A new approach for Trading based on Long-Short Term memory Ensemble technique

Zineb Lanbouri[1] and Said Achchab[2]

**[1] National School for Computer Science and Systems analysis,
Mohamed V University,
Rabat, Morocco**

**[2] National School for Computer Science and Systems analysis,
Mohamed V University,
Rabat, Morocco**

Abstract

Stock market prediction has always been crucial for stakeholders, traders and investors. We developed an ensemble Long Short Term Memory (LSTM) model that includes two time frequencies (annual and daily parameters) in order to predict next day Closing price (one-step ahead). Based on a four-step approach, this methodology is a serial combination of two LSTM algorithms. The empirical experiment is applied to 417 NY stock exchange companies. Based on Open High Low Close metrics and other financial ratios, the approach proves that the stock market prediction can be improved.

Keywords: **Times series forecasting, Prediction model, Long-Short Term Memory, Deep Learning.**

1. Introduction

Financial Times Series Forecasting is a crucial topic in finance and its prediction is an extensive ongoing research issue where computer science algorithms are very relevant. "It was estimated that, in 2012, approximately 85% of trades within the US Stock markets were performed by algorithms" [1]. To date, there has been numerous works linking Machine Learning to financial decisions and especially to trading strategies. [2] shows that "Back Propagation and Gradient Descent have been the preferred method for training finance structures due to the ease of implementation and their tendency to converge" [3], [4]. [5] Focuses on social interactions using a multi-agent based simulation. [6]'s paper introduces financial time series in a comparison between Logistic regression, MLP and Naive Bayes. Deep Learning outperforms in finance as well, and Long Short Term Memory (LSTM) is very promising in forecasting field because of its ability to memorize data. "Capturing spatio-temporal dependencies, based on regularities in the observations, is therefore viewed as a fundamental goal for Deep Learning systems."

[7] We aim to use a Long Short Term Memory ensemble method with two input sequences, a sequence of daily features and a second sequence of annual features, in order to predict the next day closing price and make a better decision in trading. LSTM is the most convenient for the following reasons:
• LSTM is considered as an improvement of Recurrent Neural Network which comes as a solution to vanishing and exploding gradient, see figure (Fig 1) [8], [9], [10]
• LSTM is suitable for sequences [11]
• LSTM can store and retrieve information using its gates [12]
• LSTM does not flow in a single way (unlike Neural Networks)
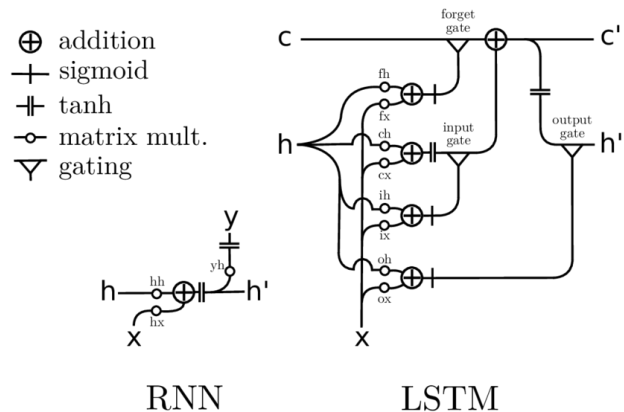• LSTM technique can distinguish between recent and early examples [13].

Figure 1: LSTM and RNN

New studies focuses on Ensemble methods where the weakness of a method is balanced out by the strength of another to produce high quality [14]. Ensemble learning can be either parallel or serial. A parallel ensemble results

from different learners, which are combined according to the schemes of Majority Voting, Weighted Majority Voting, Min, Max, etc... Serial learning arranges different base learners in sequence and selects the result of one learner as the final output. Our study relies on heterogeneous features with two temporal frequencies, it is then more appropriate to use two LSTM learners to form one ensemble method. The main contributions of this study are the following: (1) a new ensemble method where two time frequency inputs can be combined to predict the next day closing stock price; (2) evaluation of the method by comparing it to other Machine Learning techniques and financial strategies. The paper is divided into four sections. In section II we present some of the related work to this subject that can be found in the literature. Section III details the research methodology, describes the data pre-processing and develops the experiment steps. Section IV analyses the results. Finally, in section V, we draw some conclusions and future work. The objective of this project is to study the effectiveness of an LSTM ensemble method on the stock market prediction based on Open, High, Low, Close and Volume indicators as well as some financial metrics used by traders like MACD and RSI and other financial ratios. We evaluate its performance in terms of RMSE and other measures through experiments on real NYSE Stock Exchange data, analyze, and compare our results to other machine learning techniques.

2. Literature Review

Times series prediction is a field that has been related for years by statistics and Machine Learning techniques..

2.1 Pure single learner for Stock market prediction

1) Autoregressive integrated moving average (ARIMA: Researchers have exploited statistical techniques to predict and forecast financial times series like ARIMA which requires strict assumption like stationarity while finance forecasting is complex, noisy and non-stationary [8].
2) Artificial Neural Network: [15]'s paper shows that Artificial Neural Network is used for Times Series Forecasting since 1980s, more than 5000 publications has been released on ANN for Forecasting [16]. However, ANN presents issues like overfitting, impact of initial values and alpha. In order to tackle these issues, Deep Neural Network can be a solution.
3) Deep Neural Network: DNN is effectively more robust toward overfitting and can model complex nonlinear relationship between dependent and independent variables [7]. However, [17] explains some of the drawbacks of Deep Neural Network as:
-Slow Convergence time

-Vanishing or exploding gradient
-Expensive computation.
4) Deep Learning: [18] used a Deep Belief Network and gets 1, 25% test error. Deep Learning has the following advantages:
-Ability to learn complex
-Aptitude to learn with little human input with low level/intermediate/high level of abstraction
 -Strong unsupervised learning.

2.2 Ensemble and hybrid learning models for Stock market prediction

[19] Apply a hybrid model based on ARIMA and Neural Network to raise the prediction precision, [20] shows than ARIMA-ANN model outperforms statistical methods. [21] Chooses to combine neural tensor and Convolutional Neural Network to predict short and long-term influences. [8] use Stacked Auto-encoders and Long Short Term memory.

3. Methodology

Our model combines two feature categories in order to transmit the different aspects of an equity to the network. The first aspect considers historical daily stock price that reflects its trend. OHLC (Open, High, Low and Close stock prices), volume and trading indicators MACD, RSI and Signal show a share evolution based mainly on supply and demand. The second aspect consists of financial metrics that are part of the financial analysis. The objective is to take advantage of both technical analysis and fundamental-financial analysis. We adopt an Ensemble LSTM approach as explained in the figure below (Fig 2):

3.1 Data exploration

The first step in this research is Data exploration. The database used in this research contains New York Stock Exchange S&P500 data, publicly available at https://www.kaggle.com/dgawlik/nyse, for the period between 04/01/2010 and 30/12/2016. It is splitted into 2 Datasets, one for prices : Open - Low - High - Close - Volume and daily indicators : MACD - Signal - RSI.
Moving Average Convergence Divergence (MACD) is a trend-following momentum indicator and when MACD crosses its signal, it can function as a buy and sell signal. Relative Strength indicator (RSI) indicates the internal strength of an equity, it also reflects the vitality of increases over decreases. The second Dataset contains 75 financial annual ratios that shows informations about the company itself. The following ratios are some of them:
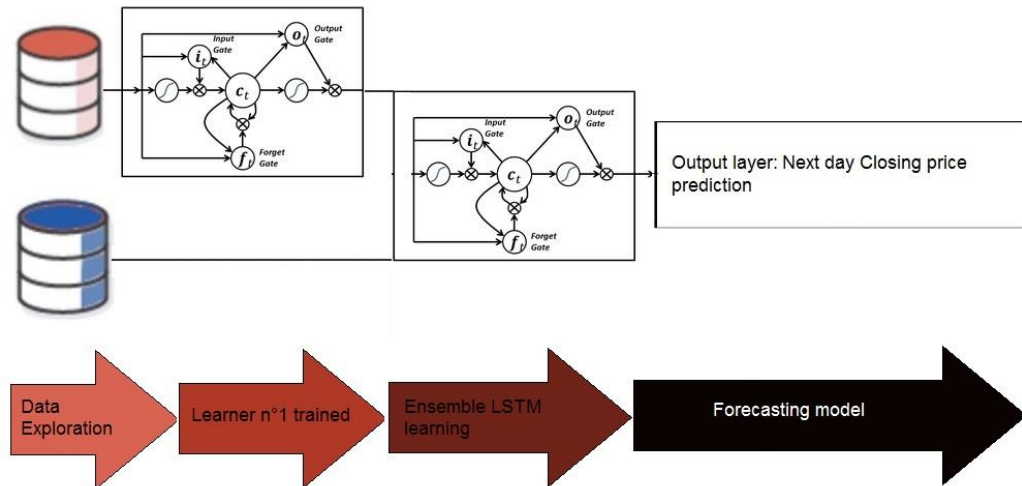
Figure 2: Ensemble LSTM approach

•EPS (Earning Per Share)
• ROE (Return On Equity)
• Payout ratio
• Dividend yield
• PER (Price Earning Ratio)
• PBR (Price to Book ratio)

Data collected contains 417 companies, the training set consists of consecutive observations from 04/01/2010 to 07/08/2015 while the test set consists of observations from 08/08/2015 to 31/12/2016 which represents sequences of 1408 time series for the training set and 352 for the test set. Before using Data as input for our Ensemble LSTM method, missing values are handled using Mean and Logistic regression functions in Matlab, then z-score normalization is applied (see equation 1).

$$z = (x - \mu)\,\sigma \quad (1)$$

MACD,the Signal and RSI are defined by the following formulas (2) and (3) :*

$$\begin{cases} MACD = \underset{12}{EMA} - \underset{26}{EMA} \\ EMA = (\underset{today}{Value} * \frac{smoothing}{(1+days)}) + (\underset{yesterday}{Value} * (1 - \frac{smoothing}{(1+days)})) \\ smoothing = \frac{2}{(selectedtimeperiod+1)} \end{cases} \quad (2)$$

$$\begin{aligned} RSI(stepone) &= 100 - \frac{100}{1 + \frac{AverageGain}{AverageLoss}} \\ RSI &= 100 - \frac{100}{1 + \frac{PreviousAverageGain*13+CurrentGain}{PreviousAverageLoss*13+CurrentLoss}} \end{aligned} \quad (3)$$

3.2 Hierarchical LSTM technique

The chosen Hierarchical LSTM technique is based on the assumption that the use of Ensemble method is a way

"such that the weakness of a method will be balanced out by the strength of another". The second step in our model is to construct the first LSTM learner also considered as the first weak learner (see fig 2). As a reminder, Long Short Term Memory (LSTM) is a special type of Recurrent Neural Network (RNN) with the capability of learning long term and short-term dependencies. Relying on gates, this Deep Learning algorithm has the ability to memorize. It is then well suited for times series prediction (see fig 1) and solve the problem of Vanishing gradient. This model contains an LSTM algorithm modeling prediction based on annual features (financial ratios) (see fig 3) for the sequence between 2012 and 2016. The input of the first learner consists of 75 features over a 5-sequence (2012 - 2013 -2014 - 2015 - 2016). The output result is a prediction for the closing price considered as one of the inputs of the second LSTM learner. Then, the serial ensemble learning is the result of the combination of the first learner's output (prediction) and the other daily features as an input for the second learner as follows:

• Open
• Highest price for the day
• Lowest price for the day
• Close
• Volume
• MACD
•Signal
• RSI
• The first learner's prediction

The training is then executed over two stages (sequentially). After training, the model is capable of forecasting and predicting the Closing price one-step ahead (see fig 2)
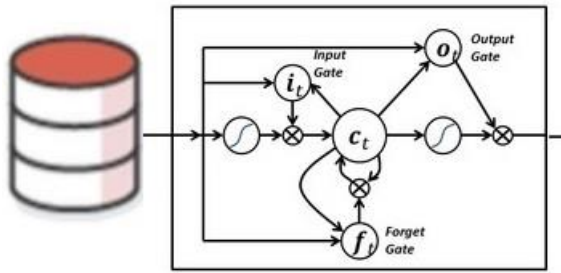
Figure 3: LSTM n°1

4. RESULTS AND EVALUATION

Experiments were carried out to predict the next day Closing price following the approach described above. As explained, the Ensemble LSTM approach is composed of one main dataset splitted into two datasets sequentially different since one is for annual features and the other for daily features. The empirical experiment is applied on 417 New York Stock Exchange companies using financial ratios for the first learner. Then the predicted value is included into the input of the second LSTM learner. This article aims to analyze the hierarchical LSTM technique in terms of Root Mean Square Error (RMSE) and Forecasting.

4.1 Results

The Ensemble LSTM is a Regression algorithm that allows predicting the next day Closing price. To fix the architecture of our method, several experiences were done to choose the number of hidden units and the learning rate in order to get the best results. The topology chosen is 20 hidden nodes for the first LSTM and an initial learning rate of 0.005, the output is then included into the input of the second LSTM, with a total of 9 features and 200 hidden nodes. Figures 5 presents the Ensemble LSTM model performance in terms of Mean Squared Error and Forecasting. Our model shows a great performance since RMSE is equal to 0.0119, which is more interesting than LSTM standalone for the first Database alone with 0.0124.

4.2 Analysis

In order to determine if there is an improvement while applying an ensemble model, we executed the LSTM algorithm to the first Database (with daily features) based on a sequence to determine the last Closing price, we then got an RMSE equal 0.0124. Then we applied LSTM standalone to the annual Database and got RMSE equals to 0.08. Finally, to get sure that LSTM is more appropriate in case of Times series, we chose to compare it to Neural Network. We used Neural Network for times series with

Matlab, LevenbergMarquardt as a training algorithm and 10 hidden nodes. We got RMSE equals to 0.067 which is better than LSTM applied to the second Database 6. Table I compare our approach to LSTM standalone, using each dataset apart and then to Neural Network.

LSTM 1	LSTM 2	Ensemble LSTM	NN
0.0124	0.08	0.0119	0.07

Table 1: RMSE for the LSTM standalone, LSTM ensemble and NN

LSTM's ability to store historic data for a long period and retrieve valuable data when required makes this technique a good technique for time series forecasting.

5. CONCLUSION AND FUTURE WORK

In this study, we established an ensemble method based on LSTM algorithm to predict the next day Closing price. This model takes into account two frequencies with two types of variables, daily and annual variables. It confirms then our hypothesis that the combination of these two databases gives a better performance. This ensemble method can be a useful tool for traders and stakeholders to determine their trading strategy. For future work, we attend to achieve a benchmark between this hierarchical technique and traditional finance methods. Furthermore, Deep Learning and LSTM are very promising, we can then expect to improve this model and achieve online Fig. 4. Forecast vs real values when LSTM is applied to DB number 1 alone LSTM.jpg Fig. 5. Forecast vs real values when Ensemble LSTM approach is used Fig. 6. Neural Network response prediction for forecasting.
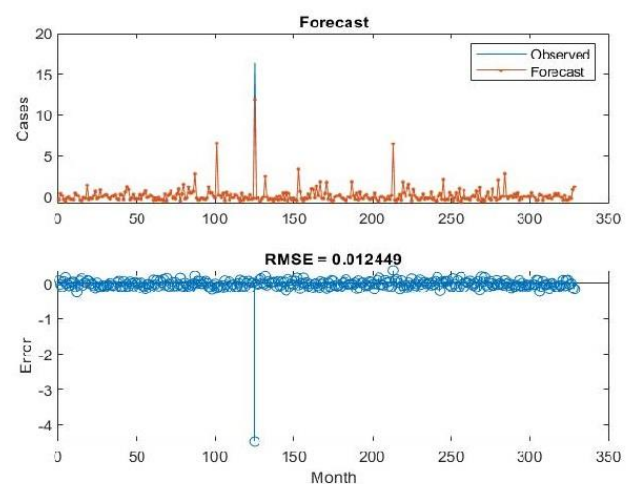

Figure 4: Forecast vs real values when LSTM is applied to DB number 1 alone
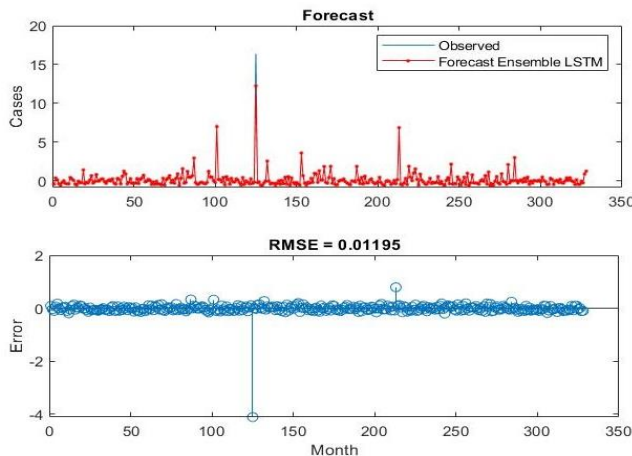
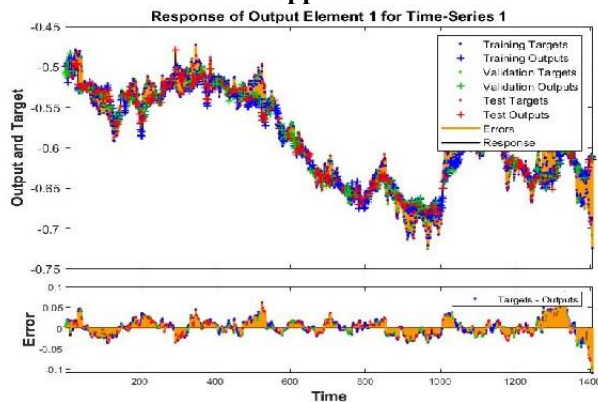**Figure 5: Forecast vs real values when Ensemble LSTM approach is used**

Figure 6: Neural Network response

References

[1] M. Glantz and R. Kissell, "Multi-asset risk modeling: techniques for a global economy in an electronic and algorithmic trading era," 2013.

[2] M. Dixon, D. Klabjan, and J. H. Bang, "Classification-based financial markets prediction using deep neural networks," Algorithmic Finance, no. Preprint, pp. 1–11, 2017.

[3] R. Fehrera and S. Feuerriegela, "Improving decision analytics with deep learning: The case of financial disclosures," stat, vol. 1050, p. 9, 2015.

[4] P. Vieira and A. Teixeira, "Are finance, management, and marketing autonomous fields of scientific research? An analysis based on journal citations," Scientometrics, vol. 85, no. 3, pp. 627–646, 2010.

[5] Z. Kodia, L. B. Said, and K. Ghedira, "A study of stock market trading behavior and social interactions through a multi agent based simulation," in KES International Symposium on Agent and Multi-Agent Systems: Technologies and Applications. Springer, 2010, pp. 302–311.

[6] B. Batres-Estrada, "Deep learning for multivariate financial time series," 2015.

[7] I. Arel, D. C. Rose, T. P. Karnowski et al., "Deep machine learninga new frontier in artificial intelligence research," IEEE

computational intelligence magazine, vol. 5, no. 4, pp. 13–18, 2010.

[8] Y. Deng, F. Bao, Y. Kong, Z. Ren, and Q. Dai, "Deep direct reinforcement learning for financial signal representation and trading," IEEE transactions on neural networks and learning systems, vol. 28, no. 3, pp. 653–664, 2017.

[9] Werbos and P.J., "Generalization of backpropagation with application to a recurrent gas market model," Neural networks, vol. 1, no. 4, pp. 339–356, 1988.

[10] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[11] I. Sutskever, L. Vinyals, and Q.V., "Sequence to sequence learning with neural networks," Advances in neural information processing systems, pp. 3104–3112, 2014.

[12] e. a. Bengio, Yoshua, "Learning deep architectures for AI," Foundations and trends R in Machine Learning, vol. 2, no. 1, pp. 1–127, 2009.

[13] D. M. Nelson, A. C. Pereira, and R. A. De Oliveira, "Stock market's price movement prediction with LSTM neural networks," International Joint Conference on Neural Networks (IJCNN), pp. 1419–1426, 2017.

[14] P.-C. Chang, C.-H. Liu, C.-Y. Fan, J.-L. Lin, and C.-M. Lai, "An ensemble of neural networks for stock trading decision making," in International Conference on Intelligent Computing. Springer, 2009, pp. 1–10.

[15] T. Kuremoto, S. Kimura, K. Kobayashi, and M. Obayashi, "Time series forecasting using a deep belief network with restricted Boltzmann machines," Neurocomputing, vol. 137, pp. 47–56, 2014. [16] S. Crone and K. Nikolopoulos, "Results of the NN3 neural network forecasting competition," in The 27th International Symposium on Forecasting, 2007, p. 129.

[17] J. Schmidhuber, "Deep learning in neural networks: An overview," Neural networks, vol. 61, pp. 85–117, 2015.

[18] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," Neural computation, vol. 18, no. 7, pp. 1527–1554, 2006.

[19] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," Neurocomputing, vol. 50, pp. 159–175, 2003.

[20] P. Werbos, "Beyond Regression:" New Tools for Prediction and Analysis in the Behavioral Sciences," Ph. D. dissertation, Harvard University, 1974.

[21] X. Ding, L. Zhang, Yue, and e. a. Ting, "Deep learning for eventdriven stock prediction," Twenty-Fourth International Joint Conference on Artificial Intelligence, 2015.