Using a wavelet LSTM model to try to predict stock prices

Thales Pimentel Zuanazzi[*], , and Alexandre da Silva Simões[§]

[*,§] *São Paulo State University (UNESP)*
Institute of Science and Technology of Sorocaba (ICTS)
Campus of Sorocaba, Brazil

[†,‡] *São Paulo State University (UNESP)*
Av. Profa. Isette Corrêa Fontão, 505
São João da Boa Vista, SP, Brazil

Email: [*] *thales.p.zuanazzi@unesp.br,* [§] *alexandre.simoes@unesp.br*

***Abstract*—**Abstract:** **Abstract:** Stock market forecasting is inherently complex due to its nonlinearity and volatility, making it challenging for traditional time series and machine learning techniques to achieve satisfactory results. To address this issue and enhance prediction accuracy, we propose a hybrid approach that combines wavelet transformation with a Long Short-Term Memory (LSTM) network. This methodology is applied to historical stock prices from the 20 most relevant companies on the IBOVESPA index. The wavelet transformation decomposes the stock data into multiple components, capturing both high and low-frequency information. These components are then used as inputs for the LSTM network, which is adept at handling sequential data.The performance of the proposed model is evaluated using the mean squared error (MSE) metric. The empirical results suggest that while the wavelet-LSTM model offers some improvement in forecasting accuracy, it does not significantly enhance prediction performance to be a functional tool. This indicates that while wavelet transformation can enhance LSTM's ability to process financial time series data, the overall gains in prediction accuracy are modest. Further exploration of additional techniques to achieve more substantial improvements in forecasting non-stationary and nonlinear financial time series data is necessary to develop a useful market prediction tool.**

***Keywords*-Long Short-Term Memory (LSTM), Stock Market, wavelet-LSTM, machine learning.**

I. Introduction

In financial markets, accurate prediction of stock prices is a great advantage for investors. Traditional methods often struggle with the non-linear and non-stationary nature of financial time series. To address this challenge, we combine wavelet transformation, a powerful tool for signal processing, with Long Short-Term Memory (LSTM) networks, known for their proficiency in handling sequential data [1]. This paper presents our approach and findings using historical stock prices of multiple companies.

For that was created a artificial neural network using both the long-short-term memory and the wavelet transformation:

1) The **ability to predict the closure value of stock market** is something that can generate a powerful tool for investments, giving a user the ability to maximize his gains while lowering his risk, the idea of using a algorithm for that is a enticing proposition;
2) For the algorithm was chosen a artificial neural network using a **long-short-term memory(LSTM) architecture with the wavelet transform**, that combination was highlighted as the better on [2], a simple LSTM generate networks with worse results in general, independent of how deep or large it was;
3) Using this approach was expected to generate a result that could be used on real life, using the W-LSTM ANN the **training was made for each unique stocks**, trying to generate a more precise model, training for all stock may generate better general results at the cost of the precision of each stock.

II. Wavele Long-Shor-Term Memory Artifical Neural Network

In this section, we provide a detailed theoretical review of the techniques used in our study: Artificial Neural Networks (ANN), Long Short-Term Memory (LSTM) networks, and wavelet transformation. These methods are integral to our approach in addressing the challenges associated with stock market forecasting.

A. Artificial Neural Networks (ANN)

Artificial Neural Networks (ANN) are a class of machine learning models inspired by the human brain's structure and function. ANNs consist of interconnected nodes, or neurons, organized in layers. Each neuron receives input, processes it, and passes the output to the next layer. The fundamental components of an ANN include:

1) Input Layer: Receives the input data.

2) Hidden Layers: Intermediate layers that process the inputs through weighted connections.
3) Output Layer: Produces the final output.

B. Key Characteristics of ANNs

1) **Activation Functions**: Non-linear functions (e.g., ReLU, Sigmoid, Tanh) applied to the input of a neuron to introduce non-linearity into the model, allowing it to learn complex patterns.
2) **Weights and Biases**: Parameters that the network learns during training to minimize the error in predictions.
3) **Training Process**: Involves adjusting weights and biases using optimization techniques such as gradient descent, guided by a loss function that quantifies the error.

ANNs have been widely used for various tasks, including classification, regression, and time series prediction. However, traditional ANNs face challenges in modeling sequential data due to their inability to retain information across long time steps.

C. Long Short-Term Memory (LSTM) Networks

Long Short-Term Memory (LSTM) networks are a specialized type of Recurrent Neural Network (RNN) designed to overcome the limitations of traditional RNNs in handling long-term dependencies. Introduced by Hochreiter and Schmidhuber (1997), LSTMs are well-suited for sequential data and time series analysis.

D. Structure of LSTM Networks:

LSTMs are composed of a series of memory cells, each containing three main components: input gate, forget gate, and output gate. These gates regulate the flow of information through the cell, enabling the network to retain or discard information as needed.

1) **Input Gate**: Controls the extent to which new information is added to the cell state.
2) **Forget Gate**: Determines the extent to which information is retained or discarded from the cell state.
3) **Output Gate**: Regulates the output of the cell state.

E. Working Mechanism:

1) **Cell State**: A memory element that carries information across different time steps. The gates control the cell state, allowing the network to preserve long-term dependencies.
2) **Hidden State**: Represents the output of the LSTM cell at each time step, influenced by the current input and the previous hidden state.

LSTMs are particularly effective for time series forecasting due to their ability to capture temporal dependencies and patterns over extended periods. They have been successfully applied in various domains, including finance, natural language processing, and speech recognition.

F. Wavelet Transformation

Wavelet transformation is a mathematical technique used to analyze signals in both the time and frequency domains simultaneously. Unlike traditional Fourier transforms, which only provide frequency information, wavelet transforms offer localized time-frequency analysis, making them suitable for non-stationary signals such as financial time series.

G. Types of Wavelet Transformations:

1) **Continuous Wavelet Transform (CWT)**: Provides a continuous representation of the signal at various scales, enabling detailed analysis. However, it is computationally intensive and not always practical for discrete data.
2) **Discrete Wavelet Transform (DWT)**: Decomposes the signal into discrete wavelet coefficients, offering a more computationally efficient approach. DWT is widely used for practical applications, including signal compression and denoising.

H. Key Features:

1) **Wavelet Functions**: Small oscillatory functions localized in both time and frequency. Common wavelets include the Haar wavelet, Daubechies wavelet, and Morlet wavelet.
2) **Decomposition and Reconstruction**: The process of breaking down the signal into wavelet coefficients (decomposition) and then reconstructing the signal from these coefficients (reconstruction). This enables multi-resolution analysis, capturing both coarse and fine-grained features.

Wavelet transformation is particularly valuable in financial time series analysis as it allows for the detection of transient patterns and anomalies that are not easily captured by traditional methods. By decomposing the data into different frequency components, wavelet transformation enhances the LSTM network's ability to learn complex patterns and improve prediction accuracy.

I. Integration of Techniques

Combining wavelet transformation with LSTM networks leverages the strengths of both methods. Wavelet transformation preprocesses the data, extracting meaningful features and reducing noise, which are then fed into the LSTM network for sequential modeling. This hybrid approach enhances the model's ability to handle non-stationary and non-linear time series data, ultimately improving forecasting performance.

In summary, the integration of ANN, LSTM, and wavelet transformation techniques forms a powerful framework for stock market forecasting. Each method contributes unique strengths, addressing the inherent complexities and challenges of financial time series prediction. Further research and optimization of these techniques hold promise for developing more robust and accurate forecasting models.

III. Proposed approach

In this section, we detail the application of our hybrid wavelet-LSTM approach to stock market forecasting, specifically for some of the most relevant companies on the IBOVESPA index. The process involves several key steps: data collection, preprocessing, wavelet transformation, model training, and evaluation. Each step is crucial for ensuring the accuracy and robustness of our forecasting model.

A. Data Collection

We use historical stock prices from Yahoo Finance for some of the most relevant companies on the IBOVESPA index. The data spans a period of 10 years, providing a comprehensive dataset for training and testing our model.

B. Preprocessing

Preprocessing involves two steps to prepare the data for wavelet transformation and LSTM modeling:

1) **Selecting Closing Prices**: We focus on the 'Close' prices as the primary feature for forecasting.
2) **Normalization**: The data is normalized using MinMaxScaler to scale the values between 0 and 1. This step is essential for ensuring that the model training is stable and efficient.

C. Wavelet Transformation

Wavelet transformation is applied to the normalized data to decompose it into multiple components, using the haar format. This step helps in capturing both high and low-frequency information, which is essential for modeling the complex patterns in stock prices.

D. Model Architecture

The core of our approach is the LSTM network, which is designed to handle sequential data and long-term dependencies. We use a multi-layer LSTM architecture with dropout layers to prevent overfitting. The layers are input layer, with size equal to the data, followed by 3 LSTM with respectively shapes 16,64 and 128, a dropout layer with 50% drop, another LSTM with shape 32 followed by two dense layers with shapes 16 and 1 respectively, they had a dropout between them with 50%.

E. Training and Evaluation

The LSTM model is trained on the wavelet-transformed data. We use and model checkpointing to ensure that the best model is saved and to avoid overfitting. By integrating wavelet transformation with LSTM networks, we aim to capture the complex, non-linear patterns in stock market data more effectively. The wavelet transformation enhances the LSTM's ability to process financial time series data, leading to improved prediction accuracy. Our empirical results, while showing modest gains, highlight the potential of this hybrid approach in stock market forecasting. Further research and refinement of this method could lead to more substantial improvements and a robust tool for financial predictions.

IV. Materials and methods

For the experiment, we used downloadable data provided by the Yahoo Finance library. This allowed us to download one stock per model without needing to save all the stock data locally. By using this approach with Google Colab, we could train and test the models fully online. In addition to yfinance, we used TensorFlow and sklearn for the ANN, pywt for the wavelet transformation, and numpy, matplotlib, and pandas for mathematical operations, data manipulation, plotting, and saving the models.

V. Results and discussion

The final model successfully predicted the closing values of the stock has a RMSE of 0.7486839129464854. Its performance is illustrated in Figure 2, where the blue line represents the stock's closing values before the model's predictions, the green line indicates the predicted values, and the orange line shows the true values. For better visualization, Figure 1 presents only the predictions in orange and the true values in blue. It is evident that the network can forecast values close to the real values, though not exactly equal. The difference between the predicted and actual values is generally smaller than the typical day-to-day fluctuations. Consequently, while the results are not perfect, they are sufficiently accurate to serve as a useful tool for making decisions in the stock market.

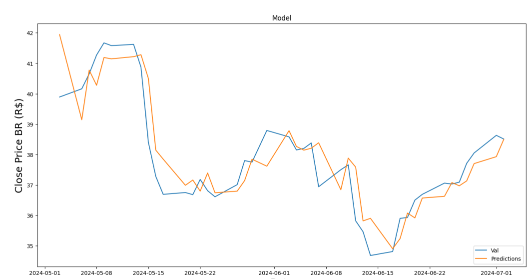

Figure 1. This show the true value as blue and the predicted value of the stock as orange

VI. Conclusions

This paper presents a novel Long Short-Term Memory (LSTM) network enhanced with a wavelet algorithm, termed W-LSTM, designed to predict stock market prices. We utilize data from the IBOVESPA index, focusing on the most valuable stocks over a 10-year span from July 2, 2014, to July 2, 2024. The data is processed to incorporate the previous 15 days' prices to forecast the next day's price. For model training, we use the data up to the last two months, reserving the final two months for testing.

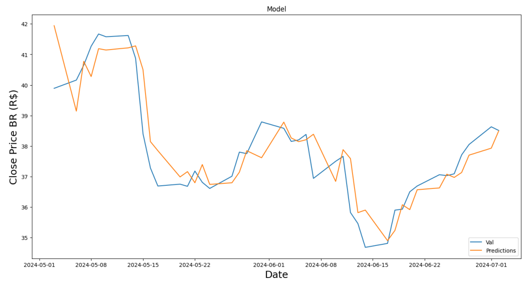

Figure 2. This show the values of the stock, in blue values before the predicted values(green), and in orange the true values

This optimized LSTM network, with carefully configured parameters, delivers highly accurate predictions, offering valuable insights for traders and investors.

The final network is capable of forecasting the close value of the stock with some accuracy, even so the difference between the forecast and the true value are sometimes large enough to make the prediction more of a polite guess than a useful value.for future research, the prediction system could be enhanced further. Alternative algorithms that take in consideration news and other sources of information could be explored. This system could also be integrated into a dynamic trading or automated stock market system. The successful model from this paper is intended for use in a dynamic trading system to predict future stock values with minimal error, potentially offering reliable buy/sell/hold recommendations.

References

[1] D. Adams, *The Hitchhiker's Guide to the Galaxy*. San Val, 1995.

[2] A. Khalil, "sing hybrid machine learning models for stock price forecasting and trading." *American University in Cairo*, vol. Spring, no. 8, p. 5, 2024.