Iosr Journal Of Economics And Finance (Iosr-Jef)
E-Issn: 2321-5933, P-Issn: 2321-5925.Volume 16, Issue 1 Ser. 3 (Jan. – Feb. 2025), Pp 59-70
www.Iosrjournals.Org

A Deep Learning Based Asset Allocation Methodology For Investment Portfolio Optimization Under Uncertainties

Marcus V. de Oliveira[1] And Gilson A. O. P. da Costa[2]
Instituto De Matemática E Estatística – Universidade Do Estado Do Rio De Janeiro (UERJ)
CEP 20550-900 - Maracanã, Rio De Janeiro – RJ – Brazil
[1]marcus.oliveira@pos.ime.uerj.br ; [2]gilson.costa@ime.uerj.br

Abstract
This paper introduces a novel methodology for determining optimal asset allocation within diversified investment portfolios by integrating Modern Portfolio Theory with a Deep Learning model. This approach enhances portfolio diversification by considering both historical asset correlations and forecasts of individual asset volatility. Distinct from traditional methods that rely solely on historical data, our methodology incorporates current trading conditions and market dynamics, as reflected in asset prices. The S&P-500 index serves as the benchmark for this study, with the primary aim of achieving returns comparable to the benchmark while minimizing risk through a strategic combination of multiple assets.

Date of Submission: 22-01-2025 Date of Acceptance: 27-01-2025

I. Introduction

The use of Deep Learning techniques for investment portfolio optimization is not an original idea itself, many researchers have already worked on that, applying very different approaches. As example is presented in [ZR20] where the authors aimed optimize an ETF portfolio Sharpe ratio. It is important to recognize that predicting future price of assets is a seemingly impossible task. Therefore, our objective is not to accurately predict the price of a stock or an ETF, commodity or any other type of asset, but to infer its probabilistic distribution, based on the model estimates and its prediction error on the test dataset. And this is relevant because, by performing an inference based on an intelligent model that considers the price dynamics, which was learned during the training process, we will be actively incorporating market conditions at the present time, not just the historical performance of the market and the assets.

In [FP18], the authors proposed a methodology that considers the dynamics of variables beyond the asset price, such as dividend yield ratio, book to market ratio, consumer price index (CPI) and "consumption wealth, income ratio". In the initial phases of this work, we strongly consider addressing variables linked to the fundamentals of assets, in addition to macroeconomic data, which certainly affect the behaviour of asset prices in the financial market. However, there is an additional difficulty in using these data, its periodicity.

Overall, the use of Machine Learning in portfolio optimization is a rapidly growing field, with many researchers and practitioners exploring new techniques and approaches to improve investment decision-making. Machine Learning has the potential to transform the investment management industry by providing portfolio managers with powerful tools to make better investment decisions and improve their overall performance. Figure 1 shows the division by subject of the articles that make up the pre-selected list for the bibliographic review by the end of 2023, not limited to those directly related to the topic of application of this work. It can be noted that the topic with the largest number of works is Trading, a category that consists of works that address applications whose objective is to deal with the modelling of high-frequency assets or derivatives, that is, those interested in the extra-short-term behaviour assets, usually with the aim of profiting from arbitrages, quick trades or market anomalies.

There are at least two reasons for the great interest in high-frequency trading using machine learning methods. High-frequency trading involves using algorithms to make trades at high speeds, often measured in microseconds or milliseconds, to take advantage of small market inefficiencies. These trades can generate significant profits in a short amount of time, making it an attractive area for research and investment. But there is another technical reason for this: the availability of large amounts of data for high-frequency modelling. If the model is trained using only daily data, the availability of moving windows will be very limited (on average, there are 252 trading days per year). However, for high-frequency trading, there are several trading steps in a day, which allows overcoming one of the biggest challenges of Deep Learning: data scarcity. Essentially an Artificial

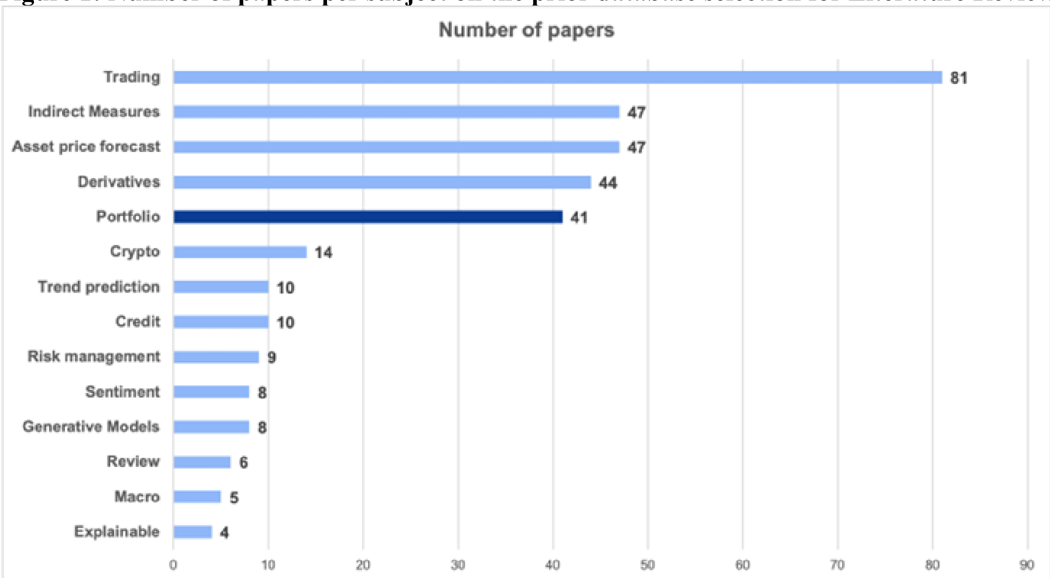
Intelligence model can learn anything, as long as two conditions are met: (i) There are recognizable patterns in the data and (ii) There is a sufficient amount of data.

The categories Indirect Measures and Asset Price Forecast have the same frequency. Asset Price Forecast includes all papers that aim to forecast the price of an asset, regardless of the methodology used. On the other hand, Indirect Measures is the category for all papers that aim to forecast something related to the future behaviour of the asset, but not its price directly. The most common indirect measure present in these papers is volatility. Many papers aim to forecast the volatility of an asset or the market as a whole, because of its relevance in tasks like Risk Management, Hedging and Portfolio Management.

Although information about price would be much more useful than volatility, it is important to consider that the closer the forecast is to the future price, the lower the forecast reliability will be. The third and fourth most frequent subjects are Derivatives and Portfolio, respectively. The original objective of derivatives is to allow investors to protect their positions on the underlying asset. There are numerous papers addressing options pricing, aiming to identify discrepancies between the fair price of an option and its market price to trade them when such discrepancies are found, from a speculative perspective.

The Portfolio group includes publications related to the optimization of investment portfolios, a subject much more related to the present work than the other topics. Despite not being the most frequently applied topic among those investigated, portfolio optimization has many publications, with different methodologies that will be better described in the specific section on the topic. Observing the Figure 1 what can be noticed is that there is a group of subjects with a relatively similar number of publications, Indirect Measures, Asset Price Forecast, Derivatives and Portfolio, all of them with much less publications than Trading. We can understand that this behaviour is greatly influenced by what has already been discussed regarding the aspects that make the use of Deep Learning attractive for algorithmic trading.

Figure 1: Number of papers per subject on the prior database selection for Literature Review



Some research has been done on the application of Data Mining techniques on the financial markets. Some of them are more directly related to time series prediction techniques, with the asset's future price being the objective function, which is not the case of this paper. In other works, like in [ZR20], the objective is to maximize the Sharpe Ratio, aiming to obtain a portfolio that maximizes the risk reward, that is, maximizes the expected increase in return due to the increase in volatility. This could be very useful, for instance, when performing a pre-screening to select the most promising assets to compose a portfolio.

II. Fundamentals

a. Fundamentals

The Modern Portfolio Theory (MPT) was developed by Harry Markowitz and its fundamentals were introduced in [Mar52] which revolutionized the way investors think about constructing portfolios and managing risk. Although the Finance research has made significant improvements on portfolio and risk management, Markowitz's modern portfolio theory remains a key concept in portfolio management today. In general, the concepts related to the trade-off between risk and return are immune to the evolution of technology, as they are based on the logic that riskier assets have a greater expectation of return, simply by the laws of the market.

When combining many assets, a very relevant question is: for a given desired expected return $\boldsymbol{r_t}$, what is the allocation (weights of each asset) that leads to the smallest portfolio variance? This is very relevant because, as mentioned earlier, investors are risk averse, so, to obtain a certain desired expected return, they are always interested in the portfolio with the lowest variance, that is, the lowest volatility (or lowest risk). Once we have, from historical data, the expected return on assets, their variances and correlation coefficients can be used to answer that question. Once we know those statistical measures, it is enough to solve a simple minimization problem, as in Equation (1), where $\boldsymbol{w} \in \mathbb{R}^{\boldsymbol{n}}$ is the vector weights of the $\boldsymbol{n}$ assets available to include in the portfolio, $\boldsymbol{\Sigma}$ is the covariance matrix and $\boldsymbol{\mu} \in \mathbb{R}^{\boldsymbol{n}}$ is the vector of mean historical returns.

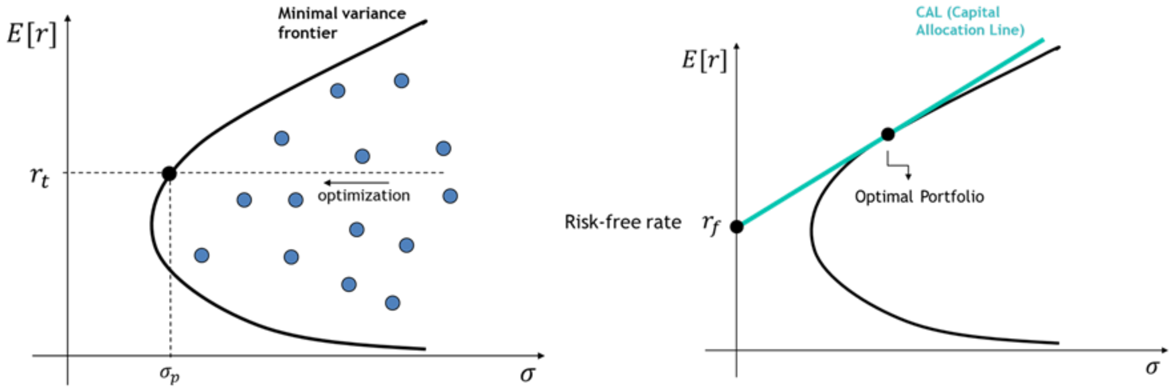
$$\min \boldsymbol{\sigma_P^2} = \boldsymbol{w^T \Sigma w} \;\; \boldsymbol{s.t.} \;\; \boldsymbol{E[r_P] = \mu^T w} \geq \boldsymbol{r_t} \,, \;\; \sum_{l=1}^{N} \boldsymbol{w_k = 1} \,, \;\; \boldsymbol{w_i} \geq \boldsymbol{0} \tag{1}$$

Alternatively, given a risk aversion coefficient $\boldsymbol{\delta} \in \mathbb{R}$ and a utility function $\boldsymbol{U} = \boldsymbol{r_p} - \frac{\boldsymbol{\delta}}{\boldsymbol{2}}\boldsymbol{\sigma_P^2}$, where $\boldsymbol{r_p}$ and $\boldsymbol{\sigma_P^2}$ are the expected value and variance of the portfolio, the portfolio optimization problem can be defined as a maximization problem of utility, as described in Equation (2).

$$\max \boldsymbol{U_P} = \boldsymbol{\mu^T w} - \frac{\boldsymbol{\delta}}{\boldsymbol{2}}\boldsymbol{w^T \Sigma w} \;\; \boldsymbol{s.t.} \;\; \sum_{l=1}^{N} \boldsymbol{w_k = 1} \,, \;\; \boldsymbol{w_i} \geq \boldsymbol{0} \tag{2}$$

Figure 2 presents Capital Allocation Line (CAL), that is a line that depicts the risk-reward tradeoff of assets that carry idiosyncratic risk. The slope of the CAL is called the Sharpe ratio, which is the increase in expected return per additional unit of standard deviation (reward-to-risk ratio). According to the MPT, rational risk-averse investors should hold portfolios that fall on the efficient frontier (since they provide the highest possible expected returns for a given level of standard deviation). Therefore, the optimal portfolio (also called the "market portfolio") is the combination of assets which combines one risk-free asset with one risky asset. The slope of CAL is called the Sharpe Ratio (SR), which was introduced by William F. Sharpe as described in **[Sha94]**. The Sharpe Ratio is defined as the excess return of a portfolio over the risk-free rate per unit of its standard deviation. It measures the incremental expected return gained per unit of volatility and is used to evaluate the risk-adjusted performance of a portfolio. It can be calculated by the Equation (3), where $E[r_p]$ and σ_p are the expected return and standard deviation of the portfolio and rf is the risk-free rate.

Figure 2: Optimization problem to generate the efficient frontier and Optimal Allocation for the risk assets portfolio and Capital Allocation Line (CAL).



$$SR = \frac{E[r_p] - r_f}{\sigma_p} \tag{3}$$

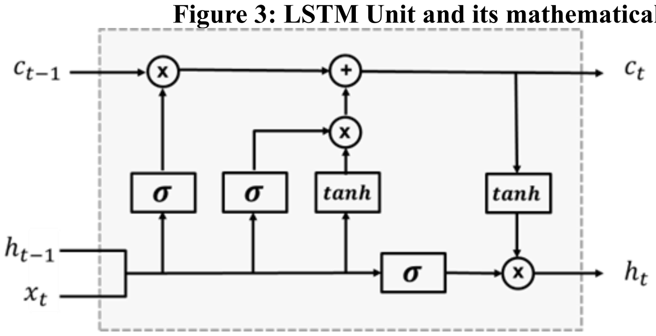
b. Long Short Term Memory (LSTM)

Unlike standard feedforward neural networks, that a given layer feeds the following layer, Recurrent Neural Networks (RNN) have feedback connections. Long Short-Term Memory (LSTM) is a type of RNN that uses the concept of gates of reinforcement and forgetting to incorporate past information and use it in future predictions, allowing to process not only single data points, but also entire sequences. The network training will be carried out through sliding windows that will generate a dataset with a certain number of historical price data (in this case 20 trading days) and only one forecast step.

Therefore, the LSTM can learn to predict based on a sequence of data instead of just one input. It is very useful for applications related speech or video. For example, LSTM is applicable to tasks such as unsegmented, connected handwriting recognition, speech recognition and anomaly detection in network traffic. In the case of the problem we are addressing, the LSTM architecture is quite adequate, as we want to predict the future prices behavior based on the dynamics of asset prices in the past.

Figure 3 shows LSTM architecture unit cell and its mathematical formulation. It consists of gates containing regular neural network operation (weighted sums and bias to be learned during training) and activation functions. In all equations, σ represents the sigmoid function and the operator o denotes the Hadamard product (element-wise product).

Figure 3: LSTM Unit and its mathematical formulation of each gate.



$$\begin{cases} f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \\ i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \\ o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \end{cases}$$

$$\begin{cases} \bar{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \\ c_t = f_t \circ c_{t-1} + i_t \circ \bar{c}_t \end{cases}$$

$$h_t = o_t \circ \tanh(c_t)$$

The forget gate, f_t, decides which information needs to be preserved and which can be ignored. The information from the current input x_t and hidden state h_{t-1} are combined like in a simple perceptron and passed through the sigmoid function. The sigmoid generates values between 0 and 1. It concludes whether the part of the old output is necessary, making its output closed to 1 or closed to 0 otherwise. The weight matrixes W_f and U_f as well as the bias vector b_f will be adjusted during the training process.

The Input gate, i_t learns how to preserve relevant information. The architecture is exactly the same of the forget gates, the operation performed by a simple perceptron. The weight matrix W_i and U_i as well as the bias vector b_i will be adjusted during the training process.

The objective of output gate, o_t, is extracting useful information from the cell state at $t-1$ and from the input at t. The architecture is exactly the same of the Input and Forget Gates. The weight matrix W_o and U_o as well as the bias vector b_o will be adjusted during the training process.

The purpose of memory cell is to consolidate what has been learned that must be forgotten and remembered. Therefore, for this step you need the results of the input and forget gates, as well as a simple multilayer perceptron type unit with tanh activation function. Usually, the last trading days are more relevant for predicting the next behavior. However, it is not necessary to do any treatment in the data for the model to learn this fact. This will be done automatically when training the gate weights and bias. Furthermore, depending on the trading volume of assets, older prices can influence more or less. However, we understand that 20 days is enough to incorporate short-term information into the price, and the long-term characteristics must be addressed *a priori*, when selecting assets based on their fundamentals.

Hidden state vector calibrates the information that will be passed to the output. This gate has no parameters to be learned during the training process (the learning parameters were learned by the output gate).

III. Methodology

a. Multi-Step Prediction under Uncertainties

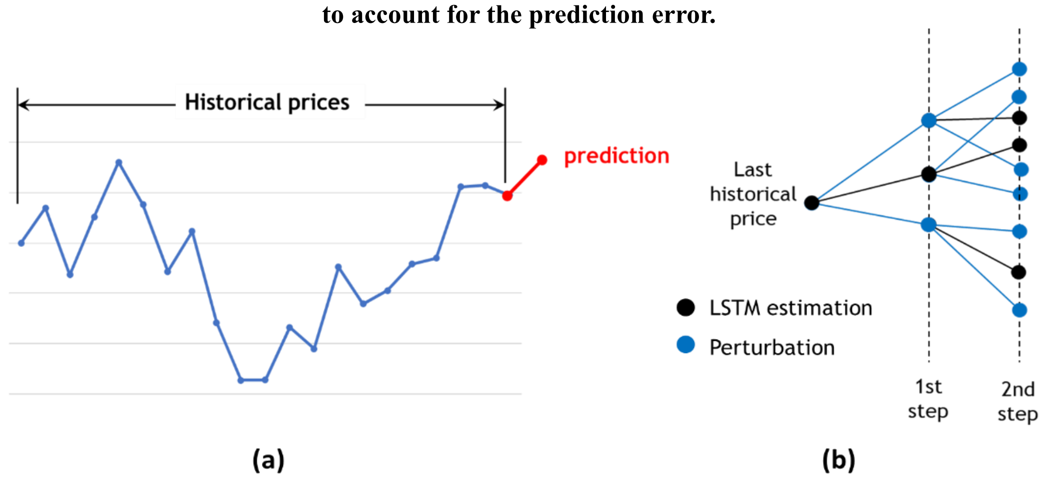
Once we have estimates of future price dispersion, we can calculate the expected return and standard deviation. These values are used to apply classical portfolio theory. The methodology proposed in this paper consists of using the LSTM to predict, based on N previous time intervals, the next price of the asset. However, as our objective is to use MPT for long-term portfolio management, it is important to predict the behavior of assets some periods beyond the historical period (1 month, since we are using a montly based approach).

Thus, an important question arises: How to turn the single step estimate provided by the trained LSTM model into a multi-step estimate? First of all we have to recognize that, although LSTM has performed very well for this kind of task, the market behavior cannot be fully predicted and some significant errors are expected for some market conditions. Therefore, the proposed methodology consists of applying LSTM multiple times and adding two additional estimates based on the LSTM performance over the test dataset.

In summary, one point at $t-1$ generates three estimates at t: (i) LSTM estimate, (ii) LSTM estimate minus one LSTM error standard deviation and (iii) LSTM estimate plus LSTM error standard deviation. By adding these two new estimates we are creating a range of possible realizations that will lead at the end to a number of different prices for the asset under evaluation. Having these many estimates we can calculate both expected return and variance to be inputted into the MPT model to calculate the optimum allocation.

Figure 4 schematically presents the proposed methodology. The starting point, which corresponds to the last point in the history, gives rise to 3 points, with one point (black) coming from the evaluation of the trained model and 2 points obtained by perturbation of this point (adding or subtracting a standard deviation). Thus a point generates 3, which generates 9, which generates 27, and so on.

Figure 4: Diagram of the multi-step prediction process based on the proposed workflow, including the forecast generated by the proposed architecture, as well as the high and low perturbations to account for the prediction error.



IV. Applications to ETF Portfolio Optimization

a. Database

Since the main objective of portfolio management is to obtain good returns while reducing risk, by reducing volatility, diversification is the fundamental tool. Mutual funds usually have hundreds and even thousands of assets in their portfolios, the fundamentals of which are carefully analyzed, as well as the associated risks involved. In this work, we will use ETFs from 10 different countries. Each ETF contains a large amount of assets (SPY has, for example, shares from 500 companies).

Thus, we understand that sector diversification is covered because we are using ETFs and geographic diversification is covered because we are using ETFs from different countries, both developed and developing. Therefore, the dataset adopted mimic the allocation step, where the analyst will decide how much capital to invest in each asset of the selected list. Table 1 shows the ETFs selected to carry out this work.

Table 1 Selected ETFs for portfolio optimization. Each ETF consists of a broad basket of equities from various countries, ensuring good geographic diversification.

Ticker	Title	Ticker	Title
SPY	SPDR S&P 500 ETF Trust	EWJ	iShares MSCI Japan ETF
EWZ	iShares MSCI Brazil ETF	EWA	iShares MSCI Australia ETF
EWN	iShares MSCI Netherlands ETF	FXU	First Trust Utilities AlphaDEX Fund
EWQ	iShares MSCI France ETF	EWU	iShares MSCI United Kingdom ETF
EWG	iShares MSCI Germany ETF	EZA	iShares MSCI South Africa ETF

It can be noted in Figure 5 that these assets have very different volatility levels, as well as historical performances. There are assets that have appreciated strongly in the past, such as SPY and FXU, and others that had great volatility and did not provide good returns for investors in the considered time horizon. Thus, it is expected that, when carrying out an allocation analysis of a portfolio that allows for long and short positions, the best allocation will have a high weight of assets that had high return and low volatility and less weight for the more volatile assets, in addition to short position in assets that had very poor returns.

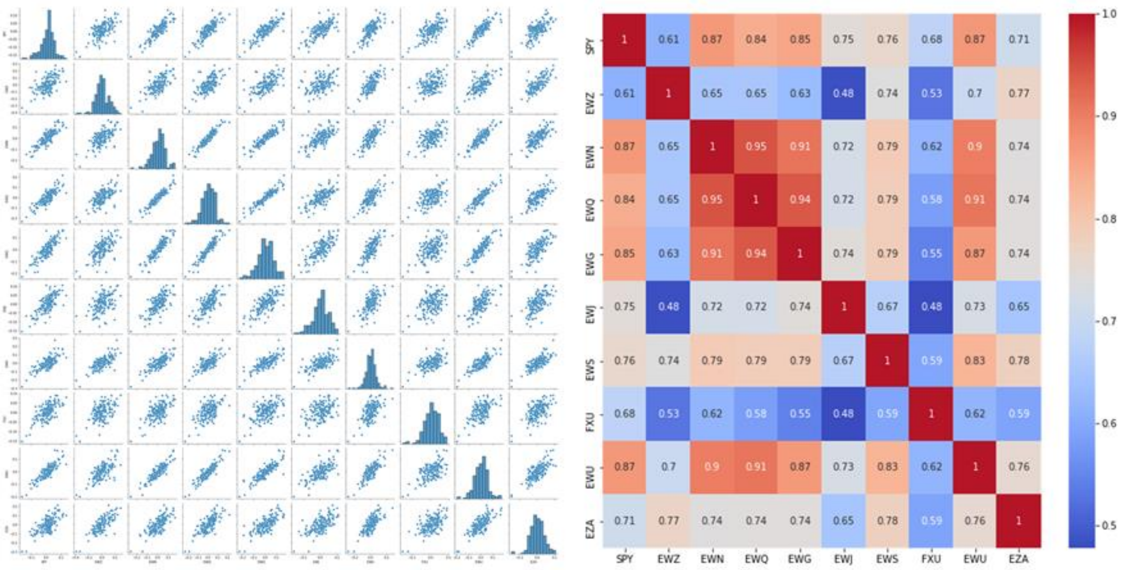
Figure 5: The time series corresponding to the 10 selected ETFs illustrate the diverse price behavior of the assets chosen to compose the portfolio.

b. Exploratory Data Analysis

The data was collected from the Yahoo Finance platform, which provides prices, trading volumes and other information at various frequencies. Since the objective of this work is not to deal with High Frequency Trading, we chose to use the closing price of related assets trades with daily frequency.

Since we are interested in diversification, it is very important to look carefully at the correlation matrix. If we look at European assets, there are very high correlations between them. One can imagine at least two reasons to explain that behavior (i) European assets have a portion of fundamentals associated to the EU and its macroeconomic dynamics is relevant to influence the behavior of the shares that compose the ETFs and (ii) those assets are denominated in Euros, and the ETF itself is denominated in USD, making the prices of those assets in USD incorporate not only the performance of European assets, but also the exchange rate, which is the same for all ETFs.

Figure 6: Correlation between assets returns for ETFs selected to compose the portfolio.



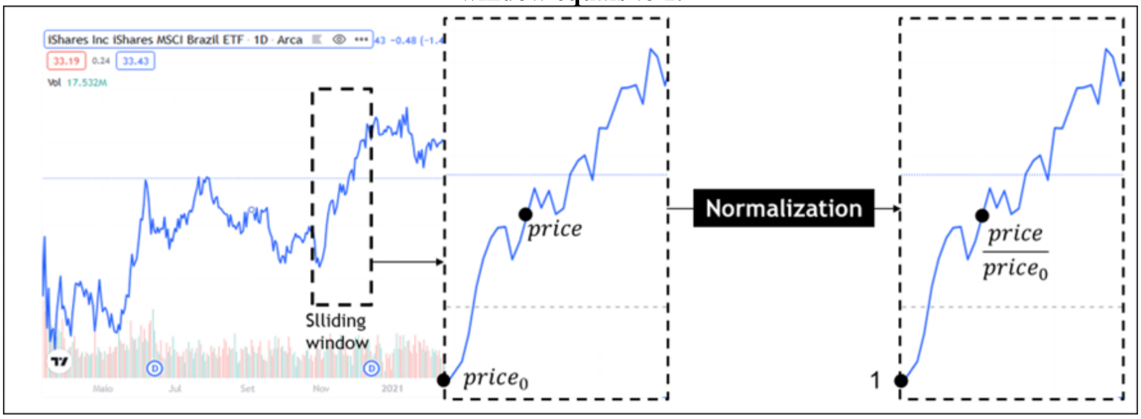
Another conclusion that we can take from the correlations presented in Figure 6 is that assets from developing countries have a lower correlation with assets from developed countries, as observed with EWZ

(Brazil) and EWJ (Japan).This tends to be positive for diversification, especially when short positioning is allowed, which is the case of the investments in which we are interested. However, assets from developing countries tend to be very volatile, both because they are less mature and riskier markets, and because these assets are originally denominated in emerging makers currencies, whose exchange rate against the dollar tends to vary widely.

c. Pre-processing

Since training will take place using sliding windows and prices vary in magnitude over time, it is important to make a normalization The proposed normalization consists of making the first price of the sliding window equal to 1. This normalization is important since the behavior of the price trajectory is what really matters in this case, not the magnitude. One option that some studies adopt is to work with returns or log-returns. However, return data rarely present identifiable patterns, which is why we chose to normalize the data starting at 1.

Figure 7: Proposed Normalization of the price time series to make the starting price of each sliding window equals to 1.



Since the assets in question are all traded in the United States, there was no need to deal with cases in which there is no data for a particular asset due to non-trading days. Furthermore, because the data came from a structured data platform and maintained by Yahoo, it was not necessary to do any missing data processing. All data on trading days of interest were available to carry out the work. Naturally, days where no trading took place are disregarded and not reported by Yahoo Finance.

d. Model's Architecture

Figure 8 presents the proposed architecture for time series forecast is composed of layers with 50 LSTM model units and applied to the input sequence of normalized prices. At the end, a dense layer was included, which outputs the desired price in the next section of the market. The model has a total of 71051 parameters.

Figure 8: Model's arquitecture for time series forecast applying LSTM model and Multilayer Perceptron

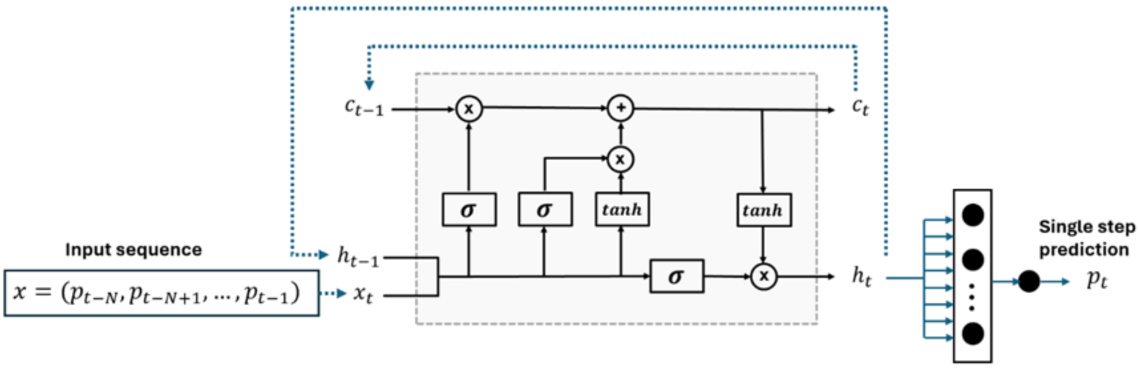
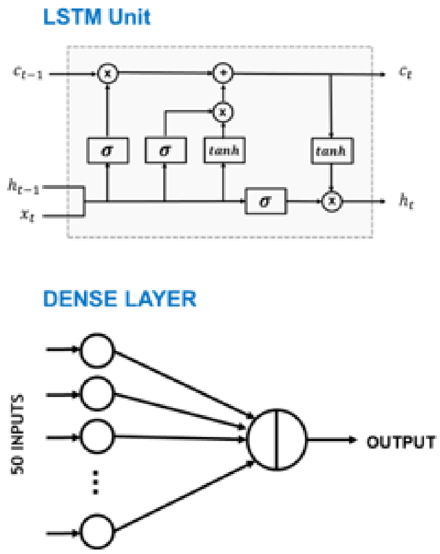


Figure 9 presents a more detailed perspective of the architecture, including all layers, output shapes and number of parameters to estimate. Since the number of parameters in the model is significant, which makes it very prone to overfitting, overfitting prevention methods such as dropout (with a probability of 20%) and L1 and L2 regularizations of the loss function were used. Although complex patterns of price behavior can be identified and many parameters are needed for this, it is essential that the model does not have prohibitive levels of overfitting, which would make the conclusions invalid.

Figure 9: Detailed architecture with the number of parameters and output shape of each layer, as well as the total number of parameters to estimate during training.

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 20, 50)	10400
dropout (Dropout)	(None, 20, 50)	0
lstm_1 (LSTM)	(None, 20, 50)	20200
dropout_1 (Dropout)	(None, 20, 50)	0
lstm_2 (LSTM)	(None, 20, 50)	20200
dropout_2 (Dropout)	(None, 20, 50)	0
lstm_3 (LSTM)	(None, 50)	20200
dropout_3 (Dropout)	(None, 50)	0
dense (Dense)	(None, 1)	51

Total params: 71,051
Trainable params: 71,051
Non-trainable params: 0



V. Results

Figure 10 shows the extrapolation step under uncertainty for the EWZ asset. Note that there are approximately 20 market sections per month and that the extrapolation was performed to 10 trading sessions (approximately two weeks). This option was motivated by the number of model evaluations that would be necessary for the extrapolation to occur over 20 trading sessions (3^{20}), which would make it computationally difficult. In this way, the expected value and standard deviation of the asset returns were estimated as proportional to the number of periods and proportional to the square root of the number of 10-day periods that exist in a month, as usual in Finance.

Figure 10: Extrapolation under uncertainties of the last window. The histogram shows the returns distribution of EWZ asset that will be used as input to the MPT workflow.

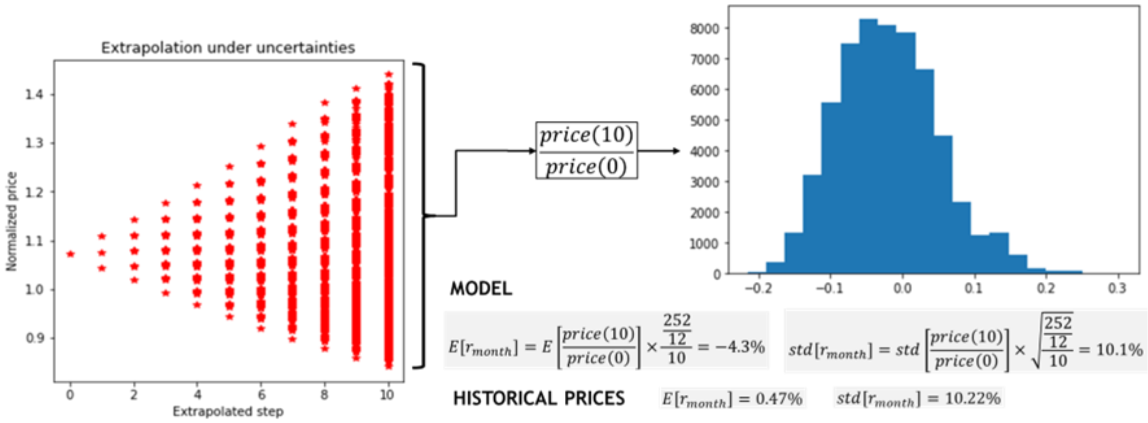


Figure 11 shows a sequence of single-step evaluations for the EWZ asset, and the model's performance on the test dataset (RMSE = 1.9%). Naturally, the figure on the left looks surprising, but we must remember that the simulation is always single-step, so there is no accumulation of errors, as would occur if the simulation were multi-step. However, for the purposes of this work, it was understood that the results were good enough to estimate the expected return and volatility of assets and allow the application of MPT to obtain the optimal allocation.

Figure 11: Example of LSTM performance to estimate EWZ prices after 1 month (about 20 market sections) over the test dataset.

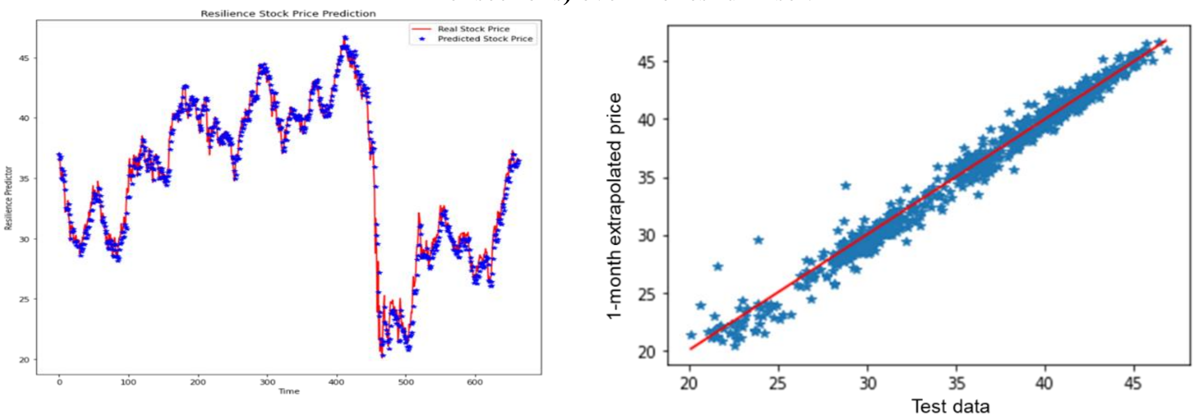
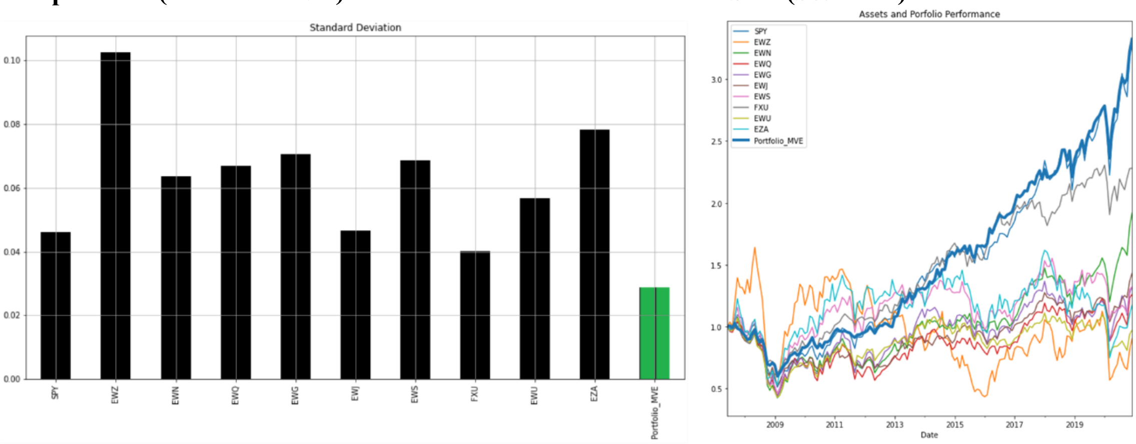


Figure 12 shows the historical performance of the optimal portfolio and its variance. Note that it is possible to obtain a portfolio (Portfolio MVE) that has a historical return very close to the return of the best asset (SPY), but with a lower variance than any individual asset, which is the ultimate goal of diversification. Naturally, this is only possible because we adopted long and short allocations, which allows us to leverage on the full long position. If it was not allowed to have short position (portfolio long only) this would not be possible.

Figure 12: Variance of each individual asset and the portfolio and historical return of the optimal portfolio (Portfolio MVE) that has overall return similar to SPY (S&P-500) and smaller variance



VI. Conclusions

This paper presented a methodology for using the LSTM to forecast time series under uncertainty and use these predictions to estimate expected return and variance for each asset in an investment portfolio, allowing the use of the Modern Portfolio Theory to define the best allocation for the portfolio in question.

Although the use of Machine Learning techniques, especially Deep Learning, in the Financial area has grown significantly in recent years, most of the work is focused on supporting trading strategies, especially in High Frequency Trading (HFT). Fewer works have been published on subjects related to the application of these techniques to longer-term portfolio management, typically used by mutual funds and hedge funds.

Although the performance of the models individually were not exceptional, it is important to consider that the extrapolation was performed under uncertainties, so that the error of the models led to an increase in uncertainty in the multi-step procedure. Thus, it is important to consider that the eventual inaccuracy of the model reflects the unpredictability of price variation, which is greater in certain assets than in others.

Consequently, we consider that, on average, the presented methodology contributes to advancing towards using Deep Learning models to forecast volatility, rather than prices, which we believe is a more useful alternative for managing longer-term investment portfolios.

We have identified an opportunity to improve the technique by incorporating trading volume data from the assets, which we believe is very important to confirm trends (higher trading volumes), or make them less relevant (lower trading volumes) during training. A major advantage of using trading volume is the fact that the sampling is the same as for prices, unlike other variables, such as macroeconomic or corporate performance data, that are available only at some dates, depending on the asset. We hope that the incorporation of the volume will significantly contribute to the improvement of the method's performance.

References

[1] [Aiu12] Fernando Antônio Lucena Aiube. Modelos Quantitativos Em Finanças: Com Enfoque Em Commodities. Bookman, 2012.
[2] [BSUM20] Eric Benhamou, David Saltiel, Sandrine Ungari, And Abhishek Mukhopadhyay. Bridging The Gap Between Markowitz Planning And Deep Reinforcement Learning. 2020.
[3] [CZHW21] Yijian Chuan, Chaoyi Zhao, Zhenrui He, And Lan Wu. The Success Of Adaboost And Its Application In Portfolio Management. 2021.
[4] [FF93] E. F. Fama And K. R. French. Common Risk Factors In The Returns On Stocks And Bonds. Journal Of Financial Economics, 33(1):3–56, 1993.
[5] [Fis18] Krauss C. Fisher, T. Deep Learning With Long Short-Term Memory Networks For Financial Market Predictions. European Journal Of Operational Research, 270:654–669, 2018.
[6] [FP18] He J. Feng, G. And J. Polson. Deep Learning For Predicting Asset Returns. Arxiv Mathematics, Computer Science, Economics. Arxiv:1804.09314v2 [Stat.ML], 2018.
[7] [HC20] J Huang And J. Chai. Deep Learning In Finance And Banking: A Literature Review And Classification. Frontiers Of Business Research In China, 2020.
[8] [HW16] N. Polson Heaton, J. And J. H. Witte. Deep Learning For Finance: Deep Portfolios. Applied Stochastic Models In Business And Industry, 33(1):3–12, 2016.
[9] [Ias20] Miquel Noguer I Alonso And Sonam Srivastava. Deep Reinforcement Learning For Asset Allocation In Us Equities. 2020.
[10] [IV22] Hugo Inzirillo And Ludovic De Villelongue. An Attention Free Long Short-Term Memory For Time Series Forecasting. 2022.
[11] [JXL17] Zhengyao Jiang, Dixing Xu, And Jinjun Liang. A Deep Reinforcement Learning Framework For The Financial Portfolio Management Problem. 2017.
[12] [KKL+23] Jiwon Kim, Moon-Ju Kang, Kanghun Lee, Hyungjun Moon, And Bo-Kwan Jeon. Deep Reinforcement Learning For Asset Allocation: Reward Clipping. 2023.
[13] [KSV16] Y. Kom Samo And A. Vervuurt. Stochastic Portfolio Theory: A Machine Learning Perspective. Proceedings Of The Thirty-Second Conference On Uncertainty In Artificial Intelligence, Page 657–665, 2016.
[14] [LPAK22] Jinho Lee, Sungwoo Park, Jungyu Ahn, And Jonghun Kwak. Etf Portfolio Construction Via Neural Network Trained On Financial Statement Data. 2022.
[15] [LYX+21] Yu Lin, Yan Yan, Jiali Xu, Ying Liao, And Feng Ma. Forecasting Stock Index Price Using The Ceemdan-Lstm Model. The North American Journal Of Economics And Finance, 57:101421, 2021.
[16] [Mar52] H. Markovitz. Portfolio Selection. The Journal Of Finance, 7(1):77–91, 1952.
[17] [Mar18] Howard Marks. The Most Important Thing Uncommon Sense For The Thoughtful Investor. Harper Business, 2018.
[18] [Men19] Xiang Meng. Dynamic Mean-Variance Portfolio Optimisation. 2019.
[19] [NE22] Hamid Nasiri And Mohammad Mehdi Ebadzadeh. Multi-Step-Ahead Stock Price Prediction Using Recurrent Fuzzy Neural Network And Variational Mode Decomposition. 2022.
[20] [OEC21] OECD. Artificial Intelligence, Machine Learning And Big Data In Finance: Opportunities, Challenges, And Implications For Policy Makers. 2021.
[21] [PCL21] Lifang Peng, Kefu Chen, And Ning Li. Predicting Stock Movements: Using Multiresolution Wavelet Reconstruction And Deep Learning In Neural Networks. Information, 12(10), 2021.
[22] [RPAGNV19] Eduardo Ramos-Perez, Pablo J. Alonso-Gonzalez, And Josåle Javier Nunez-Velazquez. Forecasting Volatility With A Stacked Model Based On A Hybridized Artificial Neural Network. Expert Systems With Applications, 129:1–9, 2019.
[23] [RWL20] Jinlong Ruan, Wei Wu, And Jiebo Luo. Stock Price Prediction Under Anomalous Circumstances. Dec 2020.
[24] [Sha94] W. Sharpe. The Sharpe Ratio. Journal Of Portfolio Management, 21(1):49—-58, 1994.
[25] [SZC+21] Srijan Sood, Zhen Zeng, Naftali Cohen, Tucker Balch, And Manuela Veloso. Visual Time Series Forecasting. In Proceedings Of The Second ACM International Conference On AI In Finance. ACM, Nov 2021.
[26] [Tal16] Nassim Nicholas Taleb. Antifragile: Things That Gain From Disorder. Random House, 2016.
[27] [War22] Xavier Warin. Deep Learning For Efficient Frontier Calculation In Finance. 2022.
[28] [ZR20] Zohren-S. Zhang, Z. And S. Roberts. Deep Learning For Portfolio Optimization. The Journal Of Financial Data Science Fall 2020, 2020.
[29] [ZSI23] Cheng Zhang, Nilam Nur Amir Sjarif, And Roslina Binti Ibrahim. Deep Learning Techniques For Financial Time Series Forecasting: A Review Of Recent Advancements: 2020-2022. 2023.