Asset Allocation and Machine Learning: a performance analysis within distressed market conditions

Auteur : Shtini, Sindi
Promoteur(s) : Hambuckers, Julien
Faculté : HEC-Ecole de gestion de l'Université de Liège
Diplôme : Master en sciences de gestion, à finalité spécialisée en Banking and Asset Management
Année académique : 2022-2023
URI/URL : http://hdl.handle.net/2268.2/16750

ASSET ALLOCATION AND MACHINE LEARNING: A PERFORMANCE ANALYSIS WITHIN DISTRESSED MARKET CONDITIONS

Jury:
Supervisor :

Julien HAMBUCKERS
Reader:
Julie JAMAR

Master thesis by
Sindi SHTINI

For a Master Degree in Economic Sciences with a specialization in Banking & Asset Management
Academic year 2022/2023

Acknowledgments

I want to start by expressing my gratitude to my promoter, Professor Julien HAMBUCKERS, for his constant support, assistance, understanding, and vital guidance about the subject matter that enabled me to successfully complete this Master's thesis.

I would like to express my thanks to Professor Julie JAMAR for her attention, invaluable advice, and the time spent reading my thesis.

Finally, my deepest thank you goes to all my family and friends that have allowed me to achieve this important milestone in my life. Words will never be sufficient to describe my gratitude.

Table of contents

List of abbreviations

AI – Artificial Intelligence

CNN – Convolutional neural networks

DL – Deep Learning

DM – Diebold-Mariano

FNN – Feedforward Neural Networks

KNN – K-Nearest Neighbors

ML – Machine Learning

MSE – Mean Squared Error

MAE – Mean-Absolute Return

MLP – Multilayer Perceptron Network

MOM – Momentum Factor

MVO – Mean-Variance Optimization

MZ – Markowitz

NYSE – New York Stock Exchange

OECD – Organization for Economic Co-operation and Development

RNN – Recurrent Neural Networks

RSI – Relative Strength Index

U.S.A. – United States of America

U.K. – United Kingdom

1. Introduction

In the first chapter, an overview of this Master's thesis is provided.

Firstly, motivations and purpose are briefly presented, by giving an overview of the main ideas and inspirations which are the fundamental basis for this research. In the following section, the research objectives, as well as the study framework, are addressed including the research questions. Finally, the chapter ends with a brief explanation of the methodology that is applied throughout this piece of work.

1.1 Motivation

Human beings hate making mistakes. This core attribute of humans leads to continuous research and development of new tools and techniques that can reduce, if not completely eliminate, perceived risk.

In recent years, the hasting development of Artificial Intelligence is shaping daily lives and is providing new means and solutions that find application in a multitude of industries and sectors likewise. One of the many fields in which Machine Learning, particularly, is gaining growing interest and attention is asset allocation. Wealth management and financial advisors are increasingly relying on artificially automated tools to carry out their activities, and robo-advisors have erupted in popularity among investors.

However, it is still an ongoing research discussion weather these new waves of innovation and techniques do provide a groundbreaking solution in optimizing asset allocation.

The motives behind the selected topic do not solely want to appraise the strength of Machine Learning within the context of asset allocation, but also aspire to reflect on the historical times the world population is currently living in. Indeed, in these times of great uncertainty and fear, individuals – thus, investors – are subject to continuous and unexpected challenges, and this includes their portfolio investment strategies.

To detail, this Master's thesis aims at capturing how the portfolio constructed by deploying Machine Learning algorithms reacts within distressed market conditions. These market conditions are impacted by extreme occurrences that fall outside the realm of what can be identified as expected or predicable scenarios of daily life. The following sections clearly captures and identifies the Research Objectives of this thesis.

1.2 Research Objective

Within the scope of this research, falls the improvement of the knowledge of Machine Learning techniques embedded in the portfolio allocation field, particularly considering the context of unforeseen catastrophic events.

The process that will allow to achieve such a goal starts with an initial stage dedicated at assessing and evaluating the Machine Learning constructed portfolio against a portfolio set-up via the universally adopted mean-variance optimization methodology, first introduced by Harry Markowitz in 1952.

This leads to the first research question which is going to be addressed throughout the analysis:

RQ 1. *Does the Machine Learning constructed portfolio deliver stronger performance for the investor in comparison to the Markowitz constructed portfolio?*

Further, the analysis will move to integrate extreme events. The events that are going to be analyzed defer in terms of origination as well as characteristics and consequences resulting in an impact on the financial market, that might variate by longevity and severity. The different sets of selected events include natural disasters, political changes, and terroristic attacks.

Within the different selected events' settings, the two constructed portfolios, which have been defined in the previous stage, will be examined and the results achieved will support in answering the second research question:

RQ 2. *Within distressed market conditions, where does the Machine Learning constructed portfolio position itself against the Markowitz constructed portfolio?*

1.3 Research Methodology

Regarding the methodology that will be implemented, a quantitative analysis will be performed.

For what concerns the classical portfolio strategy set-up, the mean-variance optimization introduced by Markowitz in 1952 will be used in order to set-up the first portfolio. Instead, the state-of-art Machine Learning algorithms – which are nowadays available and deployed in the industry – will be explored and implemented to achieve a suitable methodology used to build the second portfolio.

Moving to the second stage of the analysis, the extreme events, which were selected by the author of this thesis, will be presented and described. Within this framework, the two portfolios previously constructed will be investigated, covering the second stage of this analysis.

Further details in relation to the methodology applied will be discussed in Chapter 3 of this thesis.

2. Related literature

The chapter will present a brief set of definitions that will capture the working principles behind Artificial Intelligence and Machine Learning. This will be then followed by a section dedicated at analyzing the development of these technologies within the hat of the nowadays applications in the asset management industry.

Following, a discourse will continue touching other relevant concepts that will be implemented for the purpose of this research, namely Markowitz's Optimal Portfolio.

The last section is focused at laying out the concept of extreme adverse events which are the key of novelty for the proposed research objective.

2.1 Foreword on Artificial Intelligence

Among practitioners and professionals, the most pressing issue at hand appears to be related at achieving a unique consensus on Artificial Intelligence's (AI) definition. As Jerry Kaplan enounces in his book, there are many viable explanations and definitions of Artificial Intelligence but most of those are roughly around the concept of creating computer programs or machines which are capable of a behavior that would be regarded as clever or smart if exhibited by human beings (Kaplan, 2016).

Practically, the notion could be summarized as anything that results in machines acting more intelligently. Delving deeper into the sole definition itself, a relevant rationale provided by Kaplan and Haenlein further highlights the burdensome task of defining AI: *"…it is surprisingly difficult to define what AI is and what it is not. Or, to put it differently, there are about as many different definitions of AI as there are ways to describe Snow White's beauty"* (Kaplan and Haenlein, 2019). The authors then continue their analysis focusing on the evolution of AI implementation, initially deployed to perform simple tasks which then evolve into more complex ones. This evolution mirrors the substantial change that will affect firms: first internally – allowing for enhancement of efficiency and shrinkage of costs – and then externally – reestablishing different relationships with customers and stakeholders. In turn, it must be acknowledged that the concept of AI is broad and takes different meaning to different people, depending on who is the intended final user. No matter who is defining AI, the crucial point reverts to the fact that AI means intelligence, raising the question: how can one define intelligence?

While the effort of effectively defining intelligence, per se, does not present itself as a novelty, it still poses significant challenges to this day. Particularly, it is difficult to define human intelligence, as it is not a pure singular component that stands alone, and that is unbiased by other factors – such as conscience, empathy, and ethics. Matter of fact, the development of AI does not solely rely on progress in the engineering field (providing the software and hardware application of AI), mathematical and statistical models (providing feasible tools to assess performance) or cognitive intelligence, but also on psychology and behavioral science.

These, together with linguistics, contribute greatly to understand how AI works and to measure its effectiveness. Thus, when thinking of intelligence, the cognitive spectrum cannot be considered in isolation from the emotional intelligence, which is equivalently relevant.

Finally, the disruption brought forward by technological advancements has generated a massive amount of exchanged data and available information. As a result, significant challenges are faced by individuals attempting to navigate their way around it, culminating with human experts unable to sort the available sources of information, opening the way for an abundance of AI applications to provide efficient and practical tools to assist in such tasks. Lopez de Prado (2020) acknowledges that "*it is estimated that 90% of all recorded data have been created over the past two years, and 80% of the data is unstructured*".

As the number of fields of implementation grows, the more compelling the role of AI tools, and the greater the level of complexity and sophistication associated to the specific tasks to be carried out likewise. AI tools do not replace experts, but extend their capabilities in accomplishing tasks that would not have been possible to achieve by either human or machine, in singularity.

2.1.1 Core components of Artificial Intelligence

Any discussion on AI should be started by providing a definition of intelligence. For the purpose of this work, intelligence is defined as the ability to perceive one's environment and surrounding circumstances, to deduce and relate relevant information – while retaining it as knowledge – and apply it to the decision-making process with the final goal of providing better solutions for the intended end-user.

Going back to Kaplan, in his book, he further lays forward the following definition: "*The essence of AI – indeed, the essence of intelligence – is the ability to make appropriate generalizations in a timely fashion based on limited data. The broader the domain of application, the quicker conclusion can be drawn*" (Kaplan, 2016). The author provides an accurate and simple definition that captures the real use for Artificial Intelligence: infer conclusions from an unorganized data set.

At this stage, the concept of AI can be thought as an umbrella encapsulating and encompassing different fields, theories, and technologies. While it deviates from the scope of this master thesis to define the core components of AI, a brief focus on the most relevant ones with respect to portfolio allocation application must be delineated in order to achieve a clearer and complete comprehension of future interactions among existing technologies in a real-world application.

2.1.2 Machine Learning

Machine Learning is placed within the larger context of Artificial Intelligence. As previously acknowledged, AI is the overall field, which is concerned with endowing computers with intelligence, by any means necessary. Machine Learning (abbreviated to ML hereafter) is a sub-field of AI rooted on the idea of self-learning, that allows software systems to automatically improve their performance on a specific task by learning from data, without being explicitly programmed. It involves training a model on a dataset, and then using that trained model to make predictions or decisions without being explicitly told how to perform the task.

The desired input and output behavior is displayed to the algorithm, allowing it to progressively learn the mapping and relation between the two, through the application of statistical models. Solutions do not need to be directly programmed – depending on the type of machine learning algorithm being deployed – rather the algorithm can be continuously trained on new data.

ML algorithms focus on prediction based on already known properties and detected and extracted patterns from the data, they improve their performance over time through usage and empirical information. As a paradigm, as more data is being fed to the algorithm, better solutions are achieved or less efforts are needed to build a solution. This can be better captured by simply thinking to the information users provide every time they click or do not click on a search result, Google uses that information to improve their Google search results (Kavlakoglu, 2020).

It must be acknowledged, however, that machine learning still requires some type of human maintenance in order to properly carry out its functionalities. The inputs must be programmed outlining two specific data sets: the training data set, which will be employed to perform trials and obtain good generalization and pattern detection to achieve the desired performance; and the test data set, the actual real-life inputs on which the algorithm will be run. Conversely, programmers need to select the most appropriate algorithm, to reduce prediction error or noise (Kaplan, 2016).

ML has the potential to revolutionize many fields by enabling computers to learn and make decisions in ways that would be difficult or impossible for humans to do manually.

Covering the categorization of ML (Bishop, 2006), it can be broadly sectioned into:

- Supervised learning;
- Unsupervised learning;
- Reinforced learning.

Supervised learning is the task of providing the machine with examples, and training and teaching the model to map input data to corresponding finite target labels. To translate this to a practical example, a programmer feeds the machine with a large set of pictures of different birds and trains the algorithm to return the label "Bird" on demand whenever a picture of a bird is provided. As straight forward as this might be, it is as well of relative usefulness as the machine cannot distinguish anything that it is not a bird until taught differently. Consequently, new labels can be created, for cats and dogs for instance, and the model is again trained on a large set of pictures. The example just described is identified as classification problem, i.e., the target variable is a finite number of discrete categorical variables. When, instead, the target variable is a continuous real number value, the supervised learning algorithm is solving a regression problem. Most known supervised learning models are linear regression, decision trees, random forest, and artificial neural networks (Bishop, 2006).

Unsupervised learning algorithms are deployed when considering a scenario where the algorithm is not fed any labeled output in the training data set. The final goal of the activity is to discover and extrapolate groups of similar examples within the data, or more simply identifying patters and relationships in the data within the sample set under analysis. Commonly known algorithms of unsupervised learning are hierarchical clustering, principal component analysis, and autoencoders (Bishop, 2006).

Reinforcement learning (Sutton and Barto, 1998) is focused on the actions needed to be taken by intelligent agents within a specific situation to maximize a cumulative reward. In contrast with supervised learning, the algorithm is not fed optimal target from which to learn from, rather the algorithm must achieve the desired output through a process of trial and error whilst interacting with its environment. In reinforcement learning, the algorithm learns by interacting with its environment and receiving rewards or punishments for certain actions.

To provide a complete walkthrough of the ML process the following steps must be carried out:

- Data collection;
- Data processing;
- Pattern extraction;
- Feature selection;
- Supervised/Unsupervised/Reinforcement learning task.

The above detailed steps are of key importance in order to create appropriate feature representation throughout the ML workflow.

2.1.3 Neural Networks and Deep Learning

Neural Networks and Deep Learning are closely related techniques for implementing ML. A Neural Network is a type of machine learning model that is inspired and mimics structure and function of the human brain (Kavlakoglu, 2020). It is composed of at least four different layers of interconnected "neurons" – input, weight, bias and output – which process data and exchange information back and forth among each other.

In a Neural Network, each layer processes the input it receives and passes it on to the next layer. The process is reiterated until the final layer produces the output. The learning process involves adjusting the strengths of the connections between the neurons (called weights) based on the input data and the sought-after output, in order to improve the accuracy of the network's prediction (Bishop, 1994).

There are many different types of neural networks, specific to the type of task or activity that must be carried out, precisely Feedforward Neural Networks (FNN), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN).

Feedforward Neural Networks (FNN) are the most basic type, where the information travels in only one direction through the network, from the input layer to the output layer, and there are no loops or connections within the hidden layers that transmit the information backward.

Convolutional Neural Networks (CNN) are designed to process data that has a grid-like structure, such as an image, and are commonly used for tasks such as image classification and object detection (Lecun, Bottou, et al., 1998). In a CNN, the layers of neurons are arranged in a three-dimensional grid, with each layer processing a small region of the input data. These layers are called convolutional layers, and they are followed by one or more fully connected layers, which process the output of the convolutional layers and produce the final output.

Recurrent Neural Networks (RNN) are designed to process sequential data, such as time series or natural language, and are capable at processing data that carries within themselves a temporal feature, meaning that the output of the network can depend on data from previous time steps as well as the current input (Rumelhart, 1986). In a RNN, the hidden state of the network at each time step is a function of the input at that time step, and the hidden state of the network at the previous time step, the information which is transmitted to the following layers carries a "memory". This allows the network to retain information about the past and use it to process the current input.

Deep Learning is a subset of ML that uses Neutral Networks with many layers – hence the "deep" in the name. These deep Neural Networks are able to learn and represent very complicated patterns in data, making them particularly powerful for tasks such as image and speech recognition. In other words, Deep Learning is a specific type of ML that involves training deep Neural Networks on large datasets (Goodfellow et al., 2016).

Deep Learning has achieved significant successes in recent years, leading to many advances in a variety of fields, particularly led by a number of factors, such as a large amount of labeled data available. Deep Learning algorithms require a conspicuous amount of data to learn effectively, the increased availability of powerful computers, the significant improvements in the algorithms and techniques used for training deep Neural Networks, and, particularly, their ability to learn from unstructured data – such as audio, images and text – allow for more flexibility and ease at setting up the goals to be achieved.

To summarize, ML is a broad field that includes a variety of approaches for enabling computers to learn from data, while Neural Networks and Deep Learning are specific types of machine learning that use algorithms inspired by the structure and function of the human brain to learn from data.

2.2 Machine Learning and Asset Allocation overview

Machine Learning carries the promises to fully revolutionize all the aspects of individuals daily life, from automatizing the simplest tasks to independently carrying out activities that only recently could be performed solely by expert humans. Artificial Intelligence (AI) is bringing forth a revolution in so many different fields. The financial sector is currently experiencing an exciting period of proliferation of various ML driven activities and tools that allow for these disruptive technologies to improve their offerings to the investors. These technologies can automate manual procedures like delivering emails to customers, conducting compliance audits, or inputting data digitally, to provide the investors with tailored portfolio allocation strategies which best suit their preferences. There are many factors that can influence portfolio allocation decisions, such as the risk tolerance of the investor, the expected return on different assets, and the correlation between different assets. Machine learning algorithms can be used to analyze these factors and help make better portfolio allocation decisions. Asset Managers may boost productivity, enhance the member experience, and lower the risk of human mistake by looking for options for intelligent automation.

Lopez de Prado (2020) acknowledges that among asset managers price prediction is currently the most popular application of ML algorithms, however, he argues that considering solely price forecasting disregards equivalently relevant applications of the available technologies. Some examples are high-frequency trading firms which are deploying computerized algorithms to detect trends and analyze real-time exchange feeds identifying trade patters leveraged by informed traders at the expense of uniformed ones (Arifovic, He, et al., 2022), or credit-rating agencies which are combining the ability to automatically learn from data of ML algorithms – such as decision trees, neural networks, and support vector machines – to their credit rating models (Jiang, 2022). There are plenty of equivalently relevant applications, as recognized by Lopez de Prado (2020), "*hedging, portfolio construction, detection of outliers and structural breaks* [...]*, and many others*".

A conspicuous amount of literature is reverting around the application and the benefits for investors related to robo-advisors that can be defined as *"as financial investment services that are based on algorithms and provided to customers online"* (Maume, 2021). Particularly, the research paper "*Artificial Intelligence Alter Egos: Who might benefit from robo-investing?*" by D'Hondt, De Winne et al. (2020) investigates the use of robo-advisors, which are automated investment advisory platforms that use algorithms and machine learning to provide investment recommendations and manage portfolios. The paper discusses the benefits and drawbacks of using robo-advisors, including their ability to provide low-cost, automated investment management, their potential to improve the accuracy and consistency of investment decisions, and their limitations in terms of the complexity and sophistication of the investment strategies they can implement. The authors also discuss the risks and challenges associated with robo-advisors, including their reliance on data and algorithms, their potential to perpetuate biases and oversights, and their limited ability to provide personalized advice and support. Overall, the paper suggests that robo-advisors have the capability to be a valuable tool for investors, but also highlights the need to carefully consider the limitations and risks associated with these platforms, as well as the need for financial advisors to consider the role of robo-advisors in the investment landscape.

It is universally acknowledged by practitioners that individuals benefit from market participation, however investors exhibit low level of diversification, resulting in portfolio performance which bears idiosyncratic risk. D'Acunto and Prabhala (2018) argued that robo-advising tools offer the opportunity to investors to achieve better results while reducing their unsystematic risk exposure through a higher level of diversification. Further research highlights how the higher level of cost shrinkage feature of robo-advisors, the increased level of transparency while bypassing the behavioral biases (Gargano and Rossi, 2018), and cognitive limitations linked to human financial advisors (Rossi and Utkus, 2021) deliver a better return performance for investors. Particularly, among researchers there is a common consensus in relation to the target average investors features: low income, low diversification, and trading activity ahead of adoption of robo-advisors tools – despite the effectiveness proven by research – many end-users are reluctant to engage with automated operations (robo-advisors) as argued by Filiz, Judek et al. (2022).

Going back to D'Hondt, De Winne et al., the authors discuss the potential impact of robo-advisors on financial markets including times of market stress, such as the financial crisis of 2008. The paper suggests that robo-advisors may be particularly useful during times of market stress, as they can provide low-cost, automated investment management that is less prone to human emotions and biases. The authors also argue that the use of robo-advisors may help to reduce the impact of "herding behavior" among investors, in which investors follow the decisions of others rather than making independent ones based on their own analysis. In addition, the authors caution that the use of robo-advisors may also have negative consequences, such as increasing the concentration of assets in certain sectors or increasing the reliance on algorithms and data, which may have unintended consequences.

Overall, the paper suggests that robo-advisors may be a useful tool for investors, but also highlights the need to carefully consider the limitations and risks associated with these platforms, particularly during times of market stress.

This aspect of their research is particularly relevant for this thesis, as it is in linkage to the analysis performed, which does not account for market distressing event which are originated within financial markets – as observed by the authors with the financial crisis of 2008 – but rather to external extreme events, further analyzed in the following sections of this thesis.

2.3 Optimal Portfolio Allocation

In its seminal work paper Markowitz defined portfolio selection as "*the investor does (or should) consider expected return a desirable thing and variance of return an undesirable thing*". With its algorithm Markowitz displayed that the "efficient portfolio" can be attained from investors and that it is able to maximize expected returns given a certain level of risk (i.e., variance of returns) or, convexly, given a certain level of expected return the portfolio selection can minimize the risk.

The Markowitz portfolio allocation, also known as mean-variance optimization (MVO), is a widely used method for selecting a portfolio of assets that seeks to maximize expected return while minimizing risk. It is appealing for several reasons:

- It provides a systematic, mathematical approach to portfolio selection that helps investors make more informed decisions about their investments;
- It takes into account both the expected return and risk of each asset and allows investors to trade off these factors in order to find an optimal portfolio;
- It allows investors to diversify their portfolio in order to reduce overall risk, which can be especially important in times of market volatility;
- It has been widely tested and validated by researchers and has been shown to be effective in a variety of market conditions.

Overall, the Markowitz portfolio allocation is a widely respected and extensively used method for selecting a portfolio of assets and is considered a fundamental concept in modern investment practice.

This method is particularly appealing to portfolio managers, still to this day, thanks to its efficiency from a computational point of view. While diversification of investments – which, to be captured in essence, could be translated as not keeping all the eggs in the same basket – was a well-established investment practice long before Markowitz published its paper in portfolio selection, a multitude of new asset allocation techniques saw birth after Markowitz's publication.

Despite the general recognition and acclamation of the Markowitz framework, a great number of academics and professionals of the field criticized the mean-variance optimization – as the model presents a significant level of sensitivity with respect to the input parameters and the appropriate weight constraints to be specified in order to deliver fitting allocations (Michaud, 1989) – leading to portfolios that are less than optimal from a financial standpoint. Michaud argued that the mean-variance optimization has several limitations, including the assumption that investors have a single and well-defined risk tolerance, the reliance on historical data that may not accurately reflect future market conditions, and the inability to take into account more complex risk measures. Overall, Michaud's paper provides a critical perspective on the use of mean-variance optimization and highlights some of the drawbacks of this approach. In more current research, in strict linkage to the portfolio allocation via Machine Learning based on the mean-variance optimization model, Lopez de Prado (2016) argues that the model is prone to overfitting and does not adequately account for the risk of the out-of-sample performance.

It is undoubted, however, the significant footprint on the field of finance that the mean-variance optimization approach has left, as well it has helped fuel the ongoing debate about the most effective methods for portfolio optimization. The widespread adoption of the mean-variance optimization method is also explained by the fact that, nevertheless the vast number of proposed alternative models, very few allocation techniques succeed on a practical stance despite the theoretical appeal they initially seemed to provide.

2.4 Non-systemic extreme disruptive events

History has been many times marked by a multitude of disasters, and the last 20 years are an example of it. The Hurricane Irene in the Caribbean and in the USA, the floods in Australia as well as the earthquake in New Zealand and Japan in 2011 (Chaiechi, 2020), the 9/11 terroristic attack or the chain of terroristic attacks that put Europe to its knees for years (2015-2017), or the return to power of the Taliban in Afghanistan, and the disruption of the conflict between Russia and Ukraine early in 2022 are examples of extreme events that had severe impacts on the reference regions. These disruptive events are very different from each other but simultaneously present some similarities: they were unexpected and resulted in drastic consequences.

In their research paper, Broska et al. lay down the key aspects evolving around the concept of extreme events and provide a fitting definition, namely: "*An extreme event is a dynamic occurrence within a limited timeframe that impedes the normal functioning of a system or systems"* (Broska et al., 2020).

Further, the authors identify and isolate the following key properties of an extreme event:

- it is an outlier, as it lies outside the realm of what is recognized as normal conditions, or *"displays atypical behavior"*;
- it carries a significant impact, or extreme impact;
- in the aftermath, it appears as if it was explainable and displaying a certain level of predictability.

The extreme event definition finds fertile ground in many fields, starting with the scientific context including meteorology, climatology, and mathematics, and then always more frequently spreading to natural and social sciences, touching various aspects of life such as natural disasters, terrorist attacks, political revolutions, and major technological failures.

It must be acknowledged that an important aspect of these distressing events is strictly linked to the susceptibility of the system being impacted. Indeed, Broska et al. address how the vulnerability of the impact's recipients covers an important place in the identification of an extreme event. *"Only if stakeholders subjectively consider the event's impacts to be extreme, can they be declared as such."* (Broska et al., 2020). In addition, another key point being raised is related to the statistical financial models being applied to analyze the data related to these events that are regarded as unpredictable and which have origins related to causes or circumstances which are outside the reference framework, namely financial markets.

Following the Financial Crisis of 2008[1], Makridakis proposes various motives that explain why current statistical models were ineffective at providing accurate forecasting; among the various sources of ineffectiveness, it is highlighted the underestimation of uncertainty, due to the assumption that "*events are independent* [...], *forecasting errors are tractable* [...] and *the variance of forecasting error is finite, known and constant"* (Makridakis & Taleb, 2009).

Particularly, for those that can be identified as extreme events, their probability is most frequently disregarded as unknown or rather not easily computable and reliable; however, these events have indeed a probability of occurrence, despite how rare and unpredictable they might be. Therefore, when the events eventually occur, they are completely unexpected and carry consequences for all relevant stakeholders which are tragic.

Non-systemic extreme disruptive events can have significant implications on the financial markets, businesses, and individuals. They can lead to market volatility, asset price declines, and increased risk of financial losses. As a result, they can be challenging to manage and mitigate, and may require special strategies and resources to address, thus, it is important for investors and financial advisors as well as for policy makers to be aware of the potential risks and to have contingency plans in place to address these types of events.

Conventional financial models are not fit for computing and accounting for these events in an adequate manner to be considered reliable. For instance, there is a vast amount of literature rejecting the commonly used assumption of normally distributed prices for a variety of asset categories; for example, the work of Mandelbrot (1963), Fama (1965), Aparicio and Estrada (2001) and Estrada (2007).

Indeed, the classical assumption is that event-probabilities follow a normal distribution, but this distribution does not allow for prediction of extreme events, as these are dramatic, sudden, and remote events, outside the prediction of the normal bell-curve. As acknowledged by Makridakis *"A lack of normality means that errors, and therefore uncertainty, are much greater than those postulated by models, while in addition, there can be cataclysmic events that are impossible if the normality of error is assumed"* (Makridakis, 2009). This is an important consideration for investors and financial advisors, as it highlights the need to be aware of the potential risks and uncertainties associated with financial markets and other systems, and to have contingency plans in place to address these types of events. It is also important to be aware of the limitations of

[1] Please note that for the purpose of this Master's thesis the Financial Crisis of 2008 does not fall within the scope of unexpected extreme events, as one key component of those is that their causes of origination lay outside the environment of reference, namely financial markets. The Financial Crisis of 2008 is thus mentioned solely in reference to the quoted research article.
Makridakis, S., & Taleb, N. (2009). *Living in a world of low levels of predictability*. International Journal of Forecasting, 25(4), 840–844.

financial models and other systems, and to consider alternative approaches or strategies that may be more suitable for managing risk and uncertainty in non-normal or extreme environments.

Reverting to Broska et al., the concept of robustness must be involved as it provides approaches that are effective in meeting risks, surprises and unforeseen occurrences. These cover cautionary measures such as: designing for flexibility, testing and resilience engineering, which relate with finding ways to enhance the ability of organizations to recognize, adapt to and absorb variations, disruptions, and shocks.

Thus, a suggested solution that can be inferred from this analysis is that it is not necessarily needed to have prediction models that offer a complete and fully encompassing forecasting, rather to build robustness against their occurrence and have appropriate models which are capable of adequately responding to these distressing events and re-establish, in a timely manner, the original functioning of the reference ecosystem.

3. Data and Methodology

In relation to the methodology that will be implemented, the quantitative analysis which is performed is further detailed and the data set considered is described. The research will be based on the construction of two different portfolios: one being constructed on real returns observed on the market during a sample window of almost two years using the MVO, from here after this will be addressed as the Markowitz portfolio. The second portfolio will instead consider returns forecasted by deploying Machine Learning algorithms, and then optimized using the mean-variance approach – in line with the methodology applied by practitioners in the industry (D'Hondt et al., 2020). The construction of the two optimized portfolio will serve as ground to review the evolutions during market distressed periods.

3.1 Data set

The unique data set which is used considers a proxy market composed of 50 stocks identified from the MSCI World Index based on their market capitalization, namely the 50 greatest contributors to the index performance. Please refer to *Appendix 1* for further details. The data spans over a period of time of more than 10 years, going from January 2010 to September 2022, considering a daily quotation granularity relaying on historical data. The optimized portfolios consider a time stamp of almost 2 years rolling sample with quarterly rebalancing.

The market universe which is constructed focuses solely on equities excluding other type of financial instruments. The approach is in line with the current literature evolving around the analysis of ML techniques which are applied to the portfolio management field, see D'Hondt, De Winne et al. (2020), D'Acunto, Prabhala et al. (2019), Gu et al. (2018) and Gargano and Rossi (2020), and as well best adapts to the portfolio allocation with mean-variance analysis. The exclusion of more sophisticated products is performed in order to construct a portfolio allocation which could be presented to all types of investors, i.e. retail investors as well as more financially educated ones.

The built portfolios do not offer the possibility to hold cash, this limitation is considered in order for the hypothetical investor to be fully exposed to the market fluctuations. Further, for simplicity transactions costs are ignored to bypass adjustments related to cost management when rebalancing the constructed portfolio. No short selling is allowed within the framework analyzed in order to achieve reasonable weights.

It is observed that Healthcare (28%) and Technology (24%) represent the majority of the sectors to which the individual equities pertain to. The market universe constructed could be considered somewhat arbitrary, however the average investor which is actively trading does not invest in the full population of common stocks, which is available, but is rather undiversified. D'Acunto, Prabhala and Rossi (2019), when addressing the target investor for robo-advisors, perform an analysis in relation to the pre and post adoption and observe that the population analyzed hold less than 10 stocks. This is consistent with current literature, where it is acknowledged that the average retail investor is underdiversified.

In relation to the extreme events that will be discussed within the research, six key events have been identified within the last twenty year. The purpose of this proof of concept is to analyze events that were not originated within financial markets and to be able to provide an investigation of the effect of extreme events with different roots in different countries. Three types of extreme events will be detailed: terrorist attacks, political changes, and natural disasters.

The events to be analyzed are the following:

Extreme event	Country	Date
09/11 Terrorist Attack	U.S.A.	11/09/2001
Tsunami Thailand	Thailand	26/12/2004
Fukushima Nuclear Disaster	Japan	11/03/2011
Paris Terrorist Attack	France	13/11/2015
Brexit Referendum	U.K.	23/06/2016
2016 United States Presidential Elections	U.S.A.	08/11/2016

Table 1: *Extreme events analyzed presentation*

The 09/11 attacks refer to a series of terrorist attacks against the United States. According to the US Congress, (2006) it has been the deadliest terrorist attack in the history. Consequently, the NASDAQ and NYSE chose not to open on the day of the attacks and reopened only on September 17, 2001, marking the longest period the US markets had remained closed since 1933 (Makinen G., 2002).

The Indian Ocean Tsunami happened the day after Christmas in 2004. The 9.1 magnitude earthquake that provoked the tsunami was the strongest since 1964 and nearly 220 000 people died or were missing in 11 countries affected by the disaster, according to National Geographic. Total damages were assessed at around $508 million, while losses were estimated at $1.690 billion, thus a total loss of $2.2 billion (Telford J., Cosgrave J. and Houghton R., 2007).

Following a major earthquake, a tsunami provoked an accident at Fukushima Daiichi Nuclear Station on March 11th, 2011. In addition to the impact on humans and on the environment, the Japanese Cabinet Office (CAO) estimated damages at ¥16.9 trillion (an equivalent of $210 billion), which corresponds to about 4% of Japan's GDP.

Paris terrorist attacks on November 13th, 2015, was the deadliest terrorist attack in the history of Paris as 130 people were killed while hundreds were injured. The attacks that occurred in Paris brought a sharp decrease in tourism. Indeed, according to a Cambridge analysis, the city suffered a €750 million loss in tourism revenue in the first year after the November 2015 attacks.

In June 2016, the United Kingdom held a referendum asking to the citizens whether the country should remain a member of the European Union. The referendum ended in 51.9% votes in favor of leaving the EU. The negative outcome of the referendum arose many debates, the consequences have been widely analyzed. For example, Kierzenkowski (2016) argued: *"major negative shock to the UK economy, with eco-nomic fallout in the rest of the OECD, particularly other European countries"*.

In November 2016, the outcome of the U.S.A. Presidential election saw the victory of the Republican candidate, Donald J. Trump. The result came as a shocking one for the public primarily due to the fact that major predicting sites converged towards a Clinton's victory (Wright, 2018), and conversely for the market which initially reacted negatively concerned over Trump's unpredictable nature, which contributed uncertainty as well as at a global level.

In relation to the above listed events, the two constructed portfolios will be discussed in order to not directly infer in relation to their performance, but rather to recreate, as if it had been used in the past, the historical performance of the investment strategy developed.

3.2 Markowitz's framework

A universe composed of n stocks is considered, and $x = (x_1,\dots,x_n)$ is the vector of weights in the portfolio. As addressed in the previous section, cash holdings are not considered within the framework, thus the portfolio is fully invested, translating thus to $\sum_{i=1}^{n} x_i = 1_n^\top x = 1$.

Considering the rolling period from January 2021 to September 2022, the daily returns of the n stocks are computed. The vector of stock returns is $R = (R_1,..,R_n)$, with R_i denoting the return of asset i. The portfolio return is then represented as a function of the weights vector x and the return vector, $R(x) = \sum_{i=1}^{n} x_i R_i = x^\top R$. Let $\mu = E[R]$ and $\sum = E[(R-\mu)(R-\mu)^\top]$ be the investor's expected return vector and the covariance matrix of stock returns (Perrin and Roncalli, 2019).

The portfolio expected return is:

$$\mu = E[R(x)] = x^\top \mu$$

and the portfolio variance is:

$$\sigma^2(x) = E[\big(R(x) - \mu(x)\big)\big(R(x) - \mu(x)\big)^\top] = x^\top \sum x$$

In relation to the portfolio optimization performed, the investor's objective was aligned with the maximization of the Sharpe Ratio (Sharpe, 1994).

The Sharpe Ratio is computed as presented below:

$$Sharpe\ Ratio = \frac{\mu - R_f}{\sigma}$$

where:

- μ is the expected return of the portfolio;
- σ is the standard deviation of the portfolio;
- R_f is the risk-free rate (monthly return of the 3-month Treasury Bill).

The portfolio is rebalanced quarterly over the two-year rolling window considered. The objective formula is defined as:

$$\max x^{\top}\mu - \frac{\emptyset}{2}x^{\top}\Sigma x$$

conditional to $x \geq 0$. In the above equation ∅ is interpreted at the risk aversion parameter, which is set equal to one as it maximized the Sharpe Ratio. The expected return vector and covariance matrix are constructed as detailed above.

3.3 Machine Learning models

Within the context of supervised learning tasks, a set of four different models has been constructed relying on returns data from January 2010 to December 2020, a ten-year time sample that will be devoted to training the algorithms – the in-sample data – while the data from January 2021 to September 2022 represents the out-of-sample performance tested. Please refer to *Appendix 2* for further details.

The stock monthly returns training sample has been further enriched by including features that have been computed from the initial data set of prices and volumes of the individual stocks. This includes the Relative Strength Index (RSI) (Wilder Jr., 1978) over 21, 63 and 251 days – corresponding to one month, one quarter and one year – the momentum factor (MOM) introduced by Fama-French (1992) over 21, 63 and 251 days, and volume percentage change. The feature engineering which is performed for the purpose of this research is rather simplified, but it provides a basis for a simple estimation of prediction signaling, which improves the accuracy of the model predictions, allowing for a more precise risk management and portfolio construction. Please refer to *Appendix 3* for further details.

In order to avoid forward-looking bias, cross-validation was implemented (Gu et al., 2018). The training sample was sliced into multiple subsets composed by 750 data points – about three trading years – with an offset of 500 data points – approximately two years. Each slice of time frame was composed of 70% of the data being devoted to training and 30% of testing and validation. The mean-squared error (MSE) represents the optimized loss functions. Please refer to *Appendix 4* for further details.

Four different models are considered, K-Nearest Neighbors (KNN), Random Forest, Gradient Boosting Regressor, and Artificial Neural Networks.

K-Nearest Neighbors: the KNN algorithm is a simple and intuitive approach to classification and regression that is based on the idea of using the labels of the K examples in the training set that are most similar to the new example to predict the label of the new example. The algorithm works by calculating the distance between the new example and each example in the training set, and then selecting the K examples that are closest to the new example. The label of the new example is then predicted based on the labels of these K nearest neighbors. (Cover and Hart, 1967).

Random Forest: a Random forests model is an ensemble method, which means that it combines the predictions of multiple decision trees to make a final prediction, which is obtained by averaging the predictions of the individual trees. To create a random forest, a large number of decision trees are trained on bootstrapped samples of the training data. A bootstrapped sample is a random sample of the training data that is drawn with replacement, which means that some examples may be included multiple times in the sample. This helps to reduce the bias of the model and to improve its generalization ability as well as its accuracy, reduced overfitting, and increased stability (Breiman, 2001).

Gradient Boosting Regressor: just like a random Forest, the gradient boosting regressor is as well an ensemble method, which combines the predictions of multiple decisions trees to obtain a final prediction. In each stage of the gradient boosting process, a series of decision trees are trained on the training data, with each tree attempting to correct the errors made by the previous tree. The decision trees are trained in a sequential manner, with each tree being trained on the residual errors made by the previous tree. The prediction error is then calculated using the specified loss function – squared error – and the negative gradient of the loss function is used to update the predictions. This process is repeated until the predefined stopping criterion is reached, i.e. the minimum improvement in the error (Friedman et al., 2000).

Artificial Neural Network: the Multilayer Perceptron Network (MLP) is the selected architecture for the artificial neural network algorithm (Rosenblatt, 1958). The MLP is a feedforward network, meaning that the information is travelling in a unique direction, from the input layer throughout the hidden layers to the output layer. To train a multilayered perceptron regressor, the model is fed a set of input-output pairs and adjusts the weights of the connections between the neurons in order to minimize the error between the predicted values and the real values via the gradient descent value. The loss function is optimized using the Adam Optimizer (Kingma, 2014).

Lastly, a model blend of the above-listed models is considered, creating a voting regressor where the mean and median of the predictions of the individual regression models. Please refer to *Appendix 5* for further details.

To optimize the models, a grid search was performed to identify the best combination of the hyperparameters which were defined above. The process works by specifying a grid of hyperparameters values and then training and evaluation a model for each combination of values. The combination that produces the best results is then selected as the optimal set of hyperparameters for the model.

All the models were deployed in order to obtain the estimated return data points for all the individual stocks included in the studied market universe and identified in the *Data set* section for the period from January 2021 to September 2022. On the obtained predictions, the mean-squared error was computed.

The average mean-squared error MSE and mean-absolute return MAE of the out-of-sample forecasted values of all stocks within the rolling-winding test dataset is presented in Table 2. The four different algorithms are presented with the following notation: (KNN) is K-Nearest Neighbors, (RF) Random Forest, (XGB) Gradient Boosting Regressor, and (ANN) Artificial Neural Network. It can be recognized that the Gradient Boost regressor delivers the lowest MSE, followed up by the ensemble algorithm.

	Cross-sectional MSE and MAE				
	Ensemble	Random Forest	Gradient Boost	KNNeighbours	Multi Layer Perceptron
MSE	0,00733	0,00751	0,00694	0,00840	0,00758
MAE	0,05779	0,05796	0,05654	0,06331	0,06006

Table 2: *Out-of-sample MSE and MAE for the prediction models*

Finally, the blend algorithms along with the four citated methods were assessed in order to answer the investors Machine Learning portfolio allocation problem and identifying the optimal weight through the MVO methodology described in the above section.

3.4 Validation of prediction accuracy

The final step performed once the final predictions have been obtained is related to the validation of such estimates. In evaluating the time series analysis of stock returns, it must be acknowledged that *"...many financial time series do not exhibit stationarity, but often the changes in them, perhaps after applying a log transformation, are approximately stationary*" (Ruppert and Matteson, 2015). Further, the authors argue that the stock returns do not hold a time stamp variation, this means that the distribution of the process does not change over time and that the process does not exhibit trends or seasonality.

A stationary process is characterized by the following properties:

- Constant mean: The mean of the process does not change over time;
- Constant variance: The variance of the process does not change over time;
- Constant autocovariance: The autocovariance of the process, which measures the relationship between the process at different points in time, does not change over time.

The statistical test which is used to compare the accuracy of the obtained forecasts is the Diebold-Mariano test. First introduced in 1995 in a paper titled "Comparing Predictive Accuracy" (Diebold and Mariano, 1995), the model is used to evaluate the performance of forecasting models as well as to compare the accuracy of different models.

The Diebold-Mariano test is used to assess whether the forecasts obtained implementing the ML models described in the above section are significantly different.

Consider the g_i and e_i as the residuals for the forecasted values, and d_i is defined as the loss differential time series and is presented as:

$$d_i = g_i^2 - e_i^2 \text{ or } d_i = |g_i| - |e_i|$$

The first formula being the MSE and the second one the MAE measures.

The mean of the loss function is characterized as:

$$d' = \frac{1}{n}\sum_{i=k+1}^{n} d_i \qquad \mu = E[d_i]$$

The autocorrelation function γ_k is the autocovariance at lag k, for $n > k \geq 1$, is defined as:

$$\gamma_k = \frac{1}{n}\sum_{i=k+1}^{n}(d_i - d')(d_{i-k} - d')$$

The autocovariance at lag k is a measure of the relationship between the process at different points in time. It is defined as the covariance between the process at time i and the process at time $i-k$, where k is the lag.

The Diebold-Mariano (DM) statistics, for $h \geq 1$, is defined as follows:

$$DM = \frac{d'}{\sqrt{\left[\gamma_0 + 2\sum_{k=1}^{h-1}\gamma_k\right]/n}}$$

The parameter h in the Diebold-Mariano statistic represents the lag of the autocovariance function, which is a measure of the relationship between the process at different points in time. The value $h = n^{1/3} + 1$ is generally used as a default value.

The DM statistic is normally distributed $DM \sim N(0,1)$. To test the null hypothesis, the DM statistics is compared to the two tailed critical value, $-z_{\alpha/2}$ and $z_{\alpha/2}$ at a significance level of $\alpha = 0{,}05$.

If the Diebold-Mariano statistic is significantly different from zero, that is $|DM| > z_{\alpha/2}$, this suggests that the two forecasts are significantly different and that one of the forecasts is more accurate than the other, and thus the null hypothesis is rejected.

4. Results

The empirical results that were achieved focus on answering the research questions which were detailed in Chapter 1: (RQ. 1) *Does the Machine Learning constructed portfolio deliver stronger performance for the investor in comparison to the Markowitz constructed portfolio?* and (RQ.2) *Within distressed market conditions, where does the Machine Learning constructed portfolio position itself against the Markowitz constructed portfolio?*

4.1 Machine Learning portfolio construction

The results related to the validation of the various models are presented in the below table.

Forecast Models considered	DM	Conclusion
Ensemble vs GB Regressor	0,00633	H_0 not rejected
Ensemble vs RF	0,01997	H_0 not rejected
Ensemble vs KNN	0,01767	H_0 not rejected
Ensemble vs MLP	0,00297	H_0 not rejected

Table 3: *Diebold-Mariano test statistics results*

The Diebold-Mariano statistics have been run to assess whether significant differences could be identified among the different forecasting models, comparing each model against the ensemble algorithm, with parameter $h=18$. It is observed that in all four instances there is no statistical difference between the forecast errors of the two considered models, resulting in the null hypothesis not being rejected. Thus, it can be inferred that one model is not likely to be more accurate than the other.

Considering the characteristics of the different models, all four present a set of advantages and disadvantages, and in general, they could all benefit from feature refinement. Assessing the results of the Diebold-Mariano test, the final model which has been selected is the ensemble algorithm, which averages the results from the different models allowing for overfitting risk shrinking as well as reduction of model-specific bias.

As detailed in Chapter 3, a set of variables were included within the training data set in order to enrich the data population. In below *Table 4*, the single parameters have been ranked in contribution within the cross-validation time stamps within the training data set for the ensemble algorithm.

Rank	01/04/10 - 12/24/12	12/27/11 - 12/18/14	12/23/13 - 12/13/16	12/17/15 - 12/10/18	12/12/17 - 12/03/20
1	MOM 251	MOM 251	MOM 251	RSI 251	RSI 251
2	RSI 251	RSI 251	RSI 251	MOM 251	MOM 251
3	RSI 63	MOM 63	MOM 63	MOM 63	MOM 63
4	MOM 63	RSI 63	RSI 63	RSI 63	RSI 63
5	MOM 21	RSI 21	RSI 21	RSI 21	RSI 21
6	RSI 21	MOM 21	MOM 21	MOM 21	MOM 21
7	Volume Change	Volume Change	Volume Change	Volume Change	Volume Change

Table 4: *Features ranked based on contribution on estimation*

The ranks were computed over the 10 years training period split in about 3 years of sample adjustment. It can be observed that the contribution to the prediction does not vary greatly among sections, the Momentum factor at 251 days and the Relative Strength Index at 251 trading days. The rank validates that the larger the pool of historical data the algorithm holds, the stronger the contribution to the contribution of the features to the prediction.

Once the selected model has been validated, the forecasted values are used in order to obtain the data related to the Machine Learning portfolio, via MVO through the maximization of the Sharpe Ratio.

In *Appendix 6*, the obtained optimized weights are presented for the Machine Learning portfolio and for the Markowitz portfolio. The weights refer to the two-year rolling sample with quarterly optimization. The median of the realized returns of the portfolios over the period under review together with the Sharpe Ratio spreads are detailed in below *Table 5*.

The median of the daily realized returns provides a sense of the risk of the portfolio displaying the level of return that is achieved in the middle of the distribution of returns during the quarterly intervals of rebalancing. It is observed that a hypothetical investor deciding to allocate their wealth to the Machine Learning portfolio can benefit from a higher gain compared to the Markowitz portfolio, from a general perspective over the quasi two years investment period. The median spread between the two portfolios is rather stable, with the Machine Learning portfolio delivering a higher benefit to the investor.

The first and third quartiles are as well included in the described results, as they provide insight in relation to the distribution of the realized returns over the period under review, as well as providing information with regard to the level of risk associated with the investment. From the gathered data, it can be observed that the fluctuations between the quartiles are not strong over the different quarters, for the Markowitz portfolio it is recognized that during the first two and the last two quarters the first quartiles are negative, and the third quartiles are positive, suggesting that returns are skewed towards the negative side, with rather few returns. In addition, it is acknowledged as well that the third quartiles of the Machine Learning portfolio are consistently higher than those achieved with the Markowitz portfolio, this allows to infer that the ML Portfolio had more high returns over the period being considered.

Finally, considering the Sharpe Ratio, we observe that both portfolios can achieve a stronger level of return compared to the risk-free rate (3-month Treasury Bill). The Machine Learning portfolio delivers a stronger performance in terms of risk-adjusted returns, this is in line with the higher level of diversification that allows for a lower level of risk to which the portfolio is exposed to. The inference is in line with current literature which identifies that undiversified investors benefit from adopting strategies constructed through Machine Learning (D'Acunto et al.,2018).

	Realized returns			Sharpe Ratio
	Median	Q1	Q3	
Q1 2021				
ML	0,1166%	0,10	0,13	0,15
MZ	0,0052%	-0,01	0,01	0,12
Q2 2021				
ML	0,0901%	0,10	0,13	0,08
MZ	0,0032%	-0,01	0,01	0,04
Q3 2021				
ML	0,0682%	0,00	0,07	0,14
MZ	0,0008%	0,00	0,01	0,10
Q4 2021				
ML	0,0699%	0,07	0,07	0,32
MZ	0,0059%	0,00	0,01	0,25
Q1 2022				
ML	0,1081%	0,10	0,11	0,29
MZ	0,0044%	0,00	0,01	0,12
Q2 2022				
ML	0,0781%	0,08	0,08	0,24
MZ	-0,0008%	-0,02	0,02	0,15
Q3 2022				
ML	0,0729%	0,07	0,07	0,23
MZ	0,0033%	-0,01	0,01	0,16

Table 5: *Median and Q1 and Q3 of Realized returns at different rebalancing intervals*

4.2 Extreme Events analysis: how does the Machine Learning portfolio behave?

In this section, the single events are considered in their period pre, during, and post-event, having identified the event key date, please refer to *Table 1.* D'Hondt, De Winne et al. (2020) include in their research the financial crisis data analysis, and the outlook on the objective of the investigation is somewhat in line covering a time widow of about 2 years. The optimized weights obtained for the Markowitz portfolio and the Machine Learning portfolios are used.

The events are grouped within three sub-categories:

- Terroristic attacks: including the 09/11 and the Paris Terroristic attack;
- Natural disasters: including the Tsunami in Thailand and the Fukushima nuclear disaster;
- Political changes: including the Brexit referendum and the 2016 U.S. Presidential election.

4.2.1 Terroristic Attacks

With reference to the 09/11 terroristic attacks and the Paris terroristic attacks considered the pre during and post-event period as detailed below:

Extreme event	Pre-event period	Event Day	During period	Post-event period
09/11 Terrorist Attack	Q4 2000 - Q2 2001	17/09/2001	Q3 2001	Q4 2001 - Q2 2002
Paris Terrorist Attack	Q1 2015 - Q3 2015	16/11/2015	Q4 2015	Q1 2016 - Q3 2016

Table 6: *Period characterization for the Terroristic Attacks*

The portfolios constructed have been used to simulate and obtain factual data on how the strategy developed would have performed in the past during periods of distressed market conditions.

Please refer to below *Table 6* for the results obtained.

09/11 Terrorist Attack Period	Median MZ	Median ML	Q1 MZ	Q3 MZ	Q1 ML	Q3 ML	Sharpe Ratio MZ	Sharpe Ratio ML
Pre-crisis								
Q4 - 2000	0,07%	0,04%	-0,05	0,13	-0,02	0,08	0,14	0,11
Q1 - 2001	0,03%	0,10%	-0,02	0,08	-0,03	0,25	0,19	0,24
Q2 - 2001	0,01%	0,08%	-0,02	0,04	-0,02	0,12	0,15	0,21
During crisis								
Q3 - 2001	0,01%	0,02%	-0,01	0,05	-0,03	0,11	0,12	0,11
Post-crisis								
Q4 - 2001	0,02%	0,01%	-0,01	0,06	-0,01	0,03	0,08	0,10
Q1 - 2002	0,01%	0,01%	-0,01	0,03	-0,01	0,04	0,06	0,07
Q2 - 2002	0,03%	0,02%	-0,03	0,08	-0,02	0,06	0,10	0,09

Paris Terrorist Attack Period	Median MZ	Median ML	Q1 MZ	Q3 MZ	Q1 ML	Q3 ML	Sharpe Ratio MZ	Sharpe Ratio ML
Pre-crisis								
Q1 - 2015	0,02%	0,03%	-0,01	0,05	-0,01%	0,05	0,08	0,11
Q2 - 2015	0,01%	0,02%	-0,01	0,04	-0,01%	0,05	0,06	0,07
Q3 - 2015	0,03%	0,05%	-0,02	0,08	-0,02%	0,10	0,09	0,12
During crisis								
Q4 - 2015	0,02%	0,06%	-0,02	0,06	-0,01%	0,12	0,05	0,02
Post-crisis								
Q1 - 2016	0,05%	0,06%	-0,02	0,09	-0,03%	0,15	0,06	0,09
Q2 - 2016	0,03%	0,01%	-0,01	0,06	-0,01%	0,04	0,14	0,08
Q3 - 2016	0,01%	0,01%	-0,01	0,04	0,00%	0,03	0,17	0,13

Table 7: *Terroristic Attacks results*

Over the period under review, we observe the median and first and third quartiles of the realized returns, together with the Sharpe Ratios for the two portfolios. Focusing on the 09/11 terroristic attack, it can be observed that the investor investing via the Machine Learning allocation experiences a higher level of median return in the period before the extreme event took place, particularly in Q1 2001 with a quarterly realized return medio (0,10%) for the Machine Learning portfolio and (0,03%) for the Markowitz portfolio. As well in the period pre-crisis the first quartiles of the MZ portfolio are generally lower in comparison to the ML portfolio, signaling that the MZ portfolio has experienced a higher proportion of low return. This is as well in line with the higher values of the third quartile observable for the ML portfolio.

The during crisis period together with the post-crisis period deliver results that are more stable in terms of median realized returns for both portfolios. It is interesting to observe that the range between the first and third quartiles for the Machine Learning portfolio is particularly large during the period when the terroristic attack takes place in September 2001, suggesting that the portfolio experiences a greater range of returns. Turning the attention to the Sharpe Ratio, over the period both portfolios deliver a higher level of return with reference to the risk-free rate. The ML portfolio provided the investor with a higher risk-adjusted performance during the pre-crisis period – with a peak at (0,24) in Q1 2001 compared to (0,19) from the MZ portfolio – the result is as well maintained during the post-crisis period, with an exception for Q2 2002.

Considering the Paris terroristic attacks – focusing on the selected one that was deemed the deadliest one among those that occurred around Europe, occurred on November 2015 – it can be observed that the pre-event median of realized returns are more compelling for the investor adopting the ML portfolio. This is as well captured by a higher level of risk-adjusted return which is presented by the Sharpe Ratio, with the highest value reached during Q3 2015 with (0,12) for the ML portfolio compared to the (0,09) for the MZ portfolio.

During the quarter when the extreme event took place, the ML portfolio is delivering a higher level of median realized return and as well experienced a high volume range of returns considering the difference between the first and the third quartiles, however, the risk-adjusted return for the investor is lower compared to the MZ one. The Sharpe Ratio is at (0,02) for the ML portfolio and (0,05) for the MZ portfolio, translating to the fact that the investor experiences a better trade-off between exposure and return with the Markowitz allocation.

This is as well observed in the period following the event, where the ML portfolio sees a decrease as well in the level of median realized return, alongside a less attractive risk level compared to the achieved return.

4.2.2 Natural Disasters

In relation to the Natural Disaster extreme events, the Thailand tsunami and the Fukushima nuclear disaster have been considered, and the below period schemes have been identified:

Extreme event	Pre-event period	Event Day	During period	Post-event period
Tsunami Thailand	Q1 2004 - Q3 2004	27/12/2004	Q4 2004	Q1 2005 - Q3 2005
Fukushima Nuclear Disaster	Q2 2010 - Q4 2010	14/03/2011	Q1 2011	Q2 2011 - Q4 2011

Table 8: *Period characterization for the Natural Disasters*

Over the period considered the below results were gathered, summarizing the median, the first and third quartile of the realized returns, and the Sharpe Ratio for each sample period.

Tsunami Thailand

Period	Median MZ	Median ML	Q1 MZ	Q3 MZ	Q1 ML	Q3 ML	Sharpe Ratio MZ	Sharpe Ratio ML
Pre-crisis								
Q1 - 2004	0,03%	0,02%	-0,01	0,07	-0,01%	0,05	0,34	0,22
Q2 - 2004	0,01%	0,03%	-0,01	0,03	-0,02%	0,06	0,21	0,23
Q3 - 2004	0,01%	0,01%	-0,01	0,02	-0,01%	0,03	0,26	0,18
During crisis								
Q4 - 2004	0,01%	0,02%	0,00	0,03	0,00%	0,04	0,15	0,19
Post-crisis								
Q1 - 2005	0,02%	0,04%	-0,01	0,05	-0,03%	0,16	0,17	0,21
Q2 - 2005	0,01%	0,01%	-0,01	0,04	-0,01%	0,03	0,19	0,18
Q3 - 2005	0,02%	0,01%	-0,01	0,06	-0,01%	0,02	0,18	0,13

Fukushima Nuclear Disaster

Period	Median MZ	Median ML	Q1 MZ	Q3 MZ	Q1 ML	Q3 ML	Sharpe Ratio MZ	Sharpe Ratio ML
Pre-crisis								
Q2 - 2010	0,04%	0,03%	-0,03	0,10	-0,02%	0,06	0,14	0,05
Q3 - 2010	0,03%	0,05%	-0,01	0,07	-0,02%	0,09	0,17	0,08
Q4 - 2010	0,02%	0,02%	-0,01	0,07	-0,01%	0,04	0,18	0,12
During crisis								
Q1 - 2011	0,01%	0,03%	0,00	0,03	-0,01%	0,05	0,07	0,09
Post-crisis								
Q2 - 2011	0,01%	0,06%	-0,01	0,03	-0,02%	0,12	0,11	0,23
Q3 - 2011	0,03%	0,04%	-0,02	0,09	-0,03%	0,10	0,10	0,13
Q4 - 2011	0,02%	0,03%	-0,01	0,05	-0,01%	0,07	0,09	0,15

Table 9: *Natural Disasters results*

A great earthquake in the Indian Ocean in December 2004, provoked a tsunami that had catastrophic consequences on the impacted areas, destroying parts of Thailand. Considering the pre-event period, nor portfolios consistently deliver a higher median realized return, higher for the MZ portfolio in Q1 2004 and for the ML portfolio in Q2 2004, and with the same level of median realized return in Q3 2004, the last quarter considered in the pre-event time window. However, the Sharpe Ratio supports the investor in identifying the MZ portfolio as the one delivering a more interesting allocation option, particularly in Q3 2004, with (0,26) compared to the (0,18) achieved by the ML portfolio.

During the crisis and in the first quarter, and throughout the post-event period, it is observed that the trend is reversed. The ML portfolio achieves a higher level of median realized return as well as a stronger risk-adjusted return for the investor – (0,19) during the crisis in Q4 2004 versus the (0,15) from the MZ portfolio and (0,21) in the Q1 2005 compared to the (0,17) Sharpe Ratio reached by the MZ portfolio. Overall, in the post-event period, the trend appears to revert back to its pre-event conditions, with the MZ portfolio providing a more beneficial option for a more risk-averse investor.

The second natural disaster taken into consideration in this study takes place in Japan on March 2011. Conversely to the results observed for the Thailand tsunami, the median of realized returns does not provide significant insights in terms of a clear prevailing strategy during the pre-event period. The Sharpe Ratio value, however, allocates to Markowitz a stronger return-to-risk ratio over all three quarters, from Q2 2010 to Q4 2010.

It is interesting to observe that, during the quarter when the extreme event took place, the ML portfolio delivers a higher level of median realized return, along with a higher level of Sharpe Ratio compared to the MZ one. This was as well identified during the tsunami in Thailand analysis.

However, in contrast to what was observed before, the trend does not convolute back to the MZ portfolio as the one delivering a better result in the period post-crisis, but rather sees the ML portfolio maintain a healthier outlook. In Q2 2011, the ML portfolio achieves a (0,06%) median realized return compared to the (0,01%) delivered by the MZ portfolio. Worth noting as well that in Q2 2011 the ML portfolio delivers the highest level of difference between the first and the third quartile, going from a negative value of (-0,02) to (0,12). In the period post-crisis, the ML portfolio consistently delivers a higher level for the Sharpe Ratio compared to the MZ portfolio, with a peak value at (0,23) compared to (0,11) for the MZ portfolio in the first quarter following the extreme event, Q2 2011.

4.2.3 Political changes

For what concerns the political changes, the below three period phases have been identified in relation to the Brexit referendum and the 2016 U.S.A. Presidential Elections:

Extreme event	Pre-event period	Event Day	During period	Post-event period
Brexit Referendum	Q3 2015 - Q1 2016	24/06/2016	Q2 2016	Q3 2016 - Q1 2017
2016 United States Presidential Elections	Q1 2016 - Q3 2016	08/11/2016	Q4 2016	Q1 2017 - Q3 2017

Table 10: *Period characterization for the Political Changes*

In the table below the results related to the median, the first and third quartile of the realized returns as well as the Sharpe Ratio for the two portfolios are presented.

Brexit Referendum

Period	Median MZ	Median ML	Q1 MZ	Q3 MZ	Q1 ML	Q3 ML	Sharpe Ratio MZ	Sharpe Ratio ML
Pre-crisis								
Q3 - 2015	0,06%	0,07%	-0,01	0,10	-0,02%	0,08	0,28	0,32
Q4 - 2015	0,03%	0,08%	-0,01	0,05	-0,01%	0,10	0,17	0,29
Q1 - 2016	0,02%	0,06%	-0,02	0,06	-0,02%	0,09	0,11	0,18
During crisis								
Q2 - 2016	0,01%	0,02%	-0,01	0,04	-0,01%	0,04	0,10	0,11
Post-crisis								
Q3 - 2016	0,01%	0,03%	-0,01	0,03	0,00%	0,06	0,08	0,12
Q4 - 2016	0,02%	0,04%	-0,01	0,04	-0,01%	0,08	0,04	0,15
Q1 - 2017	0,03%	0,01%	-0,01	0,05	0,00%	0,04	0,07	0,13

2016 United States Presidential Elections

Period	Median MZ	Median ML	Q1 MZ	Q3 MZ	Q1 ML	Q3 ML	Sharpe Ratio MZ	Sharpe Ratio ML
Pre-crisis								
Q1 - 2016	0,03%	0,03%	-0,01	0,07	-0,02%	0,08	0,13	0,08
Q2 - 2016	0,02%	0,08%	-0,02	0,06	-0,02%	0,14	0,16	0,25
Q3 - 2016	0,01%	0,05%	-0,01	0,04	-0,01%	0,07	0,14	0,22
During crisis								
Q4 - 2016	0,01%	0,02%	-0,01	0,05	-0,01%	0,06	0,11	0,13
Post-crisis								
Q1 - 2017	0,01%	0,02%	-0,01	0,03	-0,02%	0,05	0,03	0,06
Q2 - 2017	0,01%	0,01%	0,00	0,02	-0,01%	0,03	0,05	0,07
Q3 - 2017	0,02%	0,02%	-0,01	0,04	-0,03%	0,05	0,08	0,09

Table 11: *Political Changes results*

The first political extreme event considered for this analysis is the Brexit referendum that took place in June 2016. A strong difference compared to the previous events analyzed is related to the fact that the occurrence of the event was not unknown, rather its outcome was rather unexpected and further fueled the uncertainty in the stock market. In the pre-event period, the results portray the ML portfolio as the one delivering a higher level of median realized return compared to the MZ one, particularly significant in Q4 2015 where the ML portfolio attains a (0,08%) value against the (0,03%) provided by the MZ one. The Sharpe Ratio as well recognizes that the allocation with the Machine Leaning strategy is factually more significant than the one achieved by the Markowitz one, with a spread between the values of (0,12) – the highest value registered for the event-specific overview.

During the crisis as well as in its aftermath period, the results do not variate drastically, with the ML portfolio maintaining a stronger level of risk-adjusted return for the end investor.

The last extreme event considered in this analysis is the 2016 U.S.A. Presidential elections, seeing Donald Trump as the winning contestant. It is observed that the results identified for the previous extreme events are somewhat confirmed. Particularly, during the pre-event period the median realized return for the ML portfolio was higher or equivalent to the MZ portfolio – Q1 2016 both portfolios deliver a (0,03%) median realized return – however, the Sharpe Ratio for Q1 allocates a higher level of achieved adjusted return to the MZ portfolio with a value of (0,13) versus the (0,08) delivered by the ML portfolio. However, in the following quarter – Q2 2016 – the ML portfolio delivers to the investor a value of (0,25) of Sharpe Ratio against the (0,16) of the MZ portfolio.

During the crisis period as well as in the post-end period, the Machine Learning portfolio is consistently delivering a higher return-adjusted value, inferred from the Sharpe Ratio – of which highest level for the during-event and post-event period is achieved in Q4 2016 with a value of (0,13) compared to the (0,11) delivered by the MZ portfolio. The median realized returns are not drastic in terms of value or evolution, in both periods reviewed.

5. Conclusions

The final discussion will bring together all the points raised along this thesis, reconnecting to the research questions and objectives stated in *Chapter 1*.

The discussion will be summarized in sections to allow for clarity. This chapter will be then closed with the conclusion and final argumentations.

The first section will focus on the pure analysis of the two portfolios set in comparison against each other. Among the aims of this thesis resided the ignition of a forthright analysis of viable Machine Learning techniques that can be adopted to build a portfolio and compare the latter against a naïve portfolio allocation set-up, this being the Markowitz portfolio allocation. This relates to the first research question: *Does the Machine Learning constructed portfolio deliver stronger performance for the investor in comparison to the Markowitz constructed portfolio?*

Four different Machine Learning algorithms, together with a blend algorithm combining the models, have been employed in order to obtain the expected returns forecasts for a about two-year period, from January 2021 to September 2022, trained on a ten-year period data set, from January 2010 to December 2020, the obtained final predictions have been leveraged to obtain the optimized Machine Learning portfolio weights, by maximizing the Sharpe Ratio.

The results obtained indicate that over the rolling period considered, the Machine Learning portfolio is able to achieve a return-to-risk ratio that is higher than the one that can be attained through Markowitz's allocation. This research thus validates the Machine Leaning allocation as being more compelling to investors, delivering a more valuable performance, conditional to investors' investment objectives and risk tolerance.

Further, the stock allocation identified for the Machine Learning portfolio construction sees a higher level of diversification in comparison to the Markowitz portfolio. D'Acunto, Prabhala et al. (2018) concluded on their research that "*Underdiversified investors increased their portfolio diversification in terms of both the number of stocks they held".* The results are as well in line with the consensus in the literature that identifies in Machine Learning an opportunity to revamp the classical financial allocation strategies which could deliver promising results in-sample which are not necessarily true in the out-sample sphere. Perrin et al. (2019) acknowledge that "*A lot of academics and professionals have proposed and alternative approach to the MVO framework, but very few of these models are used in practice. The main reason is that these competing models focus on the objective function and not on the numerical implementation*". This translates to the fact that theoretical successful models end up finding no fertile terrain among end users – asset managers for the scope of this thesis – if the computational and methodological application is not sufficiently accessible and practical.

The drawn conclusions from the analysis allow the first research question to be answered by validating that indeed the Machine Learning constructed portfolio delivers a stronger performance for the investor compared to the Markowitz constructed portfolio, both in terms of diversification and risk-adjusted return.

Moving on to the second section, the following analysis will answer the research question: *Within distressed market conditions, where does the Machine Learning constructed portfolio position itself against the Markowitz constructed portfolio?*

The events taken into consideration are clustered into three categories based on their nature: terrorist attacks, natural disasters, and political changes. The terrorist attacks category includes the 09/11 terroristic attack and the Paris terrorist attacks. The natural disasters classification comprises the Tsunami in Thailand and the Fukushima nuclear disaster. The political changes consider the outcome of the Brexit referendum and the Presidential election of Donald Trump.

The analysis of the terroristic attacks displays varying results. It can be observed that considering the 09/11 terroristic attack over the quasi-two-year period under review the Machine Learning portfolio could attain in a more consistent way results which were more significant for the investor in terms of risk-adjusted returns. The trend seems to deviate during the event data frame, where the Markowitz portfolio could achieve a more positive value for the reference investor. The Paris attack showed initially a similar outcome in terms of winning strategy, for the period pre-event, which however switches to the Markowitz portfolio for the period during the event and in its aftermath. Reviewing both events, it was not possible to identify a clear and strong predominance in terms of strategy.

From the natural disaster events analyzed, the following assertions can be extrapolated: for both the tsunami in Thailand and the Fukushima nuclear disaster, the initial strategy which is preferred by the investor is the Markowitz portfolio, during the pre-event period and post-event. In the time frame during the event, it is observed that for both scenarios the Machine Learning portfolio is capable of delivering a stronger level of return conditional to the level of exposure. The observation is sharper, particularly for the tsunami in Thailand event, where the Machine Learning portfolio remains a preferable option for the investor as well in the first quarter of the post-event period.

Finally, taking into account the political changes considered for this study, the two events that occurred in 2016 are analyzed. These two specific events – the Brexit referendum and the United States presidential election – distinguish themselves from the ones priorly discussed as they lack the unforeseeable feature. For the two events observed the Machine Learning strategy is the one delivering a stronger return level for the risk undertaken within all periods considered. In all the observed periods the Markowitz portfolio does not achieve the goal of providing a sufficiently robust allocation within the contextualized extreme events.

Going back to the second research question, the Machine Learning portfolio does not achieve a strong outperformance compared to the Markowitz portfolio within distressed market conditions. Depending on the nature of the considered extreme event, the winning strategy switches from the Machine Learning portfolio to the Markowitz portfolio, particularly experienced for terroristic attacks and natural disasters categories. A stronger footprint is however observed for the category of political changes, where the Machine Learning method achieves more significant results for the investor.

Having gathered all the points raised up to now, the Machine Learning algorithms have displayed the potential to effectively operate in the out-of-sample and to outperform the classical portfolio allocation proposed by Markowitz. The past years have undoubtfully proven that the fast-pacing development that the financial industry is experiencing in terms of Machine Learning application methods will drastically change the asset allocation market and will pave the way to fully revolutionize portfolio investment choices for investors.

6. Limitation and further research

The objective of this Master's thesis is to provide a proof of concept in relation to the application of Machine Learning techniques within the context of asset allocation categorized within the framework of market shocks related to extreme events originating outside the financial sphere. In this section, the limitations and propositions for future research will be laid forward.

6.1 Limitations

As with many pieces of research, this study does not fall short of a number of limitations being considered.

The following have been recognized as the major drawbacks:

- No transaction costs, no short selling, and available capital being fully invested;
- The population of assets selected for the construction of the portfolios is not well diversified in terms of sectors considered;
- The optimized portfolios display an overweight towards U.S.A. stock market, introducing a bias when focusing the analysis towards a more global outlook;
- The market universe considered does not offer the possibility to be compared to a relevant benchmark including the basket of stocks being considered;
- The extreme event definition and selection of events is subjective to the author's opinion and therefore it may vary from other studies considering different events and contrasting categorizations of events.

6.2 Future research

The footprint of Artificial Intelligence and Machine Learning is undeniable in all sectors, including the one of portfolio allocation. The data pool selected could be extended to a larger number of common stocks, as well as include different types of financial instruments. This could offer the opportunity to further investigate and assess the benefits that the investors could attain through Machine Learning applications while being able to compare the performance against a relevant market index.

Additionally, it would be interesting to run more ad hoc analysis in relation to extreme events, either to cover a longer period or conduct the review focused to a specific industry or company, this would allow reducing for volatility generated from external sources. Indeed, from a general economic point of view, some conclusions could be biased.

7. Reference

Akansu, A. N., Malioutov, D., Palomar, D. P., Jay, E., & Mandic, D. P. (2016a). Introduction to the Issue on Financial Signal Processing and Machine Learning for Electronic Trading. IEEE Journal of Selected Topics in Signal Processing, 10(6), 979–981. https://doi.org/10.1109/jstsp.2016.2594458

Aparicio, F. M., & Estrada, J. (2001). Empirical distributions of stock returns: European securities markets, 1990–95. The European Journal of Finance, 7(1), 1–21. https://doi.org/10.1080/13518470121786

Arifovic, J., He, X. Z., & Wei, L. (2022). Machine learning and speed in high-frequency trading. Journal of Economic Dynamics and Control, 139, 104438. https://doi.org/10.1016/j.jedc.2022.104438

Bishop, C. M. (2006). Pattern Recognition and Machine Learning (Information Science and Statistics). Springer.

Bishop, C. M. (1994). Neural networks for pattern recognition. Choice Reviews Online, 31(10), 31–5500. https://doi.org/10.5860/choice.31-5500

Breiman, L. (2001). Random Forests. Random Forests, 45(1), 5–32. https://doi.org/10.1023/a:1010933404324

Broska, L. H., Poganietz, W. R., & Vögele, S. (2020). Extreme events defined—A conceptual discussion applying a complex systems approach. Futures, 115, 102490. https://doi.org/10.1016/j.futures.2019.102490

Brown, S. J., & Warner, J. B. (1980). Measuring security price performance. Journal of Financial Economics, 8(3), 205–258. https://doi.org/10.1016/0304-405x(80)90002-1

Chaiechi, T. (2020). Economic Effects of Natural Disasters: Theoretical Foundations, Methods, and Tools (1st ed.). Academic Press.

Chiu, W. Y. (2021). Safety-first portfolio selection. Mathematics and Financial Economics, 15(3), 657–674. https://doi.org/10.1007/s11579-021-00292-3

Choueifaty, Y. (2008). Towards maximum diversification. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4063676

Clark, B., Feinstein, Z., & Simaan, M. (2020). A machine learning efficient frontier. Operations Research Letters, 48(5), 630–634. https://doi.org/10.1016/j.orl.2020.07.016

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1), 21–27. https://doi.org/10.1109/tit.1967.1053964

Dan Ryan. (2016, December 7). Ten Key Implications of Donald Trump's Electoral Victory for Financial and Securities Regulation. The Harvard Law School Forum on Corporate Governance. https://corpgov.law.harvard.edu/2016/12/07/ten-key-implications-of-donald-trumps-electoral-victory-for-financial-and-securities-regulation/

D'Acunto, F., Prabhala, N., & Rossi, A. G. (2019). The Promises and Pitfalls of Robo-Advising. The Review of Financial Studies, 32(5), 1983–2020. https://doi.org/10.1093/rfs/hhz014

DeMiguel, V., Garlappi, L., & Uppal, R. (2007). Optimal Versus Naive Diversification: How Inefficient is the 1/NPortfolio Strategy? Review of Financial Studies, 22(5), 1915–1953. https://doi.org/10.1093/rfs/hhm075

D'Hondt, C., de Winne, R., Ghysels, E., & Raymond, S. (2020). Artificial Intelligence Alter Egos: Who might benefit from robo-investing? Journal of Empirical Finance, 59, 278–299. https://doi.org/10.1016/j.jempfin.2020.10.002

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. Journal of Business & Economic Statistics, 13(3), 253-263.

Estrada, J. (2007). Mean-semivariance behavior: Downside risk and capital asset pricing. International Review of Economics & Finance, 16(2), 169–185. https://doi.org/10.1016/j.iref.2005.03.003

Fabozzi, F. J., Huang, D., & Zhou, G. (2009). Robust portfolios: contributions from operations research and finance. Annals of Operations Research, 176(1), 191–220. https://doi.org/10.1007/s10479-009-0515-6

Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. The Journal of Finance, 47(2), 427–465. https://doi.org/10.1111/j.1540-6261.1992.tb04398.x

Fama, E. F., Fisher, L., Jensen, M. C., & Roll, R. (1969). The Adjustment of Stock Prices to New Information. International Economic Review, 10(1), 1. https://doi.org/10.2307/2525569

Friedman, J., Hastie, T., & Tibshirani, R. (2000). Additive logistic regression: a statistical view of boosting (With discussion and a rejoinder by the authors). The Annals of Statistics, 28(2). https://doi.org/10.1214/aos/1016218223

FRED Economic Data. (2022). 3-Month Treasury Bill Secondary Market Rate. https://fred.stlouisfed.org/series/TB3MS#0

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning (Adaptive Computation and Machine Learning series) (Illustrated ed.). The MIT Press.

Grand, B., Boddaert, G., Louis Daban, J., Hornez, E., de Carbonnieres, A., Giral, G., Ngabou, D., Mlynski, A., Gonzalez, F., Mcbride, T., & Bonnet, S. (2017). Paris Terrorist Attack on November 13, 2015 - Applying Wartime In-hospital Triage and Damage Control Strategies. Prehospital and Disaster Medicine, 32(S1), S119. https://doi.org/10.1017/s1049023x17003399

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. The Review of Financial Studies, 33(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009

Jiang, Y. (2022). A Primer on Machine Learning Methods for Credit Rating Modeling. In (Ed.), Econometrics - Recent Advances and Applications. IntechOpen. https://doi.org/10.5772/intechopen.107317

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. Business Horizons, 62(1), 15–25. https://doi.org/10.1016/j.bushor.2018.08.004

Kaplan, J. (2016). Artificial Intelligence: What Everyone Needs to Know R (1st ed.). Oxford University Press.

Kavlakoglu, E. (2020). AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the Difference? IBM Cloud.

Kierzenkowski, R. (2016). The Economic Consequences of Brexit: A Taxing Decision. OECD Economic Policy Papers, 16.

Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. Published as a Conference Paper at ICLR 2015.

Kumaraswamy, H. D. (2021). Artificial Intelligence in Data Mining. MIT Management.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539

Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278–2324. https://doi.org/10.1109/5.726791

López de Prado, M. (2016). Building Diversified Portfolios that Outperform Out of Sample. The Journal of Portfolio Management, 42(4), 59–69. https://doi.org/10.3905/jpm.2016.42.4.059

López de Prado, M. (2018). Advances in Financial Machine Learning (1st ed.). Wiley.

López de Prado, M. (2020). Machine Learning for Asset Managers. Cambridge University Press.

Kan, R., Wang, X., & Zheng, X. (2019). In-Sample and Out-of-Sample Sharpe Ratios of Multi-Factor Asset Pricing Models. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3454628

Makinen, G. (2002). The Economic Effects of 9/11: A Retrospective Assessment. Congressional Research Service.

Makridakis, S., & Taleb, N. (2009). Living in a world of low levels of predictability. International Journal of Forecasting, 25(4), 840–844. https://doi.org/10.1016/j.ijforecast.2009.05.008

Makridakis, S., & Taleb, N. (2009). Decision making and planning under low levels of predictability. International Journal of Forecasting, 25(4), 716–733. https://doi.org/10.1016/j.ijforecast.2009.05.013

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. PLOS ONE, 13(3), e0194889. https://doi.org/10.1371/journal.pone.0194889

Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. The Journal of Business, 36(4), 394. https://doi.org/10.1086/294632

Markowitz, H. (1952). Portfolio Selection. The Journal of Finance, 7(1), 77. https://doi.org/10.2307/2975974

Markowitz, H. M. (1999). The Early History of Portfolio Theory: 1600–1960. Financial Analysts Journal, 55(4), 5–16. https://doi.org/10.2469/faj.v55.n4.2281

Markowitz, H. (1956). The optimization of a quadratic function subject to linear constraints. Naval Research Logistics Quarterly, 3(1–2), 111–133. https://doi.org/10.1002/nav.3800030110

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. https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662928/IPOL_STU(2021)662928_EN.pdf

McCargo, D., & Pathmanand, U. (2006). The Thaksinasation of Thailand. The Journal of Asian Studies, 65(3), 658–660.

Michaud, R. O. (1989). The Markowitz Optimization Enigma: Is 'Optimized' Optimal? Financial Analysts Journal, 45(1), 31–42. https://doi.org/10.2469/faj.v45.n1.31

Monthly Economic Report Executive Summary (July 2011) - Cabinet Office Home Page. (2011). (C) Cabinet Office, Government of Japan. https://www5.cao.go.jp/keizai3/getsurei-e/2011aug.html

Moscatelli, M., Parlapiano, F., Narizzano, S., & Viggiano, G. (2020). Corporate default forecasting with machine learning. Expert Systems with Applications, 161, 113567. https://doi.org/10.1016/j.eswa.2020.113567

Mullainathan, S., & Spiess, J. (2017). Machine Learning: An Applied Econometric Approach. Journal of Economic Perspectives, 31(2), 87–106. https://doi.org/10.1257/jep.31.2.87

Nordvall, A., & Heldt, T. (2017). Understanding hallmark event failure: a case study of a Swedish music festival. International Journal of Event and Festival Management, 8(2), 172–185. https://doi.org/10.1108/ijefm-11-2015-0043

Norio, O., Ye, T., Kajitani, Y., Shi, P., & Tatano, H. (2011). The 2011 eastern Japan great earthquake disaster: Overview and comments. International Journal of Disaster Risk Science, 2(1), 34–42. https://doi.org/10.1007/s13753-011-0004-9

Perrin, S., & Roncalli, T. (2019). Machine Learning Optimization Algorithms & Portfolio Allocation. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3425827

Porter, M. D., & White, G. (2012). Self-exciting hurdle models for terrorist activity. The Annals of Applied Statistics, 6(1). https://doi.org/10.1214/11-aoas513

Python. (2022). Python.Org. https://www.python.org/

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. Psychological Review, 65(6), 386–408. https://doi.org/10.1037/h0042519

Rossi, A. G., & Utkus, S. P. (2020). Who Benefits from Robo-advising? Evidence from Machine Learning. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3552671

Routledge, B. R. (2019). Machine learning and asset allocation. Financial Management, 48(4), 1069–1094. https://doi.org/10.1111/fima.12303

Rubinstein, M. (2002). Markowitz's "Portfolio Selection": A Fifty-Year Retrospective. The Journal of Finance, 57(3), 1041–1045. https://doi.org/10.1111/1540-6261.00453

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088), 533–536. https://doi.org/10.1038/323533a0

Ruppert D., & Matteson D.S. (2015) Statistics and Data Analysis for Financial Engineering: with R examples (Springer Texts in Statistics) (2023). Springer; 2nd ed. 2015 edition (2015-05-31).

Rybinski, K. (2020). Should asset managers pay for economic research? A machine learning evaluation. The Journal of Finance and Data Science, 6, 31–48. https://doi.org/10.1016/j.jfds.2020.08.001

Samitas, A., Kampouris, E., & Kenourgios, D. (2020). Machine learning as an early warning system to predict financial crisis. International Review of Financial Analysis, 71, 101507. https://doi.org/10.1016/j.irfa.2020.101507

Sharpe, W. F. (1975). Adjusting for Risk in Portfolio Performance Measurement. The Journal of Portfolio Management, 1(2), 29–34. https://doi.org/10.3905/jpm.1975.408513

Sharpe, W. F. (1994). The Sharpe Ratio. The Journal of Portfolio Management, 21(1), 49–58. https://doi.org/10.3905/jpm.1994.409501

Stuart, A., & Markowitz, H. M. (1959). Portfolio Selection: Efficient Diversification of Investments. OR, 10(4), 253. https://doi.org/10.2307/3006625

Sutton, R., & Barto, A. (1998). Reinforcement Learning: An Introduction. IEEE Transactions on Neural Networks, 9(5), 1054. https://doi.org/10.1109/tnn.1998.712192

Telford, J., & Cosgrave, J. (2007). The international humanitarian system and the 2004 Indian Ocean earthquake and tsunamis. Disasters, 31(1), 1–28. https://doi.org/10.1111/j.1467-7717.2007.00337.x

Wilder, J. W. (1978). New Concepts in Technical Trading Systems. Trend Research.

Wright, F. A., & Wright, A. A. (2018). How surprising was Trump's victory? Evaluations of the 2016 U.S. presidential election and a new poll aggregation model. Electoral Studies, 54, 81–89. https://doi.org/10.1016/j.electstud.2018.05.001

8. List of Figures

9. Appendix

9.1 Appendix 1

In the below table the population of common stocks composing the market universe considered to perform the analysis are presented, along with the ticker and sector details. The population of equities has been identified from the MSCI World Index, based on the 50 stocks with the highest market capitalization.

#	COMPANY	TICKER (NYSE)	SECTOR
1	APPLE INC.	AAPL	Technology
2	MICROSOFT CORPORATION	MSFT	Technology
3	AMAZON.COM INC.	AMZN	Consumer Cyclical
4	ALPHABET INC.	GOOGL	Communication Services
5	TESLA INC.	TSLA	Consumer Cyclical
6	JOHNSON & JOHNSON	JNJ	Healthcare
7	UNITEDHEALTH GROUP INC.	UNH	Healthcare
8	NVIDIA CORPORATION	NVDA	Technology
9	META PLATFORMS, INC.	META	Communication Services
10	EXXON MOBIL CORP	XOM	Energy
11	JPMORGAN CHASE & CO	JPM	Financial Services
12	VISA	V	Financial Services
13	CHEVRON CORP	CVX	Energy
14	NESTLE	NSRGY	Consumer Defensive
15	MASTERCARD	MA	Financial Services
16	HOME DEPOT	HD	Consumer Cyclical
17	PFIZER	PFE	Healthcare
18	BANK OF AMERICA CORP	BAC	Financial Services
19	COCA COLA (THE)	KO	Consumer Defensive
20	ABBVIE	ABBV	Healthcare
21	LILLY (ELI) & COMPANY	LLY	Healthcare
22	ROCHE HOLDING A.G.	RHHBY	Healthcare
23	BROADCOM	AVGO	Technology
24	MERCK & CO	MRK	Healthcare
25	PEPSICO	PEP	Consumer Defensive
26	ASML HOLDING N.V.	ASML	Technology
27	SHELL PLC	SHEL	Energy
28	THERMO FISHER SCIENTIFIC	TMO	Healthcare
29	ABBOTT LABORATORIES	ABT	Healthcare
30	COSTCO WHOLESALE CORP	COST	Consumer Defensive
31	ASTRAZENECA PLC	AZN	Healthcare
32	COMCAST CORPORATION	CMCSA	Communication Services
33	ADOBE	ADBE	Technology
34	NOVARTIS AG	NVS	Healthcare
35	WALMART	WMT	Consumer Defensive
36	ACCENTURE	ACN	Technology
37	CISCO SYSTEMS	CSCO	Technology
38	NOVO NORDISK	NVO	Healthcare
39	MCDONALD'S CORPORATION	MCD	Consumer Cyclical
40	INTEL CORP	INTC	Technology

41	TOYOTA MOTOR CORPORATION	TM	Consumer Cyclical
42	DANAHER CORP	DHR	Healthcare
43	LVMH MOET HENNESSY	LVMHF	Consumer Cyclical
44	SALESFORCE	CRM	Technology
45	WELLS FARGO & CO	WFC	Financial Services
46	BHP GROUP LIMITED	BHP	Basic Materials
47	ADVANCED MICRO DEVICES	AMD	Technology
48	LINDE PLC	LIN	Basic Materials
49	BRISTOL-MYERS SQUIBB COMPANY	BMY	Healthcare
50	TEXAS INSTRUMENTS	TXN	Technology

9.2 Appendix 2

The below provides a print screen from the Python coding performed in order to segregate the data set in the training sample – from January 2010 to December 2020 – and the validation set – from January 2021 to September 2022.

```
## Our model will be trained and tested on data from 2010 to 2020
## A second validation will be performed on 2020-2022 data
## We keep only our estimators + prediction

feature_names.append('21d_close_future_pct')

df = stock_data.loc['2010-01-01':'2020-12-31']
df = df[feature_names].dropna()

validation_set = stock_data.loc['2021-01-01':'2022-09-30']

# Right before we train our model we must split up the data into a train set and test set.
# However, due to the nature of time-series', we need to handle this part carefully. If we randomize our train-test set,
# we could encounter a look-ahead bias which is not good for predicting the stock market.
#  It's caused when you train your model on data it would've already seen.

# To prevent this we are going to be training the model using a different technique called cross-validation.
## We have 2769 data points
## We want cross validation for 5 segments, so each one will have 750 data points

print("Training models for ticker {} ({}/{})".format(ticker, tickers.index(ticker) + 1, len(tickers)))
ensemble_RESULTS = trainWithCrossValidation(df)
results[ticker] = ensemble_RESULTS
```

9.3 Appendix 3

Below the set-up of the parameters is detailed. The parameter tuning enriches and provides further signaling details to the prediction models. The seven parameters included are the Relative Strength Index on 21, 63 and 251 trading days, the Momentum over 21, 63 and 251 trading days and the volume percentage change.

```
for ticker in tickers:
    stock_data = pd.read_csv("prices/{}.csv".format(ticker))
    stock_data = stock_data.set_index('Date')
    stock_data.index = pd.to_datetime(stock_data.index, utc=True)
    stock_data.index.name = "Date"

    # Compute monthly returns
    stock_data['r21'] = stock_data['Adj Close'].pct_change(21)

    ## Get the MOM and RSI for our stock prices in 21, 63, 251 days. (1 month, 1 quarter, 1 year)
    feature_names = []
    for n in [21, 63, 251]:
     # type: ignore
        stock_data['rsi' + str(n)] = talib.RSI(stock_data['Adj Close'].values, timeperiod=n)
        stock_data['mom' + str(n)] = talib.MOM(stock_data['Adj Close'].values, timeperiod=n)
        feature_names = feature_names + ['rsi' + str(n), 'mom' + str(n)]


    stock_data['Volume'].replace(to_replace = 0, value = 1, inplace=True)
    stock_data['Volume_1d_change'] = stock_data['Volume'].pct_change()
    volume_features = ['Volume_1d_change']
    feature_names.extend(volume_features)
```

9.4 Appendix 4

The cross-validation performed on the model takes place within the training data set, 70% of the data points are devoted to training the algorithm and 30% are devoted to testing and validating the figures. The cross-validation is performed for all four models plus the ensemble which is combining the four.

```
# Split data into equal partitions of size len_train

## 1st iteration

    ## 2010-2013 -> 70% train 30% test
    ## 2012-2015 -> 70% train 30% test
    ## and so on until 2020

num_train = 500 # Increment of how many starting points (len(data) / num_train  =  number of train-test sets)
len_train = 750 # Length of each train-test set (~250 trading days per year, we take at least 3 years)

i = 0

## Final model results for this data
## Will be returned by this func
result = {}

while True:

    # Partition the data into chunks of size len_train every num_train days
    # i = 0
        # 0:750
    # i = 1
        # 500 : 1250
        # 1000 : 1750
        # 1500 : 2250
        # 2000 : 2750

    df = data.iloc[i * num_train : (i * num_train) + len_train]
    i += 1

    ## Interrupt when dataset does not have anymore a subset with size
    ## greater than our batch size (750)
    if len(df) < len_train:
        break

    ## Define target variable
    y = df['target']
    ## Define features and feature set
    features = [x for x in feature_names if x not in ['target']]
    X = df[features]

    ## Split train and test
    X_train, X_test, y_train, y_test = train_test_split(X, y, train_size= 7 * len(X) // 10,shuffle=False)

    # fit models
    rf.fit(X_train, y_train)
    knn.fit(X_train, y_train)
    gbt.fit(X_train, y_train)
    mlp.fit(X_train, y_train)

    ## 5th model which includes the other 4
    ## At the end of the day we can either chose this or one of the other 4
    ensemble.fit(X_train, y_train)

    ## Extract feature importance from RandomForest
    importances = rf.best_estimator_.feature_importances_
    forest_importances = pd.Series(importances, index=features)
    start_date = X.index.min().strftime("%m/%d/%y")
    end_date = X.index.max().strftime("%m/%d/%y")

    ## Append them to results
    result["{}_{}".format(start_date, end_date)] = forest_importances
```

9.5 Appendix 5

The four models used to obtain the final forecasted values are below listed. This includes Random Forest (RF), Gradient Boost regressor (GB), K-Nearest Neighbors (KNN) and the Artificial Neural Network – Multilayer Perceptron network (MLP). Finally, the ensemble algorithm is obtained averaging the results from the different models.

```
# Models which will be used
# 1. Random Forest
random_forest_params = {
        "n_estimators" : [10,50,100],
        "max_features" : [1.0, "log2", "sqrt"],
        "bootstrap"    : [True, False],
        "max_leaf_nodes" : [2,5,10,100]
    }

rf = GridSearchCV(estimator=RandomForestRegressor(n_jobs=-1), param_grid=random_forest_params, scoring="neg_mean_squared_error")

# 2. Gradient Boost
gradient_boost_params = {
    "n_estimators" : [10, 100, 500],
    "learning_rate" : [0.001, 0.1, 1.0],
    "subsample" : [0.5, 1.0],
    "max_depth" : [3, 7],
}

gbt = GridSearchCV(estimator=GradientBoostingRegressor(), param_grid=gradient_boost_params, scoring="neg_mean_squared_error")

# 3. KNN
knn_params = {
    "n_neighbors" : [1,5,13,21],
}
knn = GridSearchCV(estimator=KNeighborsRegressor(n_jobs=-1), param_grid=knn_params, scoring="neg_mean_squared_error")

# 4. MLPRegressor
mlp_params = {
    "hidden_layer_sizes": [(1,),(50,)],
    "activation": ["identity", "logistic", "tanh", "relu"],
    "solver": ["lbfgs", "sgd", "adam"],
    "alpha": [0.00005,0.0005]
}

mlp = GridSearchCV(estimator=MLPRegressor(max_iter=1000), param_grid=mlp_params, scoring="neg_mean_squared_error")

# Create a tuple list of our models
estimators=[('knn', knn), ('rf', rf), ('gb', gbt), ('mlp', mlp)]
ensemble = VotingRegressor(estimators)
```

9.6 Appendix 6

In the below tables, the optimized weights of the Markowitz portfolio and the Machine Learning portfolio are presented at each quarterly rebalancing. It can be observed that a higher level is achieved by investors which adopt Machine Learning.

ML Portfolio	Q1 2021	Q2 2021	Q3 2021	Q4 2021	Q1 2022	Q2 2022	Q3 2022
AAPL	0,00%	9,78%	0,00%	0,00%	0,00%	0,99%	0,00%
MSFT	26,36%	28,78%	29,65%	54,34%	0,00%	4,84%	10,16%
AMZN	0,00%	0,00%	0,00%	3,15%	0,00%	0,00%	0,00%
GOOGL	0,00%	0,00%	0,00%	1,43%	0,00%	0,00%	0,22%
JNJ	0,00%	0,00%	0,00%	0,00%	0,00%	0,02%	0,00%
NVDA	0,04%	0,00%	0,00%	6,83%	0,00%	0,00%	0,00%
META	0,00%	0,00%	0,00%	0,18%	0,00%	3,23%	4,19%
XOM	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,77%
JPM	8,75%	0,00%	0,00%	0,00%	0,00%	1,99%	0,58%
V	0,00%	0,00%	3,39%	1,09%	0,00%	27,06%	19,46%
CVX	0,00%	0,00%	0,00%	0,00%	0,00%	0,02%	0,39%
NSRGY	0,00%	0,00%	0,00%	0,12%	0,00%	0,72%	0,00%
MA	4,87%	0,00%	5,82%	1,88%	19,70%	8,48%	5,38%
HD	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	1,47%
PFE	0,00%	0,00%	0,00%	1,20%	0,00%	0,00%	0,00%
KO	0,00%	0,00%	0,00%	0,00%	0,00%	0,41%	5,24%
ABBV	0,00%	0,00%	0,00%	1,30%	0,00%	0,00%	0,00%
LLY	0,00%	0,00%	0,00%	0,71%	0,00%	0,00%	0,46%
RHHBY	0,00%	0,00%	2,39%	0,82%	0,00%	0,00%	0,20%
MRK	2,56%	7,89%	24,97%	0,38%	0,00%	0,00%	1,26%
PEP	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	1,31%
ASML	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,89%
SHEL	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	1,49%
TMO	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	5,52%
ABT	0,00%	0,00%	17,52%	2,81%	0,00%	0,91%	18,90%
COST	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	1,49%
AZN	55,63%	2,34%	0,00%	0,00%	0,00%	0,00%	0,07%
CMCSA	0,00%	0,00%	0,00%	2,07%	0,00%	0,20%	3,35%
ADBE	0,00%	0,00%	0,00%	0,00%	9,42%	7,68%	1,46%
NVS	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,29%
WMT	0,14%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
ACN	0,00%	0,00%	0,00%	0,00%	0,00%	0,45%	0,23%
MCD	0,00%	0,00%	0,00%	2,65%	0,00%	5,60%	2,76%
INTC	0,00%	2,51%	15,32%	3,14%	0,00%	11,86%	3,21%
TM	0,00%	0,00%	0,00%	1,73%	0,00%	0,08%	0,00%
DHR	0,00%	0,00%	0,00%	8,87%	0,00%	0,00%	1,38%
LVMHF	0,00%	5,67%	0,00%	0,00%	0,00%	15,06%	0,49%
CRM	0,00%	0,67%	0,00%	0,00%	60,58%	6,87%	0,00%
WFC	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	1,94%
BHP	0,00%	0,00%	0,00%	0,00%	5,98%	0,00%	2,04%
AMD	1,66%	42,36%	0,94%	0,57%	4,33%	0,00%	0,00%
LIN	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	3,40%
BMY	0,00%	0,00%	0,00%	3,97%	0,00%	1,23%	0,00%
TXN	0,00%	0,00%	0,00%	0,77%	0,00%	2,33%	0,00%

MZ Portfolio	Q1 2021	Q2 2021	Q3 2021	Q4 2021	Q1 2022	Q2 2022	Q3 2022
AAPL	0,00%	9,78%	0,00%	0,00%	0,00%	0,99%	0,00%
MSFT	26,36%	28,78%	29,65%	54,34%	0,00%	4,84%	10,16%
AMZN	0,00%	0,00%	0,00%	3,15%	0,00%	0,00%	0,00%
GOOGL	0,00%	0,00%	0,00%	1,43%	0,00%	0,00%	0,22%
JNJ	0,00%	0,00%	0,00%	0,00%	0,00%	0,02%	0,00%
NVDA	0,04%	0,00%	0,00%	6,83%	0,00%	0,00%	0,00%
META	0,00%	0,00%	0,00%	0,18%	0,00%	3,23%	4,19%
XOM	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,77%
JPM	8,75%	0,00%	0,00%	0,00%	0,00%	1,99%	0,58%
V	0,00%	0,00%	3,39%	1,09%	0,00%	27,06%	19,46%
CVX	0,00%	0,00%	0,00%	0,00%	0,00%	0,02%	0,39%
NSRGY	0,00%	0,00%	0,00%	0,12%	0,00%	0,72%	0,00%
MA	4,87%	0,00%	5,82%	1,88%	19,70%	8,48%	5,38%
HD	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	1,47%
PFE	0,00%	0,00%	0,00%	1,20%	0,00%	0,00%	0,00%
KO	0,00%	0,00%	0,00%	0,00%	0,00%	0,41%	5,24%
ABBV	0,00%	0,00%	0,00%	1,30%	0,00%	0,00%	0,00%
LLY	0,00%	0,00%	0,00%	0,71%	0,00%	0,00%	0,46%
RHHBY	0,00%	0,00%	2,39%	0,82%	0,00%	0,00%	0,20%
MRK	2,56%	7,89%	24,97%	0,38%	0,00%	0,00%	1,26%
PEP	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	1,31%
ASML	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,89%
SHEL	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	1,49%
TMO	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	5,52%
ABT	0,00%	0,00%	17,52%	2,81%	0,00%	0,91%	18,90%
COST	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	1,49%
AZN	55,63%	2,34%	0,00%	0,00%	0,00%	0,00%	0,07%
CMCSA	0,00%	0,00%	0,00%	2,07%	0,00%	0,20%	3,35%
ADBE	0,00%	0,00%	0,00%	0,00%	9,42%	7,68%	1,46%
NVS	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,29%
WMT	0,14%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
ACN	0,00%	0,00%	0,00%	0,00%	0,00%	0,45%	0,23%
MCD	0,00%	0,00%	0,00%	2,65%	0,00%	5,60%	2,76%
INTC	0,00%	2,51%	15,32%	3,14%	0,00%	11,86%	3,21%
TM	0,00%	0,00%	0,00%	1,73%	0,00%	0,08%	0,00%
DHR	0,00%	0,00%	0,00%	8,87%	0,00%	0,00%	1,38%
LVMHF	0,00%	5,67%	0,00%	0,00%	0,00%	15,06%	0,49%
CRM	0,00%	0,67%	0,00%	0,00%	60,58%	6,87%	0,00%
WFC	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	1,94%
BHP	0,00%	0,00%	0,00%	0,00%	5,98%	0,00%	2,04%
AMD	1,66%	42,36%	0,94%	0,57%	4,33%	0,00%	0,00%
LIN	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	3,40%
BMY	0,00%	0,00%	0,00%	3,97%	0,00%	1,23%	0,00%
TXN	0,00%	0,00%	0,00%	0,77%	0,00%	2,33%	0,00%

Executive Summary

The objective of this Master thesis is to provide a proof of concept in relation to Asset Allocation performance for two different portfolios, one constructed through Machine Learning algorithms and the other applying Markowitz's theory. The two portfolios obtained are then further assessed in the context of distressed market conditions originating from external extreme events.

The portfolios analyzed are constructed using a pool of stocks listed in the NYSE selected from the MSCI World index based on their market capitalization covering an overall sample data going from 2010 to 2022, with a testing period of about two years. Four different Machine Learning algorithms, plus a blend algorithm combining the four, were used in order to obtain forecasting data points. Those were then validated through the Diebold-Mariano test statistics, the algorithms were trained over a timeframe of ten years. The obtained results were optimized following the MVO approach with the maximization of the Sharpe Ratio as the objective function.

The constructed weighted portfolios were used within the specific extreme events context, for events that occurred in the last 20 years worldwide. Six major unpredictable events were selected for this study, having different natures and categorizations, that can be split into 3 macro-areas: terroristic attacks, natural disasters, and political changes. Further, on these events, the constructed portfolios were used in order to simulate the effectiveness of the portfolio strategies.

Machine Learning is the shiny new frontier in asset management. The increasing computational capabilities of computers and algorithms, alongside the increasing availability of conspicuous volumes of data, may smooth the way to new and unexplored avenues in the entire asset management sector. Every single piece of the investor's journey is going to change, paving a fertile ground for future research to come.

Word count = 15,254