# UNIVERSITA' DEGLI STUDI DI PADOVA

**DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI**
**"M.FANNO"**

**CORSO DI LAUREA MAGISTRALE / SPECIALISTICA IN**
**BANKING AND FINANCE**

**TESI DI LAUREA**

**"ROBO-ADVISORS AND GENERATIVE ARTIFICIAL INTELLIGENCE:**
**A COMPREHENSIVE ANALYSIS"**

**RELATORE:**

**CH.MO PROF. BRUNO MARIA PARIGI**

**LAUREANDO: MILAN ENRICO**

**MATRICOLA N. 1179679**

**ANNO ACCADEMICO 2023 – 2024**

Il candidato dichiara che il presente lavoro è originale e non è già stato sottoposto, in tutto o in parte, per il conseguimento di un titolo accademico in altre Università italiane o straniere.
Il candidato dichiara altresì che tutti i materiali utilizzati durante la preparazione dell'elaborato sono stati indicati nel testo e nella sezione "Riferimenti bibliografici" e che le eventuali citazioni testuali sono individuabili attraverso l'esplicito richiamo alla pubblicazione originale.

*The candidate declares that the present work is original and has not already been submitted, totally or in part, for the purposes of attaining an academic degree in other Italian or foreign universities. The candidate also declares that all the materials used during the preparation of the thesis have been explicitly indicated in the text and in the section "Bibliographical references" and that any textual citations can be identified through an explicit reference to the original publication.*

Firma dello studente

# Index

# Introduction

Financial advisory has traditionally implied a direct interaction with the investor in order to determine optimal resource allocations, tailored to the customers' needs. The combination of finance and technology however, has consistently tried to redefine the landscape of investment advisory services. One of the main interactions of these two worlds could be identified in 2008, with the launch of the first robo-advisor platforms, marking a paradigm shift in investment advisory practices. These platforms, driven by complex artificial intelligence algorithms, provide suitable financial advices to customers through digital platforms and periodically rebalance their portfolios, after carefully identifying a specific risk profile for each investor. Even if initially their objective was to increase access to financial advice, promising cost-efficiency, accessibility and systematic portfolio management, they never really took off as expected, due to several challenges such as adaptability to market dynamics, limited personalization and lack of trust and financial education in some nations.
In recent years, in particular starting from 2022, generative artificial intelligence models like ChatGPT, draw the attention of several companies, opening new horizons for the financial industry and presenting a possible way to enhance actual robo-advisor systems. Exploiting the power of machine learning and natural language processing, generative artificial intelligence has the potential to increase overall efficiency of robo-advisors, mitigating the limitations present today and improving current performances. This research therefore, aims at giving a new perspective on plausible future implications derived from the integration of modern robo-advisors systems with generative AI, exploring potential synergies and the impact on the financial advisory system. Hence, the chapters of the thesis will delve into a comprehensive examination of robo-advisory, generative artificial intelligence and their interactions.
The first chapter provides an analysis of robo-advisors services, discussing their main features, comparing different platforms in the US and their respective performances. Next, market size, diffusion and a general framework of the current state of regulation are analyzed.
In Chapter 2, there is a general introduction to artificial intelligence in the financial sector, followed by an in-depth analysis of the revolution that generative artificial intelligence is bringing to the financial industry with empirical evidence of its capabilities. Further, different issues of generative AI are presented like privacy and hallucinations issues as well as its regulatory framework. Chapter 3 focuses on the main quantitative approaches used in robo-advisors services, like the main algorithms and financial concepts adopted, as well as genetic algorithms interactions with different existing models. In the last part of the chapter, the implementation of ChatGPT to standard methods used in robo-advisory is discussed, showing

the potential of the technology. Throughout the whole chapter several performance analyses are conducted in order to give robustness to the thesis. Finally, in Chapter 4, different case studies are investigated, from the current state of robo-advisors service in Italy to the adoption of generative AI tools by the biggest financial companies in the world like JP Morgan and Bloomberg, among others. In summary, this thesis aims at contributing to the ongoing debate on the future of financial advisory analyzing the possible relationship between robo-advisors and generative AI, giving insights of a potential new era of personalized automated financial services.

# Chapter 1 – Robo-Advisors

## 1.1 Key concepts of Robo-Advisors

In recent years, the world of financial services has been evolving continuously.
The implementation of digital technologies, like the use of Artificial Intelligence, and innovations concerning the management of big data, like Machine Learning, has created a variety of new opportunities in this field, drawing the attention of both newcomers and experts. A new feature in the financial industry evolution has certainly been automated portfolio management, also known as robo-advice.
A robo-advisor is a financial service (usually inside a digital platform) that offers automated and algorithm-based investment decisions and financial allocation, requiring minimal or zero human oversight. It employs investment algorithms, often referred to as “algos”, to evaluate millions of data points in order to allocate an investor’s money across the most suitable investment options, making continuous adjustments to ensure an ideal portfolio, aligned with the investor’s financial objectives. The procedure begins typically with a questionnaire aimed at understanding the financial situation, the objectives and the risk tolerance of the individual. Subsequently it utilizes this information to provide advices and manage the investments.
Robo-advisory platforms offer substantial cheaper services compared to traditional advisors, since there’s no need to compensate a human financial advisor. Moreover, they are also available to anyone who has access to an internet connection. The idea is that robo-advisors should outperform human advisors due to their constant availability, speed and lack of emotional decision-making. A recent study, published on the Harvard Business Review, revealed how these algorithms achieved better results compared to average novice investors and experienced investors with cognitive biases (Machet, 2021).
Similar to standard financial advisors, robo-advisors are subject to oversight by supervisory institutions, implying they have to act in the client’s best interests when they take investment decisions.

### 1.1.1 History

In the past, only individuals with significant level of wealth could have access to customized financial advice due mainly to the high cost of investment advisors and to the minimum deposit requirement, usually greater than what an average household could afford. Starting from the global financial crisis in 2008, where several financial institutions such as Lehman Brothers declared bankruptcy, there has been a decline in trust towards banks and financial

intermediaries in general. This allowed startups to have more credibility and to provide an alternative to standard investment methods, using innovative technologies.
Robo-advisors made their debut in 2008, and in particular, Betterment and Wealthfront have been the first pioneers in this field. The former tried to help to manage the investments through an online interface, with the purpose of rebalancing assets of target-date funds (funds that are riskier while young and gets more conservative over time) while the latter, aimed to serve the tech community, recognized the potential of computer software in making more accessible investment advices.
However, the technology itself was not groundbreaking, because software for the automated asset allocations had been used by wealth managers since the early 2000s. By the way until Betterment and Wealthfront launch, this software was available exclusively to wealth managers, so that average investors had to engage a financial advisor to access this innovation. Since then, the industry has experienced an incredible growth: in 2022 the biggest robo-advisor by assets under management (AUM) was Vanguard Digital Advisor, with $140.7 billion, followed by Schwab Intelligent Portfolios with $70.3B AUM, Betterment with $36B and Wealthfront with $34B (Frankenfield, 2023; Nasr, 2023; Vanguard, 2023).

### 1.1.2 Main Features

Robo-advisors mainly utilize Modern Portfolio Theory (MPT) to construct investment portfolios for their users. MPT’s objective is to maximize portfolio returns and to minimize the risks through diversification. It can be considered as an application of the principle of not concentrating all the investments in a single financial product.
Diversifying investments across different asset types, allow MPT to increase the probability that when some investments face losses, other will appreciate. The objective is obviously to maintain the portfolio in an upward trend, even during periods of market volatility.
Once portfolios are set up, robo-advisors consistently oversee them to guarantee the optimal allocation of asset classes, regardless of market fluctuations, using rebalancing bands (Frankendiel, 2023; Lam, 2016; MCClellan, 2016). Besides diversification in fact, the majority of robo-advisors employ portfolio rebalance (through rebalancing bands) and tax-loss harvesting methods. The first one ensures that an individual hold the appropriate balance among investments in order to achieve his objectives, as market dynamic changes.
The second one instead, usually tries to reduce long term capital gains taxes.

### Rebalancing Bands

In rebalancing bands, each asset class is assigned a target weight and an associated tolerance range. For instance, an investment strategy could require to maintain 30% in emerging market equities, 30% in domestic blue chips and 40% in government bonds, with a tolerance range of ±5% for every class. Taking in consideration a ±5% corridor, it would imply that emerging market and blue-chip assets can vary from 25% to 35%, while government bonds can range from 35% to 45%. If the weight of an asset exceeds this corridor, the portfolio is rebalanced to recreate the initial target weight allocation.

Historically, this kind of rebalancing practices were discouraged due to its time-consuming nature and its associated transaction fees. However, modern low-fee robo-advisors are meant to manage rebalancing automatically. Nowadays in fact, robo-advisors fees are calculated on assets under management (AUM) and they are way cheaper than human advisors, charging around 0.3% of AUM per year, against the 1% of traditional advisors (Frankenfield, 2023; VettaFi, 2021; Wealth Harness, 2023).

### Tax-Loss Harvesting

A different kind of rebalancing frequently used in robo-advisory services, made efficient using algorithms, is tax-loss harvesting. Tax-loss harvesting is a strategy that implies selling assets at a loss to counterbalance capital-gains tax liability.

Generally, this approach is employed to reduce the realization of short-term capital gains. Robo-advisors accomplish this by holding multiple ETFs for every asset class. For instance, if the value of an S&P 500 ETF declines, a robo-advisor will automatically sell it to securer a capital loss and at the same time, it purchases another S&P 500 ETF. It is important to notice that the IRS wash-sale rule does not allow investors to buy the same asset or one that is basically identical, within 30 days after selling it. In this case, it’s crucial to ensure that the robo-investment platform comply with these regulations. Robo-advisors should in fact choose proper ETFs in order to avoid wash-sale violations (Frankenfield, 2023).

### 1.1.3 Choosing between platforms

The cost and features of robo-advisors can differ in each platform, and it should be also considered that every platform has its own strengths and weakness. Among the available robo-advisor it is possible to find hybrid robo-advisors, which integrate the advantages of automated investment with a human guidance.

There are also cases where it is possible to link the bank account with the robo-advisor account, simplifying the process of depositing funds.

Numerous digital platforms focus on specific demographics targets, in particular Millennial and Generation Z investors, since they are the most familiar with technology and they accumulate investable assets. These individuals are usually more willing to share personal information online and rely on technology for essential tasks. As a result, robo-advisory firms often use social media in their marketing strategies to attract these investors (Frankenfield, 2023; Isaia, 2022; Tan, 2023).

A great variety of robo-advisors are accessible in the US and globally, with new ones entering every year. Each of them offers different combinations of retirement planning, investment management and general financial advice. In Figure 1, the overall ratings of the best robo-advisors in the US are presented, according to Investopedia research. Here 18 companies offering robo-advisors services in the US are evaluated on a scale from 0 to 5 (through a combination of subject matter expertise, industry research and company survey data) and on a range of different categories like portfolio management, goal planning, portfolio contents, user experience and account services, individuating the best robo-advisors services (Sacchitello, 2023). In particular, Wealthfront is considered the best overall due to sophisticated financial planning and optimal portfolio construction and management.

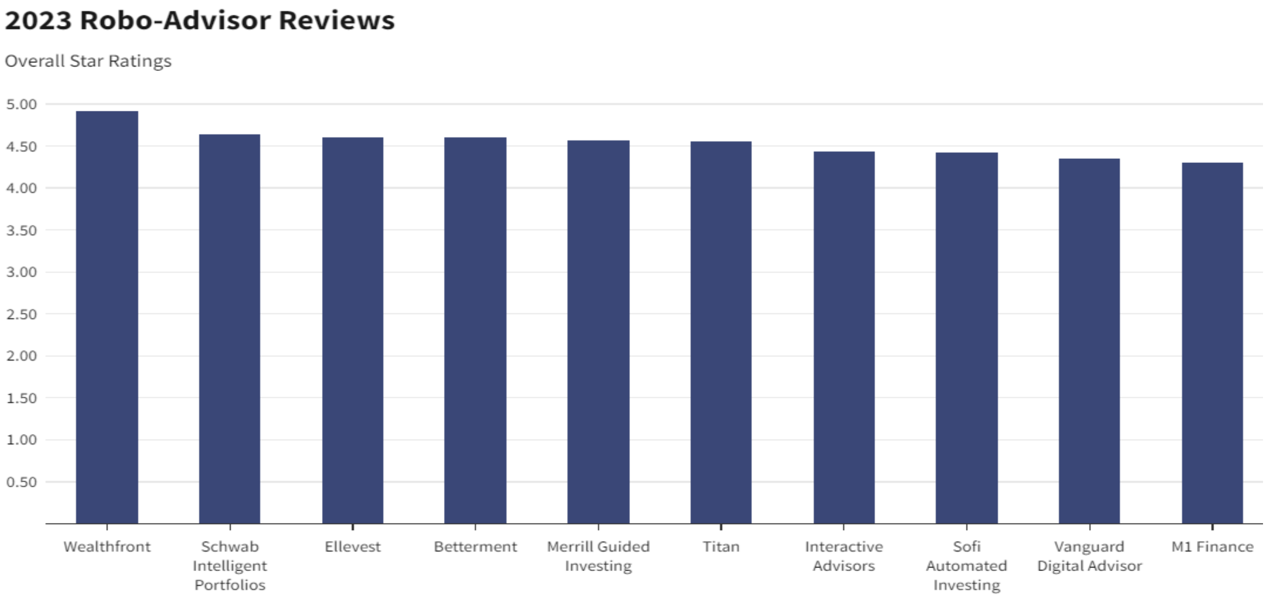



*Figure 1. 2023 Robo-Advisor Reviews, Overall Star Ratings. Source: Frankenfield, J. (2023) 'What is a Robo-Advisor?', Investopedia.*

### 1.2 Robo-Advisor Performance

In order to have a competitive service, it is important to look also at the performance. Measuring historical performance isn't immediate, since there is no regulation regarding the publication of data or the collection of performance inside a database. Moreover, returns can vary significantly depending on the investor profile and the recommended products.

For instance, Schwab doesn't divulgate information about returns of his robo-advisor service, but it is possible to have some insight from their report; in the paragraph "How did Schwab Intelligent Portfolios do?", there is a qualitative analysis on the development of different strategies but there are not many references to numbers (Charles Schwab & Co. Inc., 2020).

Wealthfront instead, tries to give more accurate information, providing customers the returns of their various portfolio, based on different timeframe, allowing investors to have an idea of past performance (Wealthfront, 2020).

However, it should be noted that portfolio created by robo-advisors are generally well diversified and therefore they could have less risk (and thus lower returns).

In the following part of this paragraph, I would like to pay attention to two different studies that try to analyze the performance of robo-advisors.

#### 1.2.1 Comparison case study

A first study here considered is a $25,000 comparison case study made by Dave (2022).

The objective is to identify the best robo-advisors (RA) and analyze their performance. The study aims to provide monthly updates on the performance of each robo-advisor chosen.

In particular these are the main features of the experiment:

- An amount of $5,000 is deposited into five different accounts of RA services;
- Every portfolio will have a moderate risk profile;
- The performance of RA will be compared monthly (here, for simplicity, are considered only the first and the last month);
- The benchmark used is the SPY 500 ETF, that is an ETF that follows the performance of the S&P 500.

The RA for the experiment are chosen taking into account the size of asset under management and the type of service. The advisors here selected are in fact exclusively RA services (except Ally Invest). In order to achieve a precise comparison, similar portfolios are chosen among the different robo-advisors. As reported in the Table 1 below, the selected RA are Wealthfront, Wealthsimple, Betterment, Acorns and Ally.

| | wealthfront | Wealthsimple | Betterment | acorns | ally |
|---|---|---|---|---|---|
| **MANAGEMENT FEE** | 0.25% | 0.50% | 0.25% | $1/mo. | 0.30% |
| **ASSET ALLOCATION** | 65% Stocks 35% Bonds | 64% Stocks 36% Bonds | 65% Stocks 35% Bonds | 60% Stocks 40% Bonds | 61% Stocks 39% Bonds |

*Table 1. Chosen RA and asset allocation. Source: Dave (2022) 'Best Robo Advisors: The $25,000 Comparison Case Study'.*

The study tries to replicate similar asset allocation for each portfolio with a 65% invested in stocks and a 35% in bonds, approximately. However, it is important to notice that the benchmark, the SPY 500 ETF, is made up of 100% stocks. It may seem that this is not a precise comparison, by the way the author gives two main explanations about this:

- The first one is that he chose to compare investment strategies instead of portfolios. He just wants to compare two simple investment strategies: creating a portfolio using RA and investing in a simple ETF. Since it is extremely easy for an investor to just invest in a single ETF, he wants to verify if RA can give any particular advantage over this simple strategy;
- The second is that the author didn't determine the asset allocation manually. He only decided the risk level and then the RA selected the right portfolio allocation. Moreover, even if these portfolios could not gain as much as a portfolio made up only of stocks, they should, on the other hand, prevent bigger losses.

The starting day in which all accounts were funded with $5,000 each was the 3rd June 2019 and since then regular updates were given every first week of the month. After roughly 3 years, exactly on the 15th November 2022, all accounts were closed. However, two accounts have been closed before this date, in particular the Wealthsimple account (since US accounts could not be accepted anymore by the company, due to their transfer to Betterment, and so the author withdrew his funds) and the Acorns account (due to an increase of the monthly management fees from $1/month to $3/month), therefore I will consider the results obtained until February 9, 2021, where every account was still open.
These were the profits on February 9, 2021 as reported on Figure 2:

- SPY(Benchmark): +$2,290.03
- Wealthfront: +$1,516.11
- Ally Invest: +$1,404.17
- Betterment: +$1,358.47
- Acorns: +$1,308.57
- Wealthsimple: +$1,064.34

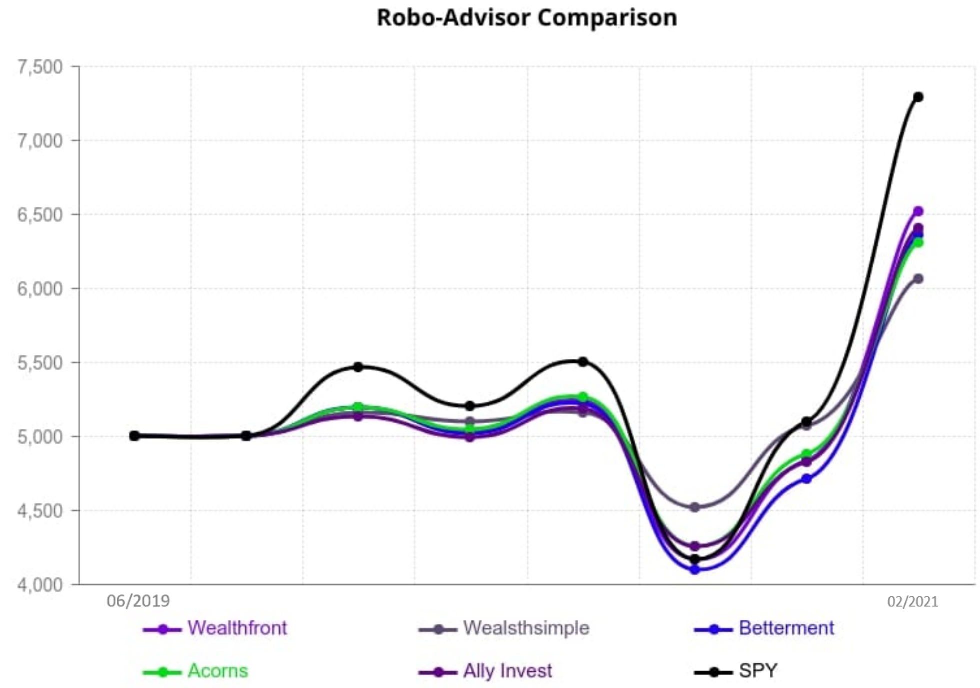



*Figure 2. RA performance comparison. Source: Dave (2022) 'Best Robo Advisors: The $25,000 Comparison Case Study'.*

The first thing to notice is that a simple investment on S&P 500 ETF outperformed every RA, both in bull and in bear market. Wealthfront RA performed quite well, while Wealthsimple and Acorns were the worst among the chosen RA. During the timeframe in which the experiment took place, the market had both stable and high volatility periods, and this showed how different RA react to market fluctuations.

In order to make the case study more plausible, I decided to add another portfolio composed by 65% of the SPY and 35% of bonds ETF (in particular I choose Vanguard Total International Bond Index Fund with ticker 'BNDX', accounting for 12%, iShares 7-10 Year Treasury Bond ETF, 'IEF', accounting for 12%, and Vanguard Mortgage-Backed Securities Index Fund, 'VMBS', accounting for 11%). I used the same timeframe, therefore starting from 03/06/2019 until 09/02/2021 and a management fee of 1% per year, calculated on the average portfolio value (considering its initial and final value). This portfolio had a return of nearly 29% and subtracting the fees, the profit was +$1,353.67. Hence, this portfolio placed behind Wealthfront, Ally Invest and Betterment RA portfolios in terms of profits. The result is pretty interesting because it shows how in the end RA are just digital financial advisors that are as efficient as human advisors and that suffer the same market risks as any other investment strategy.

### 1.2.2 Analysis of performance

In the analysis discussed by Torno and Schildmann (2020). In their analysis approach, different risk profile and time horizon are chosen in order to create distinct model customer. As shown in Table 2, there are three type of risks (low, medium and high) and two investment horizons (3 years and 15 years), thus obtaining six combinations. These characteristics have in fact the greatest influence on recommended RA portfolios.

| | |
|---|---|
| **(Lo3):** Low risk / investment horizon 3 years | **(Lo15):** Low risk / investment horizon 15 years |
| **(Me3):** Medium risk / investment horizon 3 years | **(Me15):** Medium risk / investment horizon 15 years |
| **(Hi3):** High risk / investment horizon 3 years | **(Hi15):** High risk / investment horizon 15 years |

*Table 2. Risk/investment horizon combinations of model customers. Source: Torno, Schildmann, (2020) 'What Do Robo-Advisors Recommend? - An Analysis of Portfolio Structure, Performance and Risk'.*

Subsequently, they define three categories of data availability to find proper RA portfolio recommendations:

- Category A: Clear portfolio structure and definition of asset weights and associated products;
- Category B: Semi definition of portfolio structure and weights but no associated products;
- Category C: No definition of the portfolio structure.

Obviously, Category A robo-advisors provide the best information, allowing a solid analysis of the structure and of the performances. They collected data of 20 RA of category A, 13 RA of category B and 16 RA of category C (that are not taken into account during the analysis), and they later show the findings concerning portfolio performance and risk, using only the category A of robo-advisors. Figure 3 shows the return/standard deviation graph for 138 different portfolios of these RA, based on the previous combination of model customers. Beyond these, in the figure are also shown three different benchmarks representing global equities (MSCI World), global corporate bonds (Bloomberg Barclays Global Aggregate Corporate Bond) and global government bonds (Bloomberg Barclays Global Treasury).

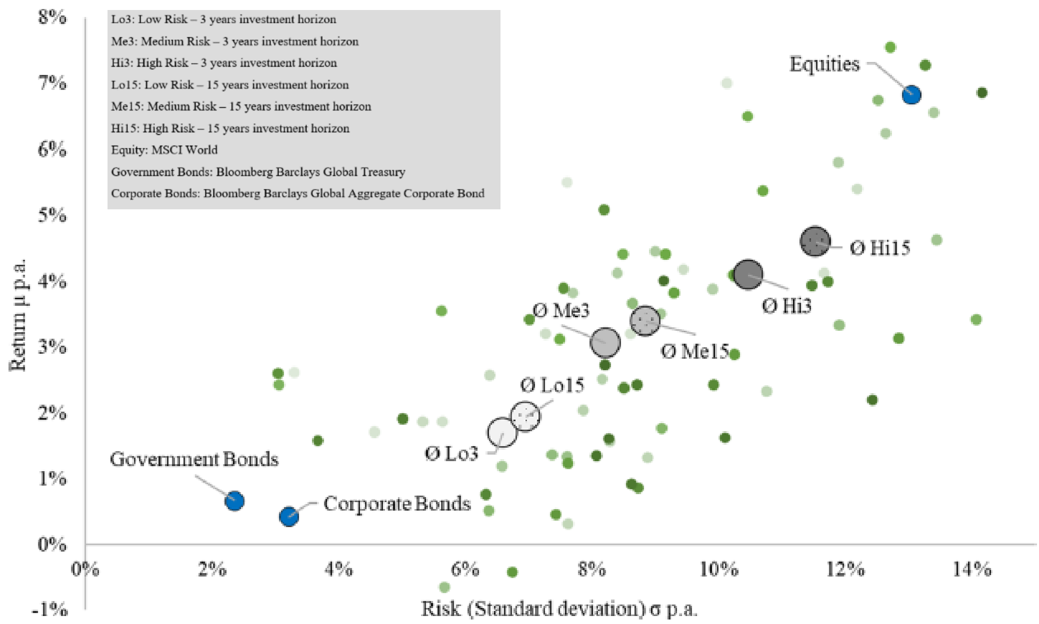



*Figure 3. RA recommended portfolios of category A in the risk-return graph with benchmarks (October 2009 to October 2019). Source: Torno, Schildmann, (2020) 'What Do Robo-Advisors Recommend? - An Analysis of Portfolio Structure, Performance and Risk'.*

It has to be noticed that there are great divergences among RA portfolio indication, because there are portfolios that have higher return but same volatility and vice versa. Usually, portfolios with an investment horizon of 3 years have an annual return of 2.95% and volatility 8.43%, while those with a 15 years horizon have an annual return of 3.3% and a standard deviation of 9.1%. They state that if during the risk assessment process the customer shows greater risk tolerance, RA usually recommend higher risk/performance portfolio.
Concluding, it is possible to say that RA investment strategies works properly, especially for higher risk individuals and longer investment horizons; instead, for shorter horizons RA propose similar risky allocations to those of 15 years horizon.

### 1.3 RA market size and diffusion

There are different studies concerning robo-advisory market size and its forecasted growth, however estimates vary significantly among them, probably due to the fact that there are still not conventional ways to measure it, the lack of a clear transparency frame and the fast development of the market. In this thesis will be taken in consideration market size estimates according to the study published by Maume (2021), that uses Statista data, and different RA market segmentation according to Grand View Research (2021).
According to Maume (2021), when dealing with robo-advisory, a market size indicator usually considered is the asset under management (AuM). The data used in this research are

taken from Statista, where it shows that starting from a market size of 297 billion USD in 2017, the market has grown to 1 trillion USD in just 3 years, implying an annual increase of about 50%. Assets under management by RA are expected to reach almost USD 2.85 trillion before the end of 2025, with an annual growth of 20% (Figure 4).

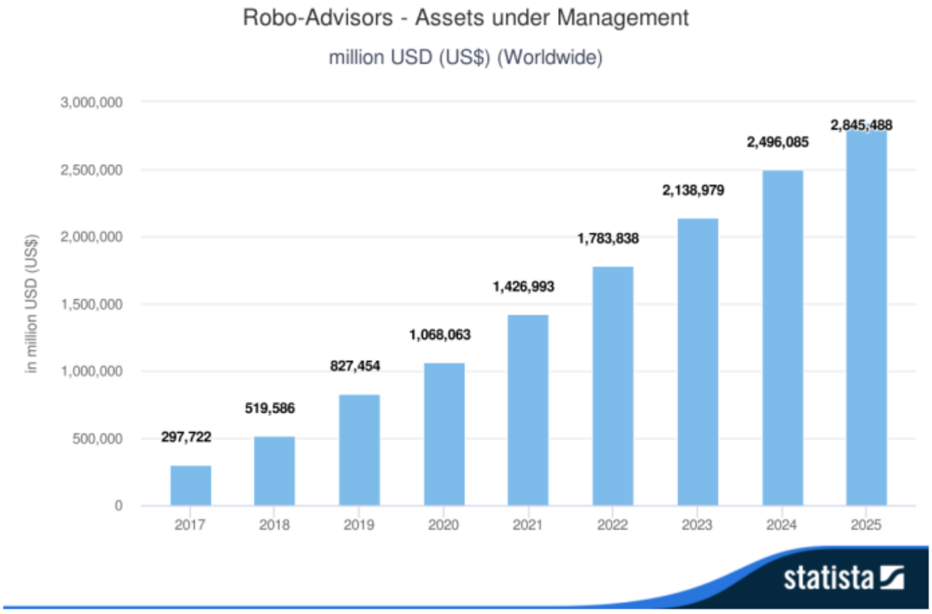



*Figure 4. Assets under management worldwide by robo-advisors. Source: Statista (2021), Maume, P. (2021) 'Robo-advisors. How do they fit in the existing EU regulatory framework, in particular with regard to investor protection?' Policy Department for Economic, Scientific and Quality of Life Policies, on request by the European Parliament committee.*

To give an idea, since the capitalization of global domestic equity market in 2020 was USD 109 trillion, it implies that robo-advisors AuM is below 1%. However, it has to be considered that AuM are not made only of equities, but also bonds and other asset classes, so the global domestic equity is only a part of the RA reference market. In conclusion, even if the market size of robo-advisors asset is growing very fast, nowadays it is just a little part of all the assets in circulation. Concerning the number of users instead, as shown on Figure 5, data taken from Statista show that in Europe there were 20.1 million users in 2020 and surpassing 40 million by 2025, implying that the potential users will double in just 5 years.

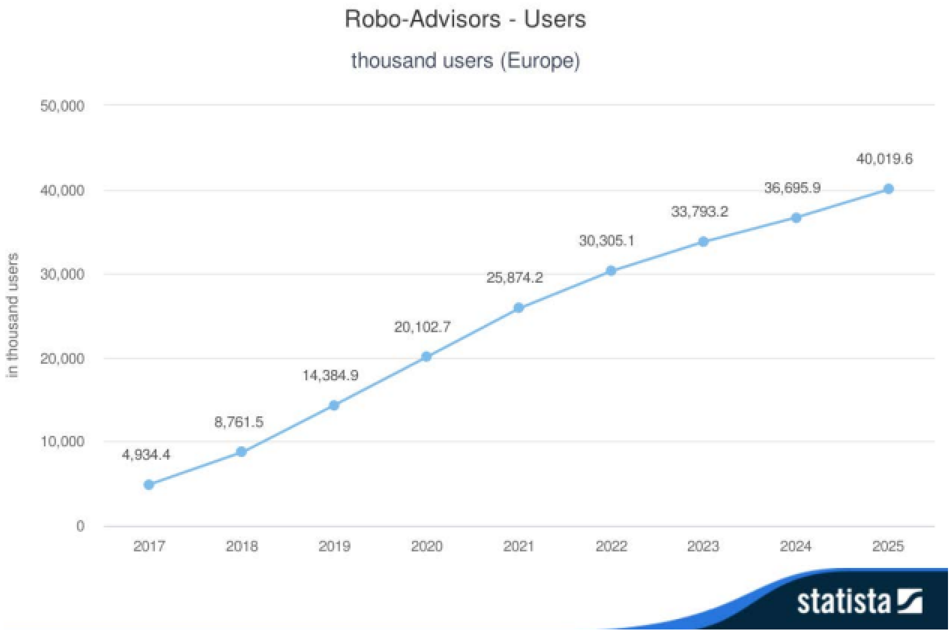



*Figure 5. Robo-advisors' users in Europe. Source: Statista (2021), Maume, P. (2021) 'Robo-advisors. How do they fit in the existing EU regulatory framework, in particular with regard to investor protection?' Policy Department for Economic, Scientific and Quality of Life Policies, on request by the European Parliament committee.*

### 1.3.1 RA Market Segmentation

According to Grand View Research analysis conducted in 2021 the market of robo-advisory can be segmented based on provider, type, service type, region and end user.

**Provider Outlook**

In 2021 robo-advisory services were distributed for more than 45% by FinTechs.
These financial institutions implement a mixture of both automated and customized advisory and adopt innovative technologies in order to provide trustworthy advisory services to every client. However, in the forecast period that goes from 2022 to 2030, the banking sector is expected to have the most rapid expansion of these services among other entities.

**Type Outlook**

Regarding the type outlook, hybrid robo advisors lead the market, representing over the 64% of the market share. Hybrid robo advisors gained lots of popularity due to the combination of efficient automated algorithms with the supervision of human financial advisor. Pure robo-advisors sector however, is expected to experience a substantial growth throughout the forecast period, thanks to the implementation of new technologies such as AI, probably leading to an increase in the demand of pure RA.

**Services Type Outlook**

The direct plan-based/goal-based segment prevailed in 2021 with a share of about 68%. Features like a robust goal planning and effective portfolio management are the most valued among investors. In the forecast period however comprehensive wealth advisory services are expected to grow the most.

**Regional Outlook**


*Figure 6. Robo Advisory Market, trends by region. Source: Grand View Research (2021) 'Robo Advisory Market Size, Share & Trends Analysis Report by Type (Pure Robo Advisors, Hybrid Robo Advisors), By Provider, By Service Type, By End-use, By Region, And Segment Forecasts, 2022 – 2030'.*

As shown on Figure 6, from a regional perspective, North America has the biggest market share in RA services, accounting for more than 29%, due to several well consolidated companies that already offer RA services, like Betterment and Vanguard Group Inc.

The Asia Pacific region is assumed to have a positive grow in the timeframe considered from 2022 to 2030. Among the several examples of the development of this sector we have TradeSmart in India and Ant Financial Services in China. In the first case TradeSmart, a prominent online brokerage company in India, entered into a partnership with Modern Algos to provide AI based advisory services, while in the second case Ant Financial Services collaborated with Vanguard Group Inc to obtain greater market share, aiming to provide simplified investment advisory to retail customer in China.

According to Maume (2021), the US market has almost USD 680 billion of AuM in 2020 while the EU market size is relatively small with USD 108 billion, divided mainly among UK (USD 18 billion), Italy (USD 15 billion), France (USD 13 billion) and Germany (USD 9 billion).

**End User Outlook**

In this part there is a main distinction between customers: High Net Worth Individuals (HNWIs) and retail investors. This market segmentation sees HNWIs as the main investors, with a share of roughly 58% and their growing demand will boost the segment growth. However, also the growth of retail investors should be considered, because since the pandemic, the creation of trading account has constantly risen. For example, in February 2021 a survey conducted by the Financial Industry Regulatory Authority (FINRA) revealed that 38% of retail customers generated more than one trading account.

## 1.4 EU Regulation

In the European Union the main regulation regarding financial intermediaries and client's protection is the Markets in Financial Instruments Directive (MiFID2). This directive generally applies to automated advice as well as to other entities that qualify as investment firms. To be classified as an investment firm within MiFID, the robo-advisor should perform either investment advice or portfolio management. These requirements apply to all kind of entities that meet these requirements, irrespective of provision by a human or an algorithm (Ringe, 2018).

### 1.4.1 Technology neutrality and regulatory uncertainties

MiFID2 adopts the principle of technology neutrality, implying that the same set of rules apply to all financial services, without considering what type of technology is used and therefore, it is applied also to robo-advisors. As a result, a technology neutral approach is considered necessary to guarantee innovation and to reduce the risk that players might circumvent regulation (Rinaldo, 2023). Obviously, this method has pros and cons: regulating behavior instead of specific technologies, offer a wide regulatory framework. For instance, when new technologies come up, the existing rules are still applicable: in fact, the introduction of new technologies usually adhere to pre-existing regulation, except in the case where the technology presents a very innovative service not covered by MiFID2. Conversely, a negative effect of technology neutrality is the rise of legal uncertainty among investors because it could not be clear how existing rules are implemented within the new digital landscape. The introduction of new technologies is often related to lack of experience on the customer side and lack of reliable guidance on the regulator side (Maume, 2021).
Moreover, new kind of financial advisors may encounter difficulties while assessing their regulatory situation, due to pre-defined categories. Innovative approaches may be hard to reconcile with standardized categories or to be categorized at all. A crucial distinction

companies have to deal with is the distinction between investment advice and investment intermediation. This determines whether the firm falls under EU regulation or domestic law. For instance, according to German legislation, investment intermediation is not obliged to adhere to a suitability requirement. According to EBA Discussion Paper (2017) examining data of robo-advice, it emerges that 35% of robo-advisors operates without any regulatory oversight, 41% are subject to EU regulations and 24% fall under national regulatory frameworks. Many consequences derive from regulatory uncertainty, such as:

- significant market barriers for newcomers in the market: FinTechs may be more cautious when introducing new technologies if the possibility of being sanctioned arise, preventing them to enter in the market in a first moment;
- regulatory arbitrage: RA may seek jurisdictions with more favorable regulatory framework, posing however greater risk for consumers, as the service may not undergo a proper evaluation by regulatory authorities;
- consumers' trust: consumers may become overly cautious about embracing new products and service if there is no clear regulatory framework, questioning the competence of regulators. (Ringe, 2018)

### 1.4.2 Other EU regulations

As pointed out by Maume (2021), MiFID does not address every aspect of financial services regulation. Specifically, the enforcement of rules is governed also by national laws and regulators. If a robo-advisor does not provide an adequate advice, regulation must ensure clients to be protected for instance through private litigation. MiFID2 in fact, provide in Art. 67-68 the necessary qualifications for national regulators, consequences for violating MiFID2 and procedural regulation. Whilst public enforcement still depends on the respective national regulator expertise, private enforcement is not regulated at all under MiFID2 and can therefore differ significantly among different states. That is an important issue in robo-advisory since the investor only gets the final investment advice, but not the process through which the software gets to that recommendation. In this sense, if a client is not sure whether a robo-advisor is acting or not in his best interests, and he cannot demonstrate and prove it in court, there is limited potential for private enforcement. As said before, MiFID2 does not cover this aspect and it should be implemented by ESMA or by specific supervision mechanisms. Although MiFID2 is the main financial regulation framework in EU, there exist also various legislative acts that regulate this field. Among the most important there are:

- Directive (EU) 2019/2034 (Investment Firms Directive, IFD) and Regulation (EU) 2019/2033 /Investment Firms Regulation, IFR), regard the supervision and resilience of non-systemic financial intermediaries;
- Regulation (EU) 2016/679 (General Data Protection Regulation, GDPR) aims to protect clients with focus on the management of their personal data;
- The proposed regulation on digital operational resilience for the financial sector (DORA), increasing the security of digital financial infrastructure.

### 1.4.3 MiFID2 investor protection regime

The definition of investor comprehends a generic customer of a financial service.
Particularly, investor protection is a key element of MiFID2, since it regards authorization and government requirements for financial services. The term 'investor protection' however has no precise definition in MiFID2, but it is intended as an objective of market regulation, achieved combining disclosure obligations (to decrease information asymmetries) and the prohibition of particular practices (such as market manipulation). The main principles of investor protection can be found in Art. 23 that regards the conflicts of interest, which has to be avoided by RA, Art. 24 that says that all services have to be in the client best interest and has to be fair and honest and lastly, Art. 25 that declares that robo-advisors must ensure they obtained all the possible information about the investor's risk tolerance and understanding of the financial product, and conduct the suitability assessment. (Maume, 2021)
Ensuring the quality of the advice, the suitability assessment is considered one of the most important requirements for investor protection and in 2017 ESMA comprehensively adjusted its guidelines on the suitability requirement for robo-advisors. (Ringe, 2018)
The investment firm must collect this information in case of investment advice or portfolio management and therefore in case of robo-advice. In MiFID2 Art.4 (1), investment advice "means the provision of personal recommendations to a client, either upon its request or at the initiative of the investment firm, in respect of one or more transactions relating to financial instruments", while in Art.4 (1)(8) MiFID2, portfolio management is defined as "managing portfolios in accordance with mandates given by clients on a discretionary client-by-client basis where such portfolios include one or more financial instruments", that is the case of standard RA (ESMA, 2018).
Firms offering robo-advice services have to obtain authorization by the Member States' competent authorities and need to be registered according to Art. 5 MiFID2. In addition, they must have sufficient initial capital with regard to the nature of the investment service according to Art. 15 MiFID2 and Art. 9 IFR (Investment Firms Regulation). These firms are

also subject to national authorities' supervision as described in Art. 21/22 MiFID2 (Rinaldo, 2023). Among all these requirements the thesis will further investigate the two most important aspects: the one regarding the suitability assessments and the one concerning conflicts of interest.

**Assessment of suitability**

According to Art. 25 MiFID2, investment firms providing advices and portfolio management must have information about the investor's experience in financial market, his understanding of specific products and his financial situation, in order to understand his possibility to bear losses, his objectives and his risk aversion (Maume, 2021). The duty to assess the suitability of investment transactions is achieved through an appropriate questionnaire that collect this information; in addition it should be explained the degree of human involvement in the process and if there are other information beyond the questionnaire. (Rinaldo, 2023)
ESMA in 2018 gave detailed recommendations on suitability requirements indicating that efficient robo advisors should ask at least 20 question to define the investor profile.
An empirical research study made in 2018 by Richter examined 21 RA to see how many questions they made to identify the investor profile. The results ranged from 5 to 24 questions, with a median of 10 (way lower from the 20 required by ESMA), however the research was published just few months after ESMA recommendations, so probably not every investment firm had the time to implement them. In practice there are difficulties to apply Art. 25, mainly due to the interaction between customers and the software; inexperienced clients could not fully understand the software questions, giving wrong input or misinterpreting advices (JCESA, 2015). Since questions are pre-defined, the software may have not a clear situation in non-standard cases, but this aspect could be overcome by new technologies like Artificial Intelligence, that may enable more customization.
Furthermore, in addition to the required information asked, there also have to be a consistency control, considering the presence of obvious inaccuracies in the information given by the investor: the investment firm must have in action appropriate policies to keep updated information about its clients. (Maume, 2021)

**Conflicts of interest**

Conflicts of interest can be considered another main problem in RA regulatory framework. Specific risks arise when the advisors providing robo-advisors services have structural links to entities that provide financial services too, such as banks or intermediates that are part of the same corporate group. In this case the advisors would have a particular interest if the client

invests in financial products issued by these entities they are affiliated to, recommending the product to the investor even if he is not really interested into it. (Rinaldo, 2023)
Maume (2021), distinguished three main cases:

- In the first one, there are banks that offer RA services in their own name or through a subsidiary, advising their own products;
- In the second case, there are funds that purchase a great number of shares of important RA firms. An example could be when in 2018, Allianz Asset management became the biggest shareholder of the Italian Fintech Moneyfarm, or when between 2017 and 2020, BlackRock invested in German market leader Scalable Capital.
- In the third and final case, there are RA that start marketing and distribution partnerships with banks, promoting their associated products.

In every case mentioned above, there is a concern that robo advisor could become just a distribution channel for banks and funds that aim at promoting their own products. Nonetheless, as reported on Better Finance (2020), there appears to be no public evidence of any robo advisor receiving some form of gain, monetary or not, for suggesting specific financial product. The lack of this evidence, as of today, does not eliminate the issue of conflict of interest within robo-advisory. Supporting this sentence there is also Brenner's research (2020), which claims that robo-advisors are a valid alternative for seeking investment advice among those investors that are more afraid of investment fraud and worry about potential conflicts of interest that appear in the context of human financial advice. Moreover, Fisch (2019) affirms that robo-advisors may be less vulnerable to potential conflicts of interest to the extent that they are independent and do not sponsor or sell the investments they recommend.
Under Art. 23(1) MiFID2, each investment company is required to undertake every necessary measure to prevent and deal with conflicts of interest. Art. 16(3) instead states that investment firm must establish organizational framework to avoid that conflicts of interest could potentially damage customers' best interests. When the previous measures are not enough to mitigate the risk of damaging the investors' interests, the provider of the service is obliged to immediately disclose the characteristics and sources of the conflict. Disclosure to investors is a measure of last resort, as defined on Art. 34(4) MiFID, which should be employed only when previous arrangements are not sufficient. Even if the investment firm disclose the conflict, it should act in any case in the investor interest, as indicated in Art. 24(1) MiFID2, implying that if a conflict of interest has emerged, the investment advice given must be still the best one for the customer.

These studies have demonstrated that MiFID2 provides a suitable framework regarding client protection using robo-advisors in EU, but there are still features that can be improved, like the information collection mechanism and features that may be adopted, such as a third-party audit for robo-advisors and the implementation of private enforcement measures.

# Chapter 2 – Fundamentals of AI in Robo-Advisory

## 2.1 Artificial Intelligence in the financial sector

Artificial Intelligence is progressively playing a fundamental role in influencing economic and financial sector productivity by enhancing efficiency and decision-based processes. Focusing on financial sector, AI is revolutionizing the nature of financial intermediation, risk management, compliance and prudential oversight (IMF, 2023).

AI and robo-advisory combined have a strong potential, going from customer related arguments to portfolio management. Regarding this one in particular, there are four main areas:

- Stock picking: AI allows investors to properly choose among stocks, analyzing different data in order to match the criteria followed by the investors;
- Trade management: the correct implementation of AI when trading can improve market entry timing and decrease mistakes linked to emotional factors;
- Portfolio optimization: AI can detect the optimal portfolio allocation for the investor needs, in terms of risk aversion and time horizon;
- Risk management: concerning this field, AI can be used for example in combination with modern portfolio theory in order to allocate the portfolio on the efficient frontier, but also to reduce over-exposure of individual stocks in a portfolio.

There exists a strong connection between investments and Artificial Intelligence and even if the retail customer is not aware of it or is not using it, financial advisors, banks and funds will utilize it in different ways on his behalf (Grossman, 2023).

The evolution brought by AI is mainly due to two factors. The first one is that data are collected in a more extensive way than ever before and are widely available. The second regards the fact that there always are more complex AI models thanks to increasing computational power at comparatively low costs. In a highly digital business like robo-advisory, the keys for success are process automation, decision optimization and customer data quality. AI can be used in robo-advisory also to develop product recommendations: for instance, scheduled financial events can trigger the advice to offer a hedging product to mitigate the risk. If an investor owns some stocks of a company whose financial results are going to be disclosed in the following days, potential robo-advisor could propose a put option to protect against possible losses. In this case Artificial Intelligence aims first at identifying clients who wants to be approached by event-driven actions and then it matches a product-event related. AI goal is to understand correctly the feature of the upcoming event and find in an autonomous way a matching product that benefits from the event or mitigate its risk, as in

the example above. AI is also useful in matching the adequate portfolio to customers which may differ in term of investment opportunities: AI can in fact provide a range of matching opportunities based on investment needs and preferences of each investor (Scholz, 2021).
On the 7th December 2022, Deutsche Bank announced a partnership with software firm NVIDIA to boost the use of AI and machine learning in the financial sector. Mixing Deutsche Bank's knowledge of financial sector with NVIDIA's leadership in AI is going to develop a broad range of AI-powered services. According to Christian Sewing, CEO of Deutsche Bank, in fact, AI and machine learning will be an innovative step forward in banking, and their partnership with NVIDIA is a concrete example, thus reimagining the way financial services are executed and delivered.
Three main results are expected to emerge from this collaboration: i) accelerated computing that enables traders to manage risk and run more scenarios faster, increasing DB leadership in risk management, ii) interactive avatars that are basically virtual assistant for employees and banking clients and iii) extract key information from unstructured data with the purpose of receiving warning signs on counterparty risk, retrieve data faster and discover data-quality problems (Deutsche Bank, 2022).
A crucial distinction should be done between AI supervised by human (human-in), AI guided by human (human-on) and AI working without human supervision (human-off).
AI human-off activities, also referred to as unsupervised learning, are present only in few tasks that analyzes different data, finds similar patterns and makes prediction without any guidance, but, as of today, including decision-making processes in this field may not be possible due to several biases. Human-in tasks, also called supervised learning, require a human to label the input data first, or more generally is defined as a model requiring human interaction, therefore it is a sort of mixed approach. AI human-on activities finally are tasks that are guided by humans which can interrupt an action in the process if necessary and usually require the human to provide feedback to the AI system to improve its performance over time (Wikipedia, 2023a; Credo, 2023).

### 2.1.1 The revolution of Generative AI

Generative AI represents a significant step forward in AI technology. It is a specific subset of AI and machine learning, characterized by the feature of generating new content. GenAI is based on large language models (LLMs), which are neural network–driven models trained on great amounts of data, text and documents and able to create coherent text, images and human languages. LLMs nowadays have various applications with substantial implications for the global economy and financial sector. According to a 2023 IMF publication, GenAI will

improve the implementation of AI in the financial sector. In recent years, there have been a rapid adoption of AI/ML in the financial sector leading to efficiency and cost savings, transformation of client interfaces, improvement of forecast accuracy. Many financial institutions explored the different capabilities and range of applications of GenAI, showing that the ability to handle extensive and different data sets and to create content in an accessible way, enhanced efficiency, improved customer experience, risk mitigation, and compliance reporting. In particular Capital One and JPMorgan Chase improved their fraud detection system, Morgan Stanley will use AI to assist financial advisor, Wells Fargo will use AI for automating document processing and Goldman Sachs and Citadel will consider GenAI to develop more advanced internal software.
According to a McKinsey & Company publication (2023), Generative AI is an important evolution of standard AI. One of the main challenges is to understand how it can give more value to the economy in order to undertake proper decisions. The study considers two different lenses, as shown in Figure 7, that determine where GenAI could have the strongest impact and how large this impact could be.

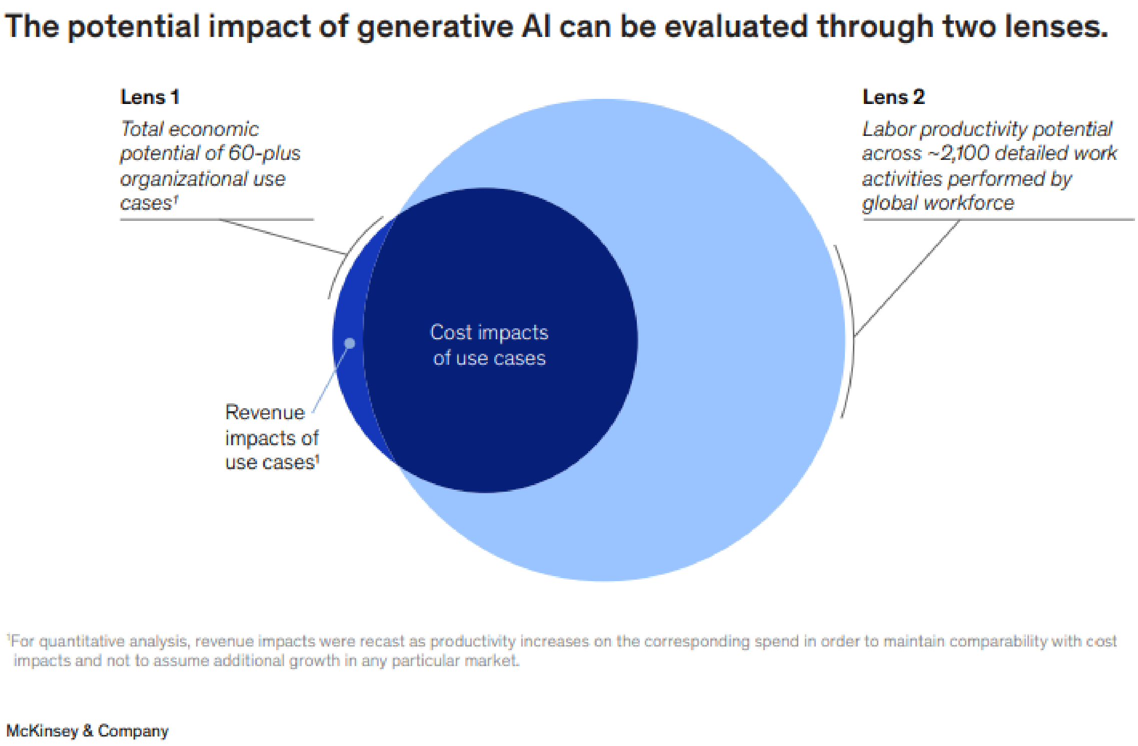



*Figure 7. The potential impact of generative AI can be evaluated through two lenses. Source: McKinsey & Company (2023) 'The economic potential of generative AI: The next productivity frontier', June 14, 2023 Report.*

The first lens identifies use cases that could be implemented by companies. A "use case" is a specific generative AI application in a determined business that produces measurable outcomes. The research considers 63 GenAI use cases through 16 different business functions that may bring from $2.6 trillion to $4.4 trillion in economic benefits annually. This will add 15% to 40% to the $11.0 trillion to $17.7 trillion of economic value that nongenerative artificial intelligence and analytics could unlock, according to their study. The second lens instead considers the impact of GenAI on more than 2.100 "detailed work activities" required in 850 occupations. Different scenarios are taken into consideration to understand when GenAI could perform these activities, in order to forecast how the current status of AI may influence labor productivity. The total economic benefits of GenAI, net of cost reduction that could overlap with cost reductions in the use case analysis, range from $6.1 trillion to $7.9 trillion annually (Figure 8).

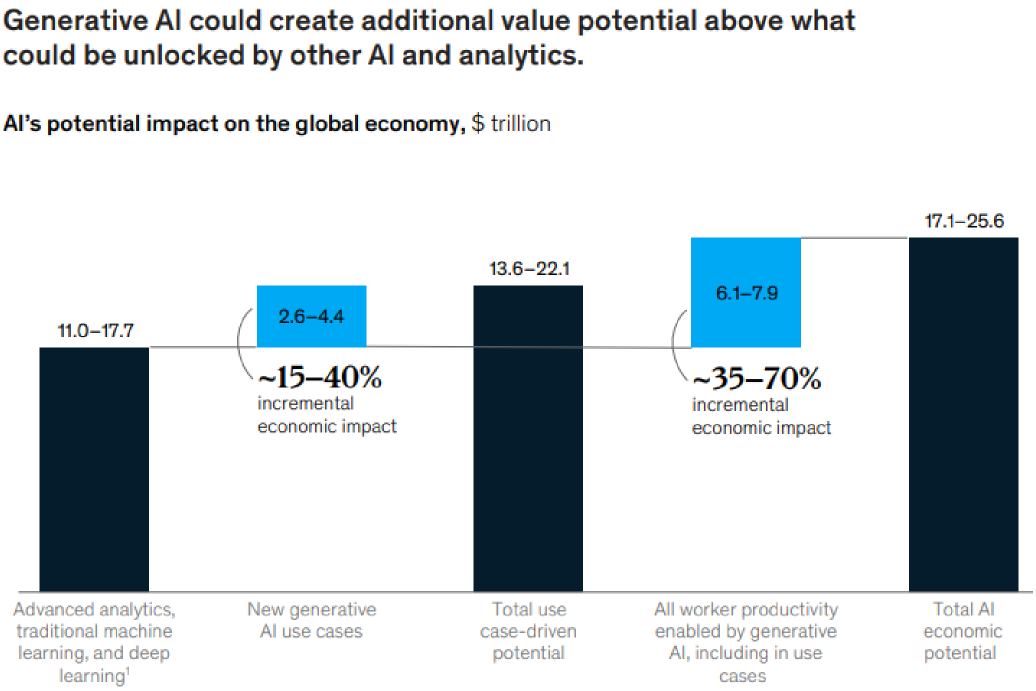



*Figure 8. Generative Ai could create additional value potential above what could be unlocked by other AI and analytics. Source: McKinsey & Company (2023) 'The economic potential of generative AI: The next productivity frontier', June 14, 2023 Report.*

In 2023 in particular, AI made consistent advancements and the entire world noticed it. The most famous GenAI bot, ChatGPT attracted 100 million monthly active users in just a couple of months from his debut in November 2022, making it the fastest growing application in history, as shown on Figure 9, where it is compared to other fast-growing applications.

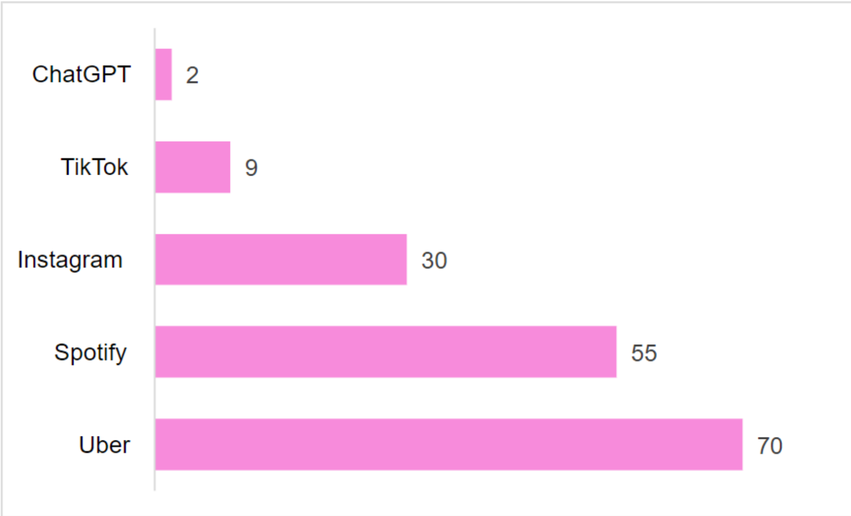



*Figure 9. Months it took companies to reach 100 million online users. Source: Spiegel, J. (2023) 'How advisors are increasing efficiency and impact with AI', BlackRock Advisor Center.*

Since demographic shifts and longer lifespans are considerably boosting the demand for advisors' services, especially with Millennials and Gen Z, that grew up with much more tailored solutions with respect to previous generation will particularly appreciate and be confident with the use of AI in investment advice. 'Schwab Advisor Services' 2022 Independent Advisor Outlook Study' (Charles Schwab & Co. Inc., 2022) in fact shows how "52% of advisors believe that investors will want more personalization over the next five years" and that "38% believe that Millennials will seek out personalization more than any other generation".

Spiegel (2023) traces a comparison between generative AI use happening now and possible use cases in the future. While at the moment GenAI can create first drafts of client communication, in the future tools may help advisors better understand clients' feelings, by analyzing e-mails for instance; if now generative AI may help in delivering information in an easy way, in the future chatbots could be employed to handle routine clients' inquiries, providing instant responses. According to a Boston Consulting Group article (2023), also CFO have to implement the use of GenAI in the near future, in order to transform core process, business partnering and mitigate risks.

Nevertheless, as reported on the Boston Consulting Group website, in the Finance Function Excellence section, financial services need to extend their objective from forecasting and preparing reports and analysis to extracting important insights that may be very helpful in business decisions, transforming insights into practical action and generative AI will probably have a crucial role in this step.

The implementation of this new technology, will likely have an S-curve pattern, like previous technology had. In Figure 10, it is possible to visualize different stages of this evolution.

Currently, with the introduction of the new Chat GPT-4 model that is way more complex than its previous engine in terms of computing and processing power, it is possible to produce content, ideas and generate basic analysis on large database. In the medium-term it will be probably possible to conduct problem solving for planning and execution with the implementation of scenario analysis. Finally, in the long-term GenAI will be able of generate forward-looking insights and learn, be autonomous and work collaboratively.

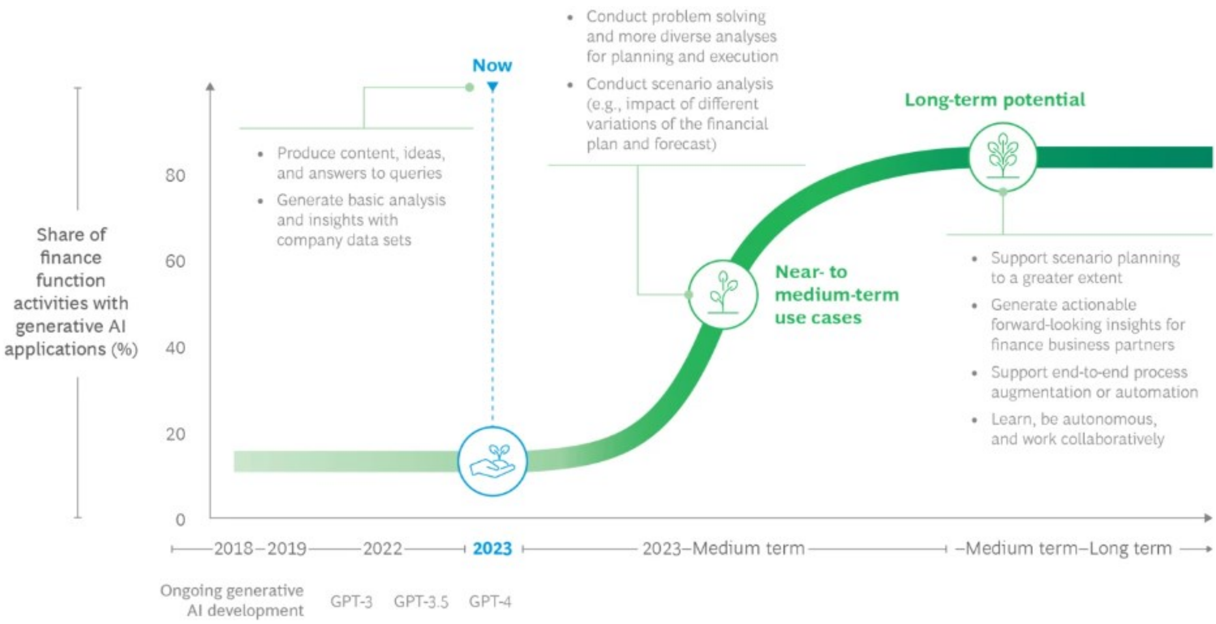


*Figure 10. Generative Ai adoption in finance will likely follow an S-Curve. Source: Demyttenaere, M., Roos, A., Sheth, H., (2023), 'Generative Ai in the Finance Function of the Future', Boston Consulting Group.*

BCG states that an important outcome in term of quality of reports, recommendations and explanation may be achieved from the interaction between GenAI and the traditional AI tools, empowering finance professionals to make better decision and increasing operational efficiency. As of today, the ability of generative AI to create numerical analysis with precise accuracy is still in its early stages and before it won't be no longer required human assistance, outcomes should become more reliable. On the other hand, traditional AI applications in finance can use data in order to make forecast and assessment, therefore, some activities are more suitable for traditional AI (like reconciliation, anomaly detection and credit scoring), other for generative AI (like writing tasks, financial scenarios generation and automated report generation) and other deriving from the combination of the two (like demand forecasting and recommendation engine). For instance, traditional AI may forecast financial data while generative AI may discuss variances and offer different advices based on different forecasted scenarios.

Differently from many other previous innovations, like robotic process automation, the barriers to implement generative AI are quite low, however as many other new technologies it may presents different challenges. The main problem concerns data accuracy, in fact in these first versions GenAI does not ensure precise calculations, lack of contextual awareness and real-time information and the generation of incorrect output usually referred to as 'hallucinations'. Even the leak of proprietary data may be another important issue since companies' proprietary data could be leaked if there is a security breach, but all these issues will be further explained later in this chapter.
GenAI is evolving at a very fast pace, due also to billion-dollar investment in this field from companies like Microsoft, Google and more. At the moment it is very useful in text and images generation but not so much in numeric analysis, however, it should be noted that this model is in its early stages and it will likely affect a wide range of sectors, in many different ways.

**2.2 AI in Financial Advisory**

According to Smith article published by Accenture (2022), financial advisors have clear in mind how to effectively use AI in their daily tasks, not only to increase service levels but also to improve relationship with customers.
In particular, the top three use case described by Accenture are:

- Increase the use of hybrid advice: some financial advisors intend to use AI to understand how they should behave in order to satisfy clients' needs. Through personalized and meaningful communication, the objective is to have longer lasting relations with customers and increase their loyalty;
- AI focus on investors' emotions: different financial advisors plan to implement AI to process and understand customers' emotions: AI can adapt portfolio offering in a personalized way, according to the investor specific needs in that precise moment of his life. Moreover, AI can be useful in identifying financial events that are of particular relevance for specific customers;
- Enable proactive customer engagements: some financial advisors aim at providing more trust with proactive engagements generated through AI: financial advisors may increase the understanding of their customers with the help of AI (using for example demographic segmentation) and intervene more frequently in order to satisfy the investors' needs at the right moment.

Going more in depth, in 2022 Accenture conducted research to interview 500 licensed financial advisors living in the United States and Canada to obtain a more detailed framework of financial advisors' interactions and attitude towards AI, exploring topics such as client experience, value and industry change. Different kind of respondents participates in the survey, differing for industry, households under management, location and asset under management. For each of these categories the main players were respectively: bank brokerage in the "Industry" category, 101-150 households under management, US based financial advisor and in the end in the category "Asset Under Management", most of the financial advisors had between $501 million - $750 million AuM. In particular, among these advisors, a large portion claimed that AI is reshaping the creation, delivery and consumption of financial advices for clients and are confident that AI will continue to play a fundamental role in the future of financial advice, enhancing client relationships. However, some advisors express that insights generated by AI are perceived as too complex to use and feel that AI client insight were not as impactful as expected. Overall, the results show that financial advisors are generally favorable on the use of this technology and that the challenges firms have to face, are matching the pace of innovation with the rate of adoption, focusing on a single AI use case (instead of trying multiple use cases) and ensuring that their priorities are aligned with financial advisors' ones.

### 2.2.1 Empirical evidence of financial advice by Generative AI

Here, a very interesting analysis presented by Fieberg, Hornuf and Streich (2023) is analyzed. In this publication, are taken into consideration the suitability of financial advice provided by the new GPT-4 engine. It is used GPT-4 instead of ChatGPT, because it leverages more data and more computation to generate more sophisticated responses. The research starts by defining four different investor profiles to verify if GPT-4 personalizes its advice: as shown on Table 3, differences in risk capacity are defined, considering investors' age (30 or 60 years old) and investment horizon (5 or 40 years), while differences in risk tolerance are implemented using two different risk aversion profile (high risk/low risk).

| Profile | Age | Investment horizon | Risk tolerance |
|---|---|---|---|
| 1 | 30 | 40 | High |
| 2 | 30 | 40 | Low |
| 3 | 60 | 5 | High |
| 4 | 60 | 5 | Low |

*Table 3. Investor profiles. Source: Fieberg, C., Hornuf, L., Streich D. (2023), 'Using GPT-4 for Financial Advice', CESifo Working Papers.*

Later, GPT-4 is used to give proper portfolio allocation for every investor profile. Requests are formulated in a hypothetical scenario because the algorithm cannot give personal investment advice probably due to legal issues. GPT-4 provided specific investment products (mainly ETFs managed by famous company like Vanguard and BlackRock) and their relative portfolio share. As a benchmark for this study, they consider the advice given by an US financial advisory firm's automated financial advice solution, unfortunately not mentioned. The composition of each portfolio, as shown on Figure 11, includes equity and fixed income in domestic (US) markets, developed markets and emerging markets, as well as alternative assets (real estate and commodities) and cash.

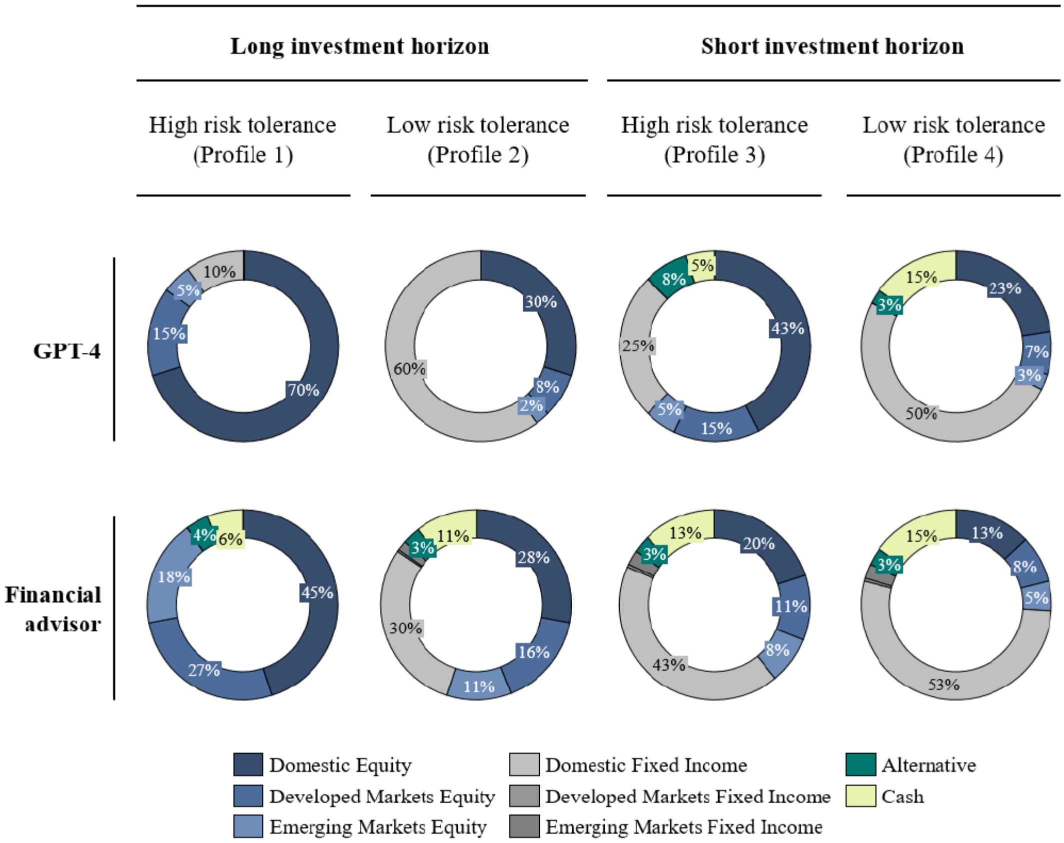
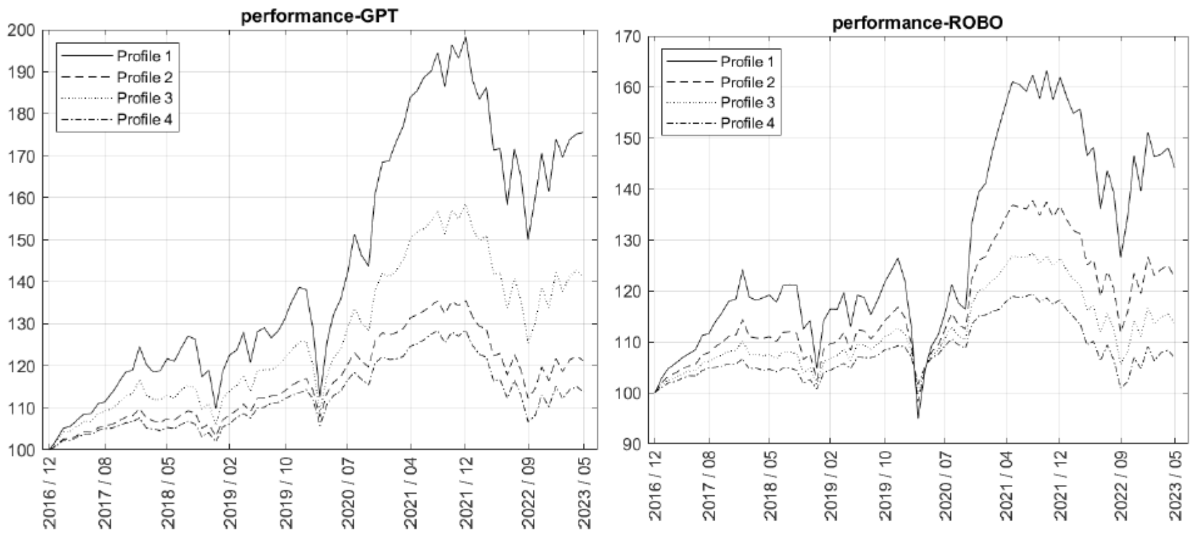



*Figure 11. Portfolio breakdowns. Source: Fieberg, C., Hornuf, L., Streich D. (2023), 'Using GPT-4 for Financial Advice', CESifo Working Papers.*

From Figure 11 is important to notice that GPT-4 recommends the same asset classes and geographies as the financial advisory firm. However, GPT-4 reports a substantial home bias (in Profile 1, 70% of domestic equity is recommended against a 45% recommended by the financial advisor), with emerging markets equities representing 5% of the GPT-4 portfolio versus 18% in the benchmark portfolio. Moreover, GPT-4 portfolio appears to be more affected by risk tolerance and less affected by the investment horizon, in fact from Profile 1 to Profile 2 there is a decrease in equity from 90% to 40%, while in the financial advisor portfolio the decrease is from 90% to 55%. When instead an investor becomes risk-averse,

equity goes from 90% to 63% in the GPT-4 portfolio and from 90% to 39% in the benchmark portfolio. To analyze both portfolio performance, the research compute monthly average return, volatility and Sharpe Ratio from December 2016 to May 2023.
The results confirms that GPT-4 average returns are more sensitive to risk tolerance and less sensitive to the investment horizon, as displayed in Figure 12, but also that profiles 1 and 3 have significantly higher returns than the benchmark portfolio, while for profile 2 and 4 returns there are not great differences. Unfortunately, the paper does not provide specific explanations about the reason why the suggested portfolios present home bias and returns are more sensitive to risk tolerance however in my opinion it should all depend on the prompt that are put into ChatGPT, therefore more refined prompts may avoid these biases.
As Figure 12 displays, GPT-4 portfolios provided equal if not better risk-return profiles: the figure shows the evolution of $100 invested in every portfolio presented in Figure 11. Even if exposed to similar market risks, like Covid-19 dip in 2020, GPT-4 portfolios had better performance in each investor profile. However, when considering risk-adjusted returns, they had comparable values, implying that better performance where probably due to greater exposure to market risks in GPT-4 portfolios.




**Note:** This table shows the evolution of $100 invested in the portfolios suggested by GPT-4 and our benchmark financial advisor for the four investor profiles from 2016/12 to 2023/05.

*Figure 12. Portfolio evolution. Source: Fieberg, C., Hornuf, L., Streich D. (2023), 'Using GPT-4 for Financial Advice', CESifo Working Papers.*

In the end, the results of this study are quite important, because GPT-4 provided advices very similar to the one provided by professional automated financial advisory services. Even if it showed relevant home bias in its decision, the risk-adjusted returns were very similar to the benchmark portfolio. However, while GPT-4 proved good competence in matching investor profiles to specific portfolios, risk profiling of investors and assistance in the use of specific platforms, must be still guided by human financial advisors (at least in the near future).

### 2.3 Generative AI risks and main issues

The rise of robo-advisors and generative AI models like GPT-4 and ChatGPT is providing efficiency as well as various regulatory challenges. While robo-advisors use transparent algorithms, large language model mechanisms are not so clear. The decision-making process in robo-advisors can be understood and regulated, but large language models generate more complex outputs, and thus monitoring and understanding their process may result complicated. RA are mainly an extension to traditional financial advisors, therefore the companies that provide these services have to follow specific rules applicable to traditional financial advisors. According to ESMA Technical Report (2023) most robo-advisors rely on transparent and clear mechanism that adhere to specific customer needs and investment goals; it is also important to notice that most robo-advisors services are part of registered financial advisors, therefore must comply with specific laws.

On the other side, GenAI applications in finance entail greater challenges to current regulations, in fact even if there are similarities between application of GenAI and RA, RA regulatory framework cannot be applied to GenAI, first because it is not made only for financial purposes and then because there exist concerns about its reasonable basis for recommendations. Moreover, hallucinations in the responses could undermine the reliability of the models and provide incorrect answers that do not follow regulatory frameworks. Therefore, integrating generative artificial intelligence in robo-advisory may on one hand improve user experiences with tailored advice but on the other it will raise concerns on compliance with regulations (Caspi, 2023).

In this period new regulatory framework are being developed in the world, in particular: Europe passed the so called “AI Act” in June 2023, setting the precedent for GenAI regulation and claiming that LLMs need to be regulated separately and developers need to prevent the generation of law-breaking content, however it will probably come into force in 2024; United States has still to provide a federal regulation on this field but the White House released indication for an AI Bill of Rights and SEC (Security and Exchange Commission) in July 2023 proposed rules to avoid conflicts of interest by advisors; China issued in July 2023 measures to every GenAI accessible for the general public, without however providing regulation to enterprise-facing GenAI (Oliver Wyman, Morgan Stanley, 2023).

Therefore, the global regulatory framework is still in its early stages and it will be probably improved and reviewed in the following years. In fact, it is important to pursue a dynamic

approach that can balance the improvements brought by technological innovation and the market stability. Moreover, in my opinion, since the development pace of this technology is very high, nations should focus intensely on how to implement as soon as possible an efficient regulatory framework in order to prevent incorrect usage of the technology and reduce the possible damages to companies, concerning mainly privacy, discrimination and intellectual property issues.
The increasing use of AI-based services is generating different concerns about its relative risks (IMF, 2023). It is possible to identify a major distinction among risk, in particular technological risks and usage risk. In the first category it is possible to include hallucination, explainability and cyberattack issues, while data privacy, biased algorithms and copyright concerns belong to the second category (Oliver Wyman, Morgan Stanley, 2023).
In the following section will be considered the main risks mentioned above.

### 2.3.1 Hallucination

Robustness of AI performance is an important issue in order to ensure financial stability and public trust, but if in a stable data environment, it can produce reliable outputs, AI faces more difficulties when previously reliable input become unreliable or when there is a change in behavioral correlation, producing lower prediction accuracy.
GenAI models can sometimes produce distorted output with high level of confidence (i.e., “hallucination”) because they tend to use pattern in previous data instead of using logical relationship and may lead to bad business decision and inefficiency.
ChatGPT, for example, is trained on a great amount of data of all sorts available on the internet, raising more concerns on its reliability and this may create wide misinformation for example in investment decisions, coding and clients’ advices (Wach, 2023).
A solution proposed by Morgan Stanley (2023) is to train models on high quality data and review the final recommendations provided by AI services, while according to Ji (2022), the ongoing efforts to address this issue should be made from a broader perspective instead of focusing on specific tasks.

### 2.3.2 Explainability

Another issue concerning GenAI, is the inability to trace underlying logic, because the outputs generated are not traceable or clearly explainable: generative AI systems are often referred to as "black box" systems for this reason (Morgan Stanley, 2023).

AI algorithms have complex architecture that is based on multiple parameters and interacting models, making input signals not easily identifiable. Explainability therefore remains a challenge for the research community, and different solution are being proposed even if as of today with unsatisfactory results (IMF, 2023).

According to Morgan Stanley (2023) instead, fine-tuning AI models for specific use cases, may create models able to explain outputs to customers. For instance, NewsTrack is a GenAI model developed by Oliver Wyman set to predict credit default events and this model is "self-explainable", indicating the reasons that contributed to a particular result. As pointed by Morgan Stanley report (2023) and EY article (Jarrell, 2023), some plausible solutions could be: having a clear understanding of the ways in which data are collected, organized, processed and prepared for the model; identifying if generative AI is the correct solution for a specific business case and reimagine operations and workflows to gain advantage of different AI possibilities; training employees to effectively use AI tools.

### 2.3.3 Data privacy

The development of generative artificial intelligence, raises concerns also about data privacy. According to GDPR (2023), personal data violation means "a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to, personal data transmitted, stored or otherwise processed". GenAI could be able to generate synthetic data that could be used to identify individuals or group of people, leading to target advertising and other forms of marketing that violate individuals' privacy (Wach, 2023).

Other concerns regarding data privacy involve data leakages from training data sets. In fact, GenAI models use inputs from users to train and adjust their statements, and this automation expose financial institutions' employee to disclosure of financial and personal data. Moreover, there is also the risk that AI may reveal anonym data through inferences and remembers users' information after the data set is used. (IMF, 2023)

Finally, everything that is prompted into ChatGPT is used to further train the model and this may leak investment strategies, trade secrets, client records and other important information. For instance, in April 2023 some Samsung employees shared some sensitive data accidentally while using ChatGPT to be helped at work, forcing Samsung to limit its use per person.

ChatGPT usage guide however informs users not to divulgate confidential information in conversations, since data are used to train its models (Mauran, 2023).
Many firms are therefore trying to implement safe environment and to use internal versions of generative AI models, in order to train them without the risk of feeding proprietary data into public models. Finally, some solutions to avoid this issue could be: maintaining trust and transparency in order to show to customers, employees and stakeholders the commitment to ethical AI use and safeguarding sensitive information according with data privacy laws (Morgan Stanley, 2023).

### 2.3.4 Embedded bias and discrimination

According to Friedman and Nissenbaum (1996), embedded bias can be defined as computer systems that systematically discriminate against certain individuals or groups of individuals in favor of others. This bias usually appears if the information used to train the models are incomplete, affected by social prejudices or influenced by human biases. In the financial sector, embedded bias may generate unethical practices, financial exclusion and damaged public trust (IMF, 2023). Biases in AI systems can have severe consequences, particularly in the management and economic domains (Dwivedi et al., 2023). For instance, biased statement when using ChatGPT in a financial institution, may lead to discriminatory treatment of customers based on their race or gender, perpetuating existing inequalities and reinforcing discriminatory practices. There is another factor to consider, in fact in case of biases responses from ChatGPT, who should be responsible? The developers, the users or the AI model itself? If for example, ChatGPT gives wrong financial advice resulting in financial losses, it could be quite complex to identify the responsible. This could result in legal disputes and challenges in establishing the guilty party, damaging also reputation of both organizations and customers. Therefore, in absence of specific regulation in this field, it is very challenging to ensure responsibilities in the use of AI services (Wach, 2023).
While the risk of bias has always been present in machine learning and standard artificial intelligence applications, in GenAI it is way more present and could lead to incorrect outputs on a larger scale (Morgan Stanley, 2023). The process to solve this problem may be more complicated for GenAI, due to the breadth and diversity of training data; in addition, while machine learning and standard AI use training data for forecasts, generative AI use training data to create textual answers, that are a sort of new content. The data bias problem could make more difficult the adoption of GenAI in financial services, because on one hand it could bring lower costs but on the other could lead to inaccurate or discriminatory client assessment (IMF, 2023). Furthermore, there is also algorithmic bias generated by bad data. Algorithms

might introduce inadvertent bias, reinforcing historical discrimination, favoring a specific political orientation or reinforcing undesired practices (Janssen, 2016). So GenAI outputs depend a lot on the quality of data it is trained on, so if these data contain biases like racial or gender stereotypes, these aspects will be presented in the statements provided. To mitigate all these risks, IMF propose to use different and high quality, pre-approved datasets to train the models. Another solution to obtain more reliable data would be to have in place human feedback, with clients providing feedbacks to help improving the models.
According to Morgan Stanley report (2023) and EY article (Jarrell, 2023), in order to mitigate the risks here discussed, several approaches may be adopted (beyond the ones already mentioned):

- use a proper and safe data environment and employ good data management practices to effectively train AI models;
- apply several tests to avoid any kind of bias and improve features and labels in the training data;
- engage third-party organization to assess requirements, specification and eventual biases.

### 2.4 The future of Generative AI

In this paragraph I would like to pay attention at two different aspects concerning the future of generative artificial intelligence. In particular I would like to investigate whether generative AI will replace human financial advisors in the first part, while in the second part I would analyze if generative AI could revive robo-advisors.

#### 2.4.1 Generative AI and financial advisors

Over the past years, lots of new technologies contributed to the evolution of financial services and AI will likely provide a greater boost to this process. However, the need for personal advices today is still strong and it will probably be the same in the near future. Every time a new technology innovation appears in the market, people is usually enthusiastic with it at first but later, in many cases, it does not live up to customers' expectations. This happened to many firms during the dot-com era, and more recently to cryptos and NFTs: while it's undeniable their potential, their adoption has not been as fast as many expected and did not generate the desired developments. Another innovation that was going to disrupt financial advisory was robo-advisors, in fact since their debut many people claimed they were going to completely replace human advisors; thing that has never happened, in particular for those individuals with higher assets value, reluctant to let algorithm manage their assets (Malamed, 2023).

As reported by Cavanaugh (2023) in fact, even robo-advisors have not decimated human financial advisors, despite their potential to manage portfolio allocation and rebalancing it, with its use that declined in 2022, especially among individuals with more than $500,000 in assets. In recent years their usage declined, with SEC in 2021 stating that even the best providers of these services may not fulfill their promises to customers. Among the main problems pointed by the commission, there are the fact that these platforms did not ask sufficient question in profiling the investors and that firms didn't re-evaluate periodically investors' circumstances and objectives, leading to trading or rebalancing mistakes.

Therefore, even if more tech-savvy investors may be willing to discuss about their finances with AI-powered services, high net worth individuals may always need a person supporting their accounts, keeping human advisors and support staff relevant to the investing public. Investors generally entrust their assets to experienced financial professionals after the creation of a certain level of confidence and of a sense of reliability, that is very difficult to create with emotionless chatbots (Malamed A., 2023). However, AI may work along financial advisor to speed up labor intensive tasks, automate repetitive ones, reviewing documents, analyzing massive data sets, therefore allowing the advisors to spend more time with customers and to add value to that time. Moreover, artificial intelligence gives detailed response as much detailed are the question asked: if an individual don't know exactly what question formulate, the recommendations provided won't probably be specific enough for his situation. According to Cavanaugh (2023), the value of financial advisors with strong interpretation and interpersonal abilities will grow; advisors understanding the difficulties of an investor and prioritizing trust and communication will keep on having a crucial role on investors' success.

In order to understand the reasons why AI can't replace human advisors for the moment, I'm going to consider some essential qualities for effective financial advice as reported by Benartzi (2023) on the Wall Street Journal:

- Debiasing: this first aspect consists in helping customers to avoid error due to behavioral tendencies. A good financial advisor can help customer with proper investments according to their specific time horizons, however a working paper by Yang Chen (2023) showed that ChatGPT has the same behavioral tendencies and biases that advisors tend to reduce. Benartzi (2023) propose the creation of metarules (a rule that governs other rules), to overcome these biases: for instance, when an AI service propose a certain advice, it should also consider the reasons why that advice could be incorrect, forcing AI to investigate what it may have missed. Debiasing AI models by applying metarules is in any case easier than debiasing human reasoning;

- Empathy: this is a crucial quality for advisors, since they have to support customer during markets' high volatility to prevent emotions to damage their long-term objectives. In this case ChatGPT perform very well and since many times advisors cannot reassure their customers due to time or ability, AI services may deliver empathetic responses customized for every client;
- Accuracy: even if AI can be debiased, it has to focus on accurate advices about investments, taxes and more. At present AI models are quite unreliable and they make several mistakes. This issue may be solved with external tools, like calculators and reliable financial databases for instance, in order to supplement its weaknesses.

On the other hand, these qualities favor artificial intelligence:

- Best interest: another quality AI may excel is best interest. Even if advisors should act in the clients' best interest, they may in some cases recommend some products to increase their gain. In theory, AI should be less subject to conflicts of interest because, differently from humans, it won't try to maximize his gains. However, in reality this aspect may be biased as it may invests in funds solely because they allocate more resources to marketing (and therefore are more famous) even if their quality is lower. Also in this case it is easier to monitor conflict of interest in AI models than in human advisors, with AI tools being able to auto-correct their decisions;
- Consistency: this point reflects the fact that investors with similar preferences should receive the same financial advices. While human advisors may show their own beliefs on the advice provided, AI should be able to achieve consistency by proposing the same recommendations to clients with similar needs.

AI will play an important role in financial advisory, however a lot of improvements are still needed. A good solution may be an hybrid model where humans collaborate efficiently with AI, helping humans to expand their thinking, giving different opinions and serving more people (Benartzi, 2023). Humans should cooperate rather than compete with AI to deliver improved services, for instance providing oversight to AI models, providing complementary services (emotional and ethical guidance for example) and feedback to these services.
In the end, although this combination of artificial intelligence and robo-advisors could produce great automation and precision, humans remain irreplaceable in particular in field like comprehensive understanding and emotional intelligence (Kubera, 2023).

A sentence I agree particularly with is the one pronounced by Karim Lakhani, a professor at Harvard Business School specialized in technology and AI, who claims that "AI is not going to replace humans, but humans with AI are going to replace humans without AI."

### 2.4.2 Generative AI as a game changer in Robo-Advisory

The previous generation of robo-advisors was groundbreaking but often lacked the comprehension required for customized financial advice. Even if they were user-friendly and accessible, their advices were usually too simple. The next generation of robo-advisors (RA) will likely be based on generative AI models, constantly trained and improved. They could be imagined as a financial advisor that not only has the knowledge to pass the most important financial exams (like Securities Industry Essentials exam, Certified Financial Planner exam and Chartered Financial Analyst, for instance) and to have a strong comprehension of regulations, but also that knows historical events and their implications in the markets.
These are defined by SEI Investments Company (2023) as generative robo-advisors: differently from traditional advisors, generative robo-advisors will be trained to stay neutral and to focus only on customers' interests. Both robo-advisors and artificial intelligence has been used for several years in the financial sector but with the introduction of generative AI models, there is a great opportunity for AI to improve traditional financial services. Even though GenAI is still in its early days, it has the potential to improve the financial industry: the combination of efficiency and scalability of RA with the creativity of generative AI, may offer a new generation of AI-based opportunities. According to Merav Ozair's article published on the Nasdaq website (2023), generative AI hype has created a strong competition between tech companies like Google, Meta, Microsoft and Apple. Microsoft in particular bet a lot on it, with a multi-billion investment plan in OpenAI, the company that created ChatGPT. However, this interest cannot be found only on big tech companies but also in major financial institutions, as described in Chapter 4. Several models employing AI and ML may have been developed in the last years, but they now seem too basic compared to ongoing advancements in generative AI and if robo-advisors didn't manage to replace financial advisors, these new applications could (Ozair, 2023).
FinGPT represents an innovative open-source framework created to apply large language models (LLMs) in the financial sector. Among its application we can find robo-advisors, quantitative trading, portfolio optimization, risk management and more. FinGPT aims to have a crucial role in data acquisition in order to enhance research, collaboration and innovation in finance, providing researchers with accessible and transparent resources to develop their

FinLLMs. Moreover, it provides an accessible solution for introducing AI in finance, facilitating therefore practical applications in the financial industry (Yang, 2023).
In the paper written by Liu (2023), is provided an example that shows the capabilities of FinGPT to deliver professional financial services, giving appropriate prompts. Crafted prompts are used to generate a new analysis of Apple stock on March 3rd 2023, focusing on potential trends and price movements. After giving an initial prompt asking to give a brief summary of the news and analyze possible trend of Apple stock price based on different assumptions, FinGPT starts with summarizing the news like company's fundamentals, cloud initiatives and new investment in India and gives its analysis on the plausible influence of this factors. In the end, it defines a positive outlook on the stock price, since there are different positive catalysts that could drive the stock price higher, stressing however that the trend may be very volatile in connection to eventual risks that investors should note, derived from the analyzed negative news articles.

# Chapter 3 – Quantitative Analysis

## 3.1 Quantitative approaches in Robo-Advisory

In order to achieve the expected results when using robo advisory services, it is important to identify risk models that have a major stability over time. Due to automated platforms, different approaches may be implemented with little cost but with higher value for customers. These approaches, reflecting different objectives, are discussed below as described by Scholz (2021). It should be noted that these principles are not new and therefore are already used by financial advisors and industry experts. Hence, even though they are not specific or only used in robo-advisory services, here is just provided an example of the main financial principles adopted in several RA algorithms.

Moreover, before delving into the chapter, it is important to briefly recall what RNN and LMST are, starting from Machine Learning (ML) and Artificial Neural Networks (ANN) definitions. Machine Learning is a branch of artificial intelligence focused on the development of statistical algorithm able to perform different tasks without explicit instructions. Among ML models, Artificial Neural Networks are built following the principles of neuron network inside a brain: it is based on a collection of nodes defined artificial neurons where each connection can provide signal to other neurons connected to it. Neurons are then grouped into layers, and signals go from the first layer (input layer) to the last layer (output layer). Recurrent neural network (RNN) is a type of ANN which uses sequential data or time series data, and in particular LSTM (Long Short-Term Memory) is a specific RNN architecture that captures long-term dependencies, making it efficient in sequence prediction tasks. Differently from traditional neural networks, LSTM employ feedback connections (a process that allow the neural network to remember past data when processing the next output) in order to process sequences of data instead of single data points and this is especially useful in predicting patterns like time series (Wikipedia, 2023b; Wikipedia, 2023c; IBM, 2021; Saxena, 2021).

### 3.1.1 Maximization of the Diversification Effect

It is now widely known that to obtain a stable portfolio, it is fundamental to diversify. This characteristic is at the base for the development of modern robo-advisors, also because it can suggest different diversification ratios to less informed clients. To obtain this diversification, there is only the need to maximize the diversification ratio *DR*, defined as:

$$\max_{w_i} DR = max \frac{\sum_{i=1}^{N} w_i \cdot \sigma_i}{\sigma_p} \tag{1}$$

with $\sum_{i=1}^{N} w_i = 1$ and $w_i \geq 0 \quad \forall \quad i = 1, \ldots, N$

where $w_i$ is the weight of the asset $i$ in the portfolio, $\sigma_i$ is the volatility of the asset, $\sigma_p$ is the portfolio volatility and $N$ is the number of assets. From the formula it is possible to see that the *DR* corresponds to the portfolio risk with no diversification divided by the portfolio risk with diversification. Therefore, in this type of portfolio optimization, the ratio of the weighted individual risks of the asset classes to the actual portfolio risk is maximized. This effect (in the case of only two assets) is illustrated in Figure 13, while inverting the two axes and plotting the diversification ratio generates the Figure 14.

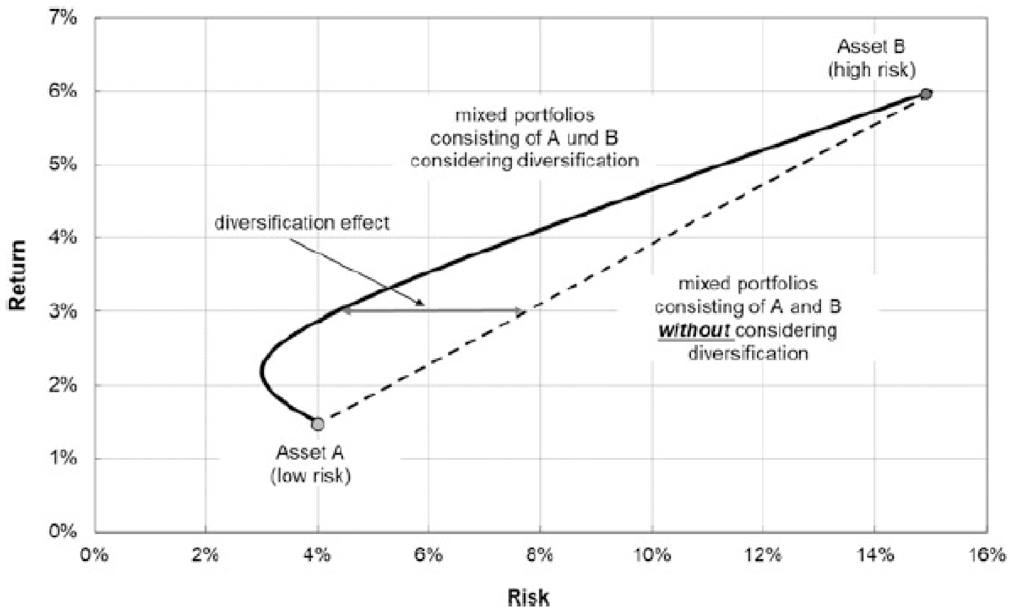



*Figure 13. Diversification effect. Source: Scholz, P. (2021) 'Robo-Advisory. Investing in the Digital Age', Palgrave Studies in Financial Services Technology.*

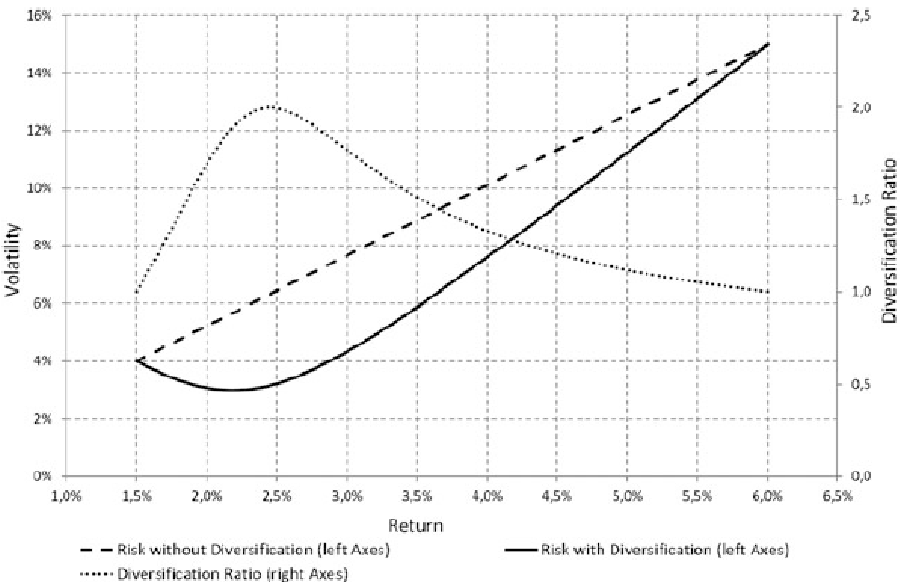



*Figure 14. Diversification ratio. Source: Scholz, P. (2021) 'Robo-Advisory. Investing in the Digital Age', Palgrave Studies in Financial Services Technology.*

### 3.1.2 Equal distribution of risks

In order to hold every desired financial instrument and keep a balance in terms of risk, it is possible to assign the same risk contribution to each instrument in the portfolio. This method, the so called 'risk parity' or 'equally weighted' portfolio is based on the fact that the portfolio is optimized so that all products have the same contribution to the total risk (Teiletche et al., 2010). Hence the percentage contribution to the total risk (PCTR) of asset *i*, is equal to 1 over the number if assets *N*:

$$PCTR_i = \frac{1}{N} \quad \forall \quad i = 1, \ldots, N \tag{2}$$

In this way it is possible to avoid structural cluster risks (risks that arise when investing in similar businesses but with different risk classifications) and the correlations between asset classes are taken into account.

### 3.1.3 Risk minimization

For more risk-averse clients also portfolios that minimize the overall risk should be considered (Clarke et al., 2011):

$$
\begin{aligned}
\min_{w_i} \sigma_p^2 &= \min \sum_{i=1}^{N}\sum_{j=1}^{N} w_i \cdot w_j \cdot \sigma_{i,j} \\
&= \min_{w_i} \underbrace{\sum_{i=1}^{N} w_i^2 \cdot \sigma_i^2}_{\text{single risk part}} + \underbrace{\sum_{i=1}^{N}\sum_{\substack{j=1 \\ i \neq j}}^{N} w_i \cdot w_j \cdot \sigma_{i,j}}_{\text{diversification part}}
\end{aligned}
\tag{3}
$$

with $\sum_{i=1}^{N} w_i = 1$ and $w_i \geq 0 \quad \forall \quad i = 1, \ldots, N$

where $\sigma_p^2$ is the portfolio variance and $\sigma_{i,j}$ is the covariance of asset *i* and *j*.

This kind of portfolio is known for outperforming in equities, however, as in the first case, it does not guarantee that every expected product is included in the portfolio and it may advise high bonds ratio in the portfolio since it considers their particular low volatility but not their low returns.

### 3.1.4 Return Forecasts

In the previous methods return forecast were omitted to avoid forecasting errors, however it may also be important to have also a look at market movements. For this reason, it is crucial to consider also Markowitz mean variance optimization (MVO) that employs expected returns:

$$U = \mu_p - \lambda \cdot \sigma_p^2 \tag{4}$$

where $\mu_p$ is the expected portfolio return, $\sigma_p^2$ is portfolio variance and $\lambda$ is the risk aversion parameter. Here the main issue is that the return estimates for each component are assumed to be the mean of the distribution and this may generate undesirable effects, like considering as perfect substitutes assets with similar expected returns. These negative aspects can be reduced using the Black-Litterman model, that integrates the MVO model, as later described, modifying the return estimators by giving greater attention to higher confidence forecast and by considering forecasts of highly correlated markets (Scholz, 2021).

### 3.2 Portfolio Optimization Methods

According to Beketov (2018) analysis based on a set of 28 RA in which 15 are the best RA services in 2017 according to Business Insider Intelligence (Business Insider, 2017) and 13 are the most relevant systems to German investors according to the Capital journal (Dohms, 2017), the typical RA's workflow is composed by the following phases:

- Preselection of ETFs;
- Identification of investors' risk profiles and objectives;
- Portfolio optimization according mainly to Markowitz Modern Portfolio Theory (MPT);
- Threshold-based rebalancing;
- Performance monitoring.

Regarding the third element, Beketov study shows how, in the analyzed robo-advisors, MPT is the most frequently adopted methodological framework with a percentage of 39,7%, followed by Sample Portfolios (27,4%) and Constant Portfolio Weights (13,7%).
Although MPT is a real quantitative method, the other two are just mere definitions provided on the companies' websites, including methods not identifiable in detail. Even though these last two frameworks are used more often, they attract lower assets under management: MPT has the highest AuM volume, followed by Full-Scale Optimization (a method that implements customer-specific utility functions and considers risk of catastrophic loss) and Black-Litterman model. Figure 15 displays the relation between the methods' occurrences (the percentage of use of these methods in RA systems) and the respective volumes of RA assets under management (million USD). For clarity, only the names with higher values are given.

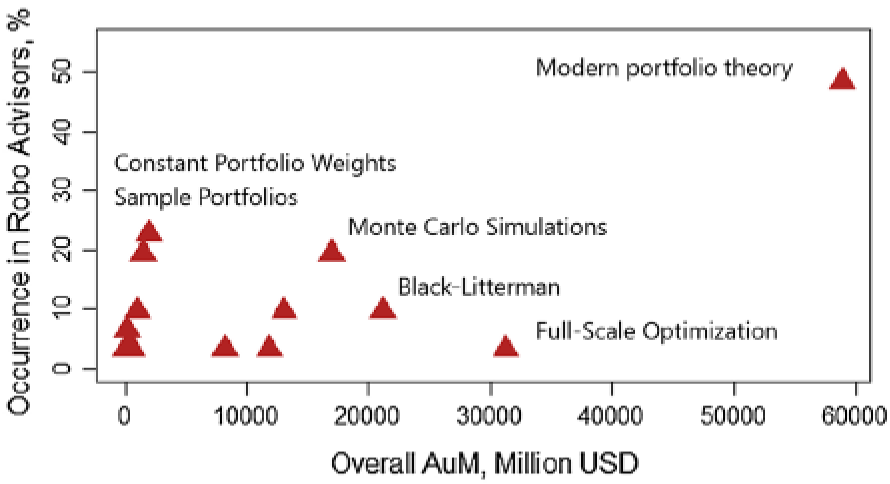



*Figure 15. Relation between methods' occurrences and RA AuM. Source: Beketov, M., Lehmann, K., Wittke, M. (2018) 'Robo Advisors: quantitative methods inside the robots'.*

Generally, there seems to be a clear gap between the methods applied in robo-advisors and newer and innovative methods. MPT presents several issues like input sensitivity and estimation error maximization, therefore many components of the original method are modified or integrated with multiple methods, defined by Beketov (2018) as 'Multidimensional improvement of Modern Portfolio Theory'. As it is possible to see in Figure 16, different methods can be applied to improve MPT framework in RA, like Black-Litterman model to increase expected returns, Monte Carlo simulation to optimize algorithms, standard deviations and Value at Risk to improve risk measures, implementing covariance to better understand asset correlations and use different constraints such as weights constraints and transaction cost minimization. Obviously, the methods mentioned are those used in RA services and are not a comprehensive list of methods that can be applied to improve the MPT framework in general. Finally, combining these methods may deviate considerably from the original MPT framework and it is a subject of debate if these combinations represent new approaches or just modifications to MPT. The use of alternative risk functions like VaR for instance, is usually considered an alternative rather than improvement of this framework, however it is presented here just to show how different methods can improve RA services. Significant changes are therefore going to affect deeply the robo-advisors' sector: innovative methods, with more complex features and approach, including also generative artificial intelligence (as we will see later), could be applied to improve performances and attract new investors.

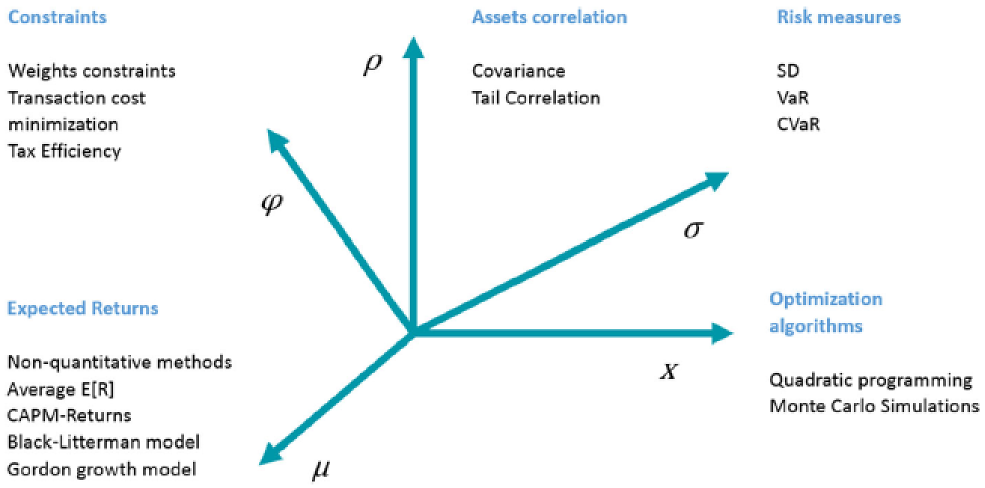



*Figure 16. Scheme of 'Multidimensional improvement of MPT'. Source: Beketov, M., Lehmann, K., Wittke, M. (2018) 'Robo Advisors: quantitative methods inside the robots'.*

### 3.2.1 Implementation of Black-Litterman Model to Modern Portfolio Theory

According to Min-Yuh Day (2018), robo-advisors system architecture is made up of different module written in Python with the aim of creating an optimal portfolio.
Different type of modules can be implemented but those used in the paper are:

- Data collection module: collects data automatically, extracting daily adjusted closing price and putting data into a database;
- Data preprocessing module: calculate log returns and reshape data into arrays that can be processed by the Long Short-Term Memory (LSTM) module;
- LSTM Training Module: trains the model to comprehend log-term tendencies and patterns to provide forecasts;
- Forecasting module: makes forecast and converts these values in usable formats;
- Data visualization module: visualizes the data of losses, root mean square errors (RMSE) and price movements;
- Portfolio optimization module: generates investor views from the result obtained by the previous module and calculates optimal portfolio weights.

The input of this research are daily adjusted closing prices of ETFs, collected for 10 years, from 03/01/2007 to 30/12/2016, extracted from Yahoo Finance.
The tickers of the chosen ETFs to build the portfolio are respectively:

- IVW: S&P 500 Growth ETF (U.S. Large Cap Growth Equities);
- IVE: S&P 500 Value ETF (U.S. Large Cap Value Equities);
- IWN: Russell 2000 Value ETF (U.S. Small Cap Value Equities);
- IWO: Russell 2000 Growth ETF (U.S. Small Cap Growth Equities);
- EFA: MSCI EAFE ETF (Foreign Developed Equities);
- EEM: MSCI Emerging Markets ETF (Emerging Markets Equities).

A manager is assumed to rebalance portfolio weights at the beginning of each quarter to obtain higher cumulative returns with respect to S&P 500 in 2016 (benchmark).
To perform portfolio optimization, investor views generated by the LSTM forecasting module are used as an input for the Black-Litterman model. This requires using the LSTM forecasting module to generate investor views about the performance of chosen assets in each quarter.
The investor views are prepared using the LSTM model to predict the quarterly return, using the past 120 daily returns as training data and forecasting the daily return for the next 60 days.
In Table 4, are analyzed the initial prices of the selected ETF and then are compared the

predicted returns and actual returns, the predicted price and the actual price. Finally, the difference between predicted and actual returns is shown.

| | IVW | IVE | IWO | IWN | EFA | EEM |
|---|---|---|---|---|---|---|
| **Initial Price** | 113.64 | 85.96 | 137.26 | 89.32 | 55.18 | 30.57 |
| **Predicted Return** | 4.99% | 5.57% | 3.10% | 3.12% | 2.84% | -0.34% |
| **Actual Return** | 0.73% | 2.41% | -5.05% | 1.77% | -1.14% | 9.39% |
| **Predicted Price** | 119.30 | 90.75 | 141.51 | 92.10 | 56.75 | 30.47 |
| **Actual Price** | 114.47 | 88.04 | 130.33 | 90.90 | 54.55 | 33.44 |
| **Difference** | 4.26% | 3.15% | 8.14% | 1.34% | 3.98% | -9.73% |

*Table 4. Forecasted cumulative returns and prices. Source: Day, M.Y., et al. (2018) 'AI Robo-Advisor with Big Data Analytics for Financial Services'.*

With forecasted returns and prices for each ETF (Table 4), they formulate investor views of each quarter (Table 5). The investor views mechanism presented in Table 5, works as follow: in investor view 1, U.S. Large Cap Growth Equities will outperform/underperform U.S. Large Cap Value Equities by X%, in investor view 2, U.S. Small Cap Growth Equities will outperform/underperform U.S. Small Cap Value Equities by X% and in investor view 3, Foreign Developed Equities will outperform/underperform Emerging Markets Equities by X%. In the end, the difference between predicted and actual values is presented.

| | View 1 | View 2 | View 3 |
|---|---|---|---|
| **Predicted** | -0.58% | -0.02% | 3.17% |
| **Actual** | -1.68% | -6.82% | -10.53% |
| **Difference** | 1.10% | 6.80% | 13.71% |

*Table 5. Predicted investor views. Source: Day, M.Y., et al. (2018) 'AI Robo-Advisor with Big Data Analytics for Financial Services'.*

After this step, the historical average returns and the covariance matrix of each asset are used as input for the Black-Litterman (B-L) model. Then, they combined investor views obtained before with historical return using the following B-L formula to generate posterior mean and posterior covariance matrix:

$$E[R] = [(\tau\Sigma)^{-1} + P'\Omega^{-1}P]^{-1}\ [(\tau\Sigma)^{-1}\Pi + P'\Omega^{-1}Q] \qquad (5)$$

where:

- E[R] is the new (posterior) combined return Vector (N x 1 column vector);
- $\tau$ is the weight-on-views scalar, set between 0 and 1 (it explains the uncertainty of the estimated equilibrium returns);
- $\Sigma$ is the covariance matrix of excess returns (N x N matrix);
- P is a matrix that identifies the assets involved in the views (K x N matrix or 1 x N row vector in the special case of 1 view);
- $\Omega$ is a diagonal covariance matrix of error terms from the expressed views representing the uncertainty in each view (K x K matrix);
- $\Pi$ is the implied equilibrium return vector (N x 1column vector);
- Q is the view vector (K x 1 column vector).

Next, they implemented these values in the Mean-Variance Optimization framework (using the same technique developed by Markowitz) to obtain the optimal portfolio weights, reported on Table 6. In other words, they first integrate LSTM investor views with the Black-Litterman model to generate posterior mean and covariance matrixes that combined with Markowitz portfolio weights theory, created the final portfolio weights.

| **Return Vectors and Resulting Portfolio Weights** | | | | |
|---|---|---|---|---|
| **Asset Class** | **Posterior Return** | **Historical Return** | **Difference** | **Black-Litterman Weight** |
| **040_IVW** | 1.35% | 2.52% | -1.17% | 2.31% |
| **041_IVE** | 1.00% | 1.19% | -0.18% | 83.65% |
| **058_IWO** | -0.64% | 0.91% | -1.55% | 0.00% |
| **057_IWN** | -0.77% | -0.21% | -0.56% | 4.58% |
| **086_EFA** | -1.00% | -0.62% | -0.39% | 9.46% |
| **108_EEM** | -3.30% | -2.13% | -1.17% | 0.00% |

*Table 6. Return vectors and portfolio weights. Source: Day, M.Y., et al. (2018) 'AI Robo-Advisor with Big Data Analytics for Financial Services'.*

These weights are then used to rebalance the portfolio at the beginning of every quarter and the performance are reported in the Table 7 and Figure 17, where it is compared to the S&P 500. As it is possible to see on Table 7, the portfolio outperformed S&P 500 Index by 6,508% (the difference between the two annual return values), maintaining almost the same annual volatility. Moreover, also indicator like the Sharpe ratio and the Max drawdown presents more performing values.

| | Black-Litterman Portfolio - the LSTM Investor Views | S&P 500 Index |
|---|---|---|
| **Annual return** | 16.151% | 9.643% |
| **Annual volatility** | 13.897% | 13.169% |
| **Sharpe ratio** | 1.14697 | 0.76492 |
| **Max drawdown** | -10.105% | -10.306% |

*Table 7. Annual portfolio statistics. Source: Day, M.Y., et al. (2018) 'AI Robo-Advisor with Big Data Analytics for Financial Services'.*

Figure 17 compares the cumulative returns of three different portfolios and the S&P 500 index. At the end of the chosen timeframe, the Black-Litterman portfolio presented the highest returns, followed by the Markowitz portfolio. In the first six months, the equally-weighted portfolio had the lowest returns, while in the following six months, it surpassed the S&P 500 index, making the latter, the one with the lowest returns in the period analyzed.

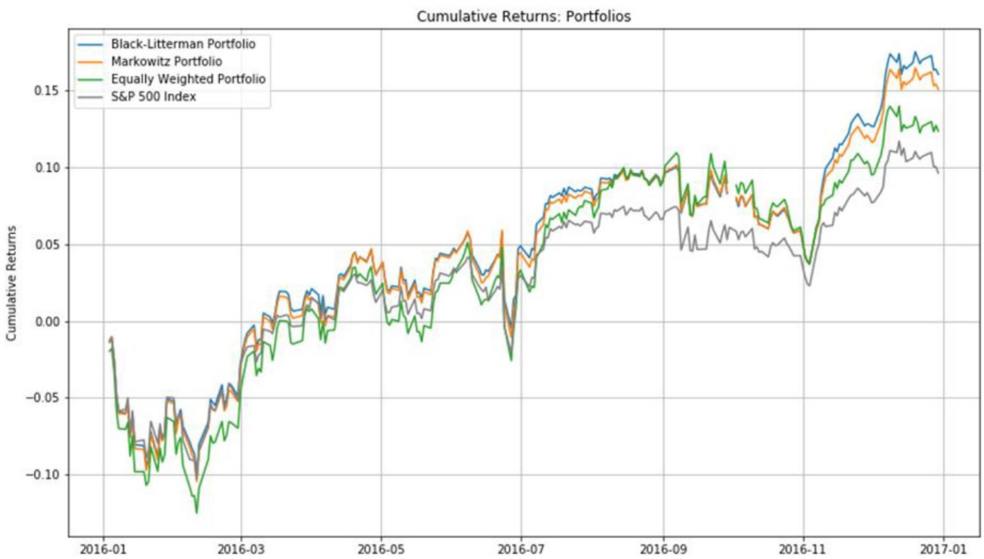



*Figure 17. Cumulated return of the LSTM based portfolio. Source: Day, M.Y., et al. (2018) 'AI Robo-Advisor with Big Data Analytics for Financial Services'.*

Beyond the study here analyzed, also Pantoja (2023) and Ta (2020) investigated how the combination of LSTM and Black-Litterman model can outperform S&P500, presenting positive results. Min (2021) managed to obtain higher cumulative returns too, but with higher volatility than S&P Index, combining B-L model with different machine learning algorithms. Each of these modular systems that advice asset allocation weights can be used as sub-systems for robo-advisors, offering customized portfolios to investors, according to their financial planning and objectives.

Hence, despite these studies found models that outperformed S&P500, it should be noted that even if robo-advisors will efficiently adopt these methods it's not certain they will beat the index. This in fact will vary significantly on the risk profile of the investor, the market conditions and the specific robo-advisor used. Some robo-advisor portfolios may outperform S&P500 in certain years or under specific conditions while others, will underperform (Hayes, 2023). The adoption of B-L model among RA services is roughly 10% when considered alone, but when considered as an additional component within the Modern Portfolio Theory framework, its rate of adoption could increase significantly (Beketov, 2018). Betterment, for instance, implement the B-L model in its RA optimization process, as well as Wealthfront and FutureAdvisor (Bjerknes, 2017), however as discussed by Pirner (2018), this method is more complex compared to classic mean-variance optimization and presents different limits such as the inability of truthful estimation and the behavioral biases views are subject to. In addition, imbalanced market weights can lead to distorted equilibrium portfolios and unclear views and confidence in the market may lead investor to avoid the complexity of the B-L model.
Despite the actual adoption of B-L model in several RA services, it's not clear if even its interaction with LSTM is used. It is likely that many existing RA are already adopting it but specific details about internal workings of RA are often proprietary, therefore without access to their internal methodologies it's quite challenging to confirm the exact techniques they employ. Finally, it is important to notice that in LSTM there are different issues that should be considered, such as the great amount of time requested in computation for the training and validation, market data noise implying that high performance on historical dataset does not guarantee the expected profits, and the complexity due to different data sources and characteristics.
Another approach to implement LSTM forecast is combining it with genetic algorithms, instead than with B-L model, as presented on the following section.

### 3.3 Genetic Algorithms

Genetic algorithms (GA) are among the most modern and best available methods used to solve complex tasks and problems associated with evolutionary processes. As species evolve and adapt to different conditions in nature, in the same way this algorithm tries to replicate this process artificially, through the use of software in which several parameter and solution are evaluated with the use of the so called "fitness functions", in order to select the best outcomes. These algorithms have been adopted in recent times to perform portfolio optimization, implemented also in robo-advisory, generating robust and dynamic portfolio optimization. The approach adopted by Soldatos (2022), employs concept of GA, including evolutionary portfolio modification and an assessment of its quality through the so called “fitness score”. In GA, fitness scores are employed to determine the acceptability of a solution to a specific problem, evaluating the quality of a portfolio across multiple factors (called “fitness factors”). The sum of the weighted fitness factors generates the fitness score: every fitness factor regards a specific aspect or financial criteria for a given financial instrument. Starting with a random initial portfolio, the optimal portfolio becomes the one with the highest fitness score after a certain number of iterations. Recently there have been a lot of implementations of GA with different existing models in order to improve robo-advisors performance, but the two I would explore in the following paragraph are GA and Recurrent Neural Networks (RNN) and GA and sentiment analysis.

### 3.3.1 RA prototype using GA and RNN

The allocation model by Muganda (2023) is trained on 10 years historical prices derived from ‘LRBc1-LCOc1 ICE RBOB Gasoline/Brent Crack Spread’, ‘LHOc1-LCOc1 ICE Heating Oil/Brent Crack Spread’ and ‘LRBc1-WTCLc1 ICE RBOB Vs WTI Crude Oil Energy Spread’. The predictive results of the RNN are put into a GA that performs the portfolio optimization, looking to minimize risk for a given return or maximize return with respect to a certain risk level. Different packages are used to implement RNN model to forecast prices of energy spread, then an LSTM model is employed to efficiently capture long-term dependencies. The predictions derived from these models are then incorporated in the genetic algorithm, designed to build a portfolio allocation strategy that optimizes investor returns, aiming at risk minimization and considering parameters like investment horizons, budget constraint and risk levels. To validate the model forecasting capability also two measures are computed: Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). These measures assess model’s accuracy verifying the deviation of the predicted values from the actual ones.

| Assets | Prices | Min | Mean | Max | Std. Dev | RMSE | MAE |
|---|---|---|---|---|---|---|---|
| Gasoline/Brent Crack Spread | Predicted | 1.290 | 9.314 | 20.347 | 6.260 | 2.3430 | 1.8846 |
| | Actual | 0.250 | 13.054 | 49.450 | 7.869 | | |
| Heating Oil/Brent Crack Spread | Predicted | 14.63 | 17.501 | 29.582 | 2.608 | 5.6275 | 4.7440 |
| | Actual | 4.864 | 18.732 | 68.051 | 10.615 | | |
| Gas /Crude Oil Energy Spread | Predicted | 9.235 | 17.417 | 26.586 | 4.700 | 2.2551 | 1.7907 |
| | Actual | 0.585 | 18.131 | 65.70 | 8.055 | | |

*Table 8. Results of LSTM network prediction performance analysis of spread prices. Source: Muganda, B.W., Kasamani, B. S. (2023), ‘Parallel Programming for Portfolio Optimization: A Robo-Advisor Prototype using Genetic Algorithms with Recurrent Neural Networks’.*

In Table 8 are summarized the actual and predicted summary statistics as well as the prediction performance metrics. A comparative analysis is presented as well as descriptive statistics, namely the minimum, mean, maximum, standard deviation, root mean square error and mean absolute error. Gasoline/Brent Crack Spread shows slightly difference with actual prices. Heating Oil/Brent Crack Spread shows higher predicted values, indicating an increase in prediction error for this asset. The Gas/Crude Oil Energy Spread instead displays similar metrics performance as those in the first case, demonstrating comparable accuracy for these assets. The results however indicate that LSTM model provides pretty accurate forecast, highlighting the utility of this RNN model as a robo-advisor prototype.

| **Portfolio Strategy** | **Portfolio Allocation** | **E(R)= 0.02 1 Week** | **E(R)=0.04, 1 Month** | **E(R)=0.08, 2 Months** | **E(R)=0.16, 3 Months** | **E(R)=0.32, 6 Months** |
|---|---|---|---|---|---|---|
| Risk Minimization by Genetic Algorithm | Heating Oil/Brent Crack Spread Allocation | 0.0034994 | 0.8244707 | 0.0010028 | 0.0045914 | 0.0034506 |
| | Gasoline/Brent Crack Spread Allocation | 0.9984682 | 0.0008905 | 0.8946378 | 0.9989343 | 0.9591886 |
| | Gas /Crude Oil Spread Allocation | 0.0020847 | 0.0003606 | 0.0038160 | 0.0030399 | 0.0042536 |
| | Portfolio Risk | 0.0018492 | 0.0067931 | 0.0216297 | 0.0227220 | 0.0346691 |

| **Portfolio Strategy** | **Portfolio Allocation** | $\sigma_p = 0.02$ **1 Week** | $\sigma_p = 0.04$, **1 Month** | $\sigma_p = 0.08$, **2 Months** | $\sigma_p = 0.16$, **3 Months** | $\sigma_p = 0.32$, **6 Months** |
|---|---|---|---|---|---|---|
| Return Maximization by Genetic Algorithm | Heating Oil/Brent Crack Spread Allocation | 0.9491266 | 0.0010667 | 0.9887244 | 0.947437 | 0.962538 |
| | Gasoline/Brent Crack Spread Allocation | 0.0039302 | 0.0007854 | 0.0016656 | 0.0041563 | 0.005626 |
| | Gas /Crude Oil Spread Allocation | 0.0019001 | 0.8986039 | 0.0034737 | 0.006856 | 0.0030394 |
| | Portfolio Return | 0.0069181 | 0.0018659 | 0.0014524 | 0.0044782 | 0.0001927 |

*Table 9. Results of Genetic Algorithm for portfolio allocation. Source: Muganda, B.W., Kasamani, B. S. (2023), 'Parallel Programming for Portfolio Optimization: A Robo-Advisor Prototype using Genetic Algorithms with Recurrent Neural Networks'.*

In Table 9 finally two different strategies mentioned above (risk minimization and return maximization) are presented, applied through GA, considering different time horizons ranging from 1 week to 6 months. In the risk minimization strategy, great part of the portfolio is allocated to the Gasoline/Brent Crack Spread, suggesting that the algorithm perceives significant less risk for this asset in those timeframes. On the other hand, in the return maximization strategy, Heating Oil/Brent Crack Spread has the biggest weight, indicating it provides the higher return for the acceptable risk levels. However, in the two-month timeframe, Gas/Crude Oil Energy Spread has a significant allocation, probably meaning that this asset has a more balanced risk-return profile and may contribute to maximize returns. The plots of these results are shown in Figure 18a, 18b and 18c where it is possible to clearly see actual and predicted spread prices.

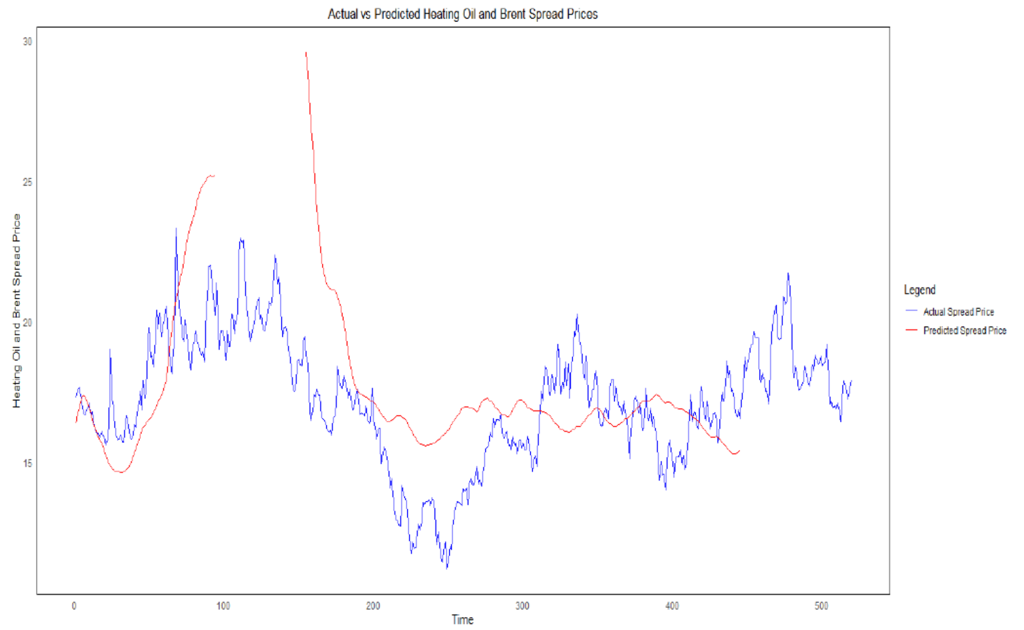



*Figure 18a.*


*Figure 18b.*


*Figure 18c.*

*Figure 18. a) LSTM network prediction plot of heating oil and brent spread prices(red line) compared with the actual spread prices (blue line); b) LSTM network prediction plot of gas and brent spread prices(red line) compared with the actual spread prices (blue line); c) LSTM network prediction plot of gas and crude oil spread prices(red line) compared with the actual spread prices (blue line). Source: Muganda, B.W., Kasamani, B. S. (2023), 'Parallel Programming for Portfolio Optimization: A Robo-Advisor Prototype using Genetic Algorithms with Recurrent Neural Networks'.*

Generally, the risk minimization strategy shows lower portfolio risk across all periods, while the other strategy demonstrates lower potential returns for higher risk targets. The risk in the first strategy increases along with the investment horizon, indicating uncertainty in longer time span investments. In the second strategy instead, the biggest returns are in the shortest timeframe horizon (1 week) and this may be due to the high volatility in Heating Oil/Brent Crack Spread that prevails in this portfolio.

The accuracy and efficiency of this model indicates potential applicability to larger dataset and use-cases like robo-advisory systems. Moreover, the genetic algorithm managed to pursue a balance between risk and return, optimizing financial returns while managing risk and highlighting its suitability in portfolio optimization tasks.

### 3.3.2 RA prototype using GA and BERT

In the analysis made by Leow (2021), a sector-based basket of ETFs is chosen, specifically ETFs of S&P500 stocks grouped by Global Industry Classification Standard (GICS), to create a diversified portfolio across sectors. Instead, for asset-based diversification, asset classes defined in the All-Weather portfolio are considered: equities, long-term bonds, intermediate-term bonds, gold and commodities. The adopted portfolio optimization techniques are

Modern Portfolio Theory and Constant Rebalanced portfolio, a simple method that rebalances the portfolio according to specific percentage for each asset in the portfolio.
Later, in order to capture overall market sentiment, tweets from specific accounts that follow financial insights are considered. Tweets are then converted into sentiments using Google's Bidirectional Transformer (BERT), in particular Araci (2019) variation that improved BERT for financial data (FinBERT), aiming at evaluating the comprehension of text and deriving negative or positive words that are used for financial sentiment classification.
Valance Aware Dictionary and sEntiment Reasoner (VADER) is also used as a method: it is a sentiment analysis tool, focused on sentiments inside social networks that produces "sentiment scores" connecting lexical features with emotion intensities. The sentiment score of a text is given by the sum of the intensity of each word in the text. The scope is to check for possible synchrony between stock data and tweet sentiments during volatile periods.
Tweets are converted to the correct time zone and sentiment scores are aggregated on a daily base. Exponential moving average (EMA) is then calculated with different time spans.
Defined $S$ as EMA-weighted sentiment and calculating Moving Average Convergence Divergence (MACD) of sentiments $S_{cd}$ by subtracting the 26-period sentiment EMA from the 12-period EMA, a bullish signal is generated if $S_{cd} > 0$ and a bearish signal if $S_{cd} < 0$.
If the estimate is bullish, an intuitive variation of MPT is to take more risk so that expected return along the efficient frontier is higher. Instead of changing directly the weights in MPT, it is considered first the optimal volatility from MPT and then the modified weights are calculated. This is displayed on Figure 19, with the Markowitz Bullet, using the All-Weather portfolio components. The efficient frontier illustrates the best possible expected level of returns for certain levels of risk. The dotted line represents the efficient frontier, the red star the Max Sharpe portfolio and the green star the minimum volatility portfolio.
DBC represents commodities, IEF intermediate-term bonds, VTI represents equities, TLT Long Term Bond and GLD is gold. The other stars represent portfolios associated with different level of expected volatility.

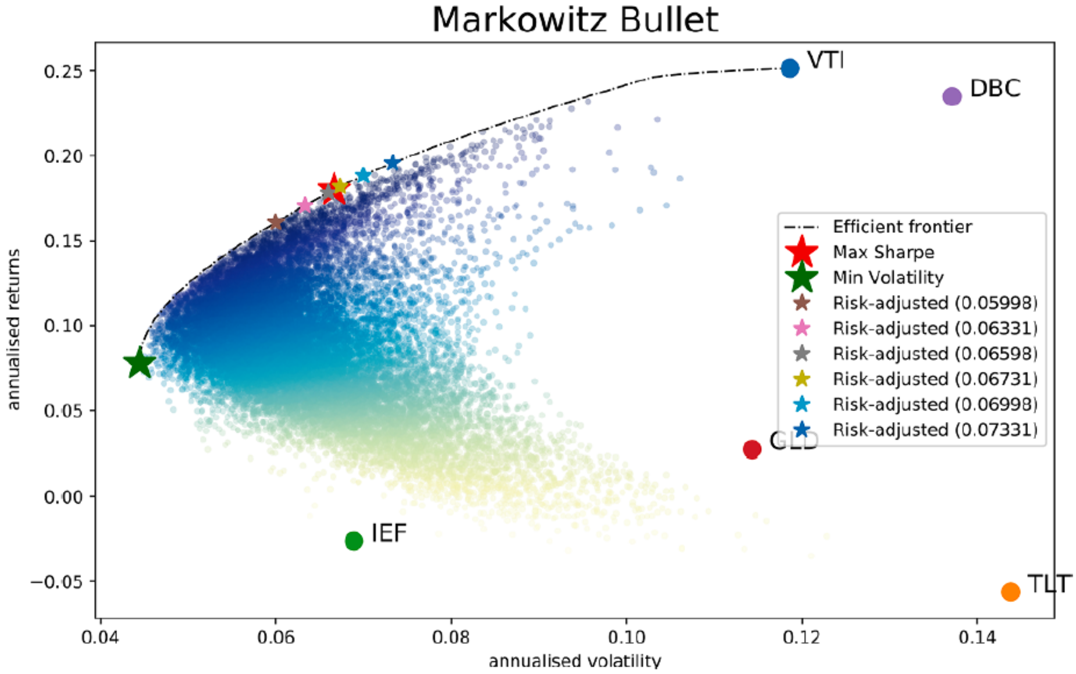



*Figure 19. Illustration of sentiment-adjusted MPT on Markowitz Bullet. Source: Leow, E.K.W, Nguyen, B.P., Chua, M.C.H. (2021), 'Robo-advisor using genetic algorithm and BERT sentiments from tweets for hybrid portfolio optimisation'.*

Optimization with genetic algorithms is then applied, using random set of weights with rebalancing at fixed intervals. A fitness score F is derived, depending on three different objectives: maximize cumulative returns, maximize Sharpe Ratio (SR) and minimize volatility. These objectives may refer to three different investors' risk profile: the first may be suitable for aggressive investors, the second for average investors who try to maximize risk-adjusted returns with SR and the third for conservative and risk averse investors.

The fitness score is considered for each person and the process is repeated for *p* population and *n* generations. The process is repeated with varying population and number of generation size. Tweet sentiment from CNBC (a news company), gave strong correlation with SPY (the ETF that tracks the performances of the S&P 500) stock data and therefore the following part is only based on their tweets. Initially, is considered only SPY and a simple modification function that include also CNBC sentiment, called "Sentimental SPY" portfolio.

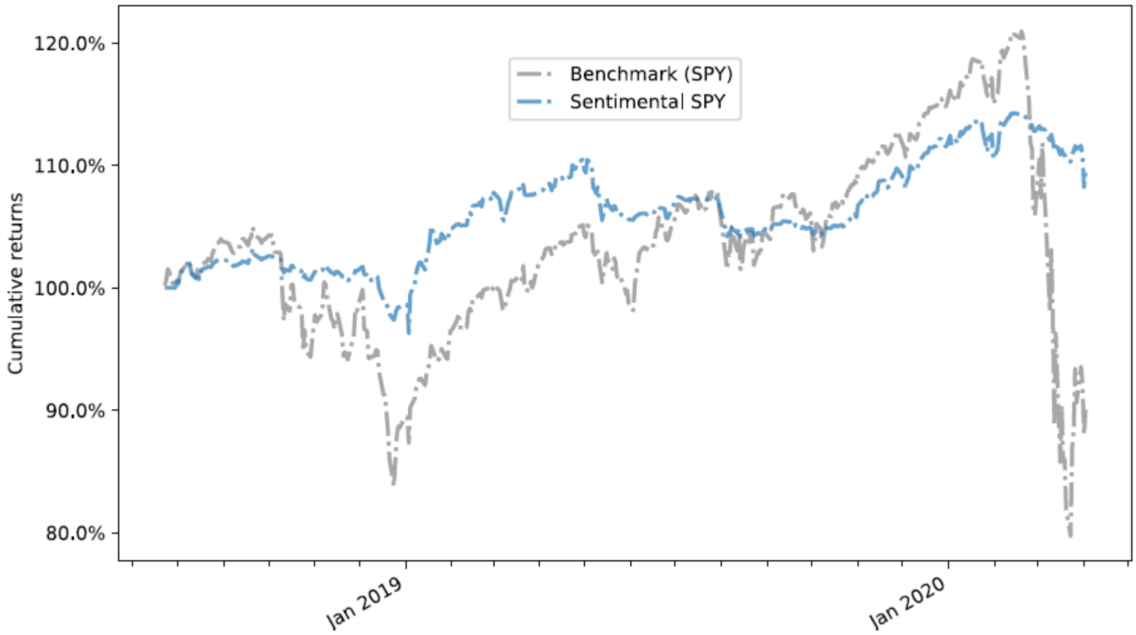



*Figure 20. Sentimental-SPY compared with SPY benchmark. Source: Leow, E.K.W, Nguyen, B.P., Chua, M.C.H. (2021), 'Robo-advisor using genetic algorithm and BERT sentiments from tweets for hybrid portfolio optimisation'.*

As displayed on Figure 20, just in this simple case, cumulative returns are better than SPY, since there is the possibility to exit the market when sentiments become negative. Unfortunately, the timeframe considered here is quite short, therefore it is not possible to know if these performances could be extended to longer timeframes.

Then a more sophisticated strategy is considered: constant-rebalanced All-Weather portfolio combined with sentiment analysis that generated investment signals, called "Sentimental All-Weather" (SAW), with the use of MACD and GA with different objectives explained above. The results are illustrated in Figure 21, where it is possible to see that in every scenario the SAW portfolio performed better than buy and hold SPY, increasing its feasibility as an improved portfolio algorithm.

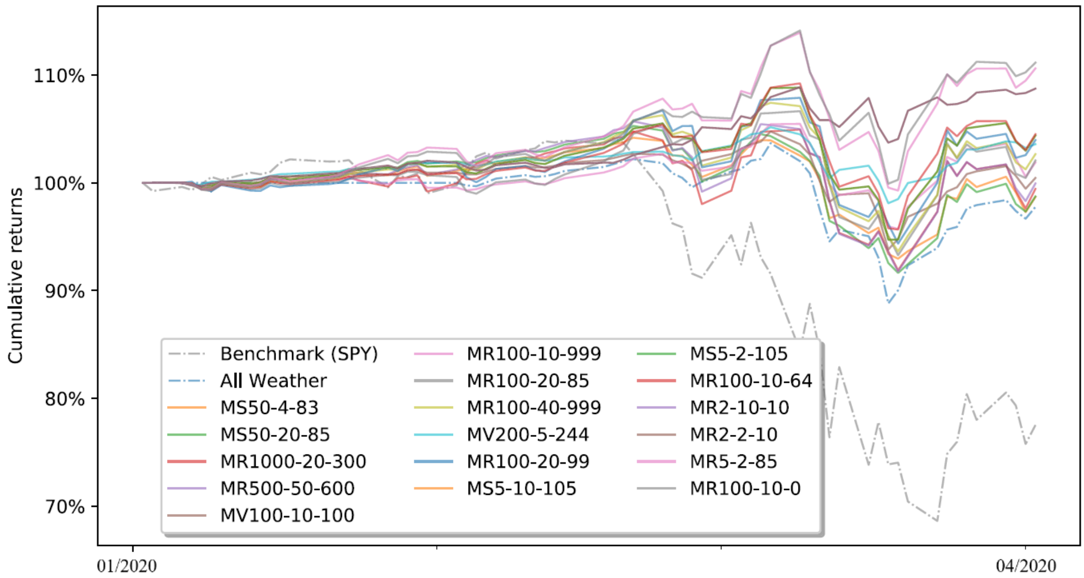



*Figure 21. Sentimental All-Weather (SAW) portfolio. Source: Leow, E.K.W, Nguyen, B.P., Chua, M.C.H. (2021), 'Robo-advisor using genetic algorithm and BERT sentiments from tweets for hybrid portfolio optimisation'.*

Another technique is to perform MPT optimization, modifying the volatility to derive sentiment-adjusted weights, generating a "Sentimental MPT" (SMPT) portfolio, always optimized with GA. The portfolio is always All-Weather but, in this case, MPT optimization is used instead of CR. As shown on Figure 22, also in this case the portfolio outperformed SPY performances, especially SAW optimized for minimum volatility and SAW optimized for maximum cumulative returns. Finally, these methods can be fully implemented into robo-advisory development framework and deployed as an end-to-end system, that means that the provider of the service delivers a complete functional solution without involving a third party.

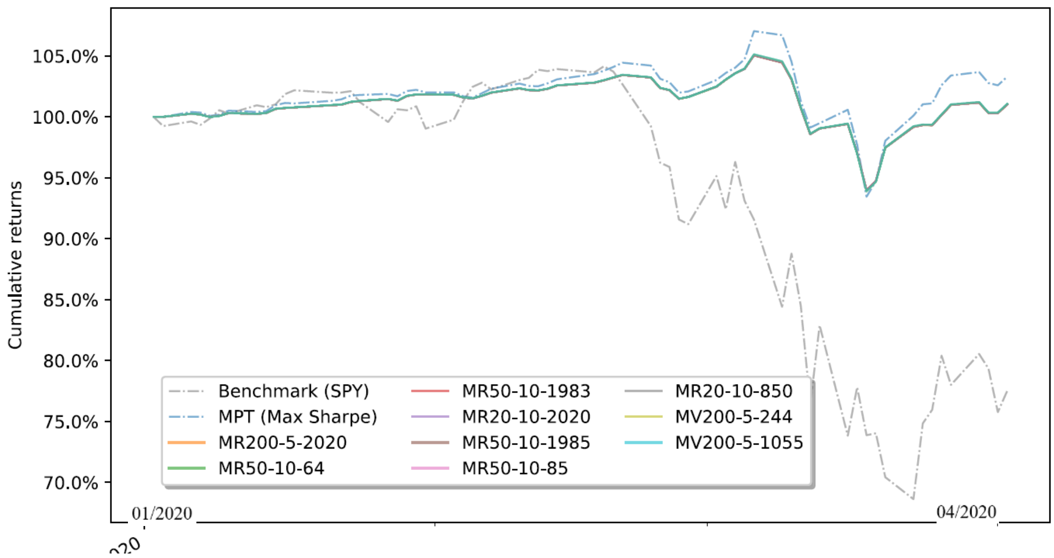
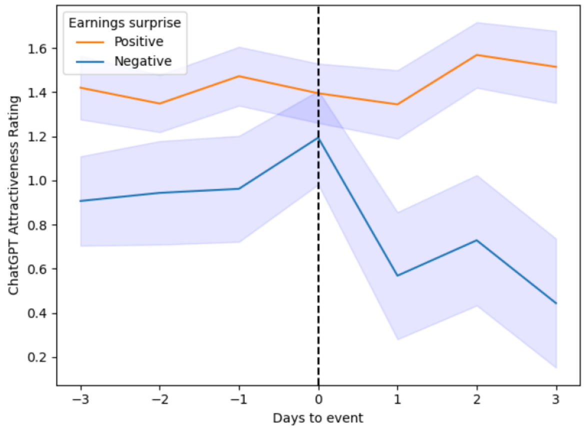


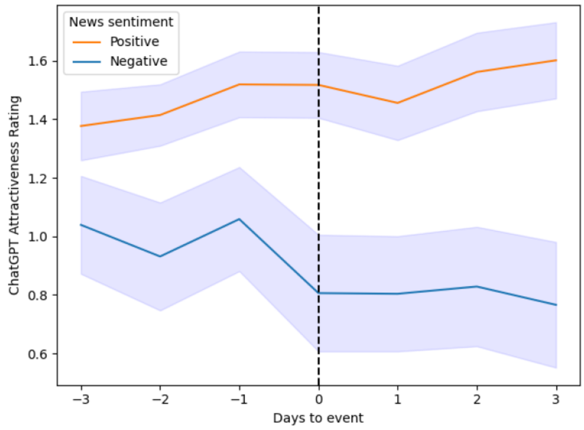

*Figure 22. Sentimental MPT (SMPT) portfolio. Note that 11 series are plotted but all SMPT portfolios overlap. Source: Leow, E.K.W, Nguyen, B.P., Chua, M.C.H. (2021), 'Robo-advisor using genetic algorithm and BERT sentiments from tweets for hybrid portfolio optimisation'.*

### 3.4 Implementation of ChatGPT to standard methods used in Robo-Advisory

The methods and quantitative models discussed in the first two sections are generally applied to standard robo-advisor services. However, in this paragraph the models illustrated above will be integrated with generative artificial intelligence, in particular ChatGPT, providing new insights and point of views concerning the applicability of this new tool. Different recent studies will be discussed below, trying to convey the efficiency of ChatGPT in forecasting. The first analysis presented is the one presented by Pelster (2023). In this paper the authors investigate the ability of ChatGPT to make informed investment decisions by conducting a live experiment in the second quarter of 2023, asking ChatGPT to forecast earnings announcements and their relative attractiveness. To fill the gap of its limited knowledge available up to September 2021, additional information from a web search engine is prompted. For every S&P 500 firm is asked to the generative AI model to provide a score on earnings forecast on a scale from -5 to +5 points, considering all available information of news and social media discussions. Then it is asked to rate the stock's attractiveness always on a scale from -5 to +5 points. The first result is that there is a positive correlation between GPT ratings and actual earnings. Moreover, on Figure 23 are plotted the average attractiveness ratings for positive and negative earnings surprises in Panel (a) and for positive and negative news events in Panel (b). The reaction of ChatGPT attractiveness ratings correlates pretty well with earnings announcements and news in general: stocks with positive earnings receive higher attractiveness ratings compared to stocks with negative earnings. In particular they showed that ratings are lowered after bad earnings announcements or bad news, while in the case of positive news/earnings, the rating remains almost identical at first with a slight increment later. This shows the engine capability to react to news and adjust the sentiment based on them and actual earnings. Interestingly, it is possible to see that ratings for negative news events are adjusted even before the news event, which could be due to news leakages or rumors.




(a) Earnings announcements




(b) News events

*Figure 23. ChatGPT rating changes in response to earnings announcements and news events. Earnings surprise are defined as the average deviation between analyst EPS consensus forecasts and actual EPS relative to the stock price. Source: Pelster, M., Val, J. (2023) 'Can ChatGPT assist in picking stocks?'.*

Then, different portfolios based on attractiveness ratings are created and the overall return of the portfolio is calculated over the following 30 days. Portfolio 1 includes stocks with attractiveness rating ≤ -3, portfolio 2 ratings > -3 and ≤0, portfolio 3 ratings > 0 and ≤ 3, portfolio 4 ratings above 3. Figure 24 shows the average next-month returns of 4 portfolios formed, together with the next-month return of an investment in the S&P500. Then are plotted the equal-weighted average of portfolio returns over the next 30 days. While the portfolio composed with higher ratings provides lower returns at the beginning, higher returns are given by the portfolios with higher rating in the final part of the sample period. The results of this strategy therefore vary over time, suggesting to be cautious when using ChatGPT in investment decisions as it might not always provide reliable financial advice.

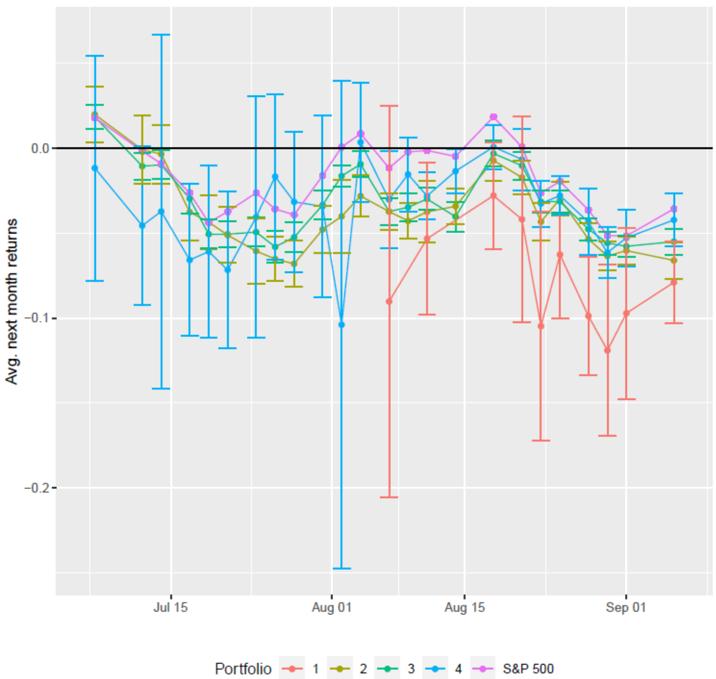



*Figure 24. Average next-month returns of portfolios based on ChatGPT-4 ratings. Source: Pelster, M., Val, J. (2023) 'Can ChatGPT assist in picking stocks?'.*

Another perspective is given by Romanko (2023), who lets ChatGPT select most performing stocks and then tests them with modern quantitative finance strategies. In the first step is asked to the GenAI model to compose three different group of stocks from S&P 500 that should outperform the index: these three groups differ only in size and they contain respectively 15, 30 and 45 stocks. Figure 25 shows the stocks for the first group.

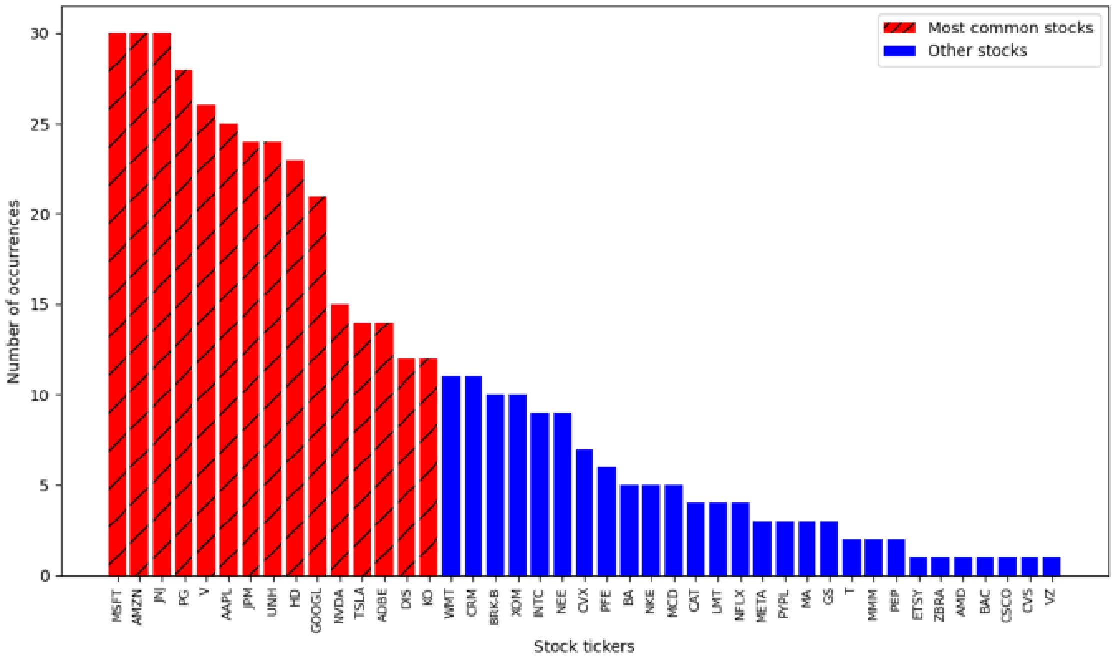



*Figure 25. Universe of 15 stocks selected by GPT-4. Source: Romanko, O., Narayan, A., Kwon, R.H. (2023) 'ChatGPT-Based Investment Portfolio Selection'.*

Later is asked to assign asset weights to each stock in the portfolios considering diversification, market capitalization, growth potential and stability. Figure 26 shows sector weights for the so called 'GPT-weighted' portfolios of 15, 30 and 45 stocks. As the portfolio size increases, it is possible to see a greater diversification across various sectors: even if 'IT' remains dominant, 'Energy' and 'Communication Services' sectors play a greater role.

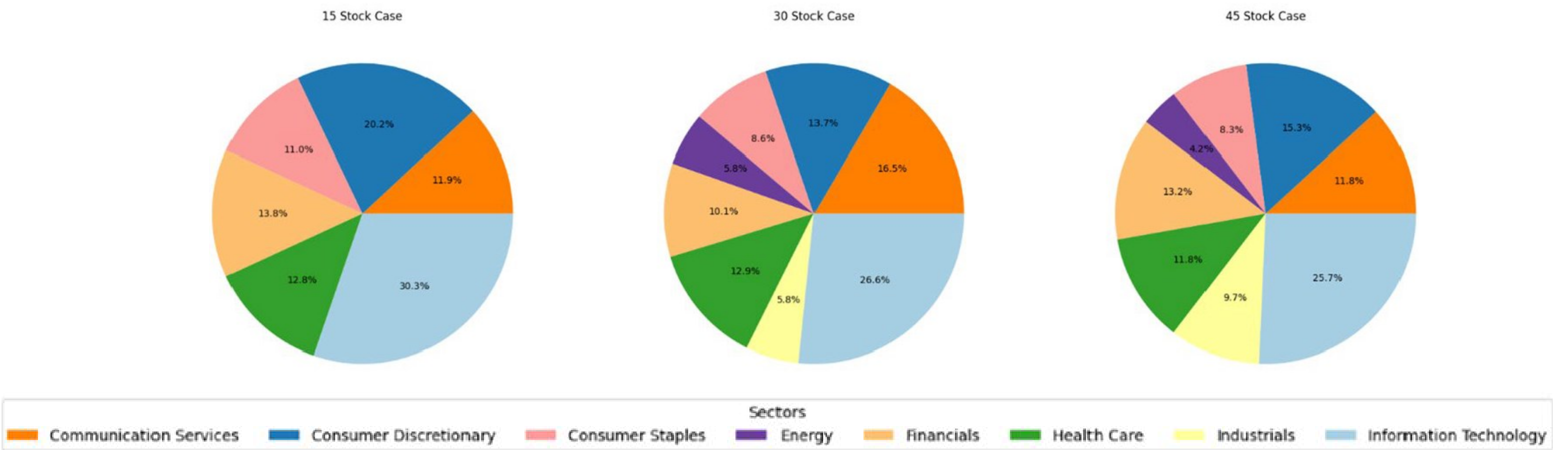



*Figure 26. Allocation of weights to sectors for 'GPT-weighted' portfolio of 15, 30 and 45 stocks. Source: Romanko, O., Narayan, A., Kwon, R.H. (2023) 'ChatGPT-Based Investment Portfolio Selection'.*

In order to compare this portfolio with other important portfolio models, also an equally weighted portfolio is built as well as three Markowitz mean-variance portfolio model, in particular: minimum variance portfolio, maximum expected return portfolio and maximum Sharpe ratio portfolio. Bound constraints are then applied to guarantee non-zero weights with lower bound equal half of the weight in the equally weighted portfolio and upper bound twice the weight of the equally weighted portfolio:

- $0.03 \leq w \leq 0.13$ for the 15 stocks portfolio;
- $0.02 \leq w \leq 0.07$ for the 30 stocks portfolio;
- $0.01 \leq w \leq 0.05$ for the 45 stocks portfolio;

where $w$ is the asset weight.

Figure 27 illustrates the efficient frontier of portfolios from the universe of 15 stocks, as described above.

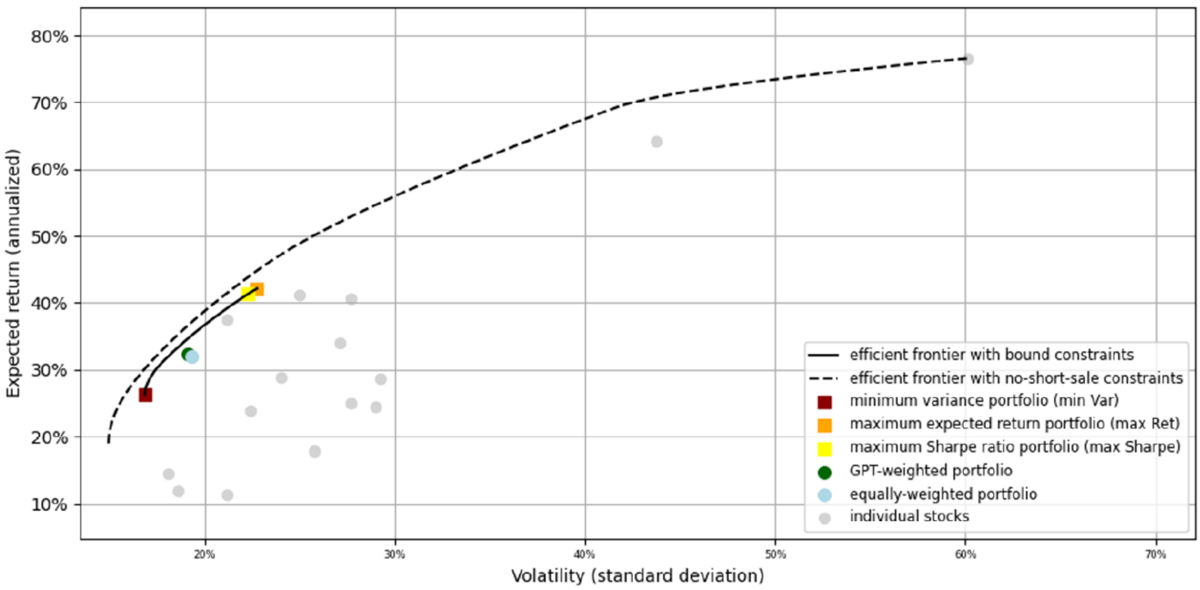



*Figure 27. Risk-reward profiles of portfolios from the group of 15 stocks. Source: Romanko, O., Narayan, A., Kwon, R.H. (2023) 'ChatGPT-Based Investment Portfolio Selection'.*

The figure indicate that the 'GPT-weighted' portfolio is close to the efficient frontier and almost equidistant to the maximum expected return and minimum variance portfolio. Moreover, the 'GPT-weighted' portfolio is slightly closer to the efficient frontier, compared to the equally weighted portfolio, showing its major efficiency.

Later, cardinality constraints are added to ensure that a portfolio has a specific number of assets (cardinality K= 15, 30, 45) and the new three efficient frontier are computed. Comparing the performance of portfolios obtained before with the three obtained with cardinality constraint (minimum variance cardinality-constrained portfolio, maximum expected return cardinality constrained and maximum Sharpe cardinality constrained), it appears that the first 3 portfolio outperform the ones with cardinality constraints. Figure 28 shows cumulative returns for the 15 stocks portfolio and Table 10 its relative risk metrics.

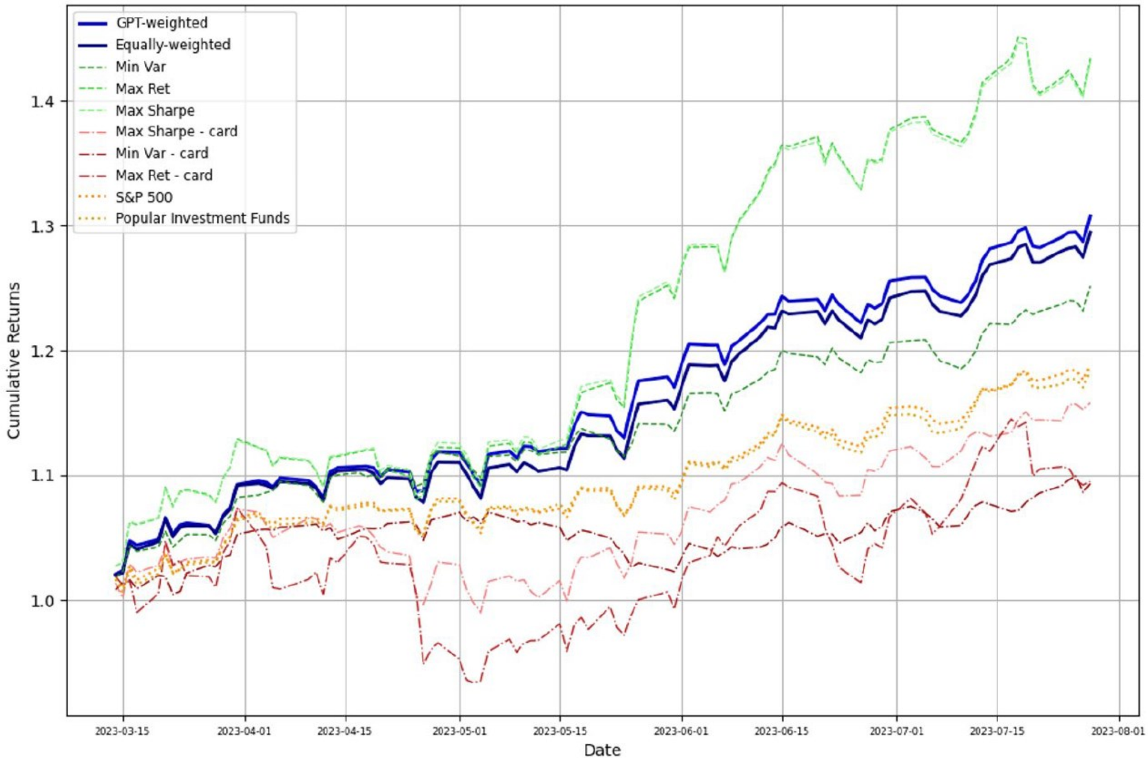



*Figure 28. Cumulative returns of portfolios with 15 assets in the period 14 March 2023- 31 July 2023. Source: Romanko, O., Narayan, A., Kwon, R.H. (2023) 'ChatGPT-Based Investment Portfolio Selection'.*

| | Cumulative Returns | Expected Return | Volatility of Return | Max Drawdown | Sharpe Ratio | VaR 99% of Return |
|---|---|---|---|---|---|---|
| GPT-weighted | 130.77% | 1.26% | 1.26% | -0.96% | 4.82 | -1.32% |
| Equally-weighted | 129.48% | 1.21% | 1.31% | -0.85% | 4.65 | -1.36% |
| Min Var | 125.18% | 1.05% | 1.07% | -1.18% | 4.74 | -1.19% |
| Max Ret | 143.39% | 1.71% | 2.21% | -2.24% | 4.58 | -1.98% |
| Max Sharpe | 143.17% | 1.70% | 2.15% | -1.77% | 4.64 | -1.90% |
| Max Sharpe - card | 115.83% | 0.69% | 1.94% | -6.61% | 2.24 | -2.05% |
| Min Var - card | 109.30% | 0.42% | 1.40% | -3.81% | 2.28 | -1.23% |
| Max Ret - card | 109.46% | 0.42% | 3.56% | -10.73% | 0.89 | -3.89% |
| S&P 500 | 118.84% | 0.79% | 1.31% | -1.39% | 3.57 | -1.59% |
| Dow Jones | 111.44% | 0.50% | 1.46% | -2.95% | 2.40 | -1.22% |
| NASDAQ | 127.95% | 1.14% | 1.58% | -1.44% | 3.92 | -1.98% |
| Popular Investment Funds | 118.38% | 0.77% | 1.40% | -1.65% | 3.44 | -1.43% |

*Table 10. Evaluation metrics for portfolios with 15 assets in the period 14 March 2023- 31 July 2023. Source: Romanko, O., Narayan, A., Kwon, R.H. (2023) 'ChatGPT-Based Investment Portfolio Selection'.*

Consistently, maximum return and maximum Sharpe portfolios have the highest cumulative returns, respectively 143,39% and 143,17%, even if paired with increased risk. The 'GPT-weighted' and equally-weighted portfolios provided solid performance, with cumulative returns equal to 130,77% and 129,48% respectively, placing between maximum returns/maximum Sharpe portfolios and the minimum variance portfolio (with cumulative

returns equal to 125,18%). Therefore, investors seeking for a balance between risk and returns should consider the ‘GPT-weighted’ or the equally-weighted portfolios.
In addition, the GPT-weighted portfolio performs better in terms of returns when compared to S&P500, Dow Jones or Nasdaq, with less volatility and drawdown. The portfolio based on GPT shows also the highest Sharpe Ratio among all the portfolios considered, indicating the best performance per unit of risk (volatility).
The fact that the GPT-weighted portfolio produces less optimal outcomes with respect to the Max Sharpe and Max Return portfolios, may suggest that while GenAI has capabilities to select profitable stocks, its capability to assign specific weights to assets can be improved.
Finally, cardinality constrained portfolios fall behind expectations, with the worst performance among the portfolios (cumulative returns from 109 to 115%), probably due to overfitting to the data: a more robust parameter estimation may therefore improve the performance of these portfolios. Despite this, it is important to notice that the majority of these strategies outperformed the benchmark (S&P 500), underlying the potential of GenAI capabilities into investment strategies. Combining ChatGPT and modern portfolio optimization generate higher performances than the isolated use of them, offering more efficient approach to practical investing and superior alternatives to current robo-advisor strategies. Nowadays, generative AI implementation is not widely adopted in finance because, in my opinion, since the technology debuted at the end of 2022, it is quite new and there is the need to test it very well on financial markets before adoption and additional checks and validation of GPT output need to be assessed. Another issue could be the lack of complete trust towards the service: the risks of hallucinations and data leakage may prevent, for the moment, many firms to implement this service.
Another evidence of ChatGPT potential is provided by Lopez-Lira’s paper (2023). In this final experiment the authors use Center for Research in Security Prices (CRSP) daily returns dataset for all stocks listed on NYSE, NASDAQ and AMEX. Next, a comprehensive news dataset is considered for all CRSP companies, considering news agencies, financial news websites and social media platforms. A ‘relevance score’, ranging from 0 to 100, is used to verify the relevance of certain news for a given company to indicate how closely the news pertains to a specific company. Prompts are then computed to ChatGPT, asking to identify if a specific headline is good or bad for a specific company stock price. Then a numerical ‘ChatGPT score’ is generated where good news for the company stock price is indicated with 1, bad news with -1 and unknown with 0.

Subsequently, portfolios that buy stocks with positive ChatGPT score and sell stocks with negative ChatGPT score, are composed. Figure 29 plots cumulative returns of 7 different trading strategies:

- Equally-weighted portfolio that buys companies with positive score based on GPT 3.5;
- Equally-weighted portfolio that sells companies with negative score based on GPT 3.5;
- A self-financing long-short strategy based on GPT 3.5;
- A self-financing long-short strategy based on GPT 4;
- Equal-weight market portfolio;
- Value-weight market portfolio;
- Equal-weight portfolio in all stocks with news the day before, regardless of the news direction.

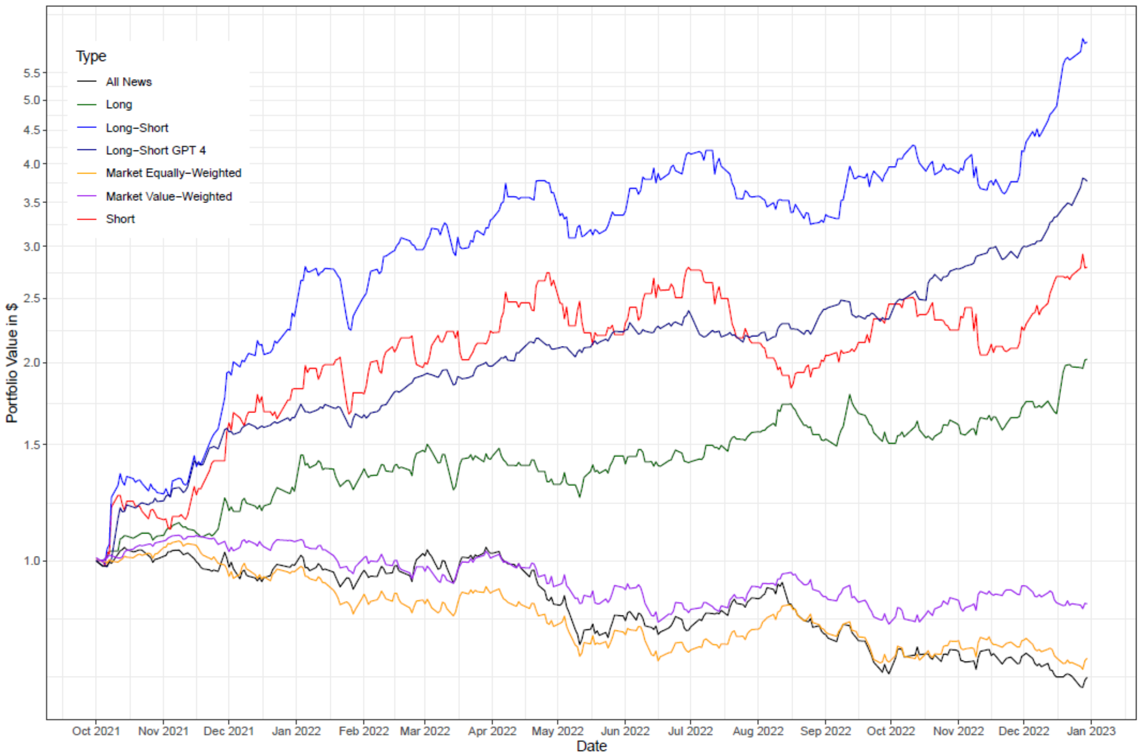



*Figure 29. Cumulative returns of investing 1$ (without transaction costs). Source: Lopez-Lira, A., Tang, Y. (2023) 'Can ChatGPT forecast stock price movements? Return predictability and Large Language Models'.*

From Figure 29, it is possible to see the potential of ChatGPT score in predicting sock returns the next day. For instance, without considering transaction costs, a self-financing strategy that buys stocks with positive score and sell the ones with a negative score, generate a cumulative return of 550% over the considered period, while equal-weight and value-weight portfolios present negative cumulative returns. Assuming instead transaction costs, the strategy still earns positive cumulative returns in the long-short strategy of GPT 3.5, but in the short leg is

more sensitive to transaction costs. GPT 4 cumulative returns are lower than the same strategy based on GPT 3.5 due to lower variations over time. The strategy based on the latest model however presents an higher Sharpe ratio and a lower maximum drawdown, showing therefore better reliability than its previous GPT 3.5 model.
Referring to the previous paragraph, is important to stress that advanced models like GPT 3.5 and GPT 4, demonstrate higher accuracy in interpreting complex headlines and generating forecasts, outperforming traditional sentiment analysis method like BERT and FinBERT models. Nevertheless, strategies based on ChatGPT 4 create the highest Sharpe ratio, indicating return predictability is an emerging capacity of complex generative AI models.
As the implementation of generative artificial intelligence evolves, more complex and accurate models can be designed to improve the efficiency of financial investment process.

# Chapter 4 – Case Studies

## 4.1 Case Study: Robo-Advisors in Italy

In Italy robo-advisors have not had the same success as in other countries. The main obstacle to their diffusion is financial education, the knowledge of financial and investment subjects, that leads an investor to be more or less willing to adopt a new technology. The slow adoption of this new service in fact could be explained by the low level of financial education in the Italian population. Italian financial literacy rate, defined as percentage of adults who are financially literate, is just 37%, higher only than Portugal among the Eurozone countries and even more critical if compared to G20 countries like Canada (68%) and Germany (66%). Financial literacy is usually measured using questions to assess basic knowledge of fundamental financial concepts like basic numeracy, diversification and inflation. Moreover, Italian population is rather averse to change, showed by the fact that in 2017 only 30% of Italians used internet banking, against 51% of European average (Milani, 2019). This situation could however change with younger generations, how pointed out by Isaia (2022).

The industry in fact is looking at customers, like young investors, who are confident to operate digitally and feel comfortable with automated investing services. According to a Vanguard survey (2020), Millennials are twice as likely as older American investors to consider the adoption of a robo-advisor and like Generation Z they are more likely to seek financial advice digitally after Covid-19 period. The Italian case, however, is particularly interesting, due to considerable amount of savings and to the fact that actual users of robo-advisor services are much less compared to other countries.

In the study presented by Isaia (2022), a sample of 1263 Italian young adults in 2020 is considered in order to investigate the potential demand of robo-advising services among Millennials and Generation Z. The estimation strategy aims at defining the characteristics of potential RA users, assuming a linear function of socio-demographic characteristics:

$$Robo-advisor_i = \beta_0 + \beta_1 Female_i + \beta_2 Age_i + \beta_3 Only\ bachelor\ degree_i + \beta_4 Master\ degree\ or\ more_i + \beta_5 Risk\ tolerant_i + \beta_6 Savings_i + \beta_7 Financial\ literacy\ index_i + \beta_8 Z_{it} + \varepsilon_i \quad (6)$$

Where $i$ is the individual identifier and $\varepsilon_i$ is an idiosyncratic error term. *Robo-advisor* is the measure of interest in robo-advising, *Only bachelor degree* and *Master degree or more* are variables that take value 1 if respondents have that education and 0 otherwise, *Risk tolerant* individuals are classified if they can bear moderate or high level of risk, *Savings* indicates the average monthly level of savings and F*inancial literacy index* is computed as the number of correct answers given to three financial questions. Finally, age and gender variable are included as well as $Z_{it}$, that corresponds to financial online behaviors like buying online, watching movies, looking for information online, working online and using digital payments. The results show that if having a bachelor degree is not associated with the probability of being a potential user of robo-advisors, having a master degree is positively correlated to the interest in RA. Women are less likely to be interested in these kinds of services and obviously more risk tolerant people are more willing to use them. These estimates suggest that a key factor to RA adoption is financial literacy. The correct answer of an additional financial literacy question is associated with higher probability of being interested in RA services, displaying thus a stronger role than education. Instead, individuals with basic financial literacy have the same interest as those without this basic knowledge. Finally, among online behaviors listed before, only buying online and using digital payments are predictive of being interested in digital platforms for financial advice like RA. Therefore, being accustomed to online financial activities can individuate potential users of RA services, as well as having an advanced level of financial literacy. The potential user of robo-advisor services is thus male, risk tolerant, highly financially literate investors and accustomed to use online financial services. Even if, up to now, Italy has not a great number of RA services users, Millennials and Generation Z are more likely to adopt this new technology, boosting the development of these functionalities in the country. According to Statista in fact, the robo-advisors market in Italy is projected to grow by 7.27% annually in the period from 2024 to 2027, resulting in a market volume of $13.90 billion in 2027.

### 4.2 Case study: Generative AI in financial companies

In this final section are briefly analyzed different financial companies that are trying to implement generative AI in their business to improve performances, in particular financial advisor performance and stock picking process. The case studies here discussed are the ones provided by the biggest investment banks and funds showing that, even if in its early stages, this technology is drawing the attention of the major players of the sector.

These implementations could represent a sort of evolution of standard robo-advisors service: in fact, rather than simply rebalancing portfolios with pre-defined algorithms, with the integration of GenAI, several stocks could be suggested in relation to current news, reports and events, creating a powerful tool for financial advisors and investors. So basically, if it seems the robo-advisor could be over, the introduction of generative AI tools may bring new life to RA or generating a brand-new product that can assist and advise customers in a similar way previous RA did.

### 4.2.1 BlackRock

Already in 2017 BlackRock Inc., one of the biggest investment management companies in the US, announced the plan to rely more on algorithms and models to pick stocks and to gradually replace human direction with AI, in order to improve performances of actively traded equity funds. BlackRock has then begun to replace human stock-pickers with Robo-Advisors 4.0, a full automated program, based on AI algorithms. BlackRock Robo-Advisor 4.0 is able to provide the SWAT framework (the basic framework of strategic management planning), evaluating specific strengths, weaknesses, opportunities and threats of a company. Moreover, they can process all available information stored in large databases to estimate value of stocks and predict business cycle to create a fully automated investment tool. BlackRock robo-advisors 4.0 was the first high-profile case of a large company trying to replace human in active investments (Tokic, 2018).

In 2015 BlackRock acquired San Francisco robo-advisor FutureAdvisor, not planning however to provide RA services but with the objective to attract mass affluent and younger investors (Toonkel, 2015). In 2023 however, BlackRock shut down this business selling it to Ritholtz Wealth Management, a registered investment advisor, and transitioning their clients to Ritholtz. This showed to many that stand-alone robo-advice are a difficult business model, according to Britton (2023). According to Welsh, Nexus Strategy founder, if a firm like BlackRock can't make a robo-advisor work, then no one can. Bruckenstein, president of consulting firm Technology Tools for Today, on the other hand claims that RA have a place in the market, but BlackRock simply could not provide successful offerings like Schwab and Vanguard (Brin, 2023). According to the Financial Times (2023), BlackRock has decided to roll out generative AI tools to clients in January 2024 to boost productivity. Clients will be able to use this technology to help them gathering information during the investment process. BlackRock aims at helping investment professionals gathering data for reports and investment proposals, reducing fixed costs and boosting margins.

### 4.2.2 Morgan Stanley

Another implementation of GenAI in financial services is provided by Morgan Stanley. On March 14, 2023 in fact, Morgan Stanley Wealth Management (MSWM) announced the launch of a strategic initiative with company OpenAI, the artificial intelligence firm owner of ChatGPT, making Morgan Stanley the only strategic client in wealth management receiving early access to OpenAI's new products. MS will use OpenAI to process and synthesize content to gain insights of companies, sectors and asset classes. Financial advisors will be able to ask questions concerning large amount of data and obtain easy and understandable answers generated from MSWM content. Moreover, MSWM plans to enhance financial advisors' offerings and their communication with clients through this tool (Morgan Stanley, 2023). Morgan Stanley has a content library with thousands of documents spanning from investment strategies to market research and analyst insights. This amount of information is stored in different internal sites, requiring advisors a great effort to find appropriate answers: with GPT-4 finding relevant information, the process may become simpler and less time consuming. To this purpose, an internal chatbot that performs comprehensive research of every content requested is adopted. Jeff McMillan, Head of Analytics, Data & Innovation at Morgan Stanley, said that in this way it is possible to have the knowledge of the most knowledgeable person in Wealth Management, in every moment, instantaneously. McMillan individuates three key points of this collaboration:

- The first is the ability to process vast amount of data almost immediately;
- The second is Morgan Stanley's intellectual capital on which the model is trained on, covering almost 100 years of published papers with insights on different topics;
- The third are Morgan Stanley's team of financial advisors that refined GPT-4 to their possible needs, enabling them to assist more people.

According to OpenAI, more than 200 Morgan Stanley employees are already using GPT-4 system on a daily basis. This case study may be an example of how humans are not replaced but rather helped by AI: in the first step, advisors talk to investors to understand their financial situation and goals, then financial advisors use GPT-4 as an instrument to gather information and make connections in order to present the client only the most valuable advices. Finally, in the process humans are also fundamental in supervising and minimizing possible error, biases and hallucinations (Witt, 2023). Testing is a very important step in order to prevent incidence of hallucinations nevertheless even if the model is trained on a wide database, they are reduced but not completely eliminated; by the way, as pointed by McMillan, this is just the first of a series of plausible solutions that could be adopted (Kinsella, 2023).

### 4.2.3 BloombergGPT

Even if general purpose language models like GPT-3 and GPT-4 have impressive capabilities, they are trained on a great variety of data sources and this generalist approach may be a limit when dealing with tasks requiring a specific knowledge in a certain domain. Therefore, in many cases a domain-specific pre-training should be adopted: this allows the model to learn terminologies and context-specific information, generating a specialized knowledge of the domain it was trained on. Domain-specific model can significantly improve performance and accuracy in a greater range of tasks (Poli, 2023).

On March 30, 2023, Bloomberg released a research paper concerning the development of BloombergGPT, a new generative AI model, trained on a vast amount of financial data to support different tasks in the financial sector, marking the first large language model, specialized in the financial domain, reported in literature (Bloomberg, 2023).

The few existing domain-specific model are trained only on specific data sources, while BloombergGPT approach is trained on both specific and general data sources, performing very well on domain-specific tasks and keeping strong performance on general-purpose benchmarks. The dataset on which the model is trained is made up of two main components:

- a financial dataset that composes 51.27% of the training and contains web content (by identifying sites with relevant financial information), news sources, company filings, press releases and Bloomberg authored news;
- a public dataset, accounting for the 48.73% of training, and containing a general dataset used in GPT engine called 'The Pile', another common dataset used to train LLMs, called 'C4' and finally Wikipedia pages to have up-to-date information (Wu, 2023).

The model aims at assisting Bloomberg in different tasks like sentiment analysis, news classification and question answering, among others. Cooperating with OpenAI has create one of the largest domain-specific datasets trained on the company's own data. Using documents over the span of forty years, the team trained a 50-billion parameter language model, validated on existing finance benchmarks (Bloomberg, 2023).

The first step to create this model is gathering a large amount of financial data from different sources, such as news articles, financial report and market data. Raw data may contain noise and inconsistencies that can lead to less accurate performance, therefore data cleaning and handling missing values are fundamental steps to improve data quality. Moreover, for models like GPT-4, data should be transformed to make it suitable for the model: this process is called tokenization, in which text data is divided into smaller pieces or tokens and these serve as the input for the model. Next, it is important to define the model architecture and, in this case, GPT-3 is chosen in a first moment: this makes it ideal for tasks involving large amount of text data, however with the introduction of the new GPT-4, the architecture has been further refined. The training process consists in giving the prepared data to the model and adjust the parameter to minimize the difference between the model's predictions and the actual data. For BloombergGPT, a particular functional relationship called 'Chinchilla Scaling Law' is applied in order to determine the optimal size of the model and the amount of data needed to train it (Wu, 2023). This relationship suggests that given a certain budget (expressed in FLOPs, a measure of computer power and performance), to achieve compute-optimal, the number of model parameters and tokens (small fraction of data used as input for the model) in the training model should scale in approximately equal proportion (Wikipedia, 2023d). Training large models may require substantial hardware resources and time and even Bloomberg training process was terminated after processing a lower number of tokens than what they supposed to use, in particular using 569 billion tokens instead of the original corpus of over 700 billion tokens (Wu, 2023).

Even though a larger model with more data can potentially lead to better performance, more resources are needed and therefore it is important to balance the size of the model, the volume of the data and the available resources to achieve the best possible performance.

After initial pre-training, the model needs to be fine-tuned and evaluated in specific tasks. This process allows the model to use its knowledge to perform specific tasks requirements and if the results are not satisfactory the model must be further fine-tuned on that task, adjusting the model's architecture or using different optimization algorithms (Poli, 2023).

In the research paper published by Bloomberg, their new model is compared to other existing models with similar size, type of training data and performance, in particular:

- GPT-NeoX: best performing model under 50B parameters;
- OPT66B: a model with similar model size and structure;
- BLOOM176B: a model with the same architecture but substantially larger.

The model is evaluated on a series of different aspects like financial tasks, sentiment analysis, knowledge assessment, reading comprehension, and linguistic tasks. Regarding financial and sentiment analysis, multiple tasks are evaluated such as, how news affect investors (positive/negative/neutral sentiment), predict sentiment in financial news and headlines, binary classification tasks to understand if news include certain information and answer conversational questions that require numerical reasoning given earnings reports. BloombergGPT performs better than all models for almost every task, showing the highest win rate among the models tested. Beyond financial tasks, BloombergGPT performs better than other models also in many general tasks: the model has thus showed abilities on general purpose data that exceed similarly sized models, outperforming them in many cases. Qualitative examples that highlight the benefits of this domain specialization can be summed up as: making interactions with financial data more natural (implementing the already existing Bloomberg Query Language that is a very powerful but complex tool used to retrieve financial data), assisting Bloomberg journalists in their daily work, and answering financial questions instantaneously, among others.

While models that are not publicly available cannot be reviewed by the community, distributing models may lead to incorrect usage, like high risk for abuse through imitation. Moreover, LLMs are subject to data leakage and sensitive information may be extracted. For this reasons Bloomberg decided not to release the model to the general public.

With this model, strong results on general LLMs benchmark are achieved and the main reasons to this success are mainly an up-to-date architecture, an accurate dataset and the choice in the number of tokens used (Wu, 2023).

Even if in the paper there are no references to stock price prediction, discussed in chapter 3, I suppose that, with further development of the model, new functions and abilities will be implemented in the model, and if it will be adopted by financial advisors it will allow investors to undertake more efficient decisions.

### 4.2.4 JP Morgan and IndexGPT

JP Morgan Chase, a multinational investment bank and financial company, has been positively embracing AI technologies too, recognizing the immense potential AI holds for the financial sector and applying it to improve efficiency and customer experience (AI Expert Network, 2023). Jamie Dimon, CEO of JPMorgan, acknowledged the advantages AI could produce, saying that this technology is groundbreaking and that is providing significantly decreased risk in their retail business, it is improving trading optimization and portfolio construction. On May 11, 2023, US financial services company JPMorgan Chase applied for a trademark for an AI-based product: IndexGPT, aiming at helping customers in financial services. This new tool combines AI and cloud computing software to analyze and select securities with respect to clients' specific needs, improving the selection process. According to the firm, IndexGPT will simplify financial investment, providing financial information and analysis services (The Economic Times, 2023). JPMorgan could be the first banking group to launch a proprietary product with financial goals: the product however has to be released within 3 years in order to guarantee the propriety of the trademark, with 1,500 engineers already working on it (Simonetta, 2023). Though the launch date remains therefore undisclosed, it should happen before 2027. While limited information is available about the tool, the program is reputed to be highly accurate and efficient in its operations (Pal, 2023). According to the US Trademark and Patent Office, three separate international classes of trademark have been filed as IndexGPT application: advertising and business services, insurance and financial services and computer and scientific services (Daniel, 2023).

The firm is nevertheless testing a variety of more than 300 use cases for AI technology across different areas like risk management, marketing and fraud prevention among others. To support this huge AI deployment, JP Morgan has invested over $2 billion in building cloud-based data centers, aiming at modernize their applications. Moreover, it is also paying attention to ethical implications: it composed an interdisciplinary team of ethicists with the tasks of preventing unintended use, anticipating regulation and promoting trust among customers, indicating a commitment to responsible use of AI. JP Morgan Chase's strategic AI implementation demonstrates the technology's potential in the banking and finance industry. Enhancing customer service, as well as leveraging AI for internal operation may inspire other companies to adopt AI, in a responsible and ethical way (AI Expert Network, 2023).

Even if different news headlines, like Fortune (2023), claimed that this new tool may put financial advisors out of business, I think that while this technology could potentially attract more retail customers, high net worth individual will always prefer human financial advisors to handle their resources. Hence, human financial advisors won't be completely substituted, but the nature of their job and tasks should probably be refined. Nevertheless, it should not be excluded the fact that, with further refinement of this technology, more people will prefer and adopt AI generated advices. This chapter aims to demonstrate how generative AI affected major US financial companies. Despite this innovative technology is relatively new and may pose different risks, its potential and possible implementations have immediately persuaded these companies to adopt it as soon as possible, to ensure they won't fall behind their competitors and miss strategic advantages that could improve their businesses. From my perspective, if firms with this magnitude are embracing the developments of this new tool, it is likely that a significant revolution in the financial sector could happen.

# Conclusions

The thesis has provided a comprehensive analysis on the potential application of generative AI to robo-advisory systems. In particular, after some descriptive analysis, valuable insights on the performances are presented: in Chapter 2, GPT-4 portfolio allocations are illustrated and compared to the ones of a human financial advisor, showing very similar trend in the final composition of the portfolio, while in Chapter 3, after analyzing experimental combination of different model to achieve better performances, there is an evidence of how GenAI may be useful in suggesting stocks and allocating them in a well-diversified portfolio, even if the examined cumulative returns are below the traditional maximum Sharpe and maximum return portfolios. GPT-4 in particular shows interesting capabilities also in stock predictions, providing insights for future research in this field. In addition, the fact that major players in the financial industry are adopting generative AI models in their business, both for employees and customers, indicates how large the potential of this new innovation can be and the multiplicity of business cases it could be applied to. It cannot be stated today if generative AI will be implemented into existing RA platforms or it will become a completely new tool used by financial advisors, but in any case, the final objective is the same: giving valuable insights to investors and improve their portfolio management. In any case, it is important to notice that this technology does not come without any risks, as hallucinations, data privacy and discrimination issues, as well as explainability problems persist. As of today, we are just in the early days of this new technology, nonetheless it already displays very interesting features; however, in the future with advancements of this technology, even more applications and functions will be available. Future research could therefore concentrate on a wide range of different aspects like improving existing empirical analysis on stock returns, integrating generative AI to already existing algorithms, as well as exploring ethical implication of AI decision-making and developing mechanism to enhance transparency, among others.

One question that arise during this research is whether financial advisor will be completely replaced by generative AI. Today, financial advisors are still fundamental in the investment process and clients usually prefer an hybrid approach; I personally think that financial advisors would implement generative AI tools as instruments to help them in their analysis and boosting their productivity, redefining their jobs but not eliminating them. Nevertheless, if this technology keeps on evolving at the current pace, it would be also a plausible hypothesis that, attracting a huge number of customers, financial advisors will become less popular than today.

# Bibliography

ACCENTURE, 2022. *AI in Wealth Management. A Financial Advisor Study* [online]. Available at: https://www.accenture.com/content/dam/accenture/final/a-com-migration/r3-3/pdf/pdf-178/accenture-wealth-management-ai-all-systems-go.pdf.

AI EXPERT NETWORK, 2023. *Case Study: Implementing AI at JP Morgan* [online]. Available at: https://aiexpert.network/case-study-implementing-ai-at-jp-morgan/.

BATRA, I., FINE, A., 2017. *An idea in brief. Robo-Advisor: The next big thing in bundled checking?* [online]. Oliver Wyman. Available at: https://www.oliverwyman.com/content/dam/oliver-wyman/v2/publications/2017/apr/Oliver_Wyman_Robo_Advisor_Idea_in_Brief.pdf.

BEKETOV, M., LEHMANN, K. & WITTKE, M., 2018. Robo Advisors: quantitative methods inside the robots. *Journal of Asset Management,* Volume 19, pp. 363-370. https://doi.org/10.1057/s41260-018-0092-9.

BENARTZI S., 2023. Can AI Replace Your Financial Adviser? Not Yet.But Wait. *The Wall Street Journal*. Available at: https://www.wsj.com/tech/ai/ai-financial-advisers-chatgpt-7f0cac60.

BETTER FINANCE, 2020. *Robo-Advice 5.0: Can Consumers Trust Robots?* [online]. Research report available at: https://betterfinance.eu/wp-content/uploads/Robo-Advice-Report-2020-25012021.pdf.

BJERKNES, L., VUKOVIC, A., 2017. *Automated Advice: A Portfolio Management Perspective on Robo-Advisors* [online]. Final thesis, Norwegian University of Science and Technology, Department of Industrial Economics and Technology Management. Available at: https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/2473732/17822_FULLTEXT.pdf?sequence=1.

BLOOMBERG, 2023. *Introducing BloombergGPT, Bloomberg's 50-billion parameter large language model, purpose built from scratch for finance* [online]. Available at: https://www.bloomberg.com/company/press/bloomberggpt-50-billion-parameter-llm-tuned-finance/.

BOSTON CONSULTING GROUP. *Finance Function Excellence* [online]. Available at: https://www.bcg.com/capabilities/corporate-finance-strategy/finance-function-excellence.

BRENNER, L., MEYLL, T., 2020. Robo-advisors: A substitute for human financial advice? *Journal of Behavioral and Experimental Finance*, Volume 25. https://doi.org/10.1016/j.jbef.2020.100275

BRIN, D.W., 2023. *Does BlackRock's Sale of FutureAdvisor to Ritholtz Deal Death Blow to Robos?* [online]. Available at: https://www.thinkadvisor.com/2023/03/01/blackrock-to-sell-futureadvisor-to-ritholtz-putting-final-nail-in-robo-advisory-coffin/.

BRITTON, D., 2023. *BlackRock sells FutureAdvisor's direct-to-consumer business to Ritholtz* [online]. Available at: https://www.wealthmanagement.com/technology/blackrock-sells-futureadvisors-direct-consumer-business-ritholtz.

BUSINESS INSIDER, 2017. *Here are the top 15 Robo Advisors making stock portfolio management easy in 2017*. In: BEKETOV, M., LEHMANN, K. & WITTKE, M., 2018. Robo Advisors: quantitative methods inside the robots. *Journal of Asset Management* [online], Volume 19, p. 365. https://doi.org/10.1057/s41260-018-0092-9.

CASPI, I., FELBER, S.S., GILLIS, T.B., 2023. *Generative AI and the Future of Financial Advice Regulation* [online]. Available at: https://genlaw.org/CameraReady/19.pdf.

CAVANAUGH, D., 2023. Why AI won't replace human financial advisors. *Financial Advisor Magazine*. Available at: https://www.fa-mag.com/news/why-ai-won-t-replace-human-financial-advisors-75237.html.

CHARLES SCHWAB & CO. INC., 2019. *Schwab Intelligent Portfolios and Q2 2019 Market Performance* [online]. Available at: https://intelligent.schwab.com/public/intelligent/insights/blog/q22019-performance-review.html.

CHARLES SCHWAB & CO. INC., 2022. *Schwab Advisor Services' 2022 Independent Advisor Outlook Study* [online]. Available at: https://www.aboutschwab.com/schwab-independent-advisor-outlook-study-2022.

CHEN, Y., et al., 2023. *A Manager and an AI Walk into a Bar: Does ChatGPT Make Biased Decisions Like We Do?* [online]. Available at: http://dx.doi.org/10.2139/ssrn.4380365.

CREDO AI, 2023. *Glossary: Human-on-the-loop* [online]. Available at: https://www.credo.ai/glossary/human-on-the-loop.

DANIEL, W., 2023. *Meet 'IndexGPT,' the A.I. stock picker JPMorgan is developing that may put your 'financial advisor out of business* [online]. Available at: https://fortune.com/2023/05/26/jpmorgan-indexgpt-a-i-stock-picker/.

DAVE, 2022. *Best Robo Advisors: The $25,000 Comparison Case Study* [online]. Available at: https://daytradereview.com/best-robo-advisor-comparison/.

DAY, M.-Y., CHENG, T.-K., LI, J.-G., 2018. AI Robo-Advisor with Big Data Analytics for Financial Services. *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM),* Barcelona, Spain, 2018, pp. 1027-1031. DOI: 10.1109/ASONAM.2018.8508854.

DEMYTTENAERE, M., et al., 2023. *Generative Ai in the Finance Function of the Future* [online]. Available at: https://www.bcg.com/publications/2023/generative-ai-in-finance-and-accounting.

DEUTSCHE BANK, 2022. *Deutsche Bank Partners with NVIDIA to embed AI into Financial Services* [online]. Media Release, December 7, 2022. Available at: https://www.db.com/news/detail/20221207-deutsche-bank-partners-with-nvidia-to-embed-ai-into-financial-services.

DOHMS, R., 2017. *Gewinnmaschinen Capital 8*. In: BEKETOV, M., LEHMANN, K. & WITTKE, M., 2018. Robo Advisors: quantitative methods inside the robots [online]. *Journal of Asset Management* [online], Volume 19, p. 365. https://doi.org/10.1057/s41260-018-0092-9.

DWIVEDI, Y.K., et al., 2023. Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management* [online], Volume 71. Available at: https://doi.org/10.1016/j.ijinfomgt.2023.102642.

EUROPEAN BANKING AUTHORITY (EBA), 2017. *Discussion Paper on the EBA's approach to financial technology (FinTech)* [online]. Available at: https://www.eba.europa.eu/sites/default/files/documents/10180/1919160/7a1b9cda-10ad-4315-91ce-d798230ebd84/EBA%20Discussion%20Paper%20on%20Fintech%20%28EBA-DP-2017-02%29.pdf.

EUROPEAN SECURITIES AND MARKETS AUTHORITY (ESMA), 2018. *Final Report – Guidelines on Certain Aspects of the MiFiD II Suitability Requirements*. In: MAUME, P., 2021. *Robo-advisors. How do they fit in the existing EU regulatory framework, in particular with regard to investor protection?* [online]. Policy Department for Economic, Scientific and Quality of Life Policies. Available at: https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662928/IPOL_STU(2021)662928_EN.pdf.

EUROPEAN SECURITIES AND MARKETS AUTHORITY (ESMA), 2021. *Questions and Answers on MIFID II and MiFIR market structure topics* [online]. Available at: https://www.esma.europa.eu/sites/default/files/library/esma70-872942901-38_qas_markets_structures_issues.pdf.

EUROPEAN SECURITIES AND MARKETS AUTHORITY (ESMA), 2023. *ESMA TRV Risk Analysis. Artificial Intelligence in Securities Markets* [online]. Available at: https://www.esma.europa.eu/sites/default/files/library/ESMA50-164-6247-AI_in_securities_markets.pdf.

FIEBERG, C., HORNUF, L., STREICH D.J., 2023. Using GPT-4 for Financial Advice. *CESifo Working Paper No. 10529*. Available at: http://dx.doi.org/10.2139/ssrn.4499485.

FINANCIAL TIMES, 2023. *BlackRock to roll out first generative AI tools to clients next month* [online]. Available at: https://www.ft.com/content/3f3431f1-d6dc-4310-9edc-3bc8cdc46caa.

FISCH, J.E., LABOURÉ, M., TURNER, J.A., 2019. The Emergence of the Robo-advisor. In book: The Disruptive Impact of FinTech on Retirement Systems, pp.13-37. *Wharton Pension Research Council Working Papers*, Volume 10. DOI:10.1093/oso/9780198845553.003.0002

FRANKENFIELD, J., 2023. *What is a Robo-Advisor?* [online]. Available at: https://investopedia.com/terms/r/roboadvisor-roboadviser.asp.

FRIEDMAN, B., NISSENBAUM H., 1996. Bias in Computer Systems. *ACM Transactions on Information Systems*, Volume 14 Issue 3, pp 330–47. In: INTERNATIONAL MONETARY FUND (IMF), 2023. *Generative Artificial Intelligence in Finance: Risk Considerations* [online]. NOTE/2023/006. Available at: https://www.imf.org/en/Publications/fintech-notes/Issues/2023/08/18/Generative-Artificial-Intelligence-in-Finance-Risk-Considerations-537570.

GDPR, 2023. *General Data Protection Regulation, art. 4 GDPR Definitions*. Available at: https://gdprinfo.eu/art-4-gdpr/.

GRAND VIEW RESEARCH, 2021. *Robo Advisory Market Size, Share & Trends Analysis Report By Type (Pure Robo Advisors, Hybrid Robo Advisors), By Provider, By Service Type, By End-use, By Region, And Segment Forecasts, 2022 – 2030* [online]. Available at: https://www.grandviewresearch.com/industry-analysis/robo-advisory-market-report.

GROSSMAN, A., 2023. *How to use Artificial Intelligence in your portfolio* [online]. Available at: https://www.investopedia.com/how-to-use-artificial-intelligence-in-your-investing-7973810.

HAYES, A., 2023. *Do Robo-Advisors beat the S&P500?* [online]. Available at: https://www.investopedia.com/do-robo-advisors-beat-the-s-and-p-500-7569726.

IBM, 2021. *What are recurrent neural networks?* [online]. Available at: https://www.ibm.com/topics/recurrent-neural-networks.

INTERNATIONAL MONETARY FUND (IMF), 2023. *Generative Artificial Intelligence in Finance: Risk Considerations* [online]. NOTE/2023/006. Available at: https://www.imf.org/en/Publications/fintech-notes/Issues/2023/08/18/Generative-Artificial-Intelligence-in-Finance-Risk-Considerations-537570.

ISAIA, E., OGGERO, N., 2022. The Potential Use of Robo-advisors Among the Young Generation: Evidence from Italy. *Finance Research Letters,* Volume 48. https://doi.org/10.1016/j.frl.2022.103046.

JANSSEN, M., KUK, G., 2016. The challenges and limits of big data algorithms in technocratic governance. *Government Information Quarterly,* Volume 33(3), pp. 371-377. https://doi.org/10.1016/j.giq.2016.08.011.

JARRELL, N. S., et al., 2023. *AI discrimination and bias in financial services* [online]. EY website. Available at: https://www.ey.com/en_us/forensic-integrity-services/ai-discrimination-and-bias-in-financial-services.

JI, Z., et al., 2022. Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*. Available at: https://doi.org/10.1145/3571730.

JOINT COMMITTEE OF THE EUROPEAN SUPERVISORY AUTHORITIES (JCESA), 2015. *Joint Committee Discussion Paper on automation in financial advice*. (No. JC 2015-080, 4 December2015), pp. 16-29. In: MAUME, P., 2021. *Robo-advisors. How do they fit in the existing EU regulatory framework, in particular with regard to investor protection?* [online]. Policy Department for Economic, Scientific and Quality of Life Policies. Available at:https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662928/IPOL_STU(2021)662928_EN.pdf.

KINSELLA, B., 2023. *Morgan Stanley rolls out OpenAI GPT-Powered assistant for financial advisors* [online]. Available at: https://synthedia.substack.com/p/morgan-stanley-rolls-out-openai-gpt.

KUBERA, 2023. *Are AI financial advisors the future?* [online]. Available at: https://www.kubera.com/blog/ai-financial-advisors.

LAKHANI K., IGNATIUS A., 2023. *AI Won't Replace Humans — But Humans With AI Will Replace Humans Without AI* [online]. Harvard Business Review. Available at: https://hbr.org/2023/08/ai-wont-replace-humans-but-humans-with-ai-will-replace-humans-without-ai.

LAM, J.L., 2016. *Robo-Advisors: A Portfolio Management Perspective.* Yale College. Available at: https://economics.yale.edu/sites/default/files/2023-01/Jonathan_Lam_Senior%20Essay%20Revised.pdf.

LEOW, E.K.W, NGUYEN, B.P., CHUA, M.C.H., 2021. Robo-advisor using genetic algorithm and BERT sentiments from tweets for hybrid portfolio optimization. *Experts Systems With Applications*, Volume 179. Available at: https://doi.org/10.1016/j.eswa.2021.115060.

LIU X., et al., 2023. *Data-centric FinGPT: Democratizing internet-scale data for financial large language models* [online]. Available at: https://doi.org/10.48550/arXiv.2307.10485.

LOPEZ-LIRA, A., TANG, Y., 2023. *Can ChatGPT forecast stock price movements? Return predictability and Large Language Models* [online]. Available at: http://dx.doi.org/10.2139/ssrn.4412788

MACHET, J., 2021. *The history of robo-advisors* [online]. Available at: https://unhedged.com.au/the-history-of-robo-advisors/.

MALAMED A., 2023. *Artificial Intelligence will drive evolution, not extinction, for financial advisors* [online]. Available at: https://www.wealthmanagement.com/technology/artificial-intelligence-will-drive-evolution-not-extinction-financial-advisors.

MARQUIT, M., 2023. *What is a Robo-Advisor? How do they Work?* [online]. Forbes Advisor. Available at: https://www.forbes.com/advisor/investing/what-is-robo-advisor/.

MAUME, P., 2021. *Robo-advisors. How do they fit in the existing EU regulatory framework, in particular with regard to investor protection?* [online]. Policy Department for Economic, Scientific and Quality of Life Policies. Available at: https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662928/IPOL_STU(2021)662928_EN.pdf.

MAURAN, C. 2023. *Whoops, Samsung workers accidentally leaked trade secrets via ChatGPT* [online]. Available at: https://mashable.com/article/samsung-chatgpt-leak-details.

MCCLELLAN, J.N., 2016. What the Evolving Robo Advisory Industry Offers. *American Association of Individual Investors Journal, October 2016.* Available at: https://www.aaii.com/journal/article/what-the-evolving-robo-advisory-industry-offers?printerfriendly=true.

MCKINSEY & COMPANY, 2023. *The economic potential of generative AI: The next productivity frontier*, June 14, 2023 Report [online]. Available at: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#introduction.

MILANI, A., 2019. *The role of risk and trust in the adoption of robo-advisory in Italy* [online]. Available at: https://www.pwc.com/it/it/publications/assets/docs/Report-robo-advisors.pdf.

MIN, L., et al., 2021. A Black-Litterman Portfolio Selection model with Investor Opinions Generating from Machine Learning Algorithms. *Engineering Letters*, Volume 29, Issue 2. Available at: https://www.engineeringletters.com/issues_v29/issue_2/EL_29_2_40.pdf.

MORGAN STANLEY, 2023. *Morgan Stanley Wealth Management announces key milestone in innovation journey with OpenAI* [online]. Available at: https://www.morganstanley.com/press-releases/key-milestone-in-innovation-journey-with-openai.

MUGANDA, B.W., KASAMANI, B. S., 2023. Parallel Programming for Portfolio Optimization: A Robo-Advisor Prototype using Genetic Algorithms with Recurrent Neural Networks. *2023 International Conference on Intelligent Computing, Communication, Networking and Services (ICCNS),* Valencia, Spain, 2023, pp. 167-176. DOI: 10.1109/ICCNS58795.2023.10193396

NASR, T., 2023. *The largest Robo-Advisors by AUM (2023)* [online]. Available at: https://investingintheweb.com/brokers/the-largest-robo-advisors-by-aum/.

OLIVER WYMAN, MORGAN STANLEY, 2023. *The AI Tipping Point* [online], Chapter 4, pp 37-52. Available at: https://www.oliverwyman.com/content/dam/oliver-wyman/v2/publications/2023/october/Oliver_Wyman_Morgan_Stanley_Global_Wealth_and_Asset_Management_report_2023_The_Generative_AI_Tipping%20Point1.pdf.

OPENAI, 2023. *Customer stories: Morgan Stanley* [online]. Available at: https://openai.com/customer-stories/morgan-stanley.

OZAIR, M., 2023. *FinanceGPT: The next generation of AI-powered robo advisors and chatbots* [online]. Available at: https://www.nasdaq.com/articles/financegpt-the-next-generation-of-ai-powered-robo-advisors-and-chatbots.

PAL, K., 2023. *AI disruption in Fintech: Exploring the impact of AI-powered tools like IndexGPT* [online]. Available at: https://www.techopedia.com/ai-in-fintech-exploring-the-impact-of-tools-like-indexgpt.

PANTOJA, J., ALEMAN, J., TELLEZ, D., 2023. *Iterative Deep Learning Approach to Active Portfolio Management with Sentiment Factors* [online]. Available at: http://dx.doi.org/10.2139/ssrn.4579068.

PELSTER, M., VAL, J., 2023. Can ChatGPT assist in picking stocks? *Finance Research Letters, 104786*. Available at: http://dx.doi.org/10.2139/ssrn.4602452.

PIRNER, D., FISCHER, M., 2018. *Asset Allocation with Black-Litterman in a case study of Robo Advisor Betterment,* Financial Research Note 04 [online]. Available at: https://www.researchgate.net/publication/326878665_Asset_Allocation_with_Black-Litterman_in_a_case_study_of_Robo_Advisor_Betterment.

POLI, C., 2023. *A case study: The development and challenges of BloombergGPT, a domain-specific Language Model* [online]. Available at: https://topframework.org/2023/07/14/a-case-study-the-development-and-challenges-of-bloomberggpt-a-domain-specific-language-model/.

RICHTER, C., 2018. MiFID2 - *Geeignetheitsprüfung bei Robo-Advisory: Rechtliche und praktische Umsetzung*. M.SC. thesis at TUM School of Management. In: MAUME, P., 2021. *Robo-advisors. How do they fit in the existing EU regulatory framework, in particular with regard to investor protection?* [online]. Policy Department for Economic, Scientific and Quality of Life Policies. Available at: https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662928/IPOL_STU(2021)662928_EN.pdf.

RINALDO, C., 2023. Automation in Investment Advice. A European Perspective. *Revista de derecho privado*, Num. 45, pp 309-340. Available at: https://doi.org/10.18601/01234366.45.11.

RINGE, W.G., RUOF, C., 2018. A Regulatory Sandbox for Robo Advice. *European Banking Institute Working Paper Series 2018 – No. 26*. Available at: http://dx.doi.org/10.2139/ssrn.3188828

ROMANKO, O., NARAYAN, A., KWON, R.H., 2023. ChatGPT-Based Investment Portfolio Selection. *Operations Research Forum*, Volume 4, Article 91. Available at: https://doi.org/10.1007/s43069-023-00277-6

SACCHITELLO, M., 2023. *How we review robo-advisors* [online]. Available at: https://www.investopedia.com/robo-advisor-ranking-methodology-4693455.

SAXENA, S., 2021. *What is LSTM? Introduction to Long Short-Term Memory* [online]. Available at: https://www.analyticsvidhya.com/blog/2021/03/introduction-to-long-short-term-memory-lstm/.

SCHOLZ, P., 2021. *Robo-Advisory - Investing in the Digital Age*. Palgrave Studies in financial services technology. Available at: https://doi.org/10.1007/978-3-030-40818-3.

SEI INVESTMENTS COMPANY, 2023. *Generative robo advisors* [online]. Available at: https://www.seic.com/pulse-future/generative-robo-advisors.

SIMONETTA, B., 2023. JP Morgan sviluppa la sua ChatGPT: fornirà consigli sugli investimenti. *Il Sole 24 Ore*. Available at: https://www.ilsole24ore.com/art/jp-morgan-sviluppa-sua-chatgpt-fornira-consigli-investimenti-AEriOYYD.

SMITH, K., 2022. *How AI could elevate financial advisor performance* [online]. Available at: https://capitalmarketsblog.accenture.com/how-ai-could-elevate-financial-advisor-performance.

SOLDATOS, J., KYRIAZIS, D., 2022. *Big Data and Artificial Intelligence in Digital Finance. Increasing Personalization and Trust in Digital Finance using Big Data and AI.* Springer, Chapter 11, pp. 199-214. Available at: https://doi.org/10.1007/978-3-030-94590-9.

SPIEGEL, J., 2023. *How advisors are increasing efficiency and impact with AI* [online]. Available at: https://www.blackrock.com/us/financial-professionals/insights/how-advisors-use-ai.

STATISTA, 2023. *Robo-Advisors - Worldwide* [online]. Available at: https://www.statista.com/outlook/fmo/wealth-management/digital-investment/robo-advisors/worldwide.

STATISTA, 2024. *Robo-Advisors Italy – Statista Market Forecast* [online]. Available at: https://www.statista.com/outlook/fmo/wealth-management/digital-investment/robo-advisors/italy.

TA, V.D., LIU, C.M., TADESSE, A., 2020. Portfolio Optimization-Based Stock Prediction Using Long-Short Term Memory Network in Quantitative Trading. *Applied Sciences*, 10 (2), 437. Available at: https://doi.org/10.3390/app10020437.

TAN, Z. Y., et al., 2023. 'The Adoption of Robo-Advisory among Millennials in the 21st Century: Trust, Usability and Knowledge Perception. *Sustainability*, 15 (7), 6016. Available at: https://doi.org/10.3390/su15076016.

THE ECONOMIC TIMES, 2023. What is IndexGPT? Know all about JP Morgan Chase's AI financial service. *The Economic Times*. Available at: https://economictimes.indiatimes.com/news/international/us/what-is-indexgpt-know-all-about-jp-morgan-chases-ai-financial-service/articleshow/101395926.cms.

TOKIC, D., 2018. BlackRock Robo-Advisor 4.0: When artificial intelligence replaces human discretion. *Strategic Change*, Volume 27, Issue 4, pp. 285-290. Available at: https://doi.org/10.1002/jsc.2201

TOONKEL, J., 2015. *BlackRock to acquire Robo Advisor* [online]. Available at: https://www.wealthmanagement.com/technology/blackrock-acquire-robo-advisor.

TORNO, A., SCHILDMANN, S., 2020. What Do Robo-Advisors Recommend? - An Analysis of Portfolio Structure, Performance and Risk. *Enterprise Applications, Markets and Services in the Finance Industry*, pp. 92-108. DOI: 10.1007/978-3-030-64466-6_6.

VANGUARD, 2020. *Generational views on financial advice, investing, and retirement. A Vanguard Digital Advisor survey.* In: ISAIA, E., OGGERO, N., 2022. The Potential Use of Robo-advisors Among the Young Generation: Evidence from Italy. *Finance Research Letters,* Volume 48. https://doi.org/10.1016/j.frl.2022.103046.

VANGUARD, 2023. *Vanguard Digital Advisor and Vanguard Personal Advisor Important Information* [online]. Available at: https://personal1.vanguard.com/pdf/vanguard-digital-advice-brochure.pdf.

VETTAFI, 2021. *Robo-Advisor Fees: What Is the True Cost* [online]. Available at: https://www.nasdaq.com/articles/robo-advisor-fees:-what-is-the-true-cost-2021-03-11.

WACH, K., et al., 2023. The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT. *Entrepreneurial Business and Economics Review*, Vol. 11 No.2. DOI:10.15678/EBER.2023.110201.

WEALTH HARNESS, 2023. *Average Fees for Financial Advisors in 2023* [online]. Available at: https://www.harnesswealth.com/articles/average-fees-for-financial-advisors/.

WEALTHFRONT, 2020. *Wealthfront's Average Annual Net-of-Fees, Pre-tax Returns. Our Historical Returns* [online]. Available at: https://www. wealthfront.com/historical-performance.

WIKIPEDIA, 2023a. *Human-in-the-loop* [online]. Available at: https://en.wikipedia.org/wiki/Human-in-the-loop.

WIKIPEDIA, 2023b. *Machine Learning* [online]. Available at: https://en.wikipedia.org/wiki/Machine_learning.

WIKIPEDIA, 2023c. *Artificial Neural Network* [online]. Available at: https://en.wikipedia.org/wiki/Recurrent_neural_network.

WIKIPEDIA, 2023d. *Neural scaling law* [online]. Available at: https://en.wikipedia.org/wiki/Neural_scaling_law.

WITT, T., 2023. *How Morgan Stanley, Stripe use GPT-4 to optimize financial customer experience (CX)* [online]. Available at: https://accelerationeconomy.com/ai/how-morgan-stanley-stripe-use-gpt-4-to-optimize-financial-customer-experience-cx/.

WU, S., et al., 2023. *BloombergGPT: A Large Language Model for Finance*. Bloomberg Research Paper [online]. Available at: https://doi.org/10.48550/arXiv.2303.17564.

YANG, H., LIU X.-Y., DAN WANG, C., 2023. *FinGPT: Open-source financial large language models* [online]. Available at: https://doi.org/10.48550/arXiv.2306.06031.