



A Comparative Study of Deep Learning Techniques for Financial Indices Prediction

Argyrios P. Ketsetsis, Konstantinos M. Giannoutakis, Georgios Spanos, Nikolaos Samaras, Dimitrios Hristu-Varsakelis, Dimitrios Thomas, Dimitrios Tzovaras

► To cite this version:

Argyrios P. Ketsetsis, Konstantinos M. Giannoutakis, Georgios Spanos, Nikolaos Samaras, Dimitrios Hristu-Varsakelis, et al.. A Comparative Study of Deep Learning Techniques for Financial Indices Prediction. 17th IFIP International Conference on Artificial Intelligence Applications and Innovations (AIAI), Jun 2021, Hersonissos, Crete, Greece. pp.297-308, 10.1007/978-3-030-79150-6_24 . hal-03287677

HAL Id: hal-03287677

<https://inria.hal.science/hal-03287677>

Submitted on 15 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

A Comparative Study of Deep Learning Techniques for Financial Indices Prediction

Argyrios P. Ketsetsis¹, Konstantinos M. Giannoutakis¹, Georgios Spanos¹, Nikolaos Samaras², Dimitrios Hristu-Varsakelis², Dimitrios Thomas³, and Dimitrios Tzovaras¹

¹ Information Technologies Institute, Centre for Research and Technology Hellas, Thessaloniki, Greece {argykets,kgiannou,gspanos,Dimitrios.Tzovaras}@iti.gr

² Department of Applied Informatics, University of Macedonia, Thessaloniki, Greece {samaras,dcv}@uom.edu.gr

³ Effect Computer Applications S.A., Athens, Greece thomasd@effect.gr

Abstract. Automated trading is an approach to investing whereby market predictions are combined with algorithmic decision-making strategies for the purpose of generating high returns while minimizing downsides and risk. Recent advancements in Machine and Deep learning algorithms has led to new and sophisticated models to improve this functionality. In this paper, a comparative analysis is conducted concerning eight studies which focus on the American and the European stock markets. The simple method of Golden Cross trading strategy is being utilized for the assessment of models in real- world trading scenarios. Backtesting was performed in two indices, the S&P 500 and the EUROSTOXX 50, resulting in relative good performance, aside from the significant downfall in global markets due to COVID-19 outbreak, which appeared to affect all models.

Keywords: Stock Market Prediction · Deep Learning · Portfolio Back-testing · Comparative Analysis.

1 Introduction

Stock market analysis and automated trading have gained significant attention, due to the inherent complexity of trading markets and to the potential financial profit. Recent advancements in Machine and Deep Learning techniques and the increasing access to high-performance computational hardware have fueled research efforts, oftentimes with results which are much better compared to more traditional statistical time series approaches.

Stock prices are modelled as discrete time series, where financial data are represented as continuous data points indexed in temporal order. Historically, statistical methods, like ARIMA and ARMA were extensively utilized for modeling and forecasting future stock prices. The gradual emergence of machine learning techniques, led to the advent of more sophisticated predictive time series models. Decision Trees (DT), Support Vector Machines (SVM) and Multi-layer Perceptrons (MLP) are some of the most prominent techniques [19]. Deep

learning techniques and more specifically the recurrent networks, which added a memory-like feature, enhanced the modelling and predictive capabilities of the proposed algorithms. These include Long Short Term Memories (LSTM) [8] and Gated Recurrent Units (GRU) [4] to name a few.

In [12], research studies that incorporated machine/deep learning techniques to predict European stock market securities were collected and reviewed. For the purpose of this study, the methods proposed in these papers were implemented and tested in order to compare them in common datasets and extract more insights. The models were trained and tested on the S&P 500 index and the EUROSTOXX 50 index. The evaluation metrics include the Mean Squared Error (MSE) and the Mean Absolute Percentage Error (MAPE). Moreover, a trading strategy was implemented, the Moving Crossing Average (Golden Cross) [2], in order to test the predictive ability of the selected methods in real-world trading scenarios. While state-of-the-art models come with a thorough comparison to previous models, there are insufficient studies that try to perform a unified comparative analysis of various machine learning and deep learning methods in American and European stock markets.

The remainder of this paper is as follows. In Section 2, some background and related studies are presented, while in Section 3, the evaluation setup implemented for the comparison of the selected studies is analytically described. In Section 4, the results of the predictions and the back-testing outcomes are illustrated, where in Section 5, a summarization of the findings are given.

2 Background

In this section, studies that either are dealing with the application/comparison of machine/deep learning techniques and traditional techniques or are related to the moving average technical indicator are presented.

Regarding the application of the moving averages as a trading strategy, Metghalchi et al. [15] examined the efficiency of this technical indicator (among others) in the stock market of 16 European countries over a period spanning from 1990 to 2006. The results highlighted the fact that the moving average can outperform the buy-and-hold strategy for most of the European stock markets.

In [17], the authors compared nine machine and two deep learning models in the prediction task of four stock market groups from the Tehran Stock Exchange. The methods reviewed were: Decision Trees, Random Forest, Boosting algorithms, Support Vector Classifier (SVC), Naive Bayes, K-nearest neighbors (KNN), Logistic Regression, Artificial Neural Networks (ANN), Recurrent Neural Networks (RNNs) and Long-Short-Term-Memories (LSTMs). Their analysis indicated the superiority of RNNs and LSTMs in predicting continuous data.

The authors of [1], implemented different deep neural network topologies, a Support Vector Regressor (SVR) model and the Random Forest (RF) algorithm to predict one month ahead stock returns in the Japanese stock market. Additionally, they examined the performance of a simple long-short portfolio strategy, noting that the best returns are achieved by Deep Neural Networks.

In [12], a systematic literature review has been conducted including primary studies that deal with the prediction of stock prices in the European Union, using machine learning and deep learning techniques. This work motivated this proposed comparative study. Although the survey of [12] included twelve papers, four papers were excluded in this work from the comparative analysis. In two of them, natural language features were utilized [3, 13], one incorporated a Hybrid Fuzzy-Neural model [7], while in [5] there was no model testing in a real-world dataset, rather than a theoretical analysis of LSTM topologies.

The works [11, 16], both suggest a MLP model with Back Propagation learning rule, while [9] proposes a similar MLP topology, whereas the learning phase is achieved with the Widrow-Hoff learning rule. In [6], a hybrid ARMA-MLP method is incorporated. [10, 14, 20] propose LSTM topologies for the prediction of closing prices, while [18] proposes a Gated Recurrent Unit (GRU). The characteristics of each work of this analysis are summarized in Table 1.

Table 1. Research works characteristics, [12]

Research work	Features	Methodologies used
Janeski et al.[11]	Time Series (OHLCVM)	ANN
Dunis et al.[6]	Time Series (Close)	ARMA-MLP
Kyoung-Sook et al.[14]	Time Series (OHLCV), Financial Indicators	LSTM
Stoean et al.[20]	Time Series (Close)	LSTM
Magnus et al.[10]	Time Series (Close)	LSTM
Shen et al.[18]	Time Series (Close)	GRU
Mourelatos et al.[16]	Financial Indicators, Gold Prices	MLP
Hanias et al.[9]	Time Series (Close)	ANN

3 Evaluation setup

The purpose of this work is the evaluation of the methodologies depicted in Table 1 in a common setting, in order to make a robust comparative analysis. The evaluation was performed on two stock indices, the Standard and Poor's 500 (S&P 500) and the EUROSTOXX 50, spanning from January 2, 2013 to December 31, 2020. S&P 500 is a market-capitalization-weighted index of the 500 largest publicly-traded companies in the United States, while EUROSTOXX 50 is a market-capitalization weighted index of 50 large, blue-chip European companies that operate within the Eurozone nations. All the approaches to be compared were adapted so as to predict closing prices during the last 650 days of the selected time range, not taking into consideration the binary problem of predicting the direction of price movement. In the case of missing information regarding topology properties and training configuration, the parameter settings were fine-tuned in order to utilize the full potential of each method.

The evaluation procedure was comprised of two distinct scenarios. First, predictive metrics were calculated, in order to gauge the models' forecasting ability. Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE) are

two common evaluation metrics utilized in the bibliography and used in this study. In addition, a portfolio back-testing scenario was implemented so as to provide a more tangible insight into each model’s potential. For this, the crossing moving average strategy was implemented, where two moving averages were applied, one shorter and one longer in time span. When the shorter-term moving average crosses above the longer-term moving average, a buy signal (buy 100 shares) is triggered, as it indicates that the asset’s price trend is shifting upwards. On the other hand, when the shorter-term moving average crosses below the long-term moving average, that indicates a sell signal (sell 100 shares), as it is taken to mean that the price trend is shifting downwards. The short-term window was set to 40 days ahead while the long-term to 100 days. Every portfolio was given a starting capital of 100,000\$ and it is assumed that it only contains shares of a stock index. The portfolio evaluation includes the Sharpe Ratio (SR) and the overall return [21]. The overall return refers to the gain or loss realized by a portfolio and depends on the price of the portfolio’s components. The Sharpe Ratio is given as $SR = \frac{R_p - R_f}{\sigma_p}$, where R_p denotes the return of portfolio, R_f the risk-free rate and σ_p the standard deviation of the portfolio’s return.

4 Numerical results

In this section, the comparative analysis findings are presented. The methods proposed in the selected studies were tested on their forecasting accuracy, and their predictions were used to construct a moving crossing average strategy for back-testing in two real-world trading scenarios. The testing phase used data from June 2018 through December 2020. It is worth mentioning that during this time period, significant disturbances in the global market took place, due to COVID-19 outbreak, thus resulting in global market instability.

Two stock indices were used to train and test the proposed algorithms. S&P500 is one of the most commonly studied indices in bibliography, and the EUROSTOXX 50 index, which was used in order to provide an insight in European markets, which are relatively less popular in bibliography. The closing price of both indices in the selected time range are presented in Fig. 1.

In Table 2 and Table 3, the MSE, MAPE, SR and Overall Returns are presented for S&P 500 and EUROSTOXX 50 respectively. In both test sets, the method proposed in [18] achieves the smallest MSE and MAPE. The satisfactory forecasting results of [10, 14], which implement LSTM models, confirm the common tendency in the bibliography, that is the superiority of recurrent models in modeling and predicting stock market securities. Taking into consideration both MSE and MAPE metrics, GRU-based method in [18] is the most predictive accurate model in both datasets. It is clear that the MLP-based models achieve greater MSE. One exception is the ARMA-MLP model proposed in [6], which scores 0.00051 MSE and 1.5477 MAPE. This result suggests that hybrid or ensemble models are greatly improved.

As SR is an indicator of risk with respect to the returns, it is clear that the more stable nature of S&P 500 index, results in less risky portfolios. This fact

can be deducted by the positive SRs in Table 2. However, this fact comes with mediocre returns, both positive and negative, as seen in the Overall Returns column in Table 2 and in Fig. 2. On the other hand, the volatile EUROSTOXX 50 index, is more risky and presents significant losses, like in [10, 14]. In general, one can notice from Fig. 3 and Fig. 5 that the trading strategy was able to detect the downward movement of indices during the COVID-19 outbreak and sell the shares. However, the amount of downfall was not accurately measured in order to add a buy signal at the time when the price was in the lowest possible level in order to achieve maximal return.

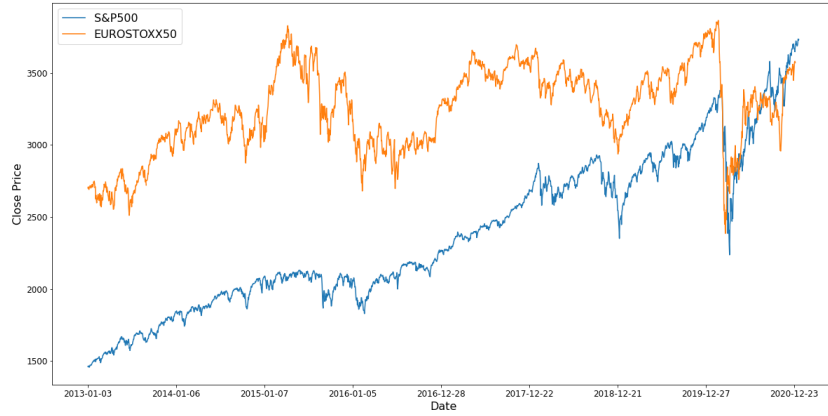


Fig. 1. S&P 500 and EUROSTOXX 50 closing prices from 02/01/2013 to 31/12/2020

Table 2. Evaluation metrics in S&P 500 dataset

Research work	MSE	MAPE	Sharpe Ratio	Overall Return
Janeski et al. (MLP) [11]	0.00159	0.02999	0.498	0.2102
Dunis et al. (ARMA-MLP) [6]	0.00051	1.54770	-1.163	-17.8186
Kyoung-Sook et al. (LSTM) [14]	0.01105	0.02201	0.974	0.3555
Stoean et al. (LSTM) [20]	0.03362	0.19827	0.964	0.3311
Magnus et al. (LSTM) [10]	0.00064	0.02545	0.871	0.3179
Shen et al. (GRU) [18]	0.00036	0.01904	0.894	0.3283
Mourelatos et al. (MLP-HONN) [16]	0.01162	73.56474	0.867	0.3323
Hanias et al. (MLP) [9]	0.26503	0.49547	0.997	-0.2107
S&P 500 Ground Truth	-	-	0.498	0.8916

In the S&P 500 index almost all methods achieve positive overall returns, 33% in approximate. On the other hand, in EUROSTOXX 50, almost all methods achieve negative returns, reflecting the fact that EUROSTOXX 50 is highly volatile. The highest returns in S&P 500 are achieved by [14, 20, 16]. The first

Table 3. Evaluation metrics in EUROSTOXX 50 dataset

Research work	MSE	MAPE	Sharpe Ratio	Overall Return
Janeski et al. (MLP) [11]	0.00762	0.06506	0.231	0.1113
Dunis et al. (ARMA-MLP) [6]	2.04E+05	0.13184	-0.029	-0.1536
Kyoung-Sook et al. (LSTM) [14]	0.00017	0.01125	-0.897	-4.4256
Stoean et al. (LSTM) [20]	0.02545	0.14148	-0.724	-3.422
Magnus et al. (LSTM) [10]	0.00038	0.01803	-0.925	-4.5571
Shen et al. (GRU) [18]	0.00017	0.01158	-0.896	-4.3709
Mourelatos et al. (MLP-HONN) [16]	0.00088	0.02527	-0.92	-4.3901
Hanias et al. (MLP) [9]	0.5623	8.44117	-0.821	-4.075
EUROSTOXX 50 Ground Truth	-	-	-0.03	-0.0471

two methods, incorporate LSTM topologies, which confirms the fact that recurrent networks present robust forecasting ability. The latter method, includes additional input features, like gold price, thus providing a more elaborate insight in stock market mechanisms. On EUROSTOXX 50, only [11] was able to provide positive return, while all other methods result in severe losses, 400% approximately. As it is illustrated in Table 3, the overall ground truth return of EUROSTOXX 50 is negative. These losses are due to the inability of the models to predict the striking downfall during COVID-19 outbreak.

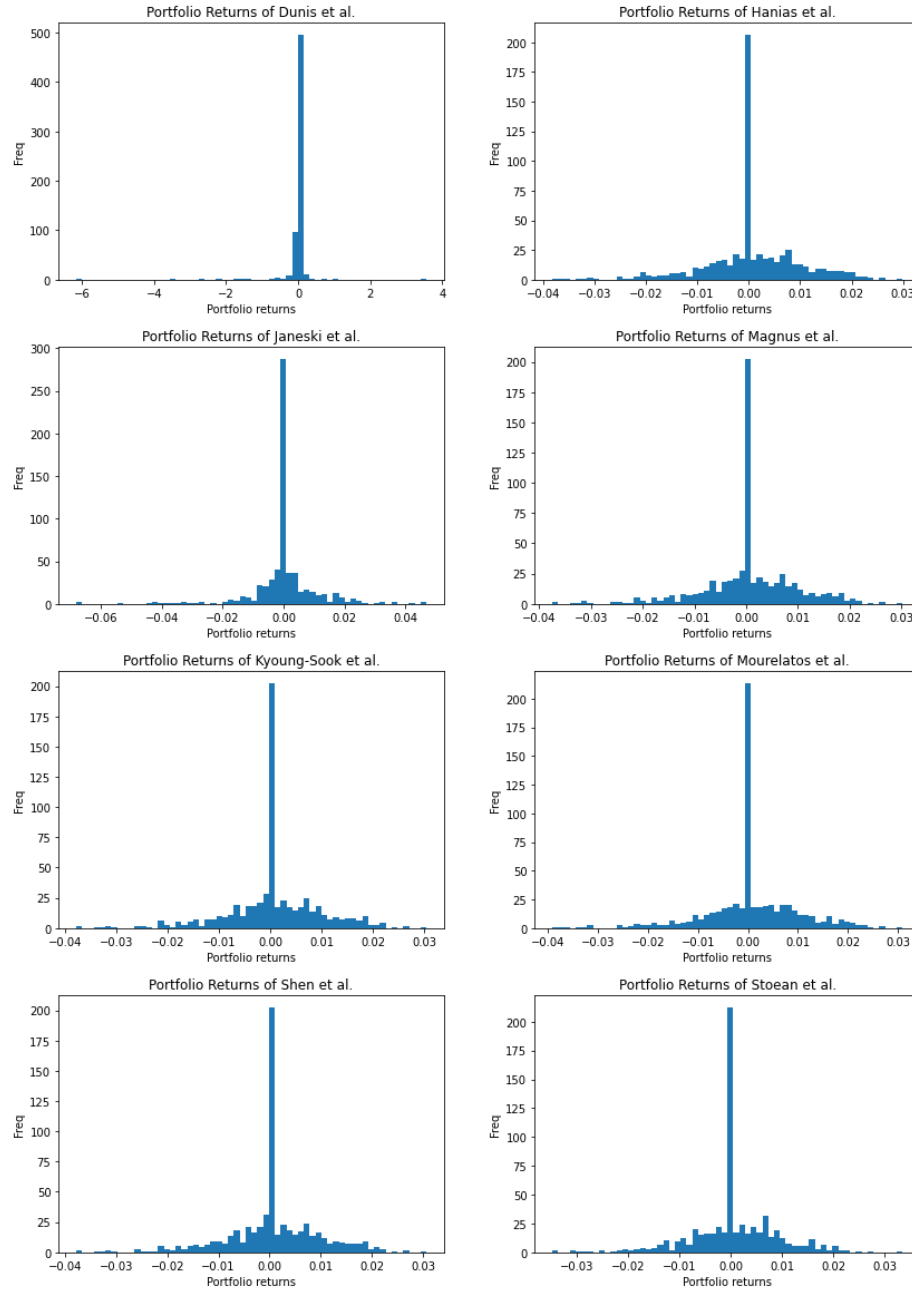
In Fig. 2 the frequency of returns of each portfolio is illustrated based on the forecasts of each model for the S&P 500 index. In Fig. 3 the predictions of each model is presented along with the short moving averages (SMA) and the long moving averages (LMA). Based on SMA and LMA, the buy and sell signals are defined, that are illustrated with the triangles. In Fig. 4, the returns of each portfolio based on the predictions of each models for the EUROSTOXX 50 index are indicated, while in Fig. 5, the trading signals for this index are presented.

An additional analysis was carried out regarding the Spearman’s rank correlation between MAPE and Overall Returns. For the S&P 500 index, the correlation was slightly positive (0.129) and a slightly negative correlation (-0.176) was observed in the EUROSTOXX 50 index. These two weak correlations for both indices state the fact that a better model with regard to the forecasting ability does not guarantee better results with respect to the profit.

Finally, it is prominent to note that the profitability of each method is dependent to the dataset, since methods that have high profits in the S&P do not perform well in the EUROSTOXX 50 (for instance [10], [14]) in comparison with methods that do not perform so well in S&P 500 ([11]). The aforementioned is validated statistically by the negative and medium correlation between the returns of the methods in each dataset provided from the result of spearman’s correlation rank test (-0.57).

5 Conclusions

In this paper, a comparative analysis of 8 studies concerning stock market prediction using a common trading strategy and time span was carried out. These

**Fig. 2.** S&P 500 Returns

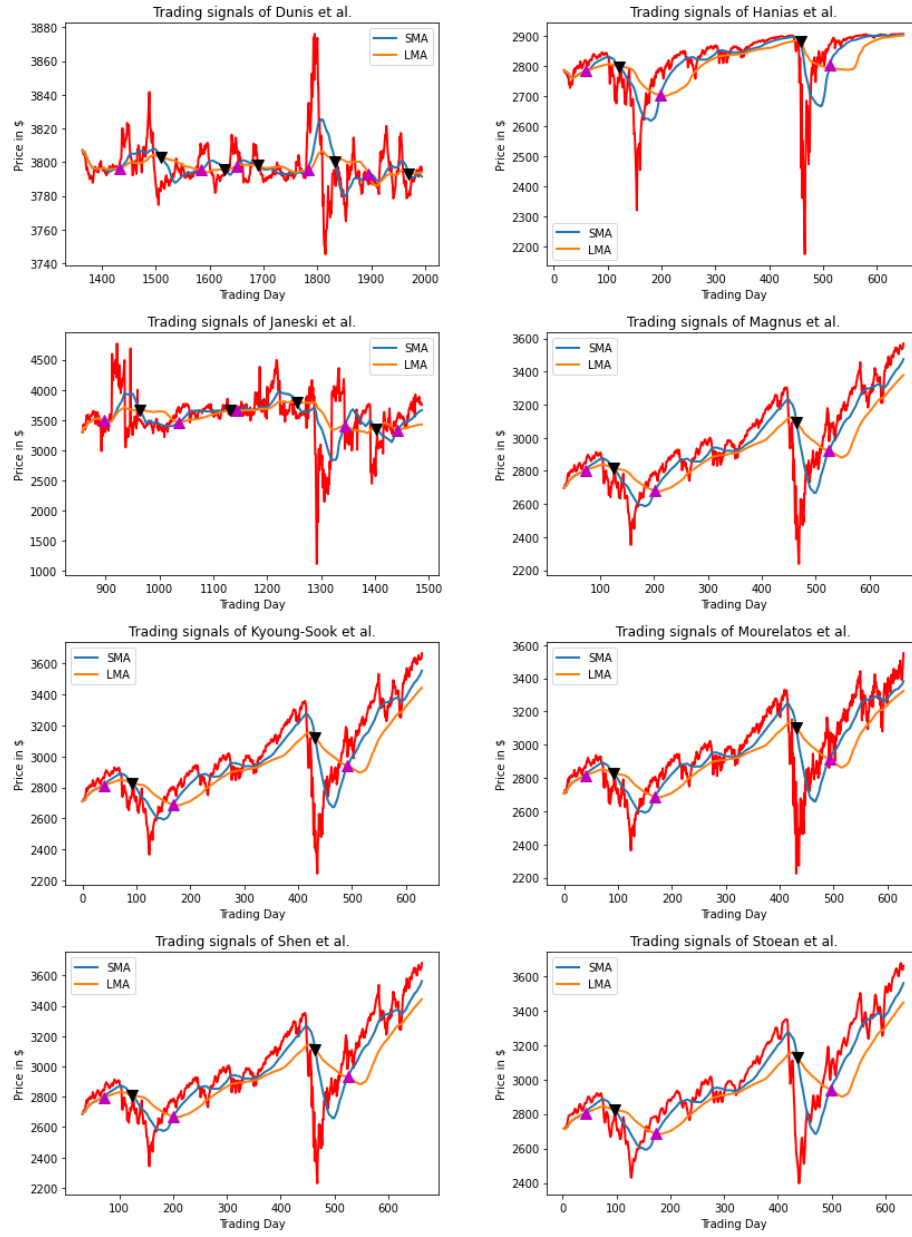
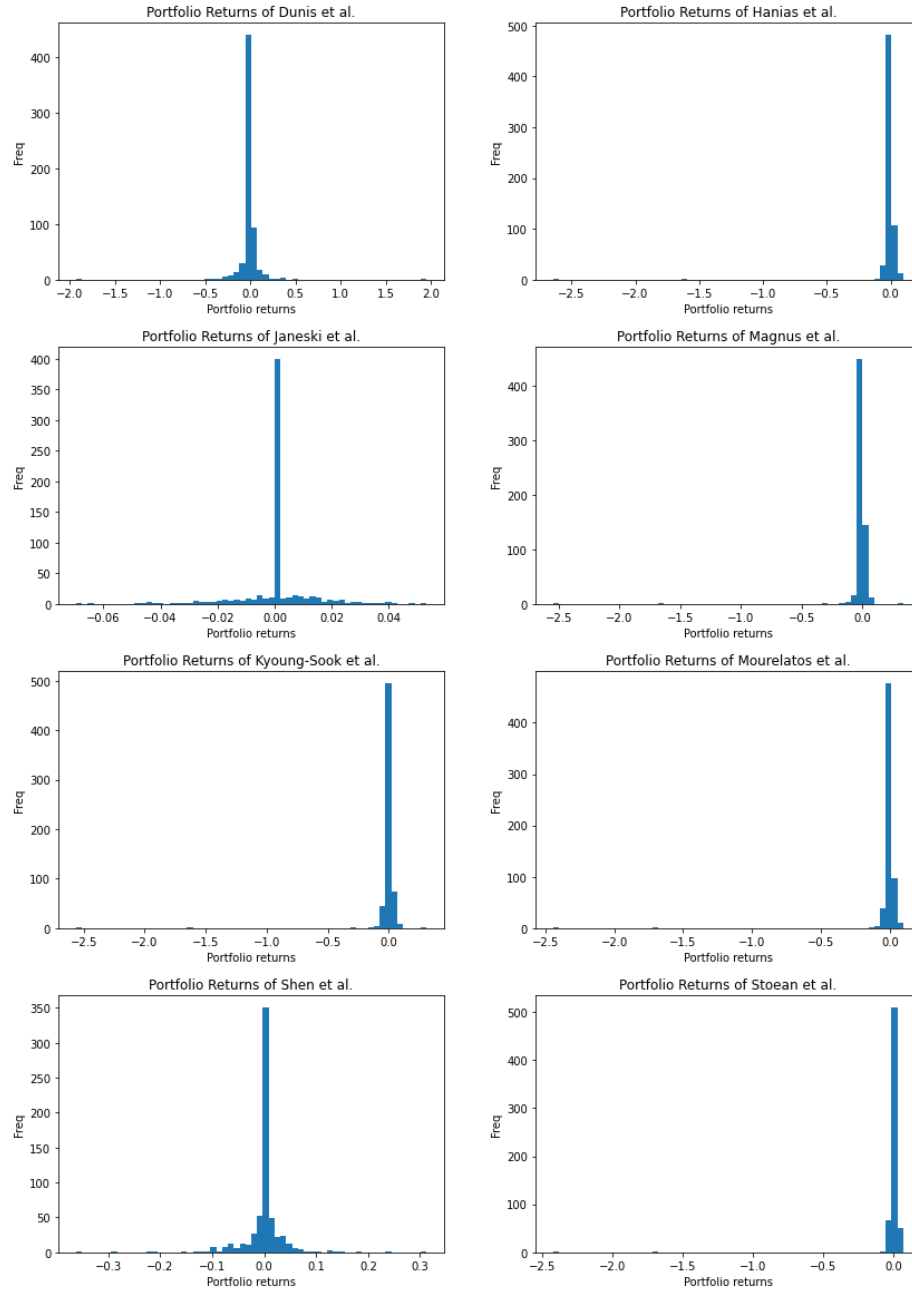


Fig. 3. S&P 500 Trading Signals

**Fig. 4.** EUROSTOXX 50 Returns

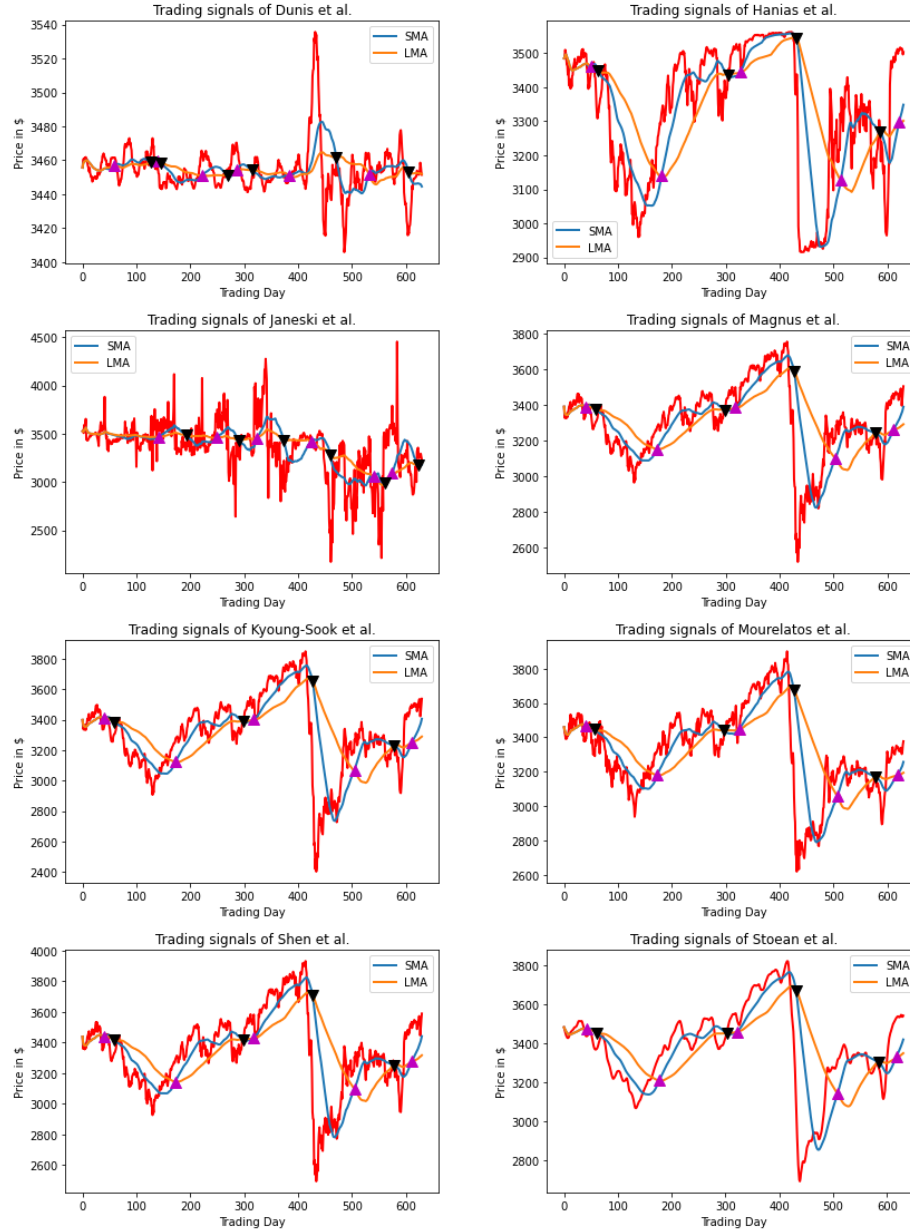


Fig. 5. EUROSTOXX 50 Trading Signals

studies were implemented and tested on two common datasets, the closing prices of the S&P 500 and the EUROSTOXX 50 indices. The evaluation metrics included the MSE and MAPE for the assessment of the price predictions. In order to come closer to a real world evaluation, the Moving Average Crossing Strategy was implemented to evaluate the performance of the studies in the two stock markets. The trading strategy evaluation metrics were the overall portfolio return and the Sharpe Ratio.

The superiority of the recurrent algorithms, and the difference in Moving Average Crossing strategy results between the two datasets were observed. This, likely indicates the difference in the stock market reaction to COVID-19 outbreak between American and European markets. Following the comparative analysis, it appears that none method proved to be universally superior on both datasets. Finally, it is worth noting that an accurate prediction model does not guarantee positive returns in the backtesting scenario as proved by the correlation analysis between MAPE and Overall Returns. Further analysis should be carried out regarding portfolio optimization, as Moving Crossing Average is just one of a multitude of possible trading strategies and also because the portfolio's components were only stock indices for the purposes of this study.

Acknowledgments

The authors would like to thank Mr. Christos Kourounis for his support on the implementation of the research works [6, 11, 16]. This research has been co-financed by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (project code:T2EDK-03743).



Co-financed by Greece and the European Union

References

1. Abe, M., Nakayama, H.: Deep learning for forecasting stock returns in the cross-section. In: Pacific-Asia conference on knowledge discovery and data mining. pp. 273–284. Springer (2018)
2. Bailey, D.H., Borwein, J., Lopez de Prado, M., Zhu, Q.J.: Pseudo-mathematics and financial charlatanism: The effects of backtest overfitting on out-of-sample performance. Notices of the American Mathematical Society **61**(5), 458–471 (2014)
3. Borovkova, S., Dijkstra, M.: Deep learning prediction of the eurostoxx 50 with news sentiment. Available at SSRN 3253043 (2018)
4. Chung, J., Gulcehre, C., Cho, K., Bengio, Y.: Empirical evaluation of gated recurrent neural networks on sequence modeling. In: NIPS 2014 Workshop on Deep Learning, December 2014 (2014)

5. Dezsi, E., Nistor, I.A.: Can Deep Machine Learning Outsmart The Market? A Comparison Between Econometric Modelling And Long- Short Term Memory. *Romanian Economic Business Review* **11**(4.1), 54–73 (december 2016)
6. Dunis, C.L., Laws, J., Karathanassopoulos, A.: Modelling and trading the Greek stock market with mixed neural network models. *Applied Financial Economics* **21**(23), 1793–1808 (December 2011)
7. García, F., Guijarro, F., Oliver, J., Tamošiūnienė, R.: Hybrid fuzzy neural network to predict price direction in the german dax-30 index. *Technological and Economic Development of Economy* **24**(6), 2161–2178 (2018)
8. Gers, F.A., Schmidhuber, J., Cummins, F.: Learning to forget: continual prediction with lstm. In: 1999 Ninth International Conference on Artificial Neural Networks ICANN 99. (Conf. Publ. No. 470). vol. 2, pp. 850–855 vol.2 (1999)
9. Hantias, M.P., Curtis, P.G., Thalassinou, E.: Time series prediction with neural networks for the athens stock exchange indicator (2012)
10. Hansson, M.: On stock return prediction with lstm networks (2017)
11. Janeski, M., Kalajdziski, S.: Neural network model for forecasting balkan stock exchanges. In: Huang, D.S., Gan, Y., Bevilacqua, V., Figueroa, J.C. (eds.) *Advanced Intelligent Computing*. pp. 17–24. Springer Berlin Heidelberg (2012)
12. Ketsetsis, A.P., Kourounis, C., Spanos, G., Giannoutakis, K.M., Pavlidis, P., Vazakidis, D., Champeris, T., Thomas, D., Tzovaras, D.: Deep learning techniques for stock market prediction in the european union: A systematic review. In: 2020 Int. Conf. on Computational Science and Computational Intelligence, to appear (2021)
13. Kraus, M., Feuerriegel, S.: Decision support from financial disclosures with deep neural networks and transfer learning. *Decision Support Systems* **104**, 38–48 (2017)
14. Kyoung-Sook, M., Hongjoong, K.: Performance of deep learning in prediction of stock market volatility. *Economic Computation & Economic Cybernetics Studies & Research* **53**(2) (2019)
15. Metghalchi, M., Marcucci, J., Chang, Y.H.: Are moving average trading rules profitable? evidence from the european stock markets. *Applied Economics* **44**(12), 1539–1559 (2012)
16. Mourelatos, M., Alexakos, C., Amorgianiotis, T., Likothanassis, S.: Financial indices modelling and trading utilizing deep learning techniques: The athens se ftse/ase large cap use case. In: 2018 Innovations in Intelligent Systems and Applications (INISTA). pp. 1–7 (2018)
17. Nabipour, M., Nayyeri, P., Jabani, H., S., S., Mosavi, A.: Predicting stock market trends using machine learning and deep learning algorithms via continuous and binary data; a comparative analysis. *IEEE Access* **8**, 150199–150212 (2020)
18. Shen, G., Tan, Q., Zhang, H., Zeng, P., Xu, J.: Deep learning with gated recurrent unit networks for financial sequence predictions. *Procedia Computer Science* **131**, 895–903 (2018)
19. Shen, S., Jiang, H., Zhang, T.: Stock market forecasting using machine learning algorithms. *Dep. of Electrical Eng., Stanford University, Stanford* pp. 1–5 (2012)
20. Stoean, C., Paja, W., Stoean, R., Sandita, A.: Deep architectures for long-term stock price prediction with a heuristic-based strategy for trading simulations. *PLOS ONE* **14**(10), 1–19 (10 2019)
21. Zhang, Z., Khushi, M.: Ga-mssr: Genetic algorithm maximizing sharpe and sterling ratio method for robotrading. In: 2020 International Joint Conference on Neural Networks. pp. 1–8. IEEE (2020)