

Portfolio Optimization with Long-Short Term Memory Deep Learning (LSTM)

Ángel Samaniego Alcántar¹   - ITESO, México

Abstract

The objective is a methodology for weighting financial assets in an investment portfolio. It is contrasted by the components of the Dow Jones Industrial Average (DJIA). For this purpose, portfolios with investment horizons between 1 and 2 years are studied using Long-Short Term Memory (LSTM) optimization. The best portfolio was with an investment horizon of 1.5 years. The neural network is trained with 1,000 observations and more than 2,777 portfolios are simulated. The model outperforms the DJIA by 73% to 85%, with a geometric mean annual return differential between 3.7% and 5%. The components of the DJIA in history are used to allocate assets to portfolios between 2008 and 2021. It is recommended that the methodology be contrasted in conjunction with another methodology for selecting financial assets. The conclusions are limited to assets that make up the DJIA. Mostly in the literature, neural networks are used for the short term; this paper contrasts the model to the long term, seeking to weigh assets and not future asset prices. The conclusion is that the LSTM model can be used for this purpose, for investment horizons of 1 to 2 years.

JEL Classification: G11, G17, C61.

Keywords: Artificial neural network, portfolio diversification, deep learning, LSTM.

Optimización de carteras con Aprendizaje Profundo de memoria a largo y corto plazo (LSTM)

Resumen

El objetivo es una metodología para ponderar los *activos financieros en una cartera de inversión*. Se contrasta con los componentes del Dow Jones Industrial Average (DJIA). Para ello, se estudian carteras con horizontes de inversión entre 1 y 2 años utilizando la optimización Long-Short Term Memory (LSTM). La mejor cartera se obtuvo con un horizonte de inversión de 1.5 años. La red neuronal se entrena con 1 000 observaciones y se simulan más de 2 777 carteras. El modelo supera al DJIA entre un 73% y un 85%, con un diferencial de rentabilidad geométrica media anual entre 3.7% y 5%. Los componentes del DJIA en la historia se utilizan para asignar activos a las carteras entre 2008 a 2021. Se recomienda contrastar la metodología junto con otra metodología de selección de activos financieros. Las conclusiones se limitan a los activos que componen el DJIA. Mayoritariamente en la literatura se utilizan redes neuronales para el corto plazo; en este trabajo se contrasta el modelo para el largo plazo, buscando ponderar activos y no precios futuros de activos. Concluyendo que el modelo LSTM puede utilizarse para este fin, para horizontes de inversión de 1 a 2 años.

Clasificación JEL: E12, C50, P10.

Palabras clave: Red neuronal artificial, diversificación de portafolios, Aprendizaje profundo, LSTM.

¹ Corresponding author. URL: https://iteso.mx/web/general/detalle?group_id=4207872

* No source of funding for research development



1. Introduction

Recurrent Neural Networks can have the ability to analyze sequences of data, connecting events that appear far apart in the input data, without their weight being diluted between the layers of the artificial neural network. An LSTM network is a type of recurrent neural network (RNN) that can learn the long-term dependencies between time steps contained in data sequentially. Studies on optimal portfolio search by models using LSTMs can be seen in the literature (Sen *et al.*, 2021a, 2021b; Istiake *et al.*, 2020; Wang & Zuo, 2021), prevailing the estimation of short-term.

Regularly used inputs can vary between their overnight returns in combination with technical indicators and macroeconomic indicators and are evaluated against other forecasting methods or portfolios. Borovkova & Tsiamas (2018) compare the performance of LSTM versus lasso, ridge logistic, the benchmark, and equally weighted portfolios. In this study, we will compare the performances of LSTM versus the DJIA and equally weighted portfolios with 1-year, 1.5-year, and 2-year investment horizons. Like Rácz & Fogarasi (2021), the goal is not to predict the returns of individual stocks, but to create a portfolio that has the highest predictability. But with the difference between using longer investment horizons, more observations in training, and more simulations for their study. In addition, these authors study the S&P500 stocks, where they study the components of the DJIA in history, seeking to outperform the DJIA.

Unlike Markowitz's (1952) modern portfolio theory, in which expected return and risk are optimized based on the historical returns of each asset using linear equations. In this paper, asset allocation portfolios use the components of the DJIA in history that have higher expected returns according to the LSTM model, leaving aside risk for asset allocation. Since the risk or volatility of the asset, it is expected to maintain its influence on the LSTM model within the artificial neural network. Hypothesis: the performance of the DJIA can be improved by weighting by return expectations that have non-linear behavior.

Selecting the underlying factors in the performance of an asset can be a complex task, since these factors can vary over time, in addition to their degree of influence. In this study, in the simulation or backtesting period, the factors and degrees of influence that were found during the LSTM training are maintained. The 48 components of the DJIA between 2000 and 2021 (under certain criteria) are used as the underlying performance factors to asset allocation the portfolios in the study period.

There is no consensus in the literature on the performance of the LSTM model in financial markets. Since we can find papers where the LSTM model does not produce above-market returns. However, the ability to predict upward and downward trends are satisfactory (Andersson & Mirkhani, 2020). Based on the previous article, two possible responses of good LSTM performance over investment horizons between 1 to 2 years are studied: the expected return and the trend that represents that return. For the first one, the investment strategy would be to weigh more of the assets with higher expectations, and for the second one, as we do not know how much each asset will rise but we know that they will rise, we use equal weight. The different LSTM models in the literature and their performance can be seen in Table 1.

Table 1. Comparison between LSTM models in the literature.

Autor	Training model	Training observations	Input variables	Output variables	Performance Measure
This paper	Adam optimization, LSTM with dropout.	252 observations using two years annual returns with rolling windows. 1,108 training sessions.	30 DJIA components using last two years annual returns.	Next two years returns of the 30 DJIA components.	Outperform the DJIA 98%, on average return differential 10.3%
Chen, K. Zhou, Y. and Dai, F. (2015)	Stochastic gradient descent.	74% of total sequence observations 1,211,361.	30 daily consecutive observations with 7 categorization of stock return (3 days return from Stocks of the Shanghai Securities Composite Index, SSE).	Next 3 days return with 5 variables (SSE Index Close, High, Low, Open, Volume).	Accuracy 27.2%
Ding, G. & Qin, L. (2020)	Adam optimization, 2 LSTM with dropout.	25% of total observations (6,112). With 50 iterations.	7 SSE Index variables (Close, High, Low, Open, Volume, Money, Change). Scaled [0, 1].	Next day price with 3 variables (SSE Index High, Low, Open). Scaled [0, 1].	Accuracy (1-MAE) 95%.
Chong, E., Han, C. and Park, F.C. (2017)	Rectified linear unit activation function.	60% of total observations (73,041). All stock returns are normalized.	10 lagged returns of the Korea KOSPI (38 stock returns, 5 minutes observations).	38 stock returns of the Korea KOSPI.	Up/Down Prediction Accuracy 62.01%.
Nelson, D.M.Q., Pereira, A.C.M. and de Oliveria, R.A. (2017)	Not mentioned.	Last 10 months. Observations from 2008-2015 with 15 minutes. Same model during 1 day.	180 technical indicators. Log-return transformation.	Up/Down [1, 0] in 15 minutes of the Brazilian stock exchange.	Up Prediction Accuracy 55.9%.
Yao, S., Luo, L. and Peng, H. (2018)	Not mentioned.	Not mentioned.	4 technical indicators.	Next 3 minutes price. Stocks randomly selected from CSI 300 components.	Better than random prediction, using 3 indicators.

Autor	Training model	Training observations	Input variables	Output variables	Performance Measure
Zhang, R., Huang, C., Zhang, W. and Chen, S. (2018)	Not mentioned.	24 months.	48 factors (details not specified).	One year CSI-300 stocks return into three categories (Up/down/neutral)	The excess earnings over the past 12 months were 4.5%.
Fisher, T. and Krauss, C. (2018)	Not mentioned. LSTM with dropout.	240 observations (daily return).	One day return (normalized).	Directional movements for the constituent stocks of the S&P 500.	Daily returns of 0.46%, Sharpe Ratio of 5.8, 54.3% accuracy with 82% annual return.

Source: own elaboration.

The first section introduces the LSTM model and explains the research methodology. The next section gives the study results with the different investment strategies and horizons. Finally, the conclusions of the study and the proposed future lines of research are given.

2. Methodology

FactSet, Matlab, and its toolbox are used to contrast this research. The methodology is explained below following Mathworks (2022), see this reference for more details. To examine the proposed asset weighting methodology, it is imperative to have pre-selected assets with historical data. The DJIA was chosen due to its availability of such information. It is important to clarify that the objective is not to establish a new index but rather to scrutinize and compare the proposed weighting approach.

Long and short-term memory networks (LSTM) are recurrent neural networks (RNN), which incorporate long and short-term event dependencies (Hochreiter & Schmidhuber, 1997). The neural network to be used has 20 hidden layers, it is trained in 500 epochs (times for optimization). The learning algorithm of the network is "Adam", the input variables are annual returns of the DJIA components in the history including DJIA ($x_t \dots x_{t-9}$), these returns are calculated from daily observations (rolling window), the output variables are the annual returns of the components and the DJIA. Only components with returns between 2000 and 2020 are used among the components in the history (see appendix). To link past events to this network, the LSTM algorithm is incorporated. This dependency is realized by the LSTM through four components (input gate, forget gate, cell candidate, output gate). The first gate (i) refers to the information to be updated. The second gate (f)

is the information to be discarded. The third gate (g) is the information to be added. The fourth gate (o) is the information to be added to the output variable (see figure 1).

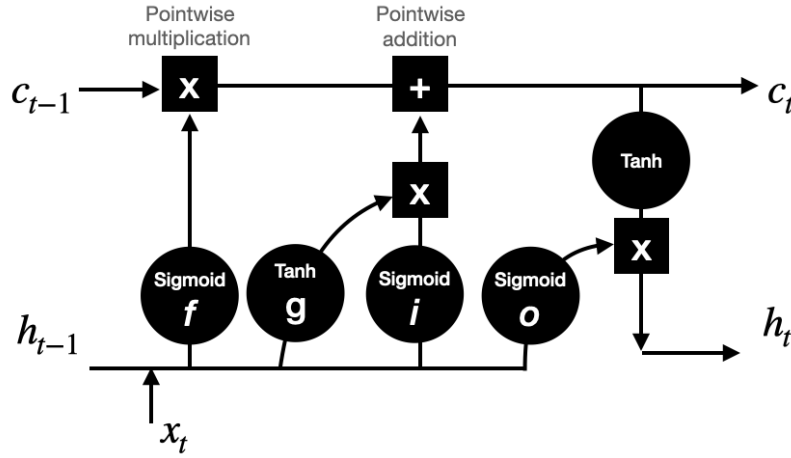


Figure 1. LSTM block.
Source: Mathworks (2022).

Likewise, for each LSTM block, we update the equations containing the weights and biases of each component, in this case, the DJIA shares in history (see equations 1 to 6). In the equations, the matrices W , R , and b are concatenations of the input weights, the recurrent weights, and the bias of each component since they represent the neural network.

$$f_t = \frac{1}{1 + e^{-(W_f x_f + R_f h_{t-1} + b_f)}} \quad (1)$$

$$g_t = \tanh(W_g x_g + R_g h_{t-1} + b_g) \quad (2)$$

$$i_t = \frac{1}{1 + e^{-(W_i x_i + R_i h_{t-1} + b_i)}} \quad (3)$$

$$o_t = \frac{1}{1 + e^{-(W_o x_o + R_o h_{t-1} + b_o)}} \quad (4)$$

It is observed that the matrix R represents the influence of the past.

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (5)$$

$$h_t = o_t \odot \tanh(c_t) \quad (6)$$

Where $\odot$ is the Hadamard product (element-wise multiplication of vectors). The equations show the activation functions of the state: hyperbolic tangent ($\tanh$) and the sigmoidal function ($((1 + e^{-x})^{-1})$).

In equation 6 we obtain the forecast within the training, compared with the real data (one-year yield for each asset) in each epoch. The weights W , R , and b are changed by the learning

algorithm used to have a lower RMSE. At the end of the 500 epochs, we have the weights and biases matrices to be able to make forecasts outside the training observations. The following section mentions the number of observations used for training and simulation, as well as the results of the study.

The output of the first block of figure 2 updates the first-time step in the sequence and continues with the next one, so successively during the whole training process of the network, the cell state and output state are updated (c_t, h_t).

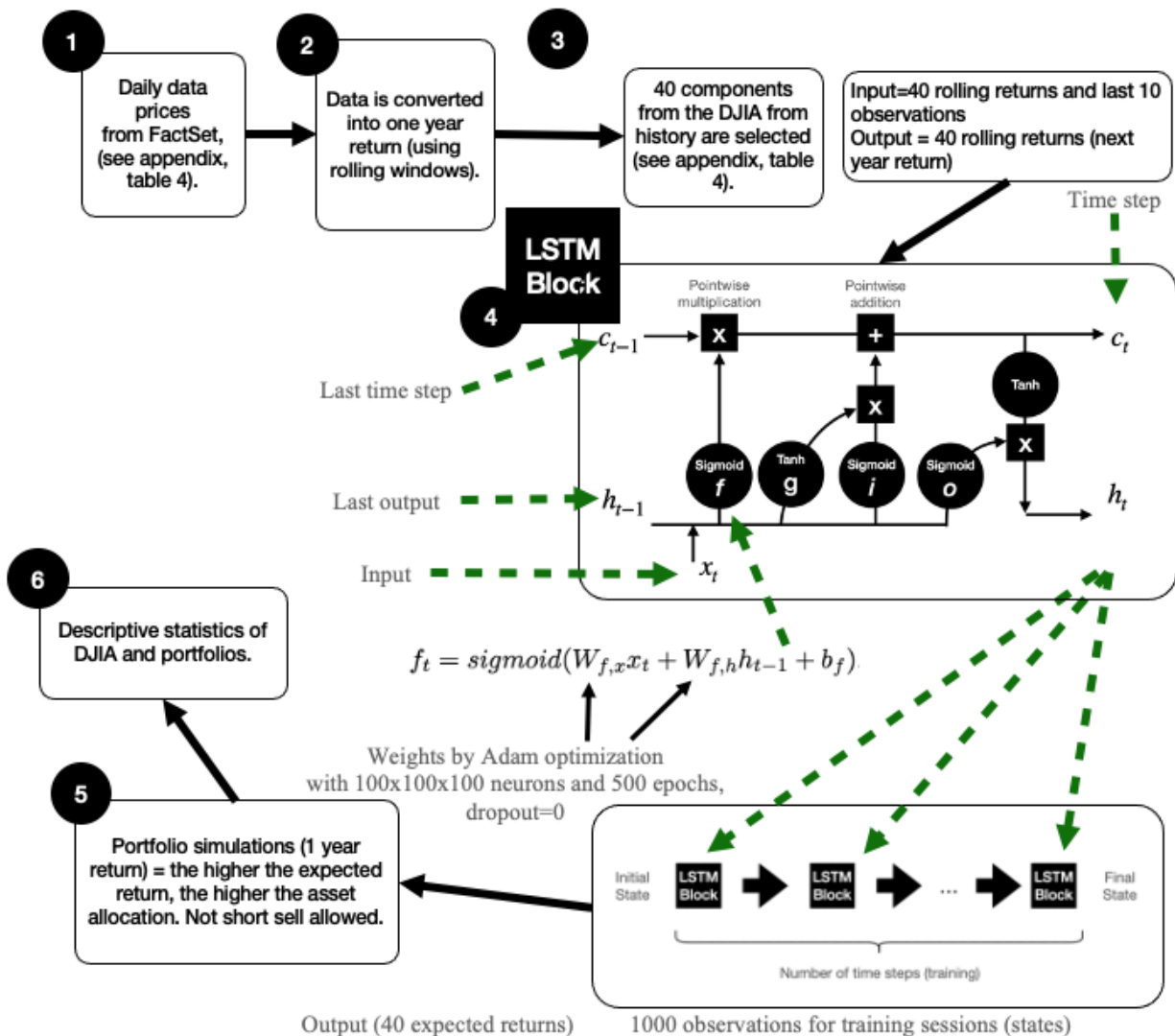


Figure 2. Training the network for one year return.

Source: own elaboration and MathWorks (2022).

3. Results

Figure 1 shows the number of assets that the three types of portfolios would have during the simulations. In general, the three types of portfolios contain enough assets to seek diversification. It is left for future studies whether the investment percentages and amount of assets achieve an efficient diversification. And not to deviate from the study's objective, to create a portfolio with the greatest predictability.

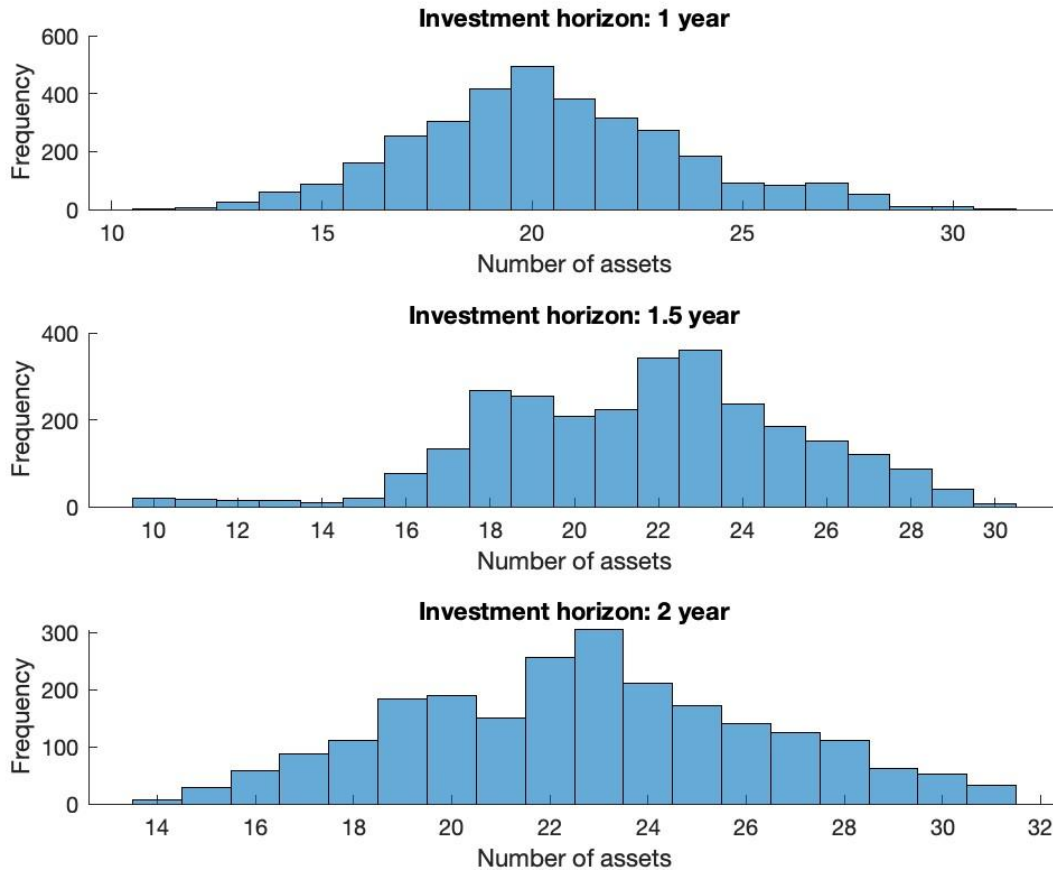


Figure 3. Number of assets in the portfolios during the simulations.
Source: own elaboration and data from FactSet.

The characteristics and performances of the three types of portfolios are shown in Table 2. Where the best strategy, which considered only purchases and weighted the assets based on the return forecast, obtained similar average returns, close to 12% (annual equivalent rate) and with an average return differential of 4% versus the DJIA. The portfolio with a 1.5-year investment horizon (378 trading days) had the highest return predictability. With 85.4% outperforming the 1.5-year returns of the DJIA.

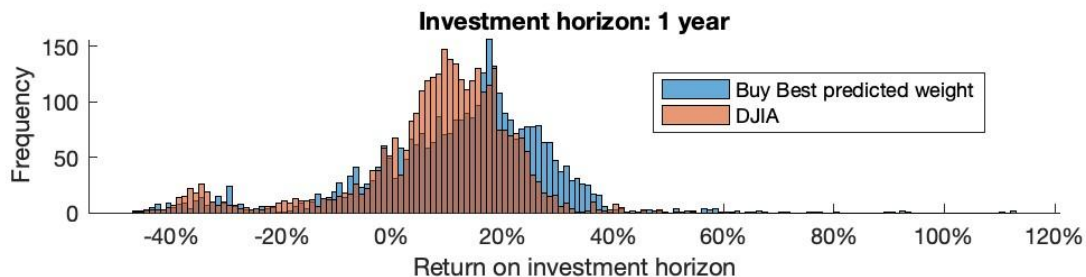
Table 2. Comparison between investment horizons.

	Investment horizon: 1 year	Investment horizon: 1.5 years	Investment horizon: 2 years
DJIA average annual return	8.04%	7.51%	8.41%
Buy-only strategies (LSTM) & equal weight: average annual return	5.36%	5.14%	8.17%
Probability of beating the DJIA	34.5%	35%	46.17%
Buy-Sell strategies (LSTM) & equal weight: average annual return	1.29%	1.06%	1.91%
Probability of beating the DJIA	2.81%	2.47%	8.17%
Buy-only strategies (LSTM) & return expectations allocation: average annual return	12.09%	12.51%	12.12%
Excess return (LSTM - DJIA)	4.05%	5%	3.71%
Probability of beating the DJIA	73.06%	85.4%	76.15%
Probability of returns > 0%	81%	86.3%	87.1%
Probability of Market direction	38%	55.6%	65%
Observations during training	1 000	1 000	1 000
Observations in the simulation	3 285	2 781	2 277

Source: own elaboration.

Note: average annual return is an annual geometric return for an investment horizon of 1.5 and 2 years.

Figure 4 shows the histograms of the performance of the three types of portfolios, where these portfolios clearly show a better distribution versus the DJIA. For diversification evaluation purposes, the results of the best portfolio are compared with portfolios using modern portfolio theory (maximum Sharpe ratio). Figure 5 shows the distribution of returns where the Sharpe ratio does not improve the proposed model.



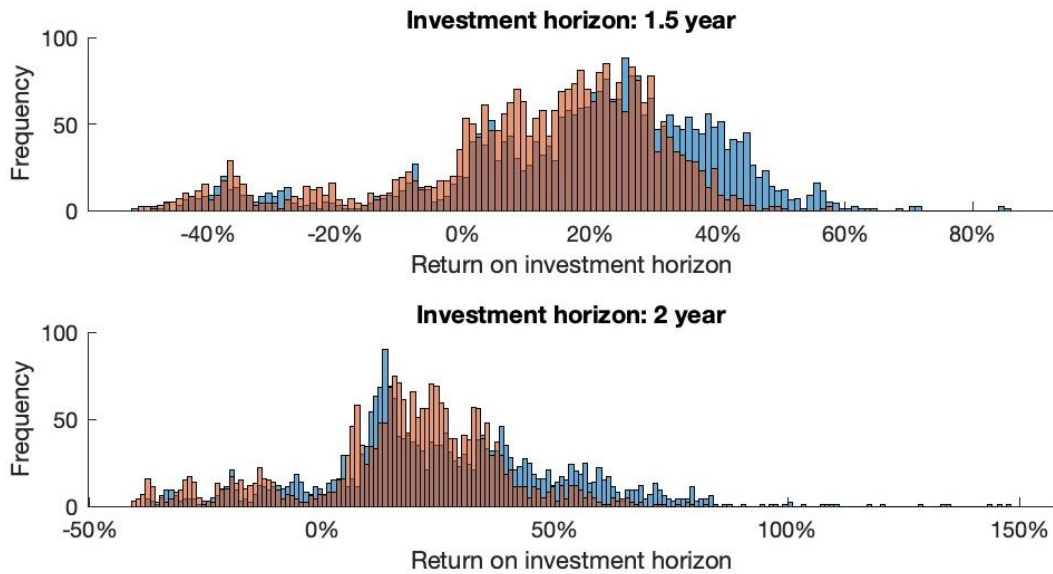


Figure 4. Histograms at different investment horizons: Buy-only strategies (LSTM): 2008-2021
Source: own elaboration and data from FactSet.

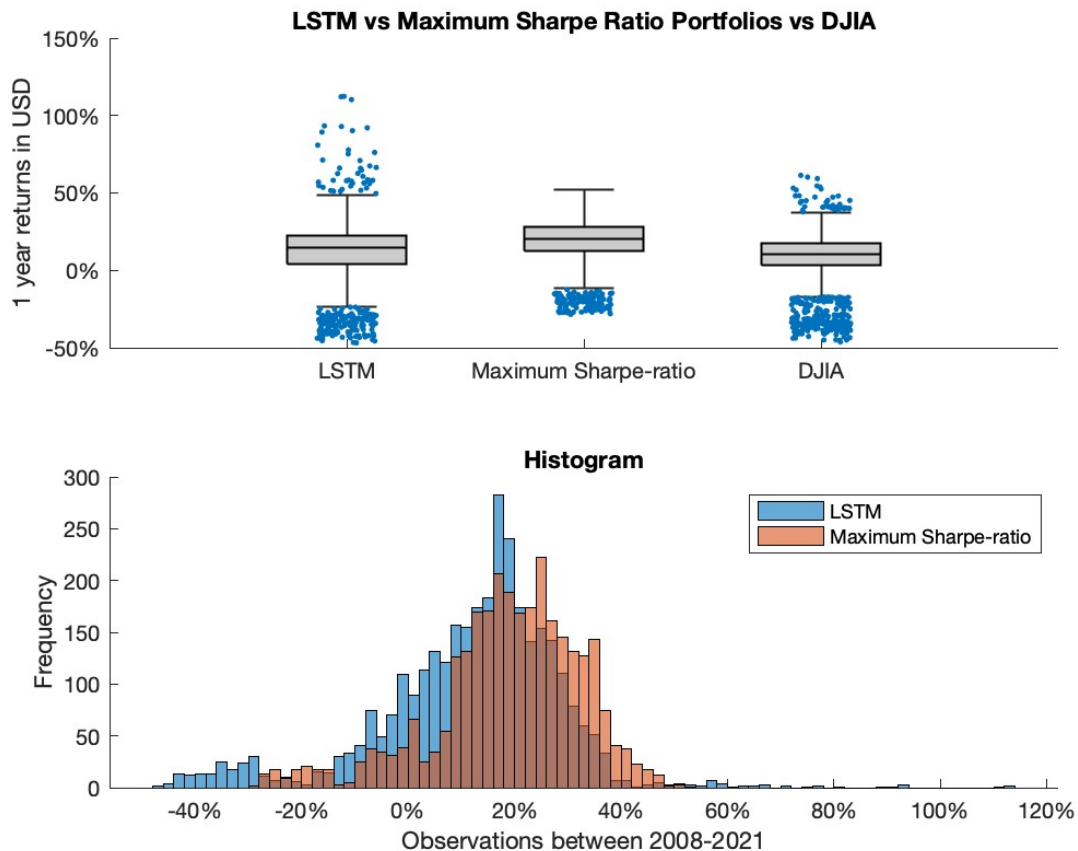


Figure 5. Sharpe portfolios vs LSTM (Buy only strategies)

Source: own elaboration.

Note: Maximum Sharpe-ratio portfolios were optimized with the restrictions of a minimum investment of 1% to 10% to ensure diversification.

Although the proposed model is shown to be superior to the DJIA (figure 4), it may have different degrees of risk. For the ranking of these portfolios, the appraisal is used (for more details see Amenc & Le Sourd, 2003). The returns are adjusted to the degree of risk to obtain the alpha (CAPM), and this is divided by the risk assumed to obtain it (residuals). The higher the appraisal, the better the ranking of the investment strategy. The Benchmark index (DJIA) has a valuation of zero. Table 3 shows that the Sharpe-ratio model is superior (appraisal=10.01).

Table 3. Performance (appraisal)

	Estimate	Standard error	T-statistic	P-value	R ²	Appraisal
Linear regression (adjusted R-square 0.7118): LSTM vs DJ, with 3 285 observations					0.88	10.01
<i>Alpha</i>	0.03	0.001	29.89	0		
<i>Beta</i>	1.09	0.006	160.80	0		
Linear regression (adjusted R-square 0.7118): Sharpe vs DJ, with 3 777 observations					0.76	7.13
<i>Alpha</i>	0.05	0.002	30.47	0		
<i>Beta</i>	0.697	0.0067	103.9	0		

4. Conclusion

The Buy-only strategy (LSTM), weighting its assets based on expected return, was the best of the three types of portfolios for the DJIA components asset allocation. The main contribution of this paper is that nonlinear models can support investment asset allocation in the medium term (2 years) using expected return. The literature uses different ways to evaluate performance (Table 1). One of the errors observed in the literature is normalizing the data and using it to measure its mean square error. The error is that normalizing the data decreases their values and thus also decreases their error. With the rest of the literature (Table 1) the performance obtained in this study is superior to most previous studies. The best investment horizon was two years. Over the study period, it had an 85.4% probability of outperforming the DJIA over 1.5 years. Compared with portfolios that use the Sharpe ratio for asset allocation, the distribution of returns has different degrees of risk. Portfolios that use Sharpe have a lower Beta and lower Appraisal, are lower risk portfolios that outperform the market but underperform LSTM portfolios. The reason is that the Share ratio portfolio is constructed based on past behaviors, assuming they will persist in the future, whereas LSTM aims to predict future returns, and portfolios are constructed based on these forecasts. For future research, it is proposed to study the short term, where LSTM predicts only one step. This is expected to increase predictability and returns even at high computational costs. The combination of LSTM with other forecasting methods can also be studied. Such as ARIMA and ARFIMA models (see Choi, 2018; Bukhari *et al.*, 2020, Shah *et al.*, 2022). Song *et al.* (2023) integrates a convolutional neural network-long short-term memory (CNN-LSTM) architecture. Yue *et al.* (2022) employ reinforcement learning to adjust the investment policy, actively optimizing the past to define it and thereby achieve improved

outcomes in short-term trading. Gülmez, B. (2023) employs various structures in an attempt to predict short-term movements of the DJIA index, with LSTM optimization using ARO proving to be the most effective (LSTM-ARO, artificial rabbits optimization).

References

- [1] Amenc, N. and V. Le Sourd (2003). Portfolio theory and performance analysis. Great Britain: Wiley.
- [2] Andersson, A., & Mirkhani, S. (2020). Portfolio Performance Optimization Using Multivariate Time Series Volatilities Processed With Deep Layering LSTM Neurons and Markowitz (Dissertation). KTH Royal Institute of Technology School of Engineering Sciences. Retrieved from <http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-273617>
- [3] Borovkova, S., & Tsiamas, I. (2018). An Ensemble of LSTM Neural Networks for High-Frequency Stock Market Classification. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.3202313>
- [4] Bukhari, A. H., Raja, M. A. Z., Sulaiman, M., Islam, S., Shoaib, M., & Kumam, P. (2020). Fractional neuro-sequential ARFIMA-LSTM for financial market forecasting. IEEE Access, Vol. 8, pp. 71326-71338. <https://doi.org/10.1109/access.2020.2985763>
- [5] Chen, K. Zhou, Y. & Dai, F. (2015). A LSTM-based method for stock returns prediction: a case study of China stock market. Paper presented at the 2015 IEEE International Conference on Big Data, pp. 2823-2824. <https://doi.org/10.1109/bigdata.2015.7364089>
- [6] Choi, H. K. (2018). Stock price correlation coefficient prediction with ARIMA-LSTM hybrid model. arXiv:1808.01560 [Online], pp. 1-22. <https://doi.org/10.48550/arXiv.1808.01560>
- [7] Chong, E., Han, C. and Park, F.C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. Experts Systems with Applications, Vol. 83, 187-205. <https://doi.org/10.1016/j.eswa.2017.04.030>
- [8] Ding, G. & Qin, L. (2020). Study on the prediction of stock price based on the associated network model of LSTM. Int. J. Mach. Learn. & Cyber. Vol. 11, pp. 1307-1317. <https://doi.org/10.1007/s13042-019-01041-1>
- [9] Fisher, T. and Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, Vol. 270, pp. 654-669. <https://doi.org/10.1016/j.ejor.2017.11.054>
- [10] Gülmez, B. (2023). Stock price prediction with optimized deep LSTM network with artificial rabbits optimization algorithm. Expert Systems with Applications, 227, 120346. <https://doi.org/10.1016/j.eswa.2023.120346>
- [11] Hochreiter, S., and J. Schmidhuber (1997). Long short-term memory. Neural computation, Vol. 9, no. 8, pp. 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [12] Istiaque Sunny, M. A., Maswood, M. M. S., & Alharbi, A. G. (2020). Deep Learning-Based Stock Price Prediction Using LSTM and Bi-Directional LSTM Model. 2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES). <https://doi.org/10.1109/niles50944.2020.9257950>
- [13] Markowitz, H. (1952). Portfolio Selection. The Journal of Finance, Vol. 7, pp. 77-91. <https://doi.org/10.1111/j.1540-6261.1952.tb01525.x>
- [14] Mathworks (2022). Long Short-Term Memory Networks. Retrieved April 13 from <https://la.mathworks.com/help/deeplearning/ug/long-short-term-memory-networks.html>
- [15] Nelson, D.M.Q., Pereira, A.C.M. and de Oliveria, R.A. (2017). Stock market's price movement prediction with LSTM neural networks. Paper presented at the 2017 International Joint Conference on Neural Networks (IJCNN 2017). <https://doi.org/10.1109/ijcnn.2017.7966019>

- [16] Rácz, A. & Fogarasi, N.(2021).Trading sparse, mean reverting portfolios using VAR(1) and LSTM prediction. *Acta Universitatis Sapientiae Informatica*, Vol. 13(2), pp. 288-302. <https://doi.org/10.2478/ausi-2021-0013>
- [17] Sen, J., Dutta, A., & Mehtab, S. (2021a). Stock Portfolio Optimization Using a Deep Learning LSTM Model. 2021 IEEE Mysore Sub Section International Conference (MysuruCon). <https://doi.org/10.1109/mysurucon52639.2021.9641662>
- [18] Sen, J., Mehtab, S., & Dutta, A. (2021b). Stock Price Prediction Using Machine Learning and LSTM-Based Deep Learning Models. *TechRxiv*. <https://doi.org/10.36227/techrxiv.15103602.v1>
- [19] Shah, J., Vaidya, D., & Shah, M. (2022). A comprehensive review on multiple hybrid deep learning approaches for stock prediction. *Intelligent Systems with Applications*, 16, 200111. <https://doi.org/10.1016/j.iswa.2022.200111>
- [20] Song, H., & Choi, H. (2023). Forecasting Stock Market Indices Using the Recurrent Neural Network Based Hybrid Models: CNN-LSTM, GRU-CNN, and Ensemble Models. *Applied Sciences*, 13(7), 4644. <https://doi.org/10.3390/app13074644>
- [21] Wang, R., & Zuo, Z. (2021). Stock Price Prediction with Long-short Term Memory Model. 2021 3rd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI). <https://doi.org/10.1109/mlbdbi54094.2021.00058>
- [22] Yao, S., Luo, L., & Peng, H. (2018). High-Frequency Stock Trend Forecast Using LSTM Model. 2018 13th International Conference on Computer Science & Education (ICCSE), pp. 1-4. <https://doi.org/10.1109/ICCSE.2018.8468703>
- [23] Yue, H., Liu, J., Tian, D., & Zhang, Q. (2022). A Novel Anti-Risk Method for Portfolio Trading Using Deep Reinforcement Learning. *Electronics*, 11(9), 1506. <https://doi.org/10.3390/electronics11091506>
- [24] Zhang, R., Huang, C., Zhang, W. and Chen, S. (2018). Multi-factor stock selection model based on LSTM. *International Journal of Economics and Finance*, Vol. 10(8), pp. 1-36. <https://doi.org/10.5539/ijef.v10n8p36>

Appendix

Table 4. Dow Jones Industrial Average components in history with more 4,789 rolling returns: 2000-2020.

1	Alcoa Inc.	AA-US	21	International Paper Company	IP-US
2	3M Company	MMM-US	22	JPMorgan Chase & Co.	JPM-US
3	Honeywell International	HON-US	23	Johnson & Johnson	JNJ-US
4	Altria Group Incorporated	MO-US	24	McDonald's Corporation	MCD-US
5	American Express Company	AXP-US	25	Merck & Co., Inc.	MRK-US
6	American International Group, Inc.	AIG-US	26	Nike, Inc.	NKE-US
7	Amgen Inc.	AMGN-US	27	Pfizer Inc.	PFE-US
8	Apple Inc.	AAPL-US	28	Raytheon Technologies Corporation	RTX-US
9	AT&T Inc.	T-US	29	The Boeing Company	BA-US
10	Bank of America Corporation	BAC-US	30	The Coca-Cola Company	KO-US
11	Caterpillar Inc.	CAT-US	31	The Goldman Sachs Group, Inc.	GS-US
12	Chevron Corporation	CVX-US	32	The Home Depot, Inc.	HD-US

13	Cisco Systems, Inc.	CSCO-US	33	The Procter & Gamble Company	PG-US
14	Citigroup Inc.	C-US	34	The Travelers Companies, Inc.	TRV-US
15	DowDuPont Inc.	DD-US	35	The Walt Disney Company	DIS-US
16	Exxon Mobil Corporation	XOM-US	36	UnitedHealth Group Inc.	UNH-US
17	General Electric Company	GE-US	37	Verizon Communications Inc.	VZ-US
18	Hewlett-Packard Company	HPQ-US	38	Walgreens Boots Alliance, Inc.	WBA-US
19	Intel Corporation	INTC-US	39	Walmart Inc.	WMT-US
20	International Business Machines Corporation	IBM-US	40	Dow Jones industrial Average	DJ

Source: own elaboration and data from FactSet.

Table 5. Annual returns statistics: 2000-2020.

						Quantile		
Ticket	Mean	Median	Std	Kurtosis	Skewness	25 %	50%	75%
AA-US	6.0%	0.2%	58.0%	21.55	3.14	-29.2%	0.2%	28.6%
MMM-US	8.4%	8.1%	19.0%	3.48	0.14	-4.1%	8.1%	21.3%
HON-US	10.4%	12.0%	23.6%	3.42	-0.42	0.1%	12.0%	24.2%
MO-US	12.9%	13.9%	26.9%	7.60	1.28	-2.9%	13.9%	24.1%
AXP-US	9.7%	11.2%	34.1%	12.62	1.74	-10.6%	11.2%	22.8%
AIG-US	-1.0%	0.4%	39.5%	11.15	0.97	-18.7%	0.4%	15.8%
AMGN-US	8.6%	6.3%	21.1%	3.64	0.47	-5.5%	6.3%	22.6%
AAPL-US	39.9%	38.7%	56.4%	4.36	0.77	0.3%	38.7%	66.9%
T.XX1-US	-13.5%	-13.9%	24.7%	2.40	0.15	-32.3%	-13.9%	2.7%
T-US	-0.3%	-0.7%	18.9%	3.28	0.14	-11.1%	-0.7%	11.5%
BAC-US	8.6%	7.3%	40.4%	15.92	1.89	-12.6%	7.3%	23.9%
CAT-US	16.6%	13.3%	34.5%	3.49	0.51	-8.1%	13.3%	38.3%
CVX-US	5.7%	6.4%	19.6%	2.82	-0.11	-6.2%	6.4%	18.5%
CSCO-US	4.2%	3.9%	31.1%	3.62	-0.00	-15.2%	3.9%	24.2%
C-US	-1.0%	0.3%	35.1%	5.20	0.14	-19.7%	0.3%	16.7%
DOW-US	15.3%	2.4%	49.6%	3.57	1.18	-15.6%	2.4%	19.1%
DD-US	7.3%	3.5%	37.9%	22.64	2.98	-13.3%	3.5%	23.4%
DD.XX1-US	5.5%	3.2%	22.2%	3.97	0.18	-7.0%	3.2%	18.7%
EKDKQ-US	-22.7%	-22.7%	39.4%	4.29	0.52	-44.6%	-22.7%	-1.1%
XOM-US	2.4%	2.0%	19.0%	3.42	-0.13	-10.3%	2.0%	14.8%
GE-US	-3.2%	-0.1%	28.7%	4.86	0.35	-22.0%	-0.1%	13.5%
MTLQQ-US	-26.5%	-25.2%	39.7%	2.30	0.02	-51.7%	-25.2%	0.4%
HPQ-US	6.3%	1.0%	35.9%	2.77	0.37	-21.3%	1.0%	34.3%
INTC-US	4.5%	4.4%	29.6%	3.62	0.33	-15.1%	4.4%	22.5%
IBM-US	2.3%	-0.2%	18.6%	3.10	0.52	-10.7%	-0.2%	14.6%
IP-US	7.0%	4.0%	40.8%	35.50	4.12	-13.3%	4.0%	18.2%
JPM-US	7.7%	4.8%	26.8%	4.50	0.76	-10.0%	4.8%	23.3%

JNJ-US	7.0%	6.7%	12.6%	2.79	0.23	-1.8%	6.7%	14.9%
MDLZ-US	5.9%	6.7%	14.4%	2.66	-0.23	-3.5%	6.7%	16.0%
MCD-US	11.6%	10.8%	21.3%	6.45	0.48	-0.5%	10.8%	23.9%
MRK-US	3.4%	3.1%	23.1%	2.64	-0.04	-11.8%	3.1%	19.9%
MSFT-US	11.7%	9.3%	23.5%	2.90	0.03	-4.2%	9.3%	29.0%
NKE-US	18.7%	20.4%	20.6%	3.08	0.02	3.8%	20.4%	33.3%
PFE-US	1.1%	2.0%	17.0%	2.44	-0.06	-11.5%	2.0%	13.3%
RTX-US	8.7%	10.1%	20.3%	3.21	0.07	-4.5%	10.1%	21.2%
CRM-US	35.0%	28.9%	41.7%	4.35	0.86	11.1%	28.9%	51.6%
BA-US	14.0%	14.3%	37.4%	2.90	0.18	-8.8%	14.3%	37.1%
KO-US	3.9%	4.4%	13.8%	2.81	-0.24	-3.9%	4.4%	13.1%
GS-US	9.4%	5.7%	32.4%	5.96	0.98	-12.5%	5.7%	28.3%
HD-US	11.7%	13.5%	25.1%	3.26	-0.19	-2.2%	13.5%	27.0%
PG-US	7.8%	7.0%	14.2%	4.18	0.27	0.8%	7.0%	14.3%
TRV-US	8.0%	8.9%	19.5%	5.41	0.50	-3.4%	8.9%	18.8%
DIS-US	10.4%	11.3%	26.1%	3.40	0.24	-7.7%	11.3%	29.1%
UNH-US	22.4%	24.1%	27.4%	4.89	-0.26	8.9%	24.1%	38.6%
VZ-US	2.3%	1.4%	15.9%	3.12	-0.09	-7.2%	1.4%	13.2%
V-US	25.1%	24.6%	20.1%	2.98	-0.10	12.9%	24.6%	39.6%
WBA-US	4.3%	-0.5%	24.7%	2.93	0.69	-14.9%	-0.5%	20.8%
WMT-US	5.7%	5.0%	14.3%	3.36	0.31	-3.5%	5.0%	14.9%
DJIA	6.2%	7.3%	14.9%	4.54	-0.48	-0.2%	7.3%	15.3%
TBILLS	1.9%	1.3%	1.8%	2.68	0.92	0.3%	1.3%	2.7%

Source: own elaboration and data from FactSet.