siano state svolte nel pieno interesse del cliente e non siano presenti dei difetti.

Per aderire agli obblighi informativi il sistema dovrà avvertire il cliente in caso di ribilanciamento specificandone il motivo e le dinamiche del mercato.

Il sistema presenterà una partecipazione attiva di consulenti per:

- Creare un rapporto con il cliente
- Verificare il corretto funzionamento della piattaforma
- Poder spiegare al cliente le operazioni effettuate

Un sistema di questo tipo ha tre tipologie diverse di utenti:

- Utenti interessati a una gestione attiva del portafoglio
- Consulenti finanziari che desiderano offrire soluzioni migliori e personalizzate ai propri clienti
- Utenti che desiderano informazioni di mercato

Le caratteristiche essenziali che deve avere sono:

- Ottenere ed analizzare dati in real time
- Offrire soluzioni per ogni livello di rischio
- Offrire prodotti che rispecchiano le esigenze del cliente
- Sicurezza

Per rendere il sistema più fruibile da parte del cliente e del consulente è necessaria l'implementazione di uno strumento di interazione, con il quale spiegando cosa desidera l'utente sia capace di ottenere le informazioni anche senza l'utilizzo di termini tecnici e senza dover imparare nomi di funzioni.

In questo caso è molto utile un sistema basato su natural language processing (NLP), che con un adeguato allenamento possa riconoscere gli intenti dell'utente.

Nel caso in cui non sia disponibile una risposta alla domanda con l'utente, questo sarà indirizzato ad un live chat con uno specialista di supporto dedicato.

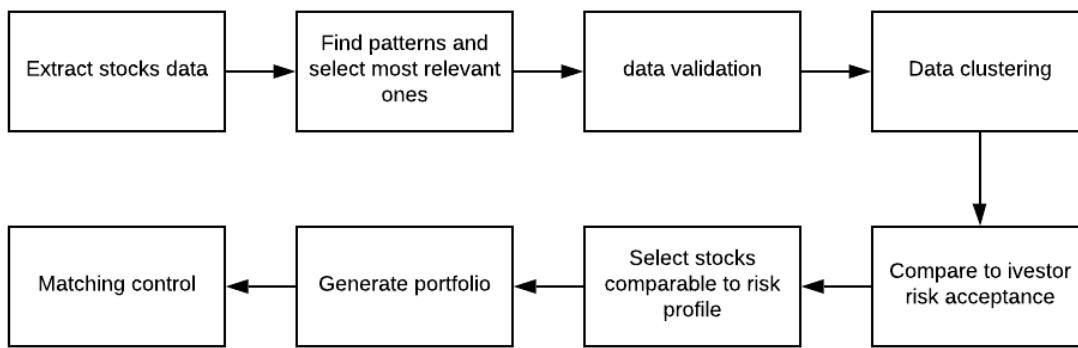


Figura 13. Diagramma rappresentativo della struttura del prodotto

## Criticità delle soluzioni attualmente in uso

L'introduzione di nuovi strumenti digitali nel campo della finanza e delle banche ha portato ad una trasformazione dei canali retail, infatti, secondo uno studio dell'Osservatorio Fintech & Insurtech della School of Management del Politecnico di Milano (autori e anno) il numero di sportelli è calato del 5% e le postazioni self-service compongono l'11% delle operazioni effettuate. Molto rilevante è l'incremento dell'utilizzo di servizi di home banking che arrivano a corrispondere al 20% delle operazioni effettuate. Nello stesso tempo emerge anche un cambiamento nei comportamenti dei clienti grazie alla diffusione in massa di smartphone e all'utilizzo di interfacce grafiche molto intuitive , Si registra un calo nell'utilizzo del PC per le operazioni finanziarie, in favore del mobile che corrisponde al 52%, arrivando ad essere la piattaforma più utilizzata dai clienti.

Cambiando e aumentando l'utilizzo di servizi fintech stanno cambiando anche le strategie distributive degli intermediari, che offrono servizi e filiali sempre più innovativi.

Gli intermediari cominciano a sentire la pressione di nuove startup innovative che offrono servizi digitali presentandosi come possibili alternative. In Italia, più di duecentocinquantamila italiani utilizzano salvadanai digitali, per accantonare piccole cifre ed impiegarle successivamente, e sono saliti a ventimila gli utilizzatori di Robo Advisor. Uno studio di Autore (2018) ha osservato su un campione di 1515 utenti tra i 18 e i 74 anni che le banche e i servizi postali restano di riferimento per la gestione dei risparmi.

Rispettivamente per il 65% e il 56% degli utenti.

I servizi fintech, nonostante si stiano diffondendo molto velocemente, trovano dei limiti sia da parte degli utenti che dal lato business.

I clienti desiderano risparmiare tempo potendo svolgere tutte le loro operazioni con sicurezza e velocità, in particolare possono essere automatizzate le operazioni di sportello e di rendicontazione. Nel settore dell'investment banking i clienti risultano ancora molto legati ai propri consulenti, I servizi di robo advising si adattano meglio alla gestione di patrimoni più contenuti, a cui i detentori sono disposti a lasciare un ampio grado di delega. Presentano inoltre dei vantaggi dal punto di vista di tassazione, che viene resa più vantaggiosa.

Le più grandi banche stanno investendo in questo settore anche se presenta dei problemi sia tecnici che di economicità. I robo advisor offrono servizi con costi di gestione nettamente più bassi rispetto a soluzioni tradizionali dal punto di vista del cliente. Questi però risultano essere poco profittevoli per le società, impegnate già nell'offrire servizi tradizionali, per i costi di sviluppo, verifica e gestione.

In base al profilo di rischio del cliente, determinato in seguito ad un questionario obbligatorio compilato all'inizio del rapporto, sono suggeriti portafogli in base alle caratteristiche di ogni singolo cliente e quest'ultimo dovrà scegliere, solo tra i suggeriti, quello che rispecchia al meglio la sua visione e necessità di investimento. A titolo esemplificativo Money Farm ha 14 tipologie, diversi di portafogli divisi in base al profilo di rischio e durata dell'investimento e, attraverso un processo automatizzato, seleziona i migliori per il profilo del cliente e questo dovrà scegliere uno tra i portafogli consigliati dal sistema.

Il problema sorge nel momento in cui vogliamo personalizzare il portafoglio. La maggior parte delle soluzioni attualmente sul mercato propongono una gestione passiva basata su ETF in modo da essere più trasparente e abbassare i costi di gestione.

Viene inoltre a mancare il supporto da parte del consulente nel prendere decisioni e di adattare il portafoglio in base a cambiamenti rispetto alle condizioni iniziali dell'investimento. Per questo motivo negli ultimi anni si sono sviluppate soluzioni di tipo phygital che riescano ad unire i vantaggi della tecnologia di analisi e supporto con la figura del consulente. In questo modo il cliente finale non si interfacerà con un tool basato su intelligenza artificiale, ma, attraverso una piattaforma o un contatto diretto, potrà parlare con un consulente che dalla sua parte avrà tutti i vantaggi dell'utilizzo di intelligenza artificiale nell'ambito della customer analysis e gestione portafoglio, ma che possa nello stesso tempo avere un rapporto di fiducia con il cliente finale.

## Conclusioni

I mercati finanziari sono sistemi molto complessi che generano enormi quantità di dati che devono essere monitorati costantemente sia dagli organi di sorveglianza, che da chiunque voglia effettuare delle previsioni. I mercati presentano inoltre una natura non deterministica, in quanto, non sono soggetti solo ad errori sistematici o umani, ma dipendono anche dai comportamenti dei singoli operatori e dalle opinioni che questi hanno in merito a prodotti finanziari o operatori.

Per questo motivo utilizzare delle tecniche di analisi tradizionale può portare a degli errori di previsione. Un sistema automatizzato, sviluppato, utilizzando algoritmi basati su intelligenza artificiale, può presentare numerosi vantaggi per tutti gli operatori di mercato. Non solo per un operatore è possibile analizzare un'enorme quantità di dati, ma è possibile creare un sistema che evolva insieme alle condizioni del mercato. Dal punto di vista operativo può diminuire i tempi di esecuzione di task ripetitive, come l'ottenimento di dati, che porta ad una riduzione del tempo di esecuzione, si ottiene una riduzione dei costi, in particolare in riferimento al personale, ed anche una velocizzazione di processi che porta ad ottenere dati più velocemente, cosa che risulta essenziale per sfruttare le imperfezioni del mercato.

L'utilizzo dell'intelligenza artificiale, proprio per le possibilità che questa offre in tutti i settori, è un trend in crescita negli ultimi anni, che continuerà anche sul lungo periodo.

Sistemi di questo tipo sono capaci di completare sempre più compiti che prima si pensavano impossibili da svolgere per una macchina, con sempre più efficienza e con possibilità di successo più alte rispetto ad un essere umano. Per questo motivo molte aziende e intermediari finanziari, come J. P. Morgan, KPMG, Deloitte, Intesa San Paolo e Credit Suisse, hanno reparti di R&D che si occupano nello sviluppare soluzioni basate su intelligenza artificiale, sia per offrire soluzioni innovative per i propri clienti, che per automatizzare processi interni, aumentandone l'efficienza. Risulta molto importante ed efficace per queste categorie essere aggiornati sugli ultimi sviluppi tecnologici, perché, nel caso vengano sviluppate dei prodotti disruptive, siano già pronti per reagire.

In questo studio sono state analizzate le principali applicazioni di tecnologie innovative basate su intelligenza artificiale nel settore della finanza, in particolare nel settore della gestione dei portafogli.

Il software progettato nell'ultimo capitolo rappresenta un'applicazione delle principali automazioni nel campo della gestione dei portafogli e del trading per offrire un servizio innovativo, da cui possano trarre vantaggio sia i gestori che gli utenti finali del servizio. Per uno sviluppo pratico si rimanda alle parti che trattano dei singoli moduli.

Particolare attenzione è da rivolgere ai singoli investitori e clienti. È molto importante riuscire a comprendere non solo il loro profilo di rischio, e creare una strategia di investimento che rispecchi solo questo dato, ma anche monitorare l'investitore costantemente, in modo da rispecchiare al meglio i suoi interessi e adeguarsi, nel caso di cambiamenti, alle sue preferenze. Per un cliente è importante non solo avere la possibilità di usufruire di un servizio che dà ottimi risultati, ma anche avere un'ottima esperienza di

utilizzo del prodotto, non solo dal punto di vista della facilità di utilizzo dello stesso, ma anche nel supporto che può dare nell'intraprendere decisioni o nel rispondere a domande che l'utente si pone

## Bibliografia

Deloitte(2019), “Artificial intelligence, The next frontier for investment management firms”.

V. Torres, T. Deason, M. Landrum, N. Lohria (2019), “A Machine Learning Model for Clustering Securities”.

S. Gautam (2019), “A Machine Learning Model for Clustering Securities”.

A. Vervuurt, Y. K. Samo (2016), “Stochastic Portfolio Theory: A Machine Learning Perspective”.

J. B. Heaton, N. G. Polson, J. H. Witte (2016), “Deep Learning for Finance: Deep Portfolios”

B. Gautam , S. Gupta , S. Awasthi , S. Gautam (2019), “Securities Analysis and Portfolio Management using Artificial Neural Networks”

Aurelien Géron (2019), “Hands-On Machine Learning With Scikit-Learn and Tensorflow: Concepts, Tools, and Techniques to Build Intelligent Systems”.

N. Albinson, D Krishna, Y. Chu (2018), “The algorithmic revolution is here”

E. Bajo (2018), “The Impact of the Media on Financial Markets”

M. Farhadloo, E. Rolland (2016), “Fundamentals of Sentiment Analysis and Its Applications”

PwC | G. Gialloreto, M. Marchiaro (2016), “Robo Advisory moves forward in Italy”

PwC | A. Milani (2019), “The role of risk and trust in the adoption of robo-advisory in Italy”

V. A. Kharde, S.S. Sonawane (2016), “Sentiment Analysis of Twitter Data: A Survey of Techniques”

A. Courville, I. Goodfellow, Y. Bengio (2015), “Deep Learning”

J. Klement, CFA (2015), “Investor Risk Profiling: An Overview”

P. Gomber, K. Zimmermann (2018), “Algorithmic Trading in Practice”

E. Colombari, R. Tedeschi (2019) Fintech e digital banking. Asset management, le nuove frontiere dell’automation

T. Erl (2013), “Cloud Computing: Concepts, Technology & Architecture”

A. B. Downey (2015), “Think Python: How to Think Like a Computer Scientist, 2nd edition”

L. Ackert, R. Deaves (2009), “Behavioral Finance: Psychology, Decision-Making, and Markets”

K. Adam, A. Marcet, J. P. Nicolini (2008), “Working Paper Series No 862 / February 2008”

A. A. Kirilenko, A. W. Lo (2013), “Moore's Law versus Murphy's Law: Algorithmic Trading and Its Discontents”

M. Meeuwis (2020), “Wealth Fluctuations and Risk Preferences: Evidence from U.S. Investor Portfolios”

T. Guida (2018), “Big Data and Machine Learning in Quantitative Investment”

F. Abraham, S. L. Schmukler, J. Tessada (2018), “Robo-Advisor, Investing through Machines”

## Sitografia

<https://builtin.com/artificial-intelligence/ai-finance-banking-applications-companies>

<https://www.investopedia.com/terms/r/roboadvisor-roboadviser.asp>

<https://docs.anaconda.com>

<https://docs.scrapy.org>

<https://www.investopedia.com/articles/personal-finance/050815/what-do-financial-advisers-do.asp>

<https://benalexkeen.com/k-means-clustering-in-python/>

<https://matplotlib.org/tutorials/introductory/pyplot.html>

<https://www.investopedia.com/articles/professionaleducation/11/bloomberg-terminal.asp>

<https://developers.google.com/machine-learning/guides>

<https://www.investopedia.com/articles/active-trading/101014/basics-algorithmic-trading-concepts-and-examples.asp>

<https://corporatefinanceinstitute.com/resources/knowledge/finance/financial-ratios/>