Tactical Asset Allocation: a Novel Approach via Machine Learning Returns Prediction

Federico Lorini

a.a. 2022/2023

Dissertation written under the supervision of professor Dan Tran

Dissertation submitted in partial fulfilment of requirements for the MSc in Finance, at the Universidade Catolica Portuguesa, 04/01/2023.

Abstract

Title: Tactical Asset Allocation: a Novel Approach via Machine Learning Returns Prediction

Author: Federico Lorini

The thesis tries to develop a novel quantitative approach for the development of trading strategies managed through Tactical Asset Allocation investment style. The work employs a returns prediction performed via supervised learning algorithms and studies whether they can help in the deliverance of superior returns compared with the classical buy-and-hold strategies. The models manage to predict the movement of the markets with an acceptable precision level and try to give hints for bet sizing, playing a key role in the weight's definitions. While most of the previously developed Tactical Asset Allocation strategies were performed through a historical analysis. The novelty of the approach comes from the incorporation of these forward-looking predicted values.
In this study, Multi-Layer Perceptron Neural Networks, Support Vector Machines, and Random Forests were used to predict the returns of two main market indexes, respectively S&P500 and Eurostoxx600. The models are used in their regressor form to get a continuous output, expected to be the true value of the returns for the following day. Based on this prediction, several trading strategies have been developed and tested.
Results indicate that the proposed approach can give positive signals for what concerns return achievement and reward-to-risk ratio improvement. Nevertheless, due to the high dynamicity of the strategy, as implied by Tactical Asset Allocation hypothesis, transaction costs play a key role in final returns deliverance. All the trading strategies are performed considering the different outcomes of the models.

Keywords: Tactical Asset Allocation, Machine Learning, Returns Prediction

Resumo

Título: Alocação Táctica de Activos: uma Nova Abordagem através da Previsão de Retornos da Machine Learning.

Autor: Federico Lorini

Esta tese tenta desenvolver uma abordagem quantitativa inovadora no desenvolvimento de estratégias de trading baseadas num estilo de investimento de Alocação Tática de Ativos. Este trabalho usa previsões de retornos a partir de algoritmos de aprendizagem supervisionada e estuda se essas previsões conseguem gerar retornos superiores que os de outras estratégias de investimento. Estes modelos conseguem prever os movimentos dos mercados com um nível de precisão aceitável, e ainda sugerem o bet sizing, contribuindo vitalmente para a definição dos pesos de cada ativo. Enquanto a maioria dos estudos alusivos à Alocação Tática de Ativos focaram numa análise histórica, a novidade que esta tese apresenta é a incorporação da previsão de valores futuros no processo de decisão. Neste estudo, foram usados Multi-Layer Perceptron Neural Networks, Support Vector Machines e Random Forests para a previsão dos retornos dos dois principais índices: S&P500 e o Eurostoxx600. Todos os modelos foram usados na sua regressor form para obter o output contínuo esperado com o intuito de prever os retornos do próximo dia de ambos os índices. Com base nestas previsões, esta tese criou e estudo várias estratégias de trading. Os resultados dessas estratégias dão sinais positivos na capacidade destes modelos em aumentar os retornos absolutos e os retornos ajustados ao risco. Contudo, devido ao elevado dinamismo deste tipo de estratégias, os custos de transação contribuem significativamente para uma discrepância entre os retornos hipotéticos e os retornos efetivos. Todas as estratégias de trading foram simuladas com base em todos os diferentes outputs dos modelos.

Palavras-chave: Alocação táctica de activos, Machine Learning, Previsão dos retornos

Contents

List of Figures

List of Tables

1 Introduction

The term Tactical Asset Allocation was born to indicate an investment style in which different securities or asset classes are constantly and medium-actively managed and adjusted. It represents one of the most dynamic levels of allocation, consisting mainly of short-term bets deriving from a previous strategic allocation. The difference between Tactical Allocation and Strategic Allocation stays purely in the initial approach: strategic allocation employs a defined strategy from the beginning which will be kept constant for the investment process while, on the other hand, the tactical approach allows a momentary deviation from the original weights, to take advantage of a certain short term market trend (Dahlquist & Harvey, 2001).

In periods of market instability, dynamic weights can represent one of the best options investors can use to generate both wealth from their portfolios, and protection from possible downsides. Recently, markets have been subject to severe episodes of volatility and consistent bear markets. This negative period, which started in 2020 with the COVID pandemic affecting the overall economy, and mined the stability also of the safest assets, such as S&P500, which recorded its highest losses since the financial crisis of 2008. Other markets suffered in similar measure, making the investing activity consistently more dangerous, making risk-free assets the only safe-haven, due to the high correlation between major markets and asset classes. Nevertheless, this relationship is not always perfectly symmetric and parallel, offering some gains opportunities when these performance asymmetries are discovered. Dynamic weight shifting might then represent an efficient measure to capture these effects and achieve the best combinations of bets to optimize reward to risk ratio.

The main purpose of this study is to develop a novel approach to Tactical Asset Allocation, with the support of machine learning-based prediction models to exploit the one-day ahead performance of the two main equity indexes representative of the two main markets, SP500 and EuroStoxx600, respectively market proxies for American and European equity markets. Subsequently, the predictions will be the base for the development of several trading strategies, to test whether the forecasting ability of the models can help in finding the best combination of risky assets, adjusting the weights under a tactical perspective beforehand events occurrence. Moreover, some strategies have included also the possibility of the investment in a risk-free asset when predictions indicate an unavoidable loss. The novelty of this approach is represented by two key steps. Firstly, in past literature, Tactical Asset Allocation problems have always been solved just by looking at the predictive power of historical short-term data, without using machine learning to generate an accurate forecast, but only relying on some technical indicators such as historical moving averages. Secondly, most of the older studies are performed by shifting between different asset classes, and not between markets. Most of the previous studies employed Tactical Asset Allocation to find the best combination of stocks, bonds, commodities, and alternative assets, testing overperformance against static asset classes combinations (Fabozzi et al., 2014), finding that in most cases practitioners do not outperform simple passive strategies, generating both lower returns and higher volatility.

Since equity returns, and in particular index returns are always located around zero, large volatility spikes can have a significant impact on the performance of the strategies. A good performance will be associated with an accurate prediction of these uncommon events, allowing a preventive weight shifting to capture either a positive effect or limit a negative one. Theoretically, Tactical Asset Allocation strategies aim should outperform significantly classical buy-and-hold ones during bear market regimes, since the latter is more sensitive to volatility because protective moves are not often performed, and this study aims to test tactical deviations' ability to foster strength during negative periods. On the other hand, a wrong evaluation of the prediction can amplify loss impacts, and decisions on single bets should be cautiously evaluated.

To conduct the study, three Machine Learning models have been involved: an Artificial Neural Network, a Support Vector Machine, and a Random Forest, considering their high versatility and good performance demonstrated in older studies. The predictions have then been incorporated together to generate an aggregate for each asset, explaining the outcome of the models into a single value. Predictions have been performed on two major market indexes respectively for the US and European ones. Thanks to their broad stocks' inclusion, they are representing good proxies for the evaluation of a whole equity market. Forecasting has been conducted considering several possible return drivers coming from the macroeconomic scenario, fixed income market, commodities market, technical indicators, and sentiment indexes. The considered risk-free asset is represented by the rate retrieved by Kenneth French Data Library.

Trading strategies have been developed to shift capital between indexes when return generative events are potentially exploited, and in some cases, when losses cannot be avoided, by divesting from these two assets and investing in the risk-free one.

The steps for the research have been performed as follows. Firstly, past literature on Tactical Asset Allocation, market timing, and return prediction have been reviewed. It follows a dataset description, with an explanation of the variable's choice. Then Machine Learning methods are introduced in the peculiarities useful for this study. Subsequently, methodologies are explained, together with results interpretation.

2 Literature Review

The advantages and disadvantages of Tactical Asset Allocation have been analyzed in past years, with evidence of their positive effects since the early years. This investment style represents a good instrument to achieve higher desired returns and limit the risk at discretionary levels, but due to the high dynamicity of the strategy, it employs also moderate and high transaction costs, which results significantly in final returns evaluation (Dahlquist & Harvey, 2001). To front this issue, usually adopted strategies focuses also on the minimization of transaction costs, or to use of derivative instruments to manage tactical deviations and optimize final outcomes. It's important to notice how economic regimes are one of the main drivers for the definition of the best Tactical strategy to adopt: weight shifting is characterized by a conditional approach depending on the overall market conditions. Equity performance, in fact, is one of the most influenced by regimes, which often result characterized by high volatility, especially during bear market periods. Research analyzed how ineffective is to exclude risk-free assets when the regime is bearish.

Equities and asset classes tend to suffer from high correlations, and investors should move to more liquid assets or directly cash in case of bearish market consistency (Ang & Bekaert, n.d.). However, these kinds of dynamic weights strategies appear in most cases to be more efficient in a bear market, while slightly underperformed buy-and-hold ones in bullish periods. Tactical Asset Allocation appears then to be a good strategy to face these market movements, for its demonstrated ability to limit risk exposure, by implying active portfolio management during more volatile periods. As consequence, research showed that losses in utility might be reduced by an average of 2% per year by adopting these kinds of approaches, especially on portfolios with a long-term scope (Moreira Muir, 2019). Nevertheless, Tactical Asset Allocation strategy, besides of good performance, shows the limitation to be performed exclusively on past trends. Research already tried to make some improvements by looking at several forward-looking metrics, and proposing some market timing strategies blended with tactical asset allocation ones. These quantitative approaches consist in the rebalancing on simple historical moving averages, which significantly outperformed simple buy-and-hold strategies. From a long-run perspective, buy and hold are showing a constant and consistent performance but suffers from bear market regimes and volatility peaks due to the non-intervention on the invested weights. On the other hand, a periodic rebalancing system allows for an increase in the risk-adjusted returns while giving protection to market drawdowns, especially when shifting from equities to different more liquid asset classes or directly to cash (Faber, 2007). This last paper underlined the importance of accurate timing in asset allocation problems, especially for its importance in drawdown prevention. Moreover, research also gave good insights about several effective market timing predictors which might help in the weight optimization process. For example, macroeconomic condition analysis is important to assess the overall economic conditions, allowing the detection of possible distressed periods, tightly linked to the safeness of the different stock markets, giving insights about how to shift capital between different markets. For example, Gross Domestic Product, Inflation, and interest rate changes can provide useful information for prediction purposes, especially in the long run (Fama & French, 2010). As a rule of thumb, research suggests that developed countries are characterized by lower volatilities, representing a safer harbor also during distressed periods (Chong Phillips, 2014). Understanding the predictive value of all these variables makes timing always more important for the good results of the allocation process, which can give insights on how to exploit future market opportunities, take advantage, or protect the wealth when things are out of the lines (Anson, n.d.). Positive dealignment exploitation often doesn't results in any systematic adverse impact on the returns generation capabilities, also when the optimization is made looking at risk profile as a primary focus. Tactical Asset Allocation optimization based on prediction acquires then evidence of the capability of systematically achieving positive returns(Colucci & Brandolini, 2011).

It's obvious that risk profile remains very central in strategy drafting, and a combined forecast of both returns and risk would improve sensitively the portfolio performance, also under a long-term perspective (Liu et al., 2003). After the analysis of the improvements that market timing and returns prediction can apport to tactical asset allocation, it's also fundamental to analyze which are the main drivers that affect market performances. The analysis of the economic environment can give accurate precision in forecasting the movement of the financial markets, which have a greater impact, particularly on indexes and aggregates more than on the single stocks, which would require a more detailed analysis of the underlying companies' financial conditions. Research showed evidence that one of the main drivers for market movements is represented by government decisions, reflected by the treasury fixed income market, yield curve offers an accurate tool for anticipating market recessions and downturns. The analysis of spreads, in fact, appears to be a good explanatory variable to forecast bear market regimes (Resnick Shoesmith, 2002). The changes in the spread between 10 Years T- Bills and 3-month interest rates can time the inversion of the yield curve and forecast a possibility of market recession(Estrella & Mishkin, n.d.). Moreover, this spread has also the power of predicting future conditions in the business cycle which results to be strictly bound to financial markets performance (Siegel, 2014), which gains relevance in the definition of some sentiment indexes. Sentiment has been found very impactful on equity returns since stocks tend to perform better in a period of high investor confidence and good business conditions (Barberis et al., 1998) (Fama French, 2010). In fact, the Aruoba-Diebold-Scotti index was born to track different underlying conditions which can affect the safe condition of businesses, which have a direct impact on some significant macroeconomic variables, such as GDP, which is defining the healthy conditions of the economic machine. In general, sentiment indexes appear to have a discriminatory power in the definition of movements of financial markets, because they are comprehensive of both general economic condition sentiment and aggregate of some technical indicators on business performance (Mascio & Fabozzi, 2019). Technical indicators' performance in market timing is also widely discussed. Evidence from past literature shows that sentiment aggregates on these variables are significantly effective on return prediction, such as for the Baker-Wurgler index (Baker & Wurgler, n.d.)(Mascio et al., 2021). When not grouped, some indicators are proven to be significant also when singularly treated. Due to their business-cycle-based nature, they can capture fluctuations in the equity risk premium, and when paired with macro they're able to achieve superior forecasting ability(Neely et al., 2014). In detail, they have been used for building trading strategies, and some champions that emerged to systematically beat buy-and-hold portfolios are Earning-Price ratio and short-term interest rates, with a particular focus on the spread between them, which appears to be a strong predictor of higher mean returns and lower variance (Shen, n.d.). Results of efficiency of those indicators, on contrary, are also proof of non-significance in the Out-Of-Sample prediction, for dividend yield and book-to-market ratio

(Welch Goyal, 2008)(Goyal et al., 2021). It's important to state that independently from these variables, earnings and growth prospects are related to stock returns, by indirectly influencing the investors' confidence, as a sign of financial health for related securities. One last predictor which allows to obtain significant results is represented by the US Market. The USA is the most powerful and active economy in the world and its dense net of commercial partnerships with other markets creates a big American influence on the global stock market. In fact, lagged US returns, T-Bill rates, and market dividend yield appear to have an intense influence in predicting industrialized markets' stock returns, putting the USA in a key role for other markets' return prediction (Rapach et al., 2013). The USA represents then a reliable proxy to have a clear overview of the overall world economic situation.

Most of the variables presented above are used as drivers for returns prediction in the machine learning workflow. In recent years, Machine Learning is gaining importance in finance research. Empirical and econometric studies, linear model-based ones, proved to achieve only a medium-low accuracy in returns prediction, besides very complex models. Research showed that returns tend to follow non-linear patterns and the impact of Machine Learning methods significantly helps in generating forecasts which improved significantly the performance of investment portfolios (Israel et al., 2020). These evolutionary gains are today implemented into Asset Management industry operations and are broadly used for prediction formulation, asset pricing, and risk management. Using machine learning is currently giving new insights on the main drivers influencing markets, linking them to several common factors, which sometimes are latent and lead to the arousal of anomalies. Machine Learning forecasts become then important for all these reasons, allowing investors to incur in large economic gains when trading strategies are based on a precise prediction. In fact, they're proven to be able to even double the performance of regression-based forecasts. Non-linear models, the ones based on unstructured data are the ones that achieved the best results, such as Artificial Neural Networks and Decisional Trees (Gu et al., 2020). The predicting power of these models is widely known and studied. For example, Neural Networks showed high accuracy in predicting stock returns daily, with excellent market direction capturing, while trees perform better on longer time spans, such as weeks, keeping high accuracy. Ensemble methods might then give some additional benefits also on daily predictions. This evidence has been verified and confirmed throughout further past research, underlying how the success of these models is highly due to their strong intrinsic regularization and ability to capture iteration effects. This latter represents a persistency in returns predictability independently from the linear combination of base features (Daul et al., 2022). Artificial Neural Networks set themselves to the first position in prediction accuracy, followed by Support Vector Machines which are demonstrated to be able to capture an accuracy higher than 60% in most cases, and both give relevant performance improvement when benchmarked with simpler models with fewer parameters, such as linear regression and logistic regression alone (Emerson et Al.). The last piece of evidence from the literature is about Tree-Based Algorithms. Most of times they appeared to be able to overcome some general problems in market predictions, like bias generated by intrinsic volatility and a sensitive reduction of forecasting error, minimizing investment risk. Moreover, with the support of technical indicators, they highlight the ability of the latter to make an accurate prediction in the medium-long term. Nevertheless, their accuracy is bound to a regional constraint of analysis: to reduce the forecasting error the analysis should be made considering a single market or country at the time and analyzing singularly their dynamics finding crucial determinants (Basak et al., 2019).

Upon this previous past research, given the importance of accurate return forecast, the aim of the thesis is to demonstrate whether Tactical Asset Allocation problems results can be improved with efficient market timing made via machine learning methods. This work aims to test the forecasting ability of three main machine learning models, such as Multi-Layer Perceptron Regressors, Support Vector Regressors and Decision Tree Regressors. This selection of models is based on past performance encountered in their usage, which identifies these models as one of the best options for returns prediction.

After this forecast, the asset allocation problem will be solved upon the predicted time series and tested with the out-of-sample realized returns. Moreover, to test whether this strategy might improve classical portfolio problems, results will be benchmarked against a buy-and-hold strategy, based on the in-sample performance of the single assets. The expected results are that decisions made upon forecast will be forward-looking for future market trends, achieving higher realized returns and hedging investments from volatility and bear market periods.

3 Dataset Description

This study is based on a time window of 10 years, starting from the 1st of January 2012 until the 31st of December 2021. Data are retrieved from several official databases: equity and futures data are retrieved from Thomson Reuters Eikon and Datastream, while bond yields, macroeconomic variables, and volatility indexes are retrieved from FRED, Eurostat, and Statistical Data Warehouse. According to their correlation with the dependent variable and evidence from previous literature, 17 independent variables have been selected to be used as predictors of index returns. The dataset is composed of a volatility index, 3 bond-related variables, 2 futures-related variables, 6 macroeconomic variables, 2 technical indicators, 2 sentiment indexes, and 1-day-lagged SP500 Returns.

3.1 Volatility Indexes

Volatility Indexes represent a strong predictor of market returns. They are an indicator of expected volatility in the market, and their value is calculated on the previous 30 day's volume of traded options. They appear to have a strong negative correlation with the index returns, and the market appears to react when the volatility index value tends to be high. These indexes are very useful also to track the sentiment of investors toward the condition of financial markets, and good timing and interpretation of their behavior can be useful to predict future market fluctuations. The analyzed variables for this study are respectively VIX for the American market and VSTOXX for the European one, even if VIX might be used as a proxy due to its good correlation also with European Market.

3.2 Macroeconomic Variables

Macroeconomic variables have been included in the predictor set based on the assumption that a healthy general environment is necessary for the good development of the overall equity market. As a direct consequence, if a country or a geographic area shows to be prosperous in terms of wealth, production, and economic stability, financial markets will directly be affected by a positive effect. The chosen variables are some of the most significant ones in describing this effect.

First considered is the GDP, under the perspective of expected variation. A well-developed and financially stable country can generate significant wealth, which is reflected by the GDP itself, which has a positive correlation with financial market performance. Another similar variable is represented by Consumer Expenditures, and its integration is made in the same way as GDP. The choice of using a nowcast instead of the true value came from both a practical issue of frequency (GDP and Consumption Expenditures are updated yearly while the nowcasts used are updated quarterly) and then to capture a more accurate sentiment and forecasting effect. On the consumer side, Consumer Price Index is another important variable, which represents the willingness of consumers to inject money into the market for the purchase of goods and services. Consumption influences market conditions, for its wealth-representative ability. The correlation appears to be moderate. It's also important to consider some variables of the monetary market, such as inflation, implemented in the model under the form of breakeven inflation, representative of the daily-updated inflation expectations. It represents financial instability in the money supply and when this rate tends to be high, market development will consequently slow down, for the negative effects inflation is bringing to the economic environment. A high rate often corresponds to bear market periods. This high negative correlation makes breakeven inflation one of the most important predictors of market returns. Moreover, it's also relevant to read this data together with the unemployment rate. As shown by the Phillips Curve relationship, it has a high cross-influence with the inflation rate and their combination allows to better capture the overall effect.

3.3 Bonds

Bond Yields represent an effective proxy for current and expected market conditions. Government bonds represent an indication of the financial health of a state, which is directly correlated with the market's performance. During the distressed time, long-term yields tend to rise, and the direct consequence is a frequent negative impact on equity prices. Moreover, it's important to underline that this information alone is not always a future prediction but more current information. To complete the picture is also important to analyze the short-term bond yields and the difference between long-term and short-term yields. This spread represents an indication of the condition of the yield curve: when the spread value is small or negative, it means that the yield curve is inverting, and the overall market is going into recession, carrying negative effects with the probability of a long bear market period.

3.4 Futures

Gold and Oil are two important commodities for their strict connection with overall market performance. Gold represents the safe-haven asset par excellence, where investors decide to put their money to be protected from distressed market periods. Gold's future price performance can then give a predictive signal of crises, or general fear. On the other hand, oil set itself as one of the most important market conductors, for its binding to the inflation rate and commodities market, which directly affects the overall economy. The correlation of these raw materials futures' prices tends to be high throughout the analyzed historical path, making oil and gold reliable predictors for market returns.

3.5 Technical Indicators

Technical indicators are intrinsic variables of the analyzed asset, which can carry some historical and actual information able to suggest a possible future scenario. This study takes into consideration the variable of the earning-to-price ratio, which is suggested by previous literature as one of the more effective technical predictors. The E/P ratio, which is the inverse of Price-Earning, describes the ability of an asset to generate earnings, which is a synonym for expected good performance. An increase or stability in its value represents a better earning-generative ability. Increased importance is defined as whether these earnings exceed a borrowing rate, which might be one of the funding sources for income generation. For this reason, another accurate predictor for returns will then be the spread between E/P Ratio and 3 Month Bond Yield, as a short-term funding source. The higher this value, the better will be the future expectations for the considered asset.

3.6 Sentiment Indexes

Evaluation of general investors' sentiment is crucial for an improved returns forecast. Better sentiment will lead to higher purchase volumes with a consequent increase in prices. Moreover, also the sentiment about the condition of the subordinated variable appears to have a significant correlation with markets. Past studies suggest incorporating sentiment indexes that are not only directly related to the asset themselves, but to the overall economic environment. Business conditions and business cycles, for example, are one of the main drivers for the economic machine and the use Aruoba index as a predictor can carry important information on their future impact on the financial markets. Moreover, Baker and Wurgler's index describes the sentiment towards some technical indicators and other intrinsic asset variables, incorporated into one value. The medium correlation with realized returns suggests their moderate forecasting ability and the importance of sentiment analysis and behavioral biases.

3.7 S&P500 Lagged Returns

As stated by past literature (Basak et al., 2019), SP500 returns show a high autocorrelation and a high influence on other related markets. Its lagged value represents then a reliable predictor of future performance for the main world market indexes. This proxy can then be used as a forecasting indicator.

4 Machine Learning

Machine Learning is a branch of artificial intelligence that allows algorithms and computers to learn autonomously without being directly and explicitly programmed (Samuel, 1956), and the big variety of developed models allows the solution of different and various types of problems, such as classification, regression, and clustering. Machine learning aims to identify patterns and recurrences in past data and make the most accurate predictions possible when new events occur. Machine Learning models are a mostly particular type of non-linear regression models, designed to be optimized to perform a particular task in the most efficient way possible. The main difference with econometrics, which is another field of study where some models such as linear and some classification regressions are widely used, is the main scope: econometrics aims to discover causal effects between variables, while machine learning focuses on the predictive value these variables can give. Machine Learning algorithms are then designed to get inputs from a defined sample of realized data and try to make an out-of-sample forecast based on the discovered patterns. The value added to these processes is given by the iterative steps allowed by computers: in fact, models can be constantly optimized with the iterative testing of internal parameters, for the achievement of superior results. This is one of the most important actions of machine learning workflow, which is composed of the following 5 steps.

4.1 Machine Learning Workflow Steps

4.1.1 Gathering Data and Data Pre-Processing

To make the machine learning workflow most efficient as possible is important to start the optimization process from the beginning, building the dataset most functionally and cleanly as possible. When working with financial data, the quality of data can sometimes be poor, with different inconsistencies such as missing values, noisy data, wrong values, and extremely large ones generating fake outliers, which can significantly confuse the model. It’s important to check the good quality of the variables from the beginning to not artificially inject bias into the results. After this step, data pre-processing will follow. Pre-processing is a key step of the data preparation process, allowing to significantly remove of a good amount of bias and noise from the collected dataset. Raw data can have several problems: they can be very unorganized, unlabeled, with mismatching frequencies, or full of missing values given also by the different sources. Training a model on raw data always generate poor results given to model confusion in understanding predictors. Some basic steps for data pre-processing are conversion, missing value handling, creation of synthetic data, outliers' detection, and correlation calculation. All these steps are important for the good performance of the models. They can only process numerical features, and all variables which define a category must be encoded or converted into dummies. Models also cannot accept missing data, which shall be removed together with the corresponding row, but this method is not efficient if there are a lot of "holes" in the dataset, because it will sensitively reduce the amount of data to feed the algorithm. A possible solution for that is to fill in the blanks manually on a rule, depending on the type of information treated. Missing data might also be predicted or generated artificially for different purposes. Sometimes there is some information that can be incorporated together in a synthetic series, which can solve most of the problems of low data availability or big data frequency mismatch. A similar approach to missing data shall be applied to outliers, which sometimes are generated by a typing error or a wrong observation, but their impact can be significant in defining their effect on the outcome, and they should be manually removed. Outliers' removal is proven to improve sensitive model accuracy. Lastly, to improve the performance of the overall model, checking the correlations between predictors and the y variable is crucial. Low correlation variables might hurt model performance, adding bias concerning the most significant ones. If a variable is found to be uncorrelated, it shall be deleted or changed with the most significant one. Sometimes significance is given by the different data frequencies and can be improved with a re-sampling.

4.1.2 Model Research

As stated above, today's Machine Learning is characterized by a huge number of models which can help to solve the widest variety of problems. Most of the problems in finance are supervised learning problems, in which all variables are labeled and structured. Regression algorithms give a numerical output and are widely used for prediction problems, while classification algorithms give a binary output, which is useful for event studies or event detection. It's important to choose the type of model that better fits the data and the desired output: the different peculiarities impact significantly the quality of the outcome, and a wrong choice can result in poor performance.

4.1.3 Model Training

Machine Learning models are constructed to analyze an in-sample dataset, find a pattern, and then make a prediction out-of-sample. The dataset must then be prepared for this process. Normally, the dataset shall be split into a training set, which will be the in-sample, and a test set, where prediction will be compared with realized data. A third set might be added, the validation set, which will be used to optimize model parameters. The train set will then be given to the model to find patterns and correlations, and prediction will be made when the test set is plugged into the model.

4.1.4 Model Optimization

Most of the time, models in the raw form result in poor performance. Models are characterized by a high number of internal parameters regulating some tasks, such as learning rate, regularization coefficient, and other features as the model increases in complexity. Often happens that models considering a high number of inputs tend to be too precise in describing the patterns of the train set, and it can easily result in overfitting, which means that the model is very precise in a single instance, but it won't work well in most cases. Another crucial step in the model building is then hyperparameter selection and tuning. This practice it's important to get the best parameters to adapt the model to the data, while avoiding overfitting at the same time, identifying a good fitting balance. This process is performed by applying cross-validation, "folding" the train set into a certain number of subsets, and evaluating the performance on a validation folding, changing over the execution time of the process.

4.1.5 Testing Results

After the model is optimized, it should theoretically be in its more efficient form. During the testing phase, the output of the optimized model is compared with the test set independent variable, which represents the true realized values. An efficient model prediction should be as close as possible to the realized result, highlighting the great precision of the prediction. The performance can be evaluated through different metrics depending on the type of model considered. At this stage is also important to evaluate the results of the loss function: all the algorithms use as a precision metric this function and try to minimize it. The lower the output of the function, the better would be the prediction.

For this study, three algorithms have been selected: the Artificial Neural Network, Support Vector Machine, and Random Forest. To fulfill the return prediction task, all the algorithms have been used in their regressor form, to have the predicted value of the returns. Classification methods have been excluded since a binary output would have not been useful for asset allocation purposes.

4.2 Neural Network

Neural Networks are a particular type of machine learning algorithm whose functioning is inspired to the human brain. Their composition is very articulated: they have a layer structure, on which each stratification is composed by a finite number of computational cells, also called neurons, interconnected with each other. There are two extreme layers, the input and output ones, and a finite quantity of computational layers, which are hidden. Neural Networks are very versatile and can be used for regression, classification, and deep learning problems. In this study, a particular class has been used, the Multi-Layer Perceptron Regressor.
The workflow starts with the input layers which get input data and pass them to the hidden layers, which are computational units. Depending on the operation to be performed, all the neurons contain an activation function that processes the signal and passes the results to the next layer. For regression operations, the most widely used activation function is Rectified Linear Unit (ReLU). The activation functions process the signals also considering the calculation of the weights on the previous layer and an artificial bias, set to reduce overfitting possibilities. The advantage of using a Neural Network is also given by its internal iterative structure, which allows the model to optimize itself during the training Backpropagation. Backpropagation aims to work on training error terms, optimizing layers backward, to reduce the differences between

the target output and the obtained one.

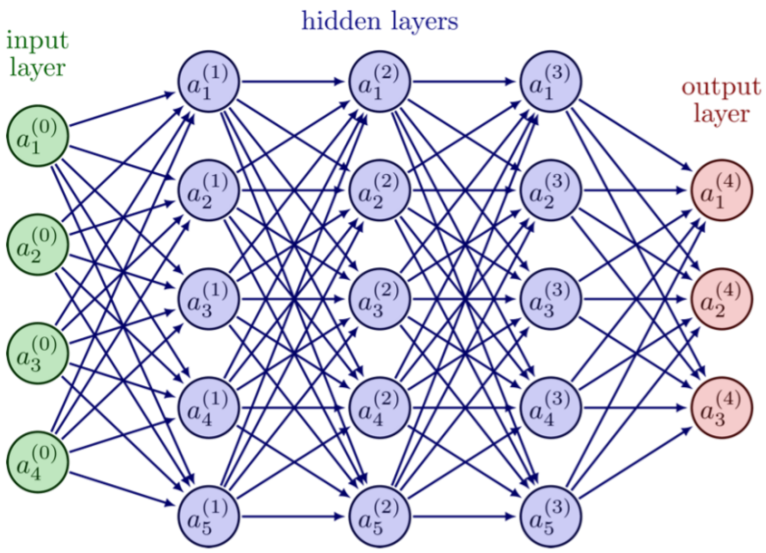

Figure 1: Neural Network Scheme

4.3 Support Vector Machine

Support Vector Machines is a versatile supervised learning algorithm that can be used both for regression and classification problems. Its objective is to find a hyperplane in a multidimensional space that can be used as a decision boundary to make the prediction. Support Vector Machines map the input into the high-dimensional space, using different linear and non-linear kernel functions, and draw a hyperplane to maximize the margin between the closest data points to it, called Support Vectors. The best solution is found when the hyperplane can fit the highest number of observations. The main characteristic of SVM is that they achieve the best solution when the best line within a threshold value is fitted. The threshold is represented by the distance between the hyperplane and the boundary lines.
Support Vector Machine is good for prediction because its robust to outliers and gives big and sudden changes in the dataset don't provide a problem for the outcome of the model, which appears to be stable and with less variance than the Neural Network output.

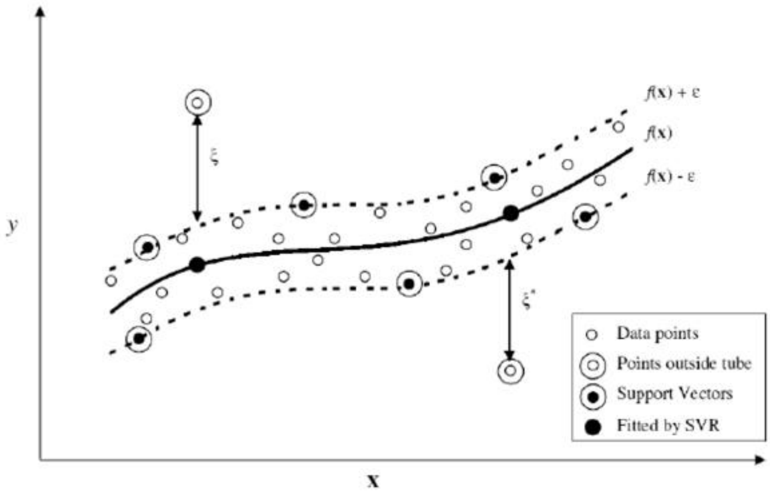

Figure 2: Support Vector Machine Functioning
Source: (Lahiri et al., 2008)

4.4 Random Forest

Random Forest is a particular kind of Machine Learning algorithm that uses ensemble learning methods. Random Forest is represented by the combination of multiple algorithms, which combined can give a more accurate prediction than a model on its own. In this case, the ensemble is made by several decision trees, which are built starting from a random subset of the data and choosing the best feature to split it in each step, using a recursive method until the tree is not completed. The random forest is characterized by the repetition of this step several times, with different data subsets. Different outputs are combined using the average of all predictions and aggregated into a single outcome.

Each tree processes the information through different nodes. The Root Node, which is represented by the entire sample, is the starting point, followed by Interior Nodes, that process information together with branches, which are the decision rules, and lastly, the leaves, which carry the output of the decisional process.

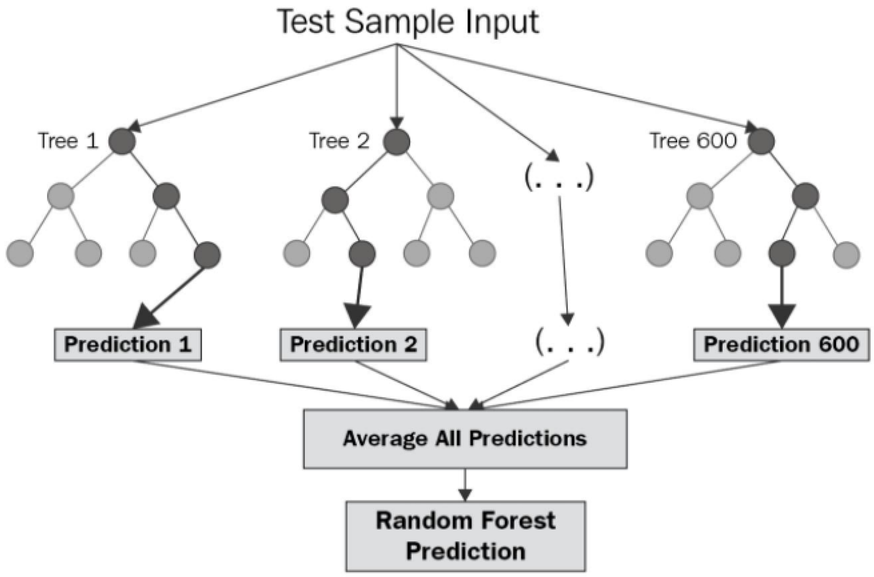

Figure 3: Random Forest Scheme
Source: https://www.keboola.com/blog/random-forest-regression

5 Methodology

The section aims to describe the steps to perform the return prediction and the subsequent trading strategies. Machine Learning procedures have been performed with scikit-learn library, using a pipeline-scheme workflow. The used models are an Artificial Neural Network, a Support Vector Machine, and a Random Forest.

5.1 Dataset Pre-Processing

As stated in the previous chapter, data preparation is one of the most important steps to improve the performance of models. The features have been considered as percentage changes between different periods, to better evaluate their effect on the dependent variable, which is expressed as daily returns of the considered index. Some features, macroeconomic and sentiment ones suffered from a frequency mismatch problem, which was generating a lower correlation. The problem has been solved by replacing the lowest frequency ones with their nowcasting, which is partially solving the problem. This change aims to keep all the frequencies most similar as possible. The missing data problem has then been handled by considering no changes in the given column for that given day. Some lines have been removed, corresponding to weekends where indexes are not traded and periods where variables were not available yet. The prediction window is set to be one day ahead, and carrying information on more previous periods is crucial since the effect of their changes might not be immediate and delayed over different periods. Each column has been lagged for one period, depending on their different frequency, and secondly, the data have been lagged for 20 days using a rolling window. By doing this, each dependent variable has been associated with the data of the 20 previous days, as if influenced by the daily changes for the previous month. This step allows carrying on information capturing possible delayed effects and improves the model performance, implying its re-training for every new considered period. All the considered variables are continuous, without any categorization, not requiring any type of encoding.

After these preliminary steps, the dataset has been normalized. Normalization has been preferred to standardization for its advantages to reduce the possible negative effect of null values and transform all the entries on a similar scale, making the study of their effect more direct, evident, and straightforward. Normalization can also constrain the effect of outliers making the model more stable and faster in processing information, by rescaling all the data between 0 and 1. Subsequently, polynomial features have been calculated. This step consists in raising all the features to an exponent, and it allows to uncover of new possible relationships between the independent variables and the desired target, improving model performance. In this case, second-order polynomial features appeared to be the best option.

5.2 Definition of inputs and valuation data

The whole dataset considers a total period of 10 years. It's important not to negatively affect the performance of the model to find a good combination of training and test set and to have enough data for the model to discover and learn the patterns. Normally, there is a positive relationship between the size of the training sample and the efficiency of the algorithm. On the other hand, the test set shouldn't be too small, otherwise, the model performance might not be evaluated correctly. In this study, the chosen sizes are 80% for the training set and 20% for the test set. The choice of these proportions comes from the evaluation of the occurrent events in the training sample: in fact, the COVID-19 period has been intentionally included in the training sample, to foster better learning on bear markets and important market shocks. Moreover, to avoid data overlapping during training and evaluation, both sets have been shortened with the elimination of the last 10 rows from the training set and the first 10 rows from the test set. Overlapping data might deviate the performance of the model from real value, incurring in some problems such as data leakage.

The definition of the sets and prediction windows follows the scheme below.

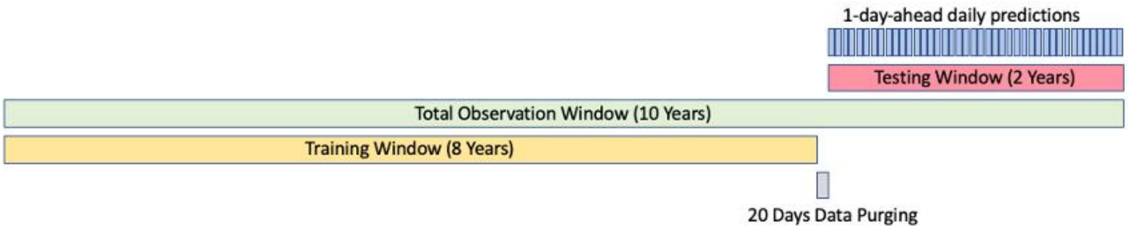

Figure 4: Train-Test Split Overview

After all the lags, the dataset structure appears as in the scheme below.

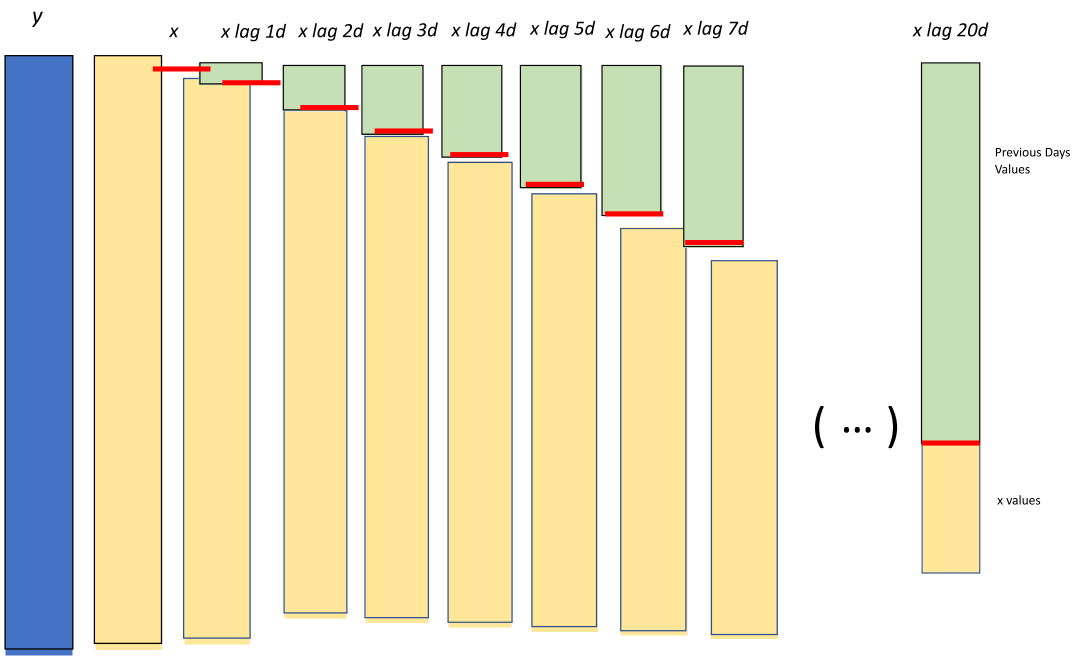

Figure 5: Dataset Structure for one single variable

5.3 Model Optimization

After a first and inaccurate prediction, models have been improved with the optimization of internal hyperparameters. As previously analyzed, this step is crucial to adapt the model to the data in the best way possible, without falling into overfitting. Regularization terms should be attentively considered. In the chosen models, as regressors, regularization parameters are represented by L1 and L2. The purpose of these values is respectively to have an impact on the weights of the regression. L1 is a penalization term corresponding to LASSO regression while L2 is to Ridge regression. For the structure of the employed models, in particular, Multi-Layer Perceptron Regressor and Support Vector Machine Regressor, L2 parameter usage is the most straightforward, for its ability to estimate the significance of the predictors and penalize the insignificant ones.
L2 parameters are defined as follows:

$$L2 = \lambda \sum_{t=1}^{n} \beta_i^2$$

To find the optimal value for regularization, the parameters have been tuned using explained variance as scoring metric and evaluated through k-folds cross-validation. The tuning has been conducted on the whole pipelines, and in general tuned parameters are the order of the polynomial features and L2 penalization coefficient. The exception has been made for the Random Forest and the Neural Network, where tuning also considered respectively the number of trees and hidden layers.

5.4 Model Evaluation

All the models have been trained over the past 8 years of data and tested out of sample over the remaining 2 years. The training set is composed by the daily variation of all 18 variables, including their rolling window lags, and the index returns are set as the dependent variable to be predicted. To test the model precision the key metrics evaluated are the following: model R^2, mean squared error, mean absolute error, and root mean squared error, to measure how large is the Euclidean distance of the prediction from the test set data points. This is the best way to evaluate the accuracy of the model prediction compared with the train set since accuracy scores are proper only for classification algorithms.

The following chart sums up all the steps used to reach the predicted values.

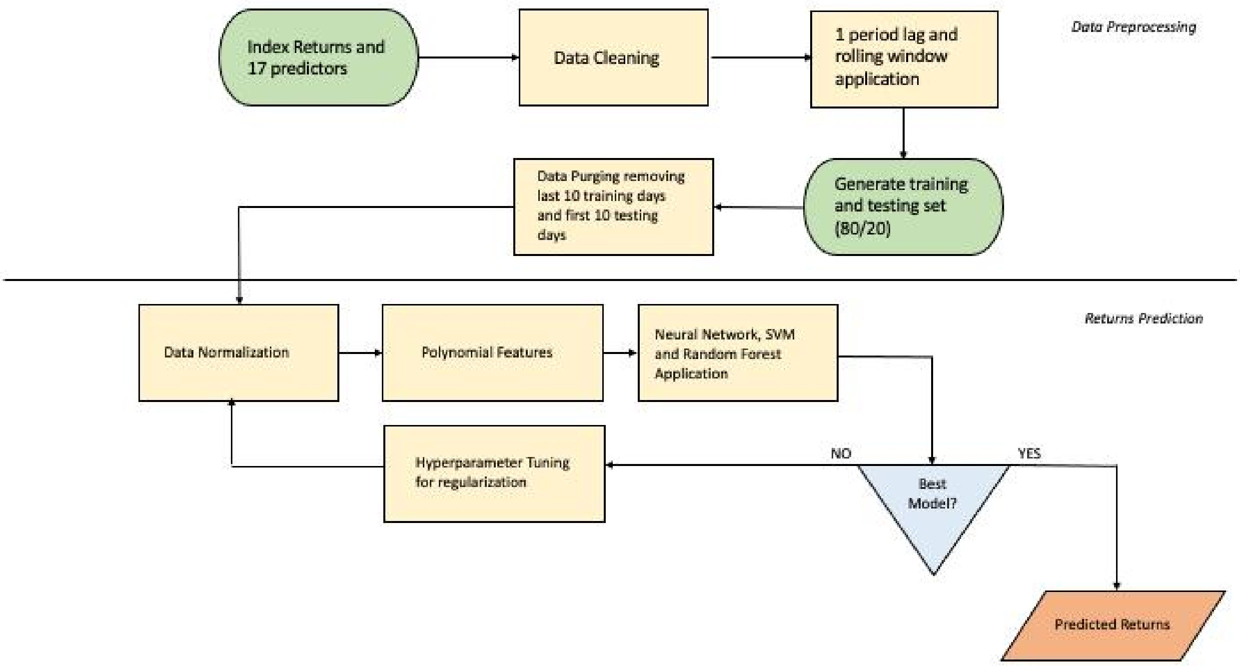

Figure 6: Model Flowchart

5.5 Trading Strategies

The outcome of each model has been incorporated into an aggregate, representing the daily average of the three predictions. Upon these series, some trading strategies have been developed. The aim of the thesis is to demonstrate whether tactical asset allocation problems can be solved in an efficient way with the support of an accurate forecast via Machine Learning. To do so, accurate model prediction plays a key role in the definition of strategies with good results. Predictions have been incorporated in different ways in the weight calculation and are the main decisional driver for the choice of investment. In general, all the portfolios are generated without any boundary: short selling is always allowed, and weight shifting is permitted without limits, to test the portfolio rebalancing without any external intervention. Moreover, some strategies are tested also under a market timing perspective, including the investment in a risk-free asset when predicted market conditions are not permitting return generation and loss avoidance.

Daily weights are calculated with the formula:

$$w_{i,t} = \frac{(\frac{1}{Prediction_{i,t}})}{(\frac{1}{Prediction_{i,t}} + \frac{1}{Prediction_{j,t}})}$$

where *i* and *j* represent the two assets and *t* represents the reference day for the prediction. Results of the evaluation are then tested on the realized returns of the asset.

All the developed strategies are the following[1]:

1. Daily weights are calculated on the predicted value of t.

2. Daily weights are calculated on the predicted value of t, keeping 20% invested on both assets.

3. Daily weights are calculated on the predicted value of t. When both predictions for t are negative, 100% on the investments is shifted on a risk-free asset.

4. Daily weights are calculated on the predicted value of t, keeping a baseline of 20% invested in each asset, including the risk free. When both predictions are negative, the baseline is kept invested into the risky asset and the remaining part is shifted into the risk-free.

5. Monthly[2] returns are calculated on the predicted value for the month t, similarly to strategy (1).

6. Monthly returns are calculated on predicted value for the month t, including a baseline of 20%, similarly to strategy (2).

The performance has been benchmarked with a buy-and-hold strategy with fixed weights, based on the historical returns average on the train set, and then tested on the same test values. More-

[1]Numeration of the strategy is indicative for returns evaluation in the table in next section

[2]Monthly returns are calculated with a mean resampling of daily portfolio returns

over, the evaluation of the number of transaction costs required to perform the strategies has been performed, with a subsequent optimization. In fact, best-performing strategies have then been subject to a leverage limit to contain the turnover. The three leverage thresholds are X3, X5, and X7, to give a better overview of the results for different levels of aggressiveness. This step has been crucial to limit the change in the composition of the portfolios, limiting transaction costs and fostering better feasibility of the different strategies. The robustness of the results has then been tested against the Fama and French 5 Factors Model and Momentum.

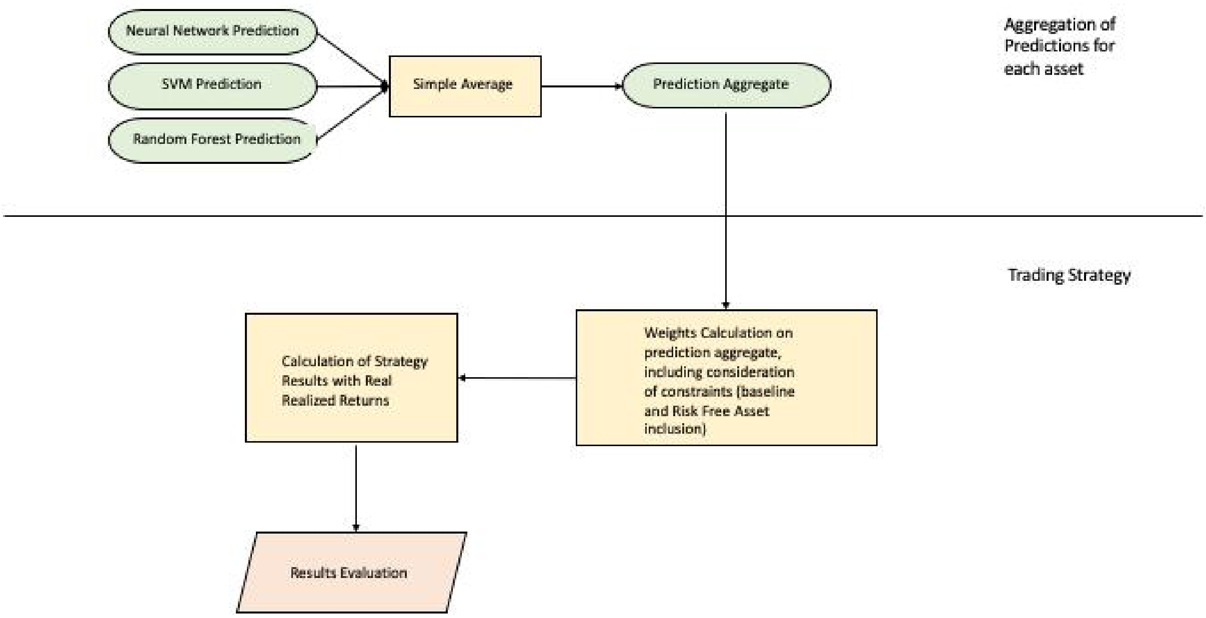

Figure 7: Trading Strategy Flowchart

6 Results

The result evaluation will be divided into two sections. The first one will analyze the performance of machine learning models in their predictive ability. The trading results will follow.

6.1 Model Performance

All the models used for this strategy have been used in their regressor form since the desired output is a continuous time series and not a classification of market trends. To do so, the considered metrics are mean squared error, root mean squared error, mean absolute error, and model R^2. Where possible, it has been evaluated also the loss function minimization trend.

Models Performance US				
	MSE	**RMSE**	**MAE**	R^2
Neural Network	0,00023349	0,015280392	0,00827284	0,8893
Support Vector Machine	0,00010654	0,010321792	0,00702071	0,1629
Random Forest	0,00010761	0,010373409	0,00704634	0,8527

Models Performance EU				
	MSE	**RMSE**	**MAE**	R^2
Neural Network	0,00029387	0,017142775	0,01034662	0,8799
Support Vector Machine	0,00016103	0,012689852	0,00859679	0,3740
Random Forest	0,00015319	0,012377109	0,00805626	0,8537

Table 1: Machine Learning Models Performance

Overall, models show a good performance. R^2 term, calculated on the training set, underlines the adaptive capacity of nonlinear models to patterns. In particular, good performance is shown by Neural Network and Random Forest, while Support Vector Machine tends to follow a more linear trend. Nevertheless, error terms between prediction and test set give good insights on the predictive performance. Neural Network showed the best trend-capturing ability, due to its high flexibility and sensibility, but it's also the one that generates higher Mean Squared Error. Looking at trend graphs is then possible to notice that the prediction of Multi-Layer

Perceptron tends to be more volatile and less stable compared to SVM and Random Forest, which response is often closer to 0. This bias might be due to the different frequencies of some variables in the dataset, which despite the lags and the rolling window implementation to better capture their delayed effect, still have an impact on the more sensible models. The best solution would be to have all the variables expressed in daily variations, but macroeconomic data often are characterized by lower frequencies. Generally, the prediction results are accurate on what concerns the capture of market trends instead of the true value of the returns to quantify which asset offers a better yield-generative ability, and this information already allows to build an investment strategy effective for short time horizons.

It's essential to underline that the good performance of the machine learning models doesn't correspond mandatorily to the good performance of the trading strategies. Capturing a trend might be useful for ponderation, even in the case of precise capturing of extreme events in the price path.

6.2 Trading Strategies Performance

After checking whether the accuracy of prediction reached a satisfactory level with model optimization, prediction played a key role in the definition of the strategy achieved returns. The evaluation has been made through the consideration of annualized mean returns, annualized standard deviation, Sharpe Ratio, Sortino Ratio, and Value at Risk with a confidence level of 5%. The metrics have been chosen to have a clear picture of the effect that prediction can have on Tactical Asset Allocation objectives: maximization of returns and minimization of undertaken risk. The table below presents the summary metrics for all the strategies. Numerical labels correspond to the numeration given in section 4.4. The last column represents the results of the benchmark strategy.

	Machine Learning Based Strategies						Buy and Hold Strategy
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(N)**
Annualized Returns	40,82%	26,93%	47,41%	21,42%	5,18%	6,76%	7,41%
Annualized Volatility	37,32%	25,82%	35,79%	17,13%	7,09%	3,79%	15,41%
Sharpe Ratio	1,094	1,043	1,325	1,250	0,731	1,784	0,481
Value at Risk (5%)	2,00%	1,84%	1,67%	1,27%	0,55%	0,30%	1,62%
Sortino Ratio	1,52	1,48	1,55	1,81	0,79	3,32	0,67
Max Drawdown	-24,27%	-15,63%	-14,88%	-10,59%	-1,91%	-0,93%	-16,86%

Table 2: Trading Strategies and Benchmark Strategy Performance

As first evaluation, it's important to underline how tactical asset allocation strategies overperform classic buy-and-hold strategies. Dynamic weight adjustments, with no short selling limitations, effectively manage to ride volatility waves and positively capture their effects. Strategies implying daily rebalancing tend to show more of this trend. High-frequency trading, on average, tends to generate higher variance due to the rapid changes in weights and a greater number of considered observations. In general Sharpe Ratio tends to be high, and besides in one case, superior to 1. These values are a good indication of the positive event exploitation given by the predictions. As stated in the previous section, calculating weights on the forecasted values is a good way to anticipate market movement, which in the positive case result in the undertaking of a long position, and the negative cases with a short one or with the shifting to a risk-free asset. The positive effects are also evidently looking at Sortino Ratio, which evidences the good volatility usage, limiting downside effects to the minimum possible. The inclusion of the risk-free is even decreasing the downside deviation, as noticeable in strategies (3) and (4), limiting the losses when prediction gave negative effects on all the assets. Strategies (4) and (6), where the weight baseline of 20% is introduced, are the ones that achieved the best metrics. Keeping always a sum invested in the assets decreases the returns in absolute terms compared with other cases but increases the worthiness and safety. Volatilities turn out to be the lowest ones and their risk-to-reward ratios are the highest ones. The surprising results are given by Sortino Ratio, as an indication of the increased stability, and a consequent better acceptance of the return variability. Another consideration is the impact of the changes in the frequency of transactions. The monthly rebalancing of strategies (5) and (6) sensitively improves the stability of the returns. This is mainly because higher frequency trading better captures all market fluctuations, and better exploits daily or general short-term volatility spikes. With returns resampling, the weight calculation is performed on the average of the previous month, and in some cases, this value might seem reductive to track all the minor changes in the returns. Nevertheless, the overall performance of the last strategies turns out to perform well and is set as the most feasible, thanks to reduced turnover given by the frequency. The inclusion of the baseline in strategy (6) shows unexpected, good effects, achieving a Sharpe Ratio of 1,784, the higher among all strategies, and the higher Sortino Ratio, which almost doubled. The losses are minimized, and returns are stable.

All the trends are shown by cumulative returns chart below, where benefits of dynamic weights are compared with benchmark strategy.

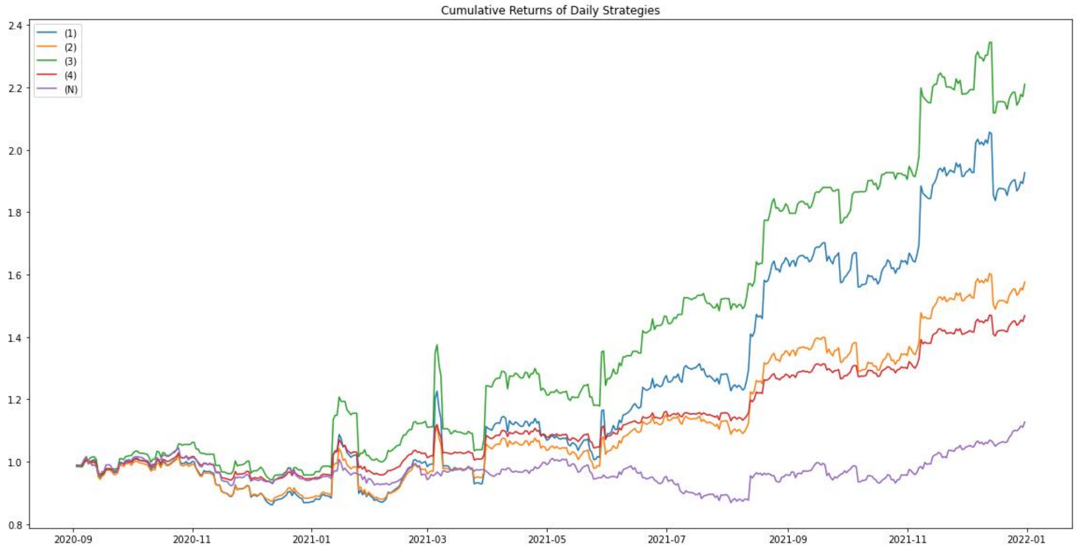

Figure 8: Cumulative Returns of Daily Strategies and Benchmark Strategies

It is evident how Tactical Asset Allocation has a positive effect in reducing volatility and protecting from bear market periods. Due to the high volume of trades, turnover evaluation is crucial for returns collection. Tactical Asset Allocation problems generally tend to have a high turnover due to the dynamicity of weights. In this case, the hypothesis appears to be partially confirmed. On average, predictions tend to keep the rebalancing of the portfolio stably except in case of needs. When predictions tend to be very positive or very negative, the portfolio tends to rebalance consistently, limiting overall drawdown and capturing positive effects, but with the trade-off of very high transaction costs.

The impact on the returns can be then significant, also affecting the worthiness and feasibility of the strategies themselves. On this assumption, it has been developed an improved version of the best-performing trading, respectively strategy (3), (4), and (6), with the limitation of the leverage to different thresholds, as stated in Section 5.

The table below shows the summary metrics of the improved strategies.

Leverage Limited Strategies									
	(3)			(4)			(6)		
	X3	X5	X7	X3	X5	X7	X3	X5	X7
Annualized Returns	25,01%	29,32%	32,57%	6,98%	17,20%	18,32%	7,12%	7,19%	6,81%
Annualized Volatility	17,60%	20,11%	22,51%	10,19%	14,18%	14,93%	3,77%	3,77%	3,79%
Sharpe Ratio	1,421	1,458	1,447	0,684	1,213	1,227	1,888	1,906	1,801
Value at Risk (5%)	1,58%	1,64%	1,67%	0,78%	1,27%	1,27%	0,30%	0,30%	0,30%
Sortino Ratio	1,675	1,744	1,776	0,807	1,733	1,771	3,438	3,445	3,348
Max Drawdown	-13,28%	-13,26%	-14,88%	-13,19%	-7,66%	-7,66%	-0,87%	-0,89%	-0,92%

Table 3: Limited Leverage Strategies Summary

Comparing results with the previous ones, it's possible to notice how the leverage has effects also on all metrics, aside from the turnover. On first instance, returns appear to be generally low. Capping the lower bound for the short selling implies reducing the positive effects that an accurate bet sizing can deliver on final returns, but at the same time fostering higher stability to portfolio fluctuations, by limiting the general volatility. Positive effects are also shown by maximum drawdown values, which are now less negative, due to the reduction of the effects on possible wrong predictions, limiting the amount of the bet and containing losses. Downside volatility also is reduced, with the consequence of higher Sortino ratios, highlighting the positivity of the "upside" standard deviation. By reflection, also Sharpe Ratios are improved, achieving greater amounts, besides in strategy (4) with leverage x3. The strategy, including both a risk-free asset and weight baseline, is designed to be the more stable as possible, and all the great benefits were derived from some positive spikes exceeding the leverage limit. A more aggressive version of the same strategy underlines this fact. In general, it's possible to state that leverage has a generally positive impact on the final deliveries, and even if the absolute values of the returns appear to be lower, the whole picture highlight a better safeness in the proposed investments, with a higher quality fostered also by the limited amount of transaction costs.

The following charts shows the cumulative returns delivered by the proposed strategies with leverage limitation. Generally, is then possible to notice the effect that the aggressivity of the

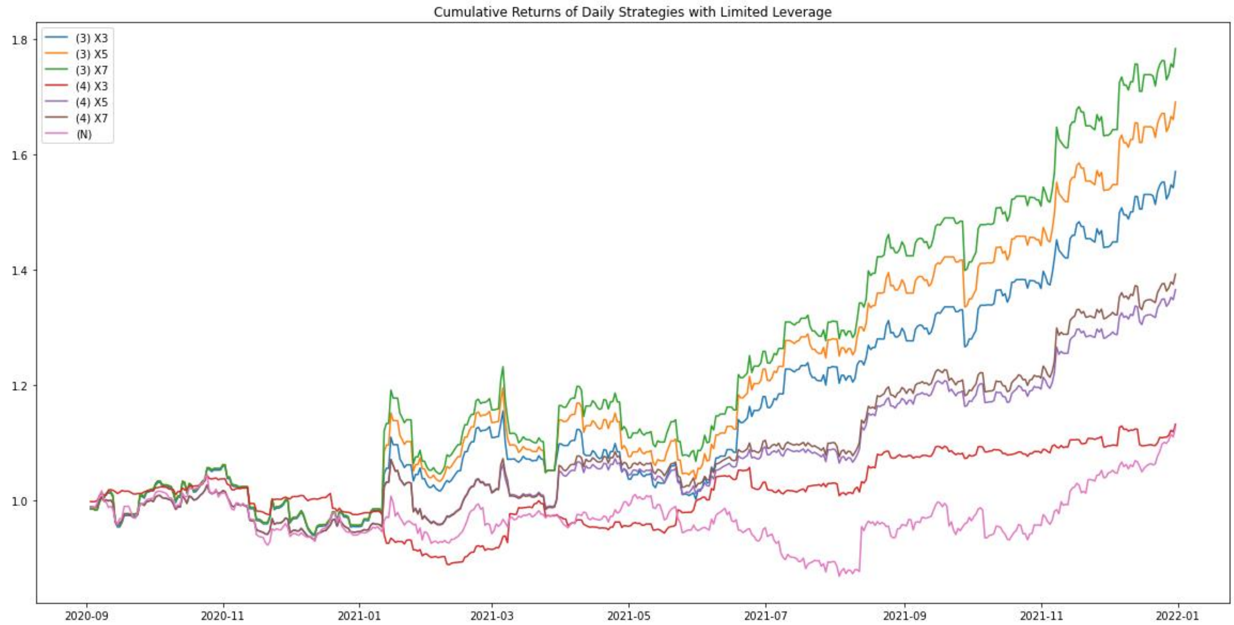

Figure 9: Cumulative Returns of Daily Strategies with Limited Leverage

strategy has on returns delivery: higher leverage allows to better exploit market events, with the side effect of the transaction costs impact. The choice of the best strategy is then discretionary to the chosen type of asset, and to the transaction costs implied for the trading.

To conclude, transaction costs need important remarks. The final turnover analysis underlines how weights tend to fluctuate during the observation window. On average, the value oscillates around zero with limited changes in portfolio composition and general stability. Nevertheless, during volatile periods weights tend to fluctuate heavily, with massive spikes which are in some cases not affordable and not convenient compared to the generated returns. The strategy will then be feasible mostly for institutional investors which are able to front those expenses and can afford to undertake the losses in case of the wrong prediction.

Nevertheless, in most cases, big spikes correspond to days of high market volatility, giving good feedback about the predictive ability of the employed models. On the other hand, monthly strategies show a sensitively lower turnover, caused by the lower frequency of rebalancing and the reduced sensitivity of the prediction, making transaction fees more affordable. It's important to consider that generated returns are way lower than in the daily strategies, making monthly ones affordable only in the lower.

With the introduced leverage limits, the fees are heavily reduced and maximum turnover is set to the given thresholds, and this step is crucial to be able to perform the strategies in real life since the limitation gives insights also on the borrowings needed to perform short sales.

6.3 Robustness Test

The robustness of the trading strategy has been tested, with a regression between the return series of the different approaches against the Fama and French 5 Factor Model and Momentum Factor. This step is important to assess whether the results are statistically significant. The outcome needs an attentive interpretation. Calculated regression betas show to be very close to zero, underlying the null effect of the considered factors on the proposed strategies. Jointly, also the obtained standard errors tend to be close to zero and of the same order of magnitude as the regression weights. As consequence, their respective t-statistic is close to one for all the factors, underlying non-significance for both the 5 Factor Model and Momentum. Also, the R^2 is very low, giving a signal of small correlation between variables. Despite that, the non-significance of the result is explanatory of the non-linearity of the proposed strategies, and the non-correlation is given by the implied linearity of the factor models. R^2 indicates that Factor models cannot explain the results, since the strategy is not in line with common risk factors and cannot be replicated with the involvement of linear models. The time span is then relevant for the explanation of the non-significance. The Benchmark model implies a long-term time span, while the experiment considers a short-term time frame, with 2 years testing set.

7 Conclusion and limitations

This work has been developed to study whether machine learning prediction can help to improve Tactical Asset Allocation strategies. The positive answer to this question is still doubtful, due to the variety of variables to be analyzed. In general terms, it's possible to state that machine learning helps in allocating resources as much as the accuracy and precision of the prediction increases. When the prediction is weaker, if the output of the models manages to track the direction of the market movement, it can suggest how to balance the assets, but with the allowance of some gain's renunciation. As the accuracy increases, the possibility of effective bet sizing can be evaluated.

Nevertheless, the statistics of portfolios highlight how this novel approach to Tactical Asset Allocation manages to reach a good reward-to-risk ratio, with a reduction of downsides, underlining how accuracy in prediction helps also with drawdown cutting. Most accurate predictions can describe with minimum error the movements one day before they occur, giving a great competitive advantage in event exploitation. On the other hand, as analyzed, good performance is also brought by a high dynamicity in portfolio rebalancing, which makes transaction costs high and, in some cases, barely unaffordable if compared with gains generated by the strategies. Limiting the leverage appeared to be a quite good strategy to solve this problem, but further studies will be necessary to find the optimum threshold to achieve higher Sharpe Ratios and deliver superior returns. It's also important to underline how the robustness test didn't respond to the significance of the study: the good outcomes might be limited only to the period taken into consideration or to the chosen assets, requiring additional tests on longer training and testing sets and broader portfolios. Notwithstanding this non-significance, it's important to notice how the presented work delivers signals on the possibility to build efficient systematical trading strategies over longer periods.

Improvements can be performed also on the models themselves, with the inclusion of new predictors and the choice of other regressors to achieve a more precise and accurate outcome. Also, different windows can be evaluated, to assess whether the strategy shows systematical robustness, with constant model retraining over constant timeframes, using a rolling window approach. Future research might also be focused on the development of some strategies with the secondary objective to minimize transaction costs without refusing to return achievement.

8 Bibliography

1. Ang, A., Bekaert, G. (n.d.). How do Regimes Affect Asset Allocation? 28.

2. Anson. (n.d.). Strategic versus Tactical Asset Allocation. 2004.

3. Baker, M., Wurgler, J. (n.d.). Investor Sentiment in the Stock Market. 114.

4. Barberis, N., Shleifer, A., Vishny, R. (1998). A model of investor sentiment. Journal of Financial Economics.

5. Basak, S., Kar, S., Saha, S., Khaidem, L., Dey, S. R. (2019). Predicting the direction of stock market prices using tree-based classifiers. The North American Journal of Economics and Finance, 47, 552–567. https://doi.org/10.1016/j.najef.2018.06.013

6. Chong, J., Phillips, G. M. (2014). Tactical Asset Allocation with Macroeconomic Factors. The Journal of Wealth Management, 17(1), 58–69. https://doi.org/10.3905/jwm.2014.17.1.058

7. Colucci, S., Brandolini, D. (2011). A Risk Based Approach to Tactical Asset Allocation. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.1965423

8. Dahlquist, M., Harvey, C. R. (2001). Global Tactical Asset Allocation. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.795376

9. Daul, S., Jaisson, T., Nagy, A. (2022). Performance attribution of machine learning methods for stock returns prediction. The Journal of Finance and Data Science, 8, 86–104.

https://doi.org/10.1016/j.jfds.2022.04.002

10. Estrella, A., Mishkin, F. S. (n.d.). The Yield Curve as a Predictor of U.S. Recessions. 6.

11. Faber, M. T. (2007). A Quantitative Approach to Tactical Asset Allocation. The Journal of Wealth Management, 9(4), 69–79. https://doi.org/10.3905/jwm.2007.674809

12. Fama, E. F., French, K. R. (2010). Luck versus Skill in the Cross-Section of Mutual Fund Returns. The Journal of Finance, 65(5), 1915–1947. https://doi.org/10.1111/j.1540-6261.2010.01598.x

13. Goyal, A., Welch, I., Zafirov, A. (2021). A Comprehensive Look at the Empirical Performance ofEquity Premium Prediction II. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3929119

14. 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

15. Israel, R., Kelly, B. T., Moskowitz, T. J. (2020). Can Machines 'Learn' Finance? SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3624052

16. Lahiri, Sandip Ghanta, Kartik. (2008). The Support Vector Regression with the parameter tuning assisted by a differential evolution technique: Study of the critical velocity of a slurry flow in a pipeline. Chemical Industry and Chemical Engineering Quarterly. 14. 10.2298/CICEQ0803191L.

17. Liu, J., Longstaff, F. A., Pan, J. (2003). Dynamic Asset Allocation with Event Risk. The

Journal of Finance, 58(1), 231–259. https://doi.org/10.1111/1540-6261.00523

18. Mascio, D. A., Fabozzi, F. J. (2019). Sentiment indices and their forecasting ability. Journal of Forecasting, 38(4), 257–276. https://doi.org/10.1002/for.2571

19. Mascio, D. A., Fabozzi, F. J., Zumwalt, J. K. (2021). Market timing using combined forecasts and machine learning. Journal of Forecasting, 40(1), 1–16. https://doi.org/10.1002/for.2690

20. Moreira, A., Muir, T. (2019). Should Long-Term Investors Time Volatility? Journal of Financial Economics, 131(3), 507–527. https://doi.org/10.1016/j.jfineco.2018.09.011

21. Neely, C. J., Rapach, D. E., Tu, J., Zhou, G. (2014). Forecasting the Equity Risk Premium: The Role of Technical Indicators. Management Science, 60(7), 1772–1791. https://doi.org/10.1287/mnsc.2013.1838

22. Rapach, D. E., Strauss, J. K., Zhou, G. (2013). International Stock Return Predictability: What Is the Role of the United States?: International Stock Return Predictability. The Journal of Finance, 68(4), 1633–1662. https://doi.org/10.1111/jofi.12041

23. Resnick, B. G., Shoesmith, G. L. (2002). Using the Yield Curve to Time the Stock Market. Financial Analysts Journal, 58(3), 82–90. https://doi.org/10.2469/faj.v58.n3.2540

24. Shen, P. (n.d.). Market-Timing Strategies That Worked? 39.

25. Welch, I., Goyal, A. (2008). A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. Review of Financial Studies, 21(4), 1455–1508. https://doi.org/10.1093/rfs/hhm014

9 Appendix

Table 4: Summary of employed variables

Name of the Variable	Group	Dependent/Independent	Source
S&P500 Returns	Index Returns	Dependent	FRED
STOXX600 Returns	Index Returns	Dependent	Refinitiv Eikon
VIX	Volatility Index	Independent	FRED
VSTOXX	Volatility Index	Independent	EUROSTAT
10Y Bond Yield	Bond Yield	Independent	FRED
3M Bond Yield	Bond Yield	Independent	FRED
Spread	Bond Yield	Independent	Calculated
Gold Futures	Futures Returns	Independent	Refinitiv Eikon
Oil Futures	Futures Returns	Independent	Refinitiv Eikon
GDP	Macroeconomics	Independent	FRED
PCE	Macroeconomics	Independent	FRED
Unemployment Rate	Macroeconomics	Independent	FRED
Industrial Production	Macroeconomics	Independent	FRED
CPI	Macroeconomics	Independent	FRED
Breakeven Inflation	Macroeconomics	Independent	FRED
E/P Ratio	Technical Indicator	Independent	FRED, Statistical Data Warehouse
E/P – 3M Bond Yield	Technical Indicator	Independent	Calculated
Baker-Wurgler Index	Sentiment Index	Independent	Jeffrey Wurgler Website
Aruoba Index	Sentiment Index	Independent	Philadelphia FED
S&P500 Lag Returns	Index Returns	Independent	FRED

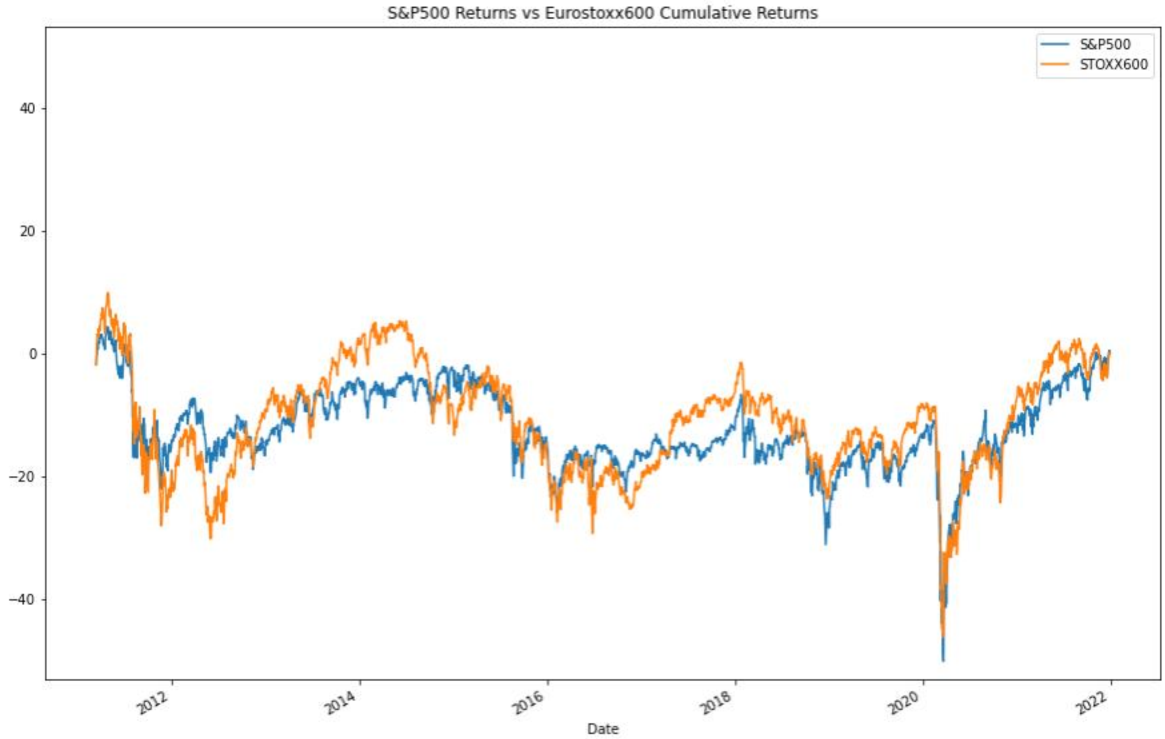

Figure 10: SP500 and VIX Trends

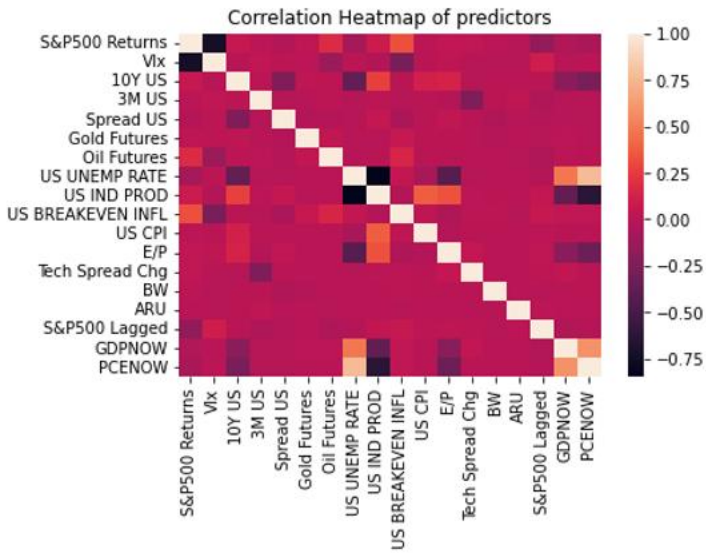

Figure 11: Correlation Heatmap of variables

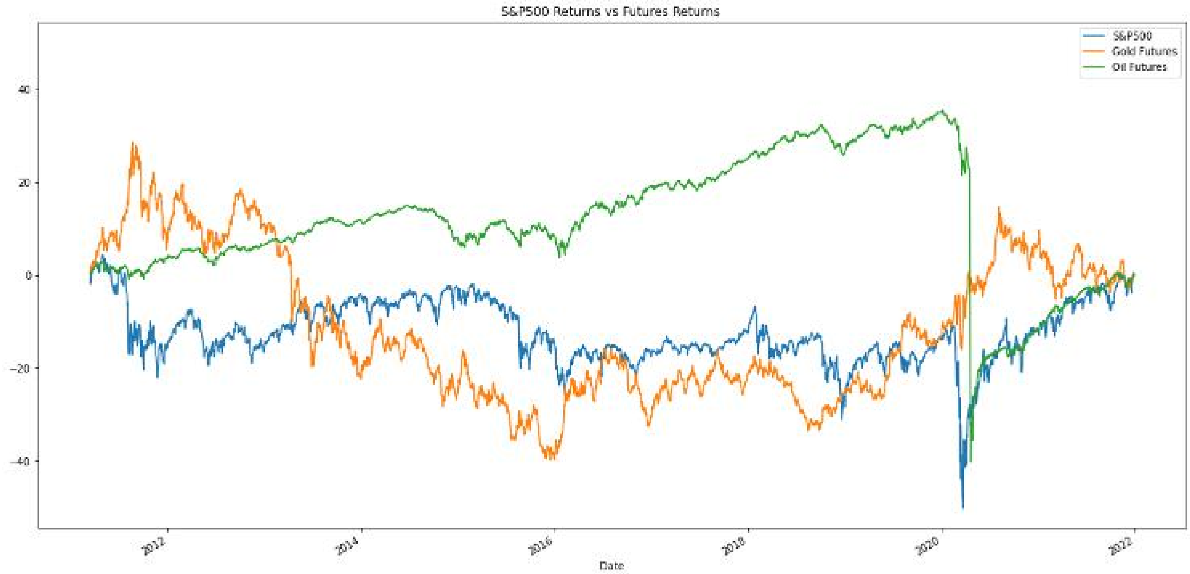

Figure 12: Future Price Trends and SP500

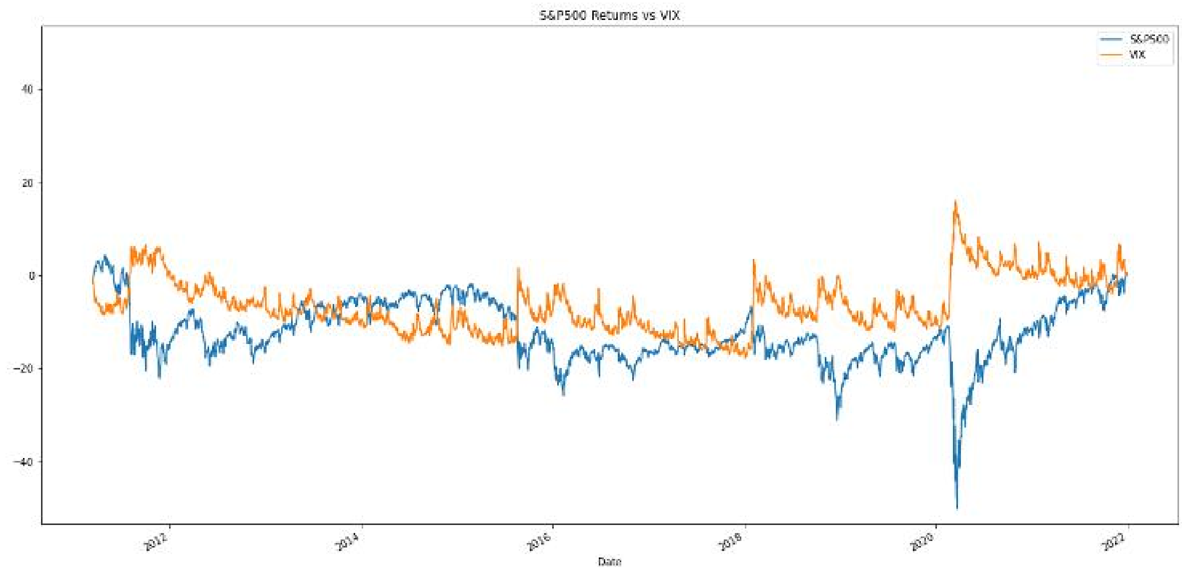

Figure 13: VIX and SP500 Price Trends

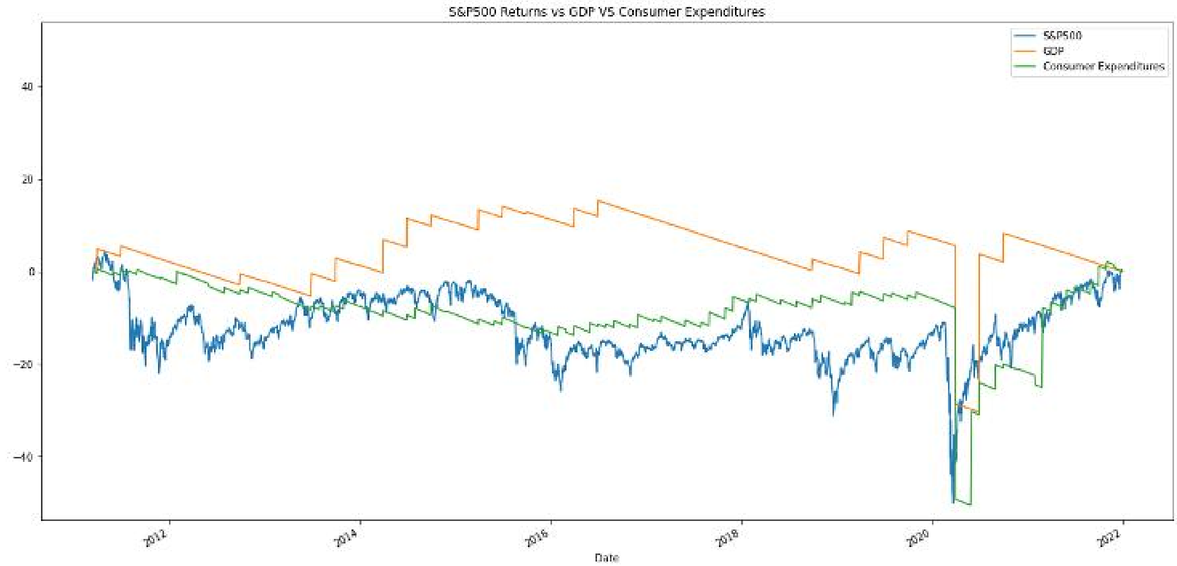

Figure 14: SP500, GDP and Consumer Expenditures Trends

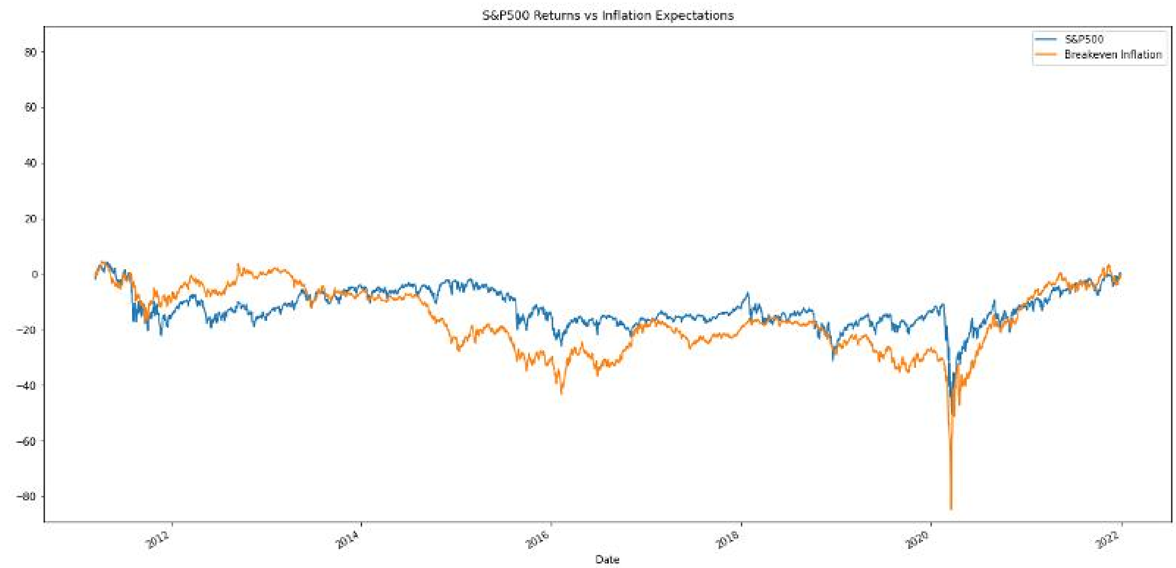

Figure 15: SP500 Price and Inflation Expectations Trends

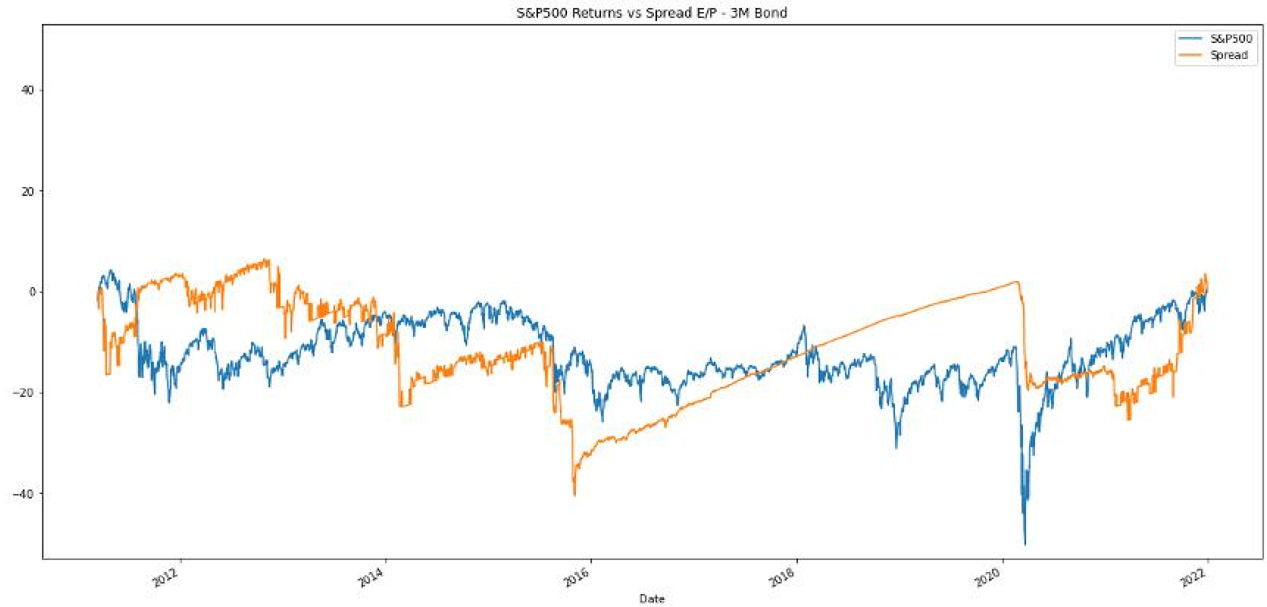

Figure 16: SP500 and Technical Indicators

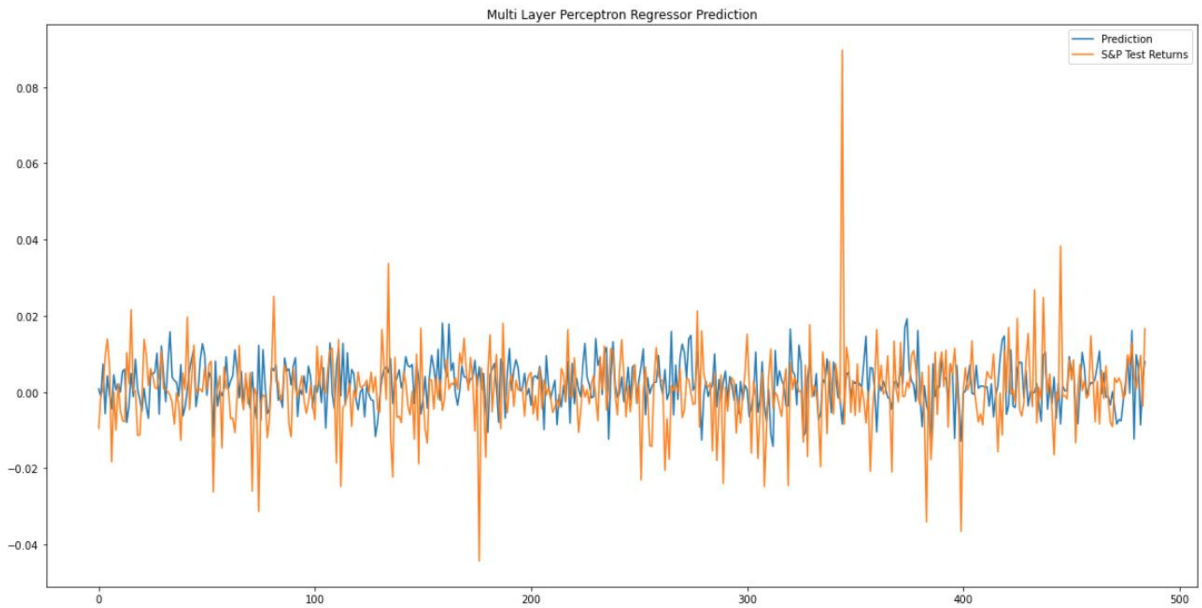

Figure 17: Neural Network Prediction on US Returns

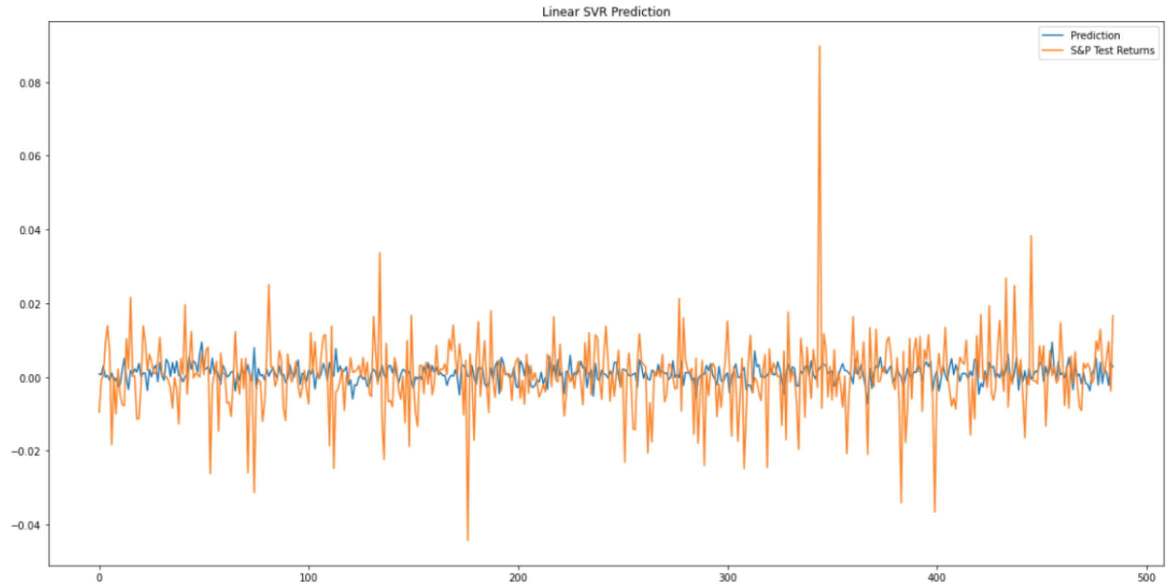

Figure 18: Support Vector Machine Prediction on US Returns

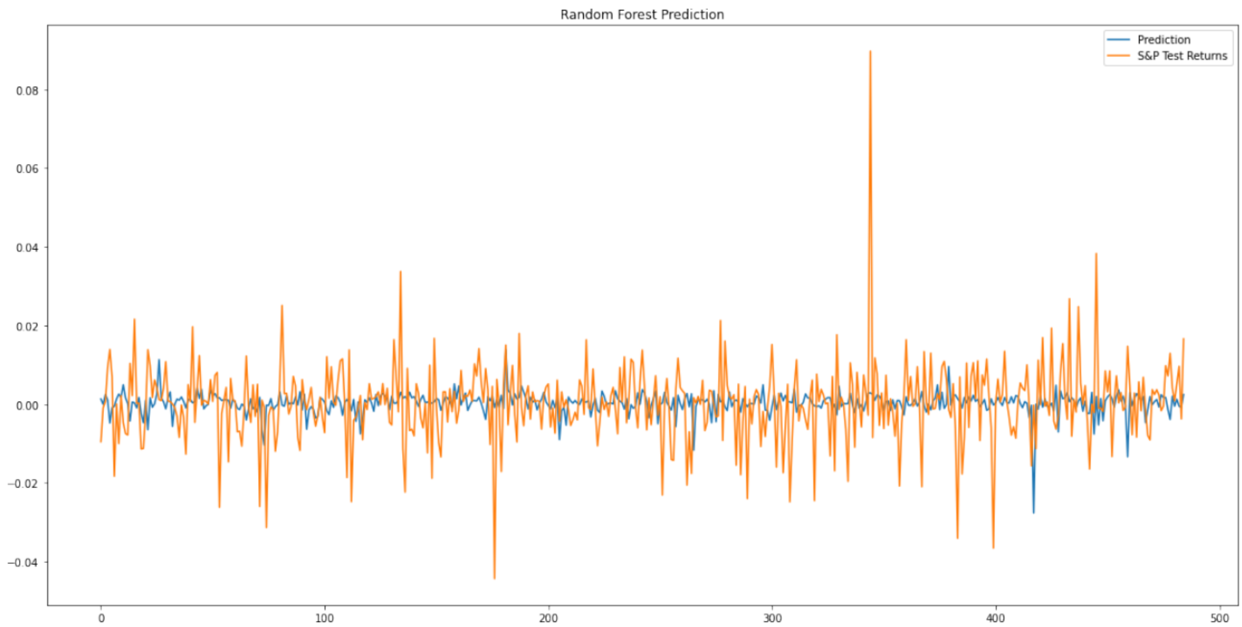

Figure 19: Random Forest Prediction on US Returns

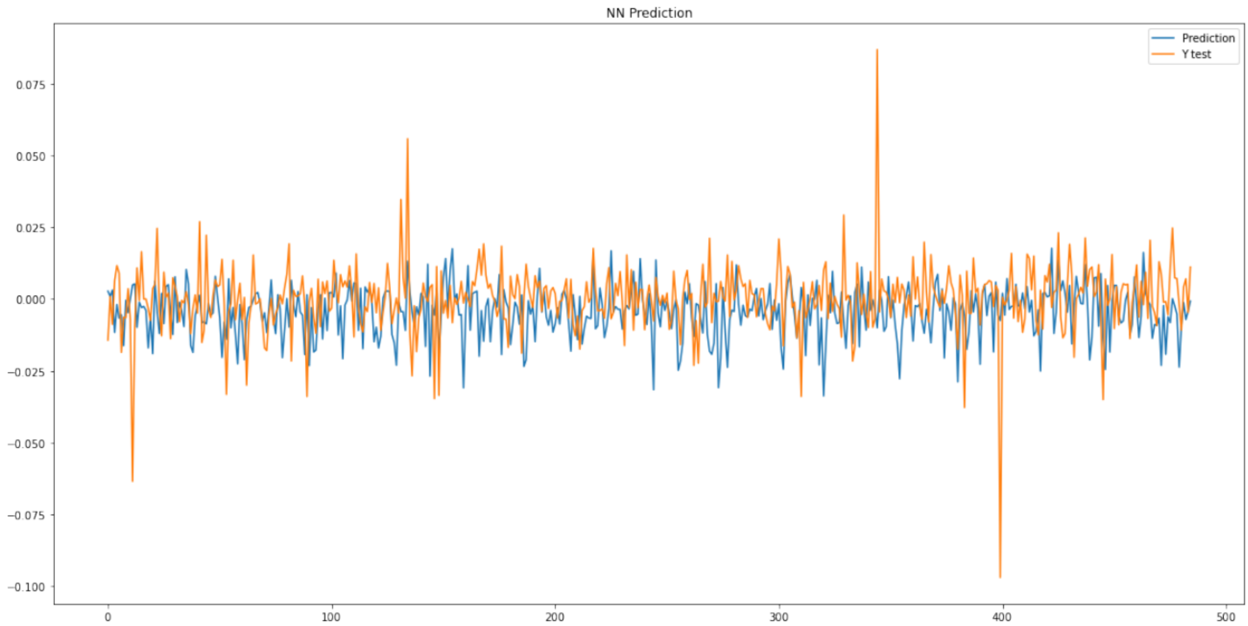

Figure 20: Neural Network Prediction on EU Returns

Figure 21: Support Vector Machine Prediction on EU Returns

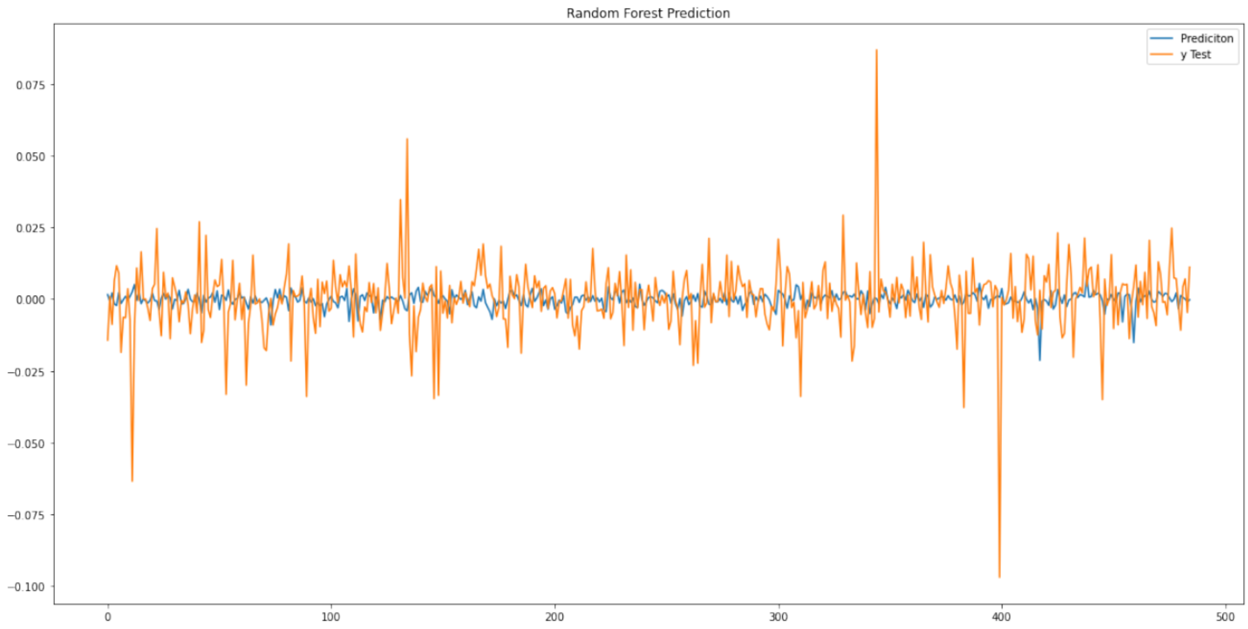

Figure 22: Random Forest Prediction on EU Returns