Comparing the Performance of Long Short-Term Memory Architectures (LSTM) in Equity Price Forecasting: A Research on the Mexican Stock Market

Comparación del desempeño de arquitecturas de memoria a corto y largo plazo (LSTM) en el pronóstico de precios de acciones: una investigación sobre el mercado bursátil mexicano

Samuel García

EGADE Business School, Tecnológico de Monterrey, Mexico

160

Abstract

This study compares the performance of univariate and multivariate Long Short-Term Memory (LSTM) to predict next-day closing prices on four stocks in the consumer retail sector of the Mexican Stock Exchange. Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Median Absolute Percentage Error (MdAPE), and Root Mean Squared Error (RMSE) are used to test the networks' performance. Results show a better performance on multivariate price forecasts when using 20-day and 15-day length sequences, generating consistent results for the sample, including illiquid and liquid stocks. On the other hand, univariate LSTM discloses lower forecast performance when predicting the price of illiquid stocks.

Keywords: *forecast, stocks, univariate, multivariate, LSTM.*
JEL Classification: *G1, G15, G20, C6.*

Resumen

Este trabajo compara el desempeño de la memoria de corto y largo plazo (LSTM, por sus siglas en inglés) univariada y multivariada en la predicción de los precios de cierre del día siguiente de cuatro acciones del sector de consumo minorista en la Bolsa Mexicana de Valores. El error absoluto medio (MAE, por sus siglas en inglés), el error porcentual absoluto medio (MAPE, por sus siglas en inglés), la mediana del error porcentual absoluto (MdAPE, por sus siglas en inglés) y la raíz del error cuadrático medio (RMSE, por sus siglas en inglés) se utilizan para probar el desempeño de las redes. Por un lado, los resultados muestran un mejor desempeño en el pronóstico multivariado de precios cuando se utilizan secuencias de 20 y 15 días de duración, generando resultados coherentes para la muestra, incluidas las acciones líquidas e ilíquidas. Por otro lado, la LSTM univariada revela un desempeño de pronóstico menor para la predicción del precio de acciones ilíquidas.

Palabras clave: *predicción, acciones, univariada, multivariada, LSTM.*
Clasificación JEL: *G1, G15, G20, C6.*

1. Introduction

Given the growing complexity of the global financial industry and the unstable nature of the financial markets, pricing analysis of financial assets—like stocks—and predicting future prices and returns in the financial market is a complex and challenging activity, highly valued in the financial sector. Since noise and non-parametric and non-linear dynamics are characteristic of the stock market, its traditional statistical tools to analyze historical data—where past events have great importance in predicting future states (e.g., price and returns) and trends—may struggle to model those dynamics on stock prices over time (Pramod & Mallikarjuna, 2020; Bhandari et al., 2022).

In recent years, developments in ML, AI, and Deep Learning (DL) have played a central role in enhancing stock price prediction. A case in point is that academics have noticed the advantages of DL models when capturing non-linear features of data sequences through Recurrent Neural Networks (RNN) and LSTM networks (Tianxiang & Zihan, 2020). ML is a sub-field of AI, which tries to emulate some human cognitive features like the learning process to identify patterns and/or classify specific sets of objects and is currently used in the financial sector because of its analytical capabilities to analyze and manage big data (Lu, 2017; Liebergen, 2017).

Artificial Neural Networks (ANN) are part of deep learning, which attempt to recreate the logic of the human brain to perform cognitive tasks. These models are mainly based on the interconnection of individual neurons, which creates a network (Nielsen, 2015; Krenker et al., 2011; Tirozzi et al, 2007); RNNs and LSTMs are a subset of Neural Networks, mainly designed to capture information on historical data.

According to the literature reviewed, there is no extensive research on the price forecasting capabilities of LSTM in Latin American markets. Given the importance of AI and DL techniques in the analysis and prediction of financial assets' prices, as well as the growing presence of Latin American financial markets on the global financial landscape, this research is focused on comparing the performance of univariate and multivariate LSTM when predicting next day closing price of four Mexican stocks from the consumer retail sector in the Mexican Stock Exchange (BMV, in Spanish), as well as analyzing the impact of the size of the sequence length used for prediction accuracy. As mentioned, the sample used for this work includes four stocks, two of them liquid when comparing the 3-month and 10-day average traded volume with the other two stocks in the sample. The contribution of this research is to test

Samuel García

Comparing the Performance of Long Short-Term Memory Architectures (LSTM)
in Equity Price Forecasting: A Research on the Mexican Stock Market
*Comparación del desempeño de arquitecturas de memoria a corto y largo plazo (LSTM) en el pronóstico
de precios de acciones: una investigación sobre el mercado bursátil mexicano*

the performance of LSTM when predicting stock prices, using different historical timeframes, to predict the prices of four Mexican stocks from the consumer retail sector.

As part of the results, it can be observed that the size of the rolling window impacts the performance of the univariate model when predicting the price of illiquid stocks, assessed through four performance metrics to measure the magnitude of errors: Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Median Absolute Percentage Error (MDAPE), and Root Mean Squared Error (RMSE), whereas the performance of the multivariate output shows consistency for both illiquid and liquid stocks.

This paper is ordered as follows: Section 2 discloses works related to LSTMs prediction capabilities for a stock process; Section 3 provides an overview of the tested architectures and their features; Section 4 contains methodology and model implementation; Section 5 includes performance metrics; Section 6 details preliminary results; and Section 7 contains conclusions.

2. Related Works

Pricing prediction of financial products is a significant issue in the financial sector and the academy. Currently, several ML and AI models are used to enhance price prediction accuracy; most are based on RNNs, LSTM, and other DL. As mentioned in the previous section, RNNs were mainly designed to capture information on historical data. Focusing on RNNs and LSTM DL models, academic research is centered on analyzing the prediction capabilities of the DL through plain and mixed deep learning models, testing different variables, architectures, and levels of model parameters to obtain a better analysis and prediction accuracy. For instance, Nourbakhsh and Habibi (2023) combined Convolutional Neural Network and LSTM as well as specific variables used in fundamental analysis, to enhance the model's accuracy measured through MAE and MAPE. Also, Zaheer et al. (2023) explored the capabilities of a hybrid deep-learning model based on single and mixed RNN, LSTM, and CNN architectures to predict closing and high prices on the next trading day of the Shanghai Composite Index, where they found that a single layer RNN outperforms the other tested models showing the lowest MAE and RMSE metrics.

Another interesting option to enhance prediction capabilities is found in Tianxiang and Zihan (2020), who proposed a method to predict the West Texas Intermediate oil price, from January 1986 to January 2020, with the LSTM and GM (1,1) model, based on a multi-step prediction method. The model showed significant prediction accuracy measured through MAPE and RMSE. The model effectively captured long-term effects with lower frequency and also price trends. The work performed by these authors can show how mixing different DL models enriches the existing literature on price prediction of financial assets.

The model proposed in this paper analyses how the length of historical data would contribute to better price forecasting, as mentioned by Bhandari et al (2022) who used 15 years of market and macroeconomic data, as well as technical indicators to predict the closing price for the S&P 500 index through a multivariable LSTM. They found the best performance results were based on RSME, MAPE, and a correlation coefficient obtained through a single-layer model.

Although this research evaluates price prediction for the next trading day, other studies examined the prediction accuracy for longer periods. For example, Ghosh et al. (2019) employed LSTM techniques on historical stock price data of five companies from some pre-decided sectors in the Indian market, to infer future trends. The authors proposed a framework based on LSTM models to calculate the best time length to forecast the future share price of a company from a particular sector, as well as predict the future growth of a company for periods of 3 and 6 months, and 1 or 3 years. They found a decrease in the error level when using test data for longer periods, dependencies, and the same growth rate in companies from a certain sector.

Published studies on price prediction have also explored the impact of transformed variables, the number of layers in models, parameter levels, and the length of historical data used, for better model learning, and enhancing prediction capabilities on DL models when compared to other statistical tools. For instance, Andi (2021) normalized variables on a data set to compare the performance of LSTM with other prediction models, like linear regression and the Lasso algorithm, concluding that the first model obtained the most accurate forecast on the bitcoin price based on accuracy, precision, recall, and sensitivity because using common variation ranges on the variables allow capturing trends.

Finally, Pramod and Mallikarjuna (2020) explored predicting Tata Motors Limited's stock price using LSTM. The output produced a low loss and low error rate. They also found that increases in layers and epoch batch rates had a positive impact on the performance.

Samuel García
Comparing the Performance of Long Short-Term Memory Architectures (LSTM)
in Equity Price Forecasting: A Research on the Mexican Stock Market
Comparación del desempeño de arquitecturas de memoria a corto y largo plazo (LSTM) en el pronóstico
de precios de acciones: una investigación sobre el mercado bursátil mexicano

Overall, there is an extensive body of research focused on measuring the accuracy of DL models to enhance forecasting capabilities. For example, LSTM architectures combined with other DL models have been used, as well as other techniques like transformed variables to improve model performance, all these evaluated under different financial markets in Asia, Europe, and North America. However, there is no extensive body of research assessing the performance of LSTM models in predicting stock prices in the Latin American financial markets. The contribution of this research is to test the performance of LSTM, using different timeframes to predict the prices of four stocks issued by Mexican firms, with different liquidity attributes, in the Mexican Stock Exchange.

3. A Brief on LSTM

Recurring Neural Networks (RNNs) have loops to feedback other neurons in the architecture, hence the output of a neuron in the network impacts the input of another neuron, resulting in closed paths for the transmission of information in the network (Haykin, 2010). LSTMs are a type of RNN architecture, used to find patterns in data, where the occurrence of events of interest is uncommon in time and frequently mixed with other events (Bhandari et al., 2022; Pramod & Mallikarjuna, 2020).
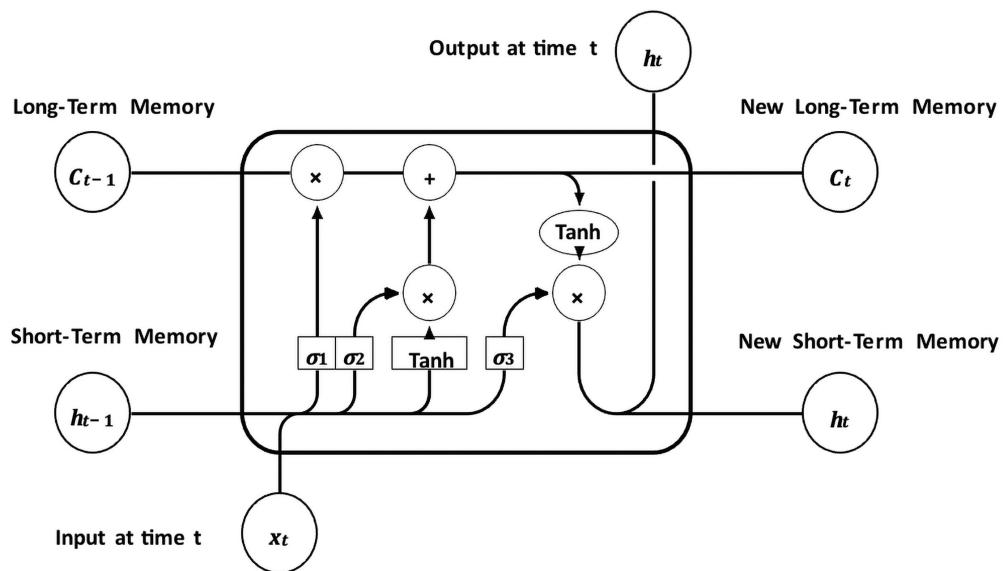
LSTMs deal with the problem of "long-term dependencies", present in RNNs, by retaining information from past inputs contained in a variable number of time steps, so they can manage to learn and allow facts of interest to persist over time while overcoming the vanishing and exploding gradient problem. As mentioned previously, this network can find relationships in historical data where the existence of the event of interest is scarce in the data set (Benchaji et al, 2021a; Yu et al, 2019; Benchaji et al., 2021b).

In general, a LSTM architecture (see Figure 1) has explicit memory blocks containing different states: a hidden state (h) and a cell state (C), which allow to store and manage both, short and long-term information through three gates (stages), each one performing an individual function:

1. Forget gate, which chooses, through a sigmoid function (σ_1), whether information coming from h_{t-1} and current input (x_t) needs to be remembered (values near to 1) or is irrelevant and can be forgotten (values near to 0).

2. Input (update) gate, this allows learning from the input x and h_{t-1} to update C, which contains the long-term information; the layer includes two parts: first, a sigmoid layer it will decide which new values will be stored in the cell state and second, a tangent layer creates a vector of new candidate data with values between -1 and 1 to rate relevant data. Then, the output of the input gate is obtained through multiplying the values of sigmoid layer and the tangent layer.

3. Output Layer determines the new hidden state (h_t), based on h_{t-1}, x_t and the tanh of the current cell state (C_t).

Figure 1. General Representation of a LSTM Cell

Source: Prepared by the author.

Where:

σ = Sigmoid Function

Tanh = Hyperbolic Tangent Function

Samuel García
Comparing the Performance of Long Short-Term Memory Architectures (LSTM)
in Equity Price Forecasting: A Research on the Mexican Stock Market
Comparación del desempeño de arquitecturas de memoria a corto y largo plazo (LSTM) en el pronóstico
de precios de acciones: una investigación sobre el mercado bursátil mexicano

4. Methodology and Model Implementation

4.1 Coding & Data Overview

The analysis for the Research was done using Scikit-learn, a Python-related library to create and implement ML models and perform statistical analysis and modelling; TensorFlow, a high-level open-sourced end-to-end platform to create DL and AI models, and Keras is a high-level open-sourced library which takes the underlying operations provided by other platforms like TensorFlow.

For this research, I used daily market data obtained from Yahoo Finance. The dataset contains stock transactions executed in the Mexican stock market, in the consumer retail sector from January 1, 2020, to February 9, 2024 (1036 workdays). This sector is important not only because it includes companies selling several retail products related to the basic needs of the Mexican population, distributed across Mexico, but the sector was resilient during the pandemic, presenting the smallest drop in value in the Mexican financial market. This sector had the speediest recovery in comparison to other sectors (Landazuri Aguilera & Ruíz Pérez, 2021).

The analyzed stocks were the following:

Grupo Comercial Chedraui, S.A.B. de C.V. (ticker: CHDRAUIB.MX)

- La Comer, S.A.B. de C.V. (ticker: LACOMERUBC.MX)
- Organización Soriana, S. A. B. de C. V. (ticker: SORIANAB.MX)
- Wal-Mart de México, S.A.B. de C.V. (ticker: WALMEX.MX)

The four stocks were selected, since all of them are nationwide supermarkets, selling comparable retail products with similar target markets, making them comparable in terms of business models.

Table 1 shows some market data for the four stocks used for the research (see Table 1). WALMEX and LACOMERUBC would be considered the most liquid stocks in the sample because both have the greater number of shares outstanding and average traded volume on a three-month and ten-day timeframes, which allow the stocks to be easily traded in the stock market at the current fair market price (Armitage et al., 2014).

Table 1. Market Statistics for the Analyzed Stocks

Statistics	Walmex	Soriana	Chedraui	La Comer
Average Volume on a 3-Month timeframe	15.08 M	65.88 k	311.35 k	645.09 k
Average Volume on a 10-Day timeframe	15.27 M	2.56 k	295.92 k	224.8 k
Shares Outstanding	17.46 B	1.8 B	959.82 M	1.09 B
Implied Shares Outstanding	17.46 B	1.85 B	959.82 M	N/A
Intraday Market Cap	1.19 T	62.76 B	120.30 B	N/A
Enterprise Value	1.21 T	81.19 B	163.16 B	N/A

Source: Prepared by the author with data from Yahoo Finance as of February 20, 2024.

The research was based on six variables extracted from the data set, including open, high, low, closing, and adjusted closing prices, as well as volume. During a trading day, open and close are prices at which the stock began and ended trading in the stock market, high and low prices are the highest and lowest traded prices for that stock, during a trading day. Adjusted (Adj) close price is the closing price after considering any splits and dividend distributions. Finally, volume indicates the total quantity of stocks traded during a day.

All variables in the dataset were normalized considering a 0 to 1 range to maintain a common scale and to contribute to the model's accuracy.

4.2 Model Implementation and Training

As mentioned in Section 1, this research aims to compare the prediction accuracy of univariate *vs*. multivariate LSTMs on stocks related to the consumer retail sector in Mexico. The research was performed through the following LSTM core architectures:

- Two hidden layers with 50 units each. For every network, the output h_t (see Figure 1), is an input for X_t, at time *t*, as shown before in Figure 1.

- A Dense layer with five neurons, to convert the output of the final layer into a vector.

- Finally, the vector flows to a linear activation, used to predict the next day's closing price.

Samuel García
Comparing the Performance of Long Short-Term Memory Architectures (LSTM)
in Equity Price Forecasting: A Research on the Mexican Stock Market
Comparación del desempeño de arquitecturas de memoria a corto y largo plazo (LSTM) en el pronóstico
de precios de acciones: una investigación sobre el mercado bursátil mexicano

Comparison between univariate and multivariate LSTM networks is performed using sequence lengths of 20, 15, and 10 historical stock prices and volume (described in Section 4.1). For example, Figure 2 discloses an architecture with a sequence length of 20 daily data (see Figure 2).

Figure 2. Architecture of a LSTM with a Sequence Length of 20 and 50 units

```
_____
Layer (type)                Output Shape
===============================================
lstm_2 (LSTM)               (None, 20, 50)

dropout_2 (Dropout)         (None, 20, 50)

lstm_3 (LSTM)               (None, 50)

dropout_3 (Dropout)         (None, 50)

dense_2 (Dense)             (None, 5)

dense_3 (Dense)             (None, 1)

===============================================
```

Source: Prepared by the author.

Both LSTMs predicted the close price for the next trading day. Multivariate, forecasting was based on the six features mentioned in section 4.1. Univariate forecasting was run using the close price.

4.3 Hyperparameters

Following Wiese and Omlin (2009), several test runs were executed to find the best combination of hyperparameters before executing the runs of the research. The model was compiled using the following hyperparameters:

- Dropout technique of 30% to avoid overfitting and to allow the network to get a better generalization.

- Learning rate to adjust the weights in response to changes in the gradient. For this research, the learning rate is 0.0018.

- The m square error (MSE) Loss Function is commonly used on regression tasks. It calculates the magnitude of the average error between the model's prediction ($\hat{y}_i$) and the target value (y_i) by taking the average of the squared

difference between these two values. Squaring differences results in a higher penalty for material deviations from the target value.

$$MSE \; = \; \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n} \tag{1}$$

Where:

- n is the total sample size.
- $\hat{y}_i$ is model's prediction.
- y_i is the target value.

5. Performance Metrics

To compare both architectures, the prediction accuracy was evaluated through four different metrics:

- MAE shows the arithmetic mean over the absolute difference between $\hat{y}_t$ and y_t (residuals) at time t in the analyzed timeframe.

$$MAE \; = \; \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t| \tag{2}$$

Where "n" is the total sample size.

- MAPE. This indicator measures prediction accuracy as a percentage based on the average of the ratios of individual absolute errors, at each point in time. Defining the error between $\hat{y}_t$ and y_t at time t as a ratio, as follows:

$$e_t = \; \frac{y_t - \hat{y}_t}{y_t} \tag{3}$$

MAPE is represented as:

$$MAPE = \; \frac{\sum_{t=1}^{n}|e_t|}{n} * 100 \tag{4}$$

- MDAPE. It is a performance metric used to evaluate the accuracy of forecasts in time series analysis. Unlike MAPE, MDAPE uses the median of the absolute percentage errors. This property enables MDAPE to be less sensitive to outliers than MAPE. Mathematically, MDAPE is represented as follows:

Samuel García
Comparing the Performance of Long Short-Term Memory Architectures (LSTM)
in Equity Price Forecasting: A Research on the Mexican Stock Market
Comparación del desempeño de arquitecturas de memoria a corto y largo plazo (LSTM) en el pronóstico
de precios de acciones: una investigación sobre el mercado bursátil mexicano

$$MDAPE = median\left(\frac{|e_t|}{y_t}\right) * 100 \tag{5}$$

- RMSE measures the difference between $\hat{y}_t$ and targets y_t at time t, through squaring the errors, taking the mean, and finally calculating the square root. RMSE is used to quantify the error on $\hat{y}_t$, when y_t is a continuous number and gives a friendly view of the model's performance, since it shows data on the same scale/units as the Target variable. RMSE is calculated as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \tag{6}$$

6. Preliminary Results

Both architectures described in Section 4.2 were tested using the hyperparameters described in Section 4.3; early stopping was used to prevent overfitting. All tests were run considering a dataset from January 1st, 2020, to February 9th, 2024, encompassing 1036 trading days.

6.1 First Test

The test was run using a historical timeframe (sequence length) of 20 days. Table 2 shows the results for both architectures, after replicating 10 times the test over the same stock to provide model reliability (see Table 2).

Results suggest that the multivariate architecture has more consistent performance results (i.e., MAE, MAPE, MDAPE, and RMSE) on the four stocks than the univariate, where results for WALMEX and LA COMER differ significantly from those for SORIANA and CHEDRAHUI.

When comparing performance results between univariate and multivariate models, it can be observed that the multivariate model outperforms when forecasting prices of less liquid stocks: the univariate MAPE for SORIANA and CHEDRAUI is 357% and 1021% higher than the same multivariate metric (see Table 2). Additionally, the univariate MDAPE for the two cases is 457% and 1319% higher than the same metric obtained through the multivariate model. Similar differences are observed for RMSE where the univariate results are 92% and 1252% higher.

When comparing MAPE and MDAPE metrics obtained from the two models for WALMEX and LA COMER, the univariate results outperform the multivariate in almost all indicators (see Table 2). By way of illustration, the univariate MAE for WALMEX is 0.8500 and the multivariate is 1.1400. The results may imply that stock liquidity impacts the forecast capability of the univariate LSTM. Finally, the multivariable average MAE (1.0325) is -80.77% compared to the univariate (5.3700), MAPE -77.42%, MDAPE -81.14 %, and RSME -76.67% respectively.

6.2 Second Test

The sequence length was changed from 20 to 15 days. Table 3 shows the performance results (see Table 3).

Observing performance metrics, a shortened sequence length shows a positive impact on the univariate architecture when compared with the first test, lowering differences in performance metrics: as Table 3 shows, the average MAE is 1.7650, MAPE 3.0450%, MDAPE 2.5375% and RMSE 2.5305 among the four stocks, however the results are higher than those obtained with the multivariate.

Multivariate LSTM discloses more consistent and accurate results, showing small differences in the four indicators, on average MAE is 1.64, MAPE 2.42%, MDAPE 2.24%, and RMSE 1.93. Additionally, when comparing performance metrics between both architectures for the analyzed stocks, the multivariable average results are more accurate than the univariate: MAE is -6.94%, MAPE -20.36 %, MDAPE -11.53 %, and RSME -23.61% respectively.

Table 2. Long Short-Term Memory (LSTM), 20 days

Timesteps	20	Architecture	2 HL / 50 HU
Batch size	16	Data	From Jan 1, 2020 to Feb 9, 2024
Early stopping	yes	# Days	1036
Learning rate	0.0018	Frequency	Daily

		Univariate				Multivariate			
Issuer	Ticker	Median Absolute Error (MAE)	Mean Absolute Percentage Error (MAPE) %	Median Absolute Percentage Error (MDAPE) %	Difference between actual and predicted values (RMSE)	MAE	MAPE (%)	MDAPE (%)	RMSE
WALMEX	WALMEX	0.8500	1.2600	0.8900	1.1796	1.1400	1.6800	1.4000	1.4259
SORIANA	SORIANA	2.1700	6.8600	6.3500	2.5707	1.0100	1.5000	1.1400	1.3365
CHEDRAUI	CHDRAUI	17.8800	17.4900	17.3100	18.2680	1.0500	1.5600	1.2200	1.3514
LA COMER	LACOMER	0.5800	1.4900	1.0600	0.7953	0.9300	1.3800	1.0700	1.2085
Average		5.3700	6.7750	6.4025	5.7034	1.0325	1.5300	1.2075	1.3306
Δ Univariate vs multivariate						-80.77%	-77.42%	-81.14%	-76.67%

Source: Prepared by the author.

Table 3. Long Short-Term Memory (LSTM), 15 days

Timesteps	15	Architecture	2 HL / 50 HU
Batch size	16	Data	From Jan 1, 2010 to Feb 9, 2024
Early stopping	yes	# Days	1036
Learning rate	0.0018	Frequency	Daily

		Univariate				Multivariate			
Issuer	Ticker	Median Absolute Error (MAE)	Mean Absolute Percentage Error (MAPE) %	Median Absolute Percentage Error (MDAPE) %	Difference between actual and predicted values (RMSE)	MAE	MAPE (%)	MDAPE (%)	RMSE
WALMEX	WALMEX	2.6000	4.0500	3.1900	3.4742	1.6000	2.3600	2.1400	1.9060
SORIANA	SORIANA	1.4600	4.5500	4.8600	1.7166	1.5400	2.2800	2.0900	1.8417
CHEDRAUI	CHDRAUI	2.4700	2.2400	1.1200	4.1867	1.7600	2.6000	2.5300	2.0231
LA COMER	LACOMER	0.5300	1.3400	0.9800	0.7444	1.6700	2.4600	2.2200	1.9611
Average		1.7650	3.0450	2.5375	2.5305	1.6425	2.4250	2.2450	1.9330
Δ Univariate vs multivariate						-6.94%	-20.36%	-11.53%	-23.61%

Source: Prepared by the author.

Table 4. Long Short-Term Memory (LSTM), 10 days

Timesteps	10	Architecture	2 HL / 50 HU
Batch size	16	Data	From Jan 1, 2010 to Feb 9, 2024
Early stopping	yes	# Days	1036
Learning rate	0.0018	Frequency	Daily

		Univariate				Multivariate			
Issuer	Ticker	Median Absolute Error (MAE)	Mean Absolute Percentage Error (MAPE) %	Median Absolute Percentage Error (MDAPE) %	Difference between actual and predicted values (RMSE)	MAE	MAPE (%)	MDAPE (%)	RMSE
WALMEX	WALMEX	0.8700	1.2800	0.9400	1.1899	1.2900	1.8800	1.5500	1.6678
SORIANA	SORIANA	2.5600	8.1100	7.6200	2.9482	1.5700	5.0200	4.4100	1.8782
CHEDRAUI	CHDRAUI	27.2200	26.6300	26.8500	27.7901	13.0100	12.8700	12.5600	13.6385
LA COMER	LACOMER	0.5500	1.4000	1.0200	0.7522	0.7100	1.7900	1.3700	0.9799
Average		7.8000	9.3550	9.1075	8.1701	4.1450	5.3900	4.9725	4.5411
Δ Univariate vs multivariate						-46.86%	-42.38%	-45.40%	-44.42%

Source: Prepared by the author.

Samuel García
Comparing the Performance of Long Short-Term Memory Architectures (LSTM)
in Equity Price Forecasting: A Research on the Mexican Stock Market
Comparación del desempeño de arquitecturas de memoria a corto y largo plazo (LSTM) en el pronóstico
de precios de acciones: una investigación sobre el mercado bursátil mexicano

6.3 Third Test

The third test was performed using the same database used for the first and second tests, the sequence length was changed to 10 trading days.

Table 4 shows the performance test results for both architectures (see Table 4). The average performance metrics for the univariate deteriorated when compared with the results in the second test (see Table 3), mainly for less liquid stocks (SORIANA and CHEDRAUI, In particular, SORIANA MAE varied from 1.45 in the second test to 7.8; MAPE from 4.55% to 8.11%, RMSE from 1.7166 to 2.9482, and MDAPE from 4.86% to 7.62%. Additionally, the performance metrics for the multivariate show the worst results when compared with tests one and two.

Although the multivariable performance results deteriorate when compared to those obtained for the same model in the first and second tests, these numbers are better than those in the univariate model. On average, multivariate (4.1450) MAE is -46.86% than univariate (7.8000), MAPE -42.38 %, MDAPE -45.40 % and RSME -44.42% respectively.

7. Conclusions

Stock price prediction is a very researched and complex area because all variables involved in trading activities have a nonlinear behavior. Thus, there is an interest in developing models that will allow more accurate and consistent forecasts. This study focuses on comparing the performance of univariate and multivariate LSTM in predicting prices for stocks in the consumer retail sector in Mexico, as well as the impact of the size of the sequence length on the models. The performance results under different sequence lengths were analyzed in Section 6.

In general, results show that the univariate LSTM works better when predicting prices over liquid stocks, although the performance in this model was less consistent among the four stocks in the sample than the multivariate. Multivariate LSTM shows accurate and consistent performance metrics when predicting prices for liquid and illiquid stocks, producing minor errors, measured through the performance metrics.

Sequence length impacts the accuracy of price prediction on both tested models. For instance, the univariate model disclosed a better performance with a sequence

length of 15 trading days, whereas the multivariate shows a better performance with a sequence length of 20 days and 15 days.

Finally, it is worthwhile to continue exploring in future works the impact of price volatility and trends on predicting prices of illiquid stocks traded in developing economies.

Samuel García
Comparing the Performance of Long Short-Term Memory Architectures (LSTM)
in Equity Price Forecasting: A Research on the Mexican Stock Market
*Comparación del desempeño de arquitecturas de memoria a corto y largo plazo (LSTM) en el pronóstico
de precios de acciones: una investigación sobre el mercado bursátil mexicano*

References

Andi, H. (2021). An accurate bitcoin price prediction using logistic regression with LSTM machine learning model. *Journal of Soft Computing Paradigm*, *3*(3), 205–217. https://doi.org/10.36548/jscp.2021.3.006

Armitage, S., Brzeszczyński, J. & Serdyuk, A. (2014). Liquidity measures and cost of trading in an illiquid market. *Journal of Emerging Market Finance, 13*(2), 15–5196. https://doi.org/10.1177/0972652714541340

Benchaji, I., Douzi, S. & Ouahidi, B. E. (2021a). Credit card fraud detection model based on lstm recurrent neural networks. *Journal of Advances in Information Technology*, *12*(2), 113–118. https://doi.org/10.12720/jait.12.2.113-118

Benchaji, I., Douzi, S., Ouahidi, B. E. & Jaafari, J. (2021b). Enhanced credit card fraud detection based on attention mechanism and lstm deep model. *Journal of Big Data*, 8(1), 1–21. https://doi.org/10.1186/s40537-021-00541-8

Bhandari, H. N., Rimal, B., Pokhrel, N. R., Rimal, R., Dahal, K. R. & Khatri, R. K. (2022). Predicting stock market index using LSTM. *Machine Learning with Applications*, *9*, article 100320. https://doi.org/10.1016/j.mlwa.2022.100320

Ghosh, A., Bose, S., Maji, G., Debnath, N. & Sen, S. (2019). Stock Price Prediction Using LSTM on Indian Market Share. *International Conference on Computer Applications in Industry and Engineering, 63*, 101–110. https://doi.org/10.29007/qgcz

Haykin, S. (2010). *Neural Networks and Learning Machines.* Pearson Education India.

Krenker, A., Bešter, J. & Kos, A. (2011). Introduction to the Artificial Neural Networks. In K. Suzuki (Ed.), *Artificial Neural Networks: Methodological Advances and Biomedical Applications.* InTech. https://doi.org/10.5772/15751

Landazuri Aguilera, Y. & Ruíz Pérez, R. (2021). El desempeño del sector productos de consumo frecuente en época de pandemia. *Memorias del Coloquio Nacional de Investigación en las Ciencias Económicas y Administrativas*, 114–132. https://shorturl.at/eEX78

Liebergen, B. V. (2017). Machine learning: a revolution in risk management and compliance? *Journal of Financial Transformation*, *45,* 60–67. https://www.capco.com/Capco-Institute/Journal-45-Transformation

Lu, Y. (2017). *Deep neural networks and fraud detection* [Master's thesis], Uppsala University. https://shorturl.at/ozEGI

Nielsen, M. (2015). Neural networks and deep learning. *Determination Press*, 25, 15–24. http://neuralnetworksanddeeplearning.com/

Nourbakhsh, Z., & Habibi, N. (2023). Combining LSTM and CNN methods and fundamental analysis for stock price trend prediction. *Multimedia Tools and Applications*, *82*,17769–17799. https://doi.org/10.1007/s11042-022-13963-0

Pramod, B. & Mallikarjuna, S. P. (2020). Stock Price Prediction Using LSTM. *Test Engineering and Management*, *83*, 5246-5251. https://shorturl.at/bjtA2

Tianxiang, Y. & Zihan, W. (2020). Crude oil price prediction based on LSTM network and GM (1,1) model. *Grey Systems: Theory and Application*, *11*(1), 80–94. https://doi.org/10.1108/GS-03-2020-0031

Tirozzi, B., Puca, S., Pittalis, S., Bruschi, A., Morucci, S., Ferraro, E. & Corsini, S. (2007). *Neural networks and sea time series: reconstruction and extreme-event analysis.* Springer Science & Business Media. https://doi.org/10.1007/0-8176-4459-8

Wiese, B. & Omlin, C. (2009). Credit card transactions, fraud detection, and machine learning: Modelling time with LSTM recurrent neural networks. In Bianchini, M., Maggini, M., Scarselli, F. & Jain, L.C. (Eds.), *Innovations in neural information paradigms and applications.* Springer, 231-238. https://doi.org/10.1007/978-3-642-04003-0_10

Yu, Y., X. S., Hu, C. & Zhang, J. (2019). A review of recurrent neural networks: Lstm cells and network architectures. *Neural Computation*, *31*(7),1235–1270. https://doi.org/10.1162/neco_a_01199

Zaheer, S., Anjum, N., Hussain, S., Algarni, A., Iqbal, J., Bourouis, S. & Ullah, S. (2023). Multi Parameter Forecasting for Stock Time Series Data Using LSTM and Deep Learning Model. *Mathematics*, *11*(3), 590. https://doi.org/10.3390/math11030590

Samuel García
Comparing the Performance of Long Short-Term Memory Architectures (LSTM)
in Equity Price Forecasting: A Research on the Mexican Stock Market
Comparación del desempeño de arquitecturas de memoria a corto y largo plazo (LSTM) en el pronóstico
de precios de acciones: una investigación sobre el mercado bursátil mexicano

About the author

Samuel García is Mexican and has earned two master's degrees: the first in Economics and Technological Change and the second in Finance, and a Bachelor's in Business Administration. He has 25 years of experience in national and international financial institutions. He is a senior executive with knowledge of market surveillance, compliance, risk management, structured finance, financial products, and strategic management. Samuel worked at Citi for 16 years, where he was responsible for building and providing strategic direction for the ICRM Surveillance program, covering sales and trading as well as banking businesses across the Latin American Region—with more than 20 countries, including Mexico.

a1018966@tec.mx
https://orcid.org/0000-0001-7366-1406