A Predictive Sales System Based on Deep Learning

Jean Paul Luyo Ballena, Cristhian Pool Ortiz Pallihuanca, Ernesto Adolfo Carrera Salas
Facultad de Ingeniería de la Universidad Peruana de Ciencias Aplicadas, Lima, Perú

***Abstract*—There are several techniques for predictive sales systems, in this study, a system based on different machine learning algorithms is developed for a trading company in Lima. As any company, it needs to be accurate in its sales calculations to manage the volume of production or product purchases. With the system, the trading company has a mechanism to order products from its supplier based on the predictions and estimates of the needs according to the projection of its sales. For the sales predictive system, Deep Learning technology and the neural network architectures GRU (Gated Recurrent Unit), LSTM (Long Short Term Memory) and RNN (Recurrent Neural Network) were used, 10 products were sampled, and the sales quantities of the last 12 months were obtained for the evaluation. The study found that the LSTM architecture excels in accuracy, significantly outperforming GRU and RNN in terms of Mean Absolute Percentage Error (MAPE), achieving an average MAPE of 7.07%, in contrast to the MAPE of 27.14% for GRU and the MAPE of 36.17% for RNN. These findings support the effectiveness and versatility of LSTM in time series prediction, demonstrating its usefulness in a variety of real-world applications.**

***Keywords*—Deep learning; neural network architectures; sales prediction; neural networks**

I. Introduction

Effective sales management is essential to achieving business objectives but is often hampered by the availability of incomplete or outdated information, making it difficult to make informed decisions. Most of the business organizations heavily depend on a knowledge base and demand prediction of sales trends. The accuracy in sales forecast provides a big impact in business [1]. Despite the existence of conventional sales management systems, few take advantage of recent advances in artificial intelligence, and even fewer apply deep learning to improve accuracy in sales predictions. If the sales forecast is not accurate, situations of shortage or excess of stock may occur, which can have a direct and immediate impact on the profitability of the company. This effect is not only limited to the performance of profitability, but also to the customer service, therefore, service can be affected by an inefficient sales forecasting system. For example, if a customer is faced with a stock-out situation, they might decide to shop in a different company [2].

In this study, a Deep Learning based sales prediction system is proposed using three neural network architectures provided by the Brain.js library in JavaScript: GRU, LSTM and RNN. This research was conducted in a company located in Lima, using historical sales data to train and model each of these architectures. The research used 10 products specifically selected due to their relevance and analyzed data collected over a 12-month period to explore varied approaches to prediction.

To carry out this study, a series of key steps were followed. It started with the recording of historical sales data provided by the supplier. Then, different neural network architectures were used in JavaScript to perform sales analysis and prediction. Then, the data was loaded for neural network training, using the sales history as input data and the number of sales predicted for the next month as the expected output value. Subsequently, the neural network was modeled and trained with the architectures provided by the Brain.js library. Once the prediction for the next month was completed using each trained model, the corresponding predicted sales were displayed.

As part of the comparative analysis of the models, the predicted sales, the MAE (Mean Absolute Error) and MAPE (Mean Absolute Percentage Error) performance indicators and the inference time obtained from the three evaluated architectures are calculated. These steps allow to evaluate and select the most appropriate prediction model for the system, and to effectively apply the sales predictions.

Once the prediction for the next month was completed using each trained model, we show the corresponding predicted sales. As part of the comparative analysis of the models, we calculated the performance indicators MAE (Mean Absolute Error) and MAPE (Mean Absolute Percentage Error). Of the three architectures evaluated, the strongest result is the choice of LSTM due to its outstanding prediction accuracy, which significantly outperforms GRU and RNN in terms of MAPE, achieving a MAPE of 7.07%, in contrast to 27.14% for GRU and 36.17% for RNN.

This article is organized as follows: Section II shows the related works, where different neural network architectures are used for data analysis. It conceptualizes the different neural network architectures used in this study, as well as other relevant concepts used in the research. Section III explains the proposed methodology for the implementation of the Predictive Sales System, developing the following steps: Recording historical sales data, use of neural networks in JavaScript, Loading data for neural network training, Modeling and training the neural network, Model prediction and Comparison of results. In Section IV we show the results and discuss the findings, interpreting and analyzing their implications and limitations, and finally Section V concludes the paper.

In research [3], a novel and effective method for multi-channel retail demand forecasting is proposed. This method combines long-term memory (LSTM) and random forest (RF) neural networks to model complex temporal as well as regression relationships, which makes it more accurate than other forecasting methods. In addition, the proposed method has been shown to be statistically significantly better than other forecasting methods, including neural networks, multiple regression, ARIMAX and LSTM. The proposed method also improves the interpretability of the model by ranking the relative importance of the explanatory variables, which makes it useful for decision making. An empirical evaluation was conducted on 192 time series and their corresponding information signals from two different channels: an online channel for 16 products and an offline channel covering the same 16 products sold in 11 different physical stores. For this reason, the analysis shows that the proposed method outperformed the other methods with a 95% confidence level in predicting the weekly demand for a product in a store. In conclusion, the proposed method is a valuable tool for multi-channel retailers in accurately and reliably predicting the demand for their products. This approach proves to be a valuable tool for retailers, providing accurate and reliable forecasts of the demand for their products, and contributing to our research on the performance of the LSTM algorithm.

In study [4], the authors used an LSTM neural network to evaluate its performance on point-of-sale data from a large retail chain. It was observed that bottom-up forecasts are more accurate than top-down forecasts when point-of-sale information is used for forecasting. The proposed framework helps to produce consistent short- and long-term forecasts for each level of decision making in a retail supply chain. This feature of the proposed framework helps align decision making at all levels of the organization and reduces the cost of decision misalignment. The authors conducted an empirical analysis using 141 different demand series for 10 different products offered through 1 online and 10 offline stores. Based on the results, it is safe to assume that at a confidence level of at least 95%, the bottom-up approach is better than the top-down approach in the case of retail demand forecasting. A top-down approach only generates a base forecast for the top principal node, and the top-down heuristic disaggregates the top-level forecast to obtain forecasts for all secondary nodes. In contrast, in the bottom-up approach, base forecasts are made for all lower-level nodes and then aggregated to obtain forecasts for all higher-level nodes. The article is included because of the similarity of the use of LSTM neural networks in the context of retail demand forecasting. This approach is critical to better understand how these networks can improve forecasting accuracy in these retail businesses.

In study [5], a meta-learning technique using deep convolutional neural networks, which have the feature of enhancing image recognition tasks and computer vision, is proposed for the purpose of automatically learning feature representation from sales time series. The meta-learner is used to automatically learn from a feature representation from raw sales time series data and then link the learned features to a common data set that is used to combine a set of methods from the base forecaster; the learned knowledge is used to select an optimal forecasting method for each time series according to its data features. Then, these characteristics are combined with a dataset to combine a set of forecasting methods to improve the accuracy of retail sales forecasting. The results indicate that the proposed meta-learner has a superior forecasting performance compared to several benchmark methods. The proposed meta-learner (M0) has significantly superior forecasting performance over all baseline forecasters and simple average combinations of forecasters. On average, M0 has 3.2% improvements over the best performing base predictor. In addition, the meta-learner is particularly effective during promotion weeks, with a 5% improvement over the best-performing base predictor, compared to 1.6% improvements in non-promotion weeks. This approach helps us revolutionize retail sales forecasting. The deep convolutional neural network-based meta-learning technique proposed in this paper represents a valuable contribution to the field. Its ability to automatically learn time series characteristics and select optimal forecasting methods has demonstrated substantial improvements in forecasting accuracy. This not only benefits the trading company by enabling better planning and decision making, but also highlights the potential of applying deep learning and meta-learning in real-world business scenarios, as demonstrated in our sales forecasting research.

In study [6], the authors propose economic models based on exports and imports to forecast world trade using Deep Learning. The study presents a trade forecasting framework characterized by a theoretical integration of time series and economic structures, considering the uncertainty of international trade trends. The proposed approach demonstrates the power of trade data modeling using hybrid deep neural networks. In addition, the authors use prediction performance criteria to evaluate predictions using Root Mean Square Error (RMSE) and Mean Percentage Error (MAPE). The validity of this framework has been validated using commercial data from 10 countries: Brazil, Canada, China, France, Germany, India, Italy, Japan, UK, and USA. The predictive power of the proposed approach is compared with that of ten predictive models: ARIMA, ETS, TBATS, MLR, LASSO, RFR, XGB, SVR, MLP and HDL. The highest average values for (MAPE) and (RMSE), were observed in the time series model, reaching 2979 and 2933 for exports, and 4194 and 4475 for imports, respectively. In contrast, the models based on the economic structure of the time series showed lower average values for MAPE and RMSE, registering 2861 and 3198 for exports, and 3527 and 3716 for imports, respectively. The lowest MAPE values in their study stand at 2.861% for exports and 3.527% for imports, which contrasts significantly with our research, which obtains a minimum MAPE of 0.40%. These results underscore the ability of our model using LSTM to outperform in terms of accuracy the figures presented in the article, reinforcing the relevance and effectiveness of our methodology in accurately predicting future sales.

In study [7], the authors propose a study of electricity price estimation using deep learning approaches: an empirical study on Turkish markets in normal and Covid-19 periods. In recent decades, several methods have been used to estimate electricity bills. In the early days of these techniques, econometric models were used, but it was clear that these models were not successful enough to predict peak power prices. Furthermore, no transition period, such as COVID-19, is expected to affect electricity prices. However, machine learning and deep learning models are becoming increasingly important. Many researchers have shown that these models can also successfully predict extreme fluctuations in electricity prices. The proposed method is the original Transformer Encoder Decoder with Self-Attender (TEDSE) for electricity rate estimation. It is organized in such a way that the model groups the data by a dummy variable and estimates the 24-hour electricity price of the Turkish electricity market. In this study, predictions were made using 12 different models based on Recurrent Neural Networks (RNN). The RNN-based estimates used to estimate electricity tariffs corroborated the results of the TEDSE model and underlined the success of the TEDSE model. However, due to market changes, these models individually could not accurately predict price estimates during COVID-19. For this reason, the TEDSE model forecasts sub-periods and allows electricity market participants to work more efficiently. This approach helps us understand the importance of deep learning, in particular, Recurrent Neural Networks (RNNs), in predicting complex phenomena such as electricity prices or as in our research on demand forecasting. In the study RNNs and advanced models, such as the Transformer Encoder-Decoder with Self-Attender (TEDSE), can provide more accurate forecasts even in challenging situations such as the COVID-19 period, where abrupt changes in markets are inevitable.

Table I presents a comparison of the works related to our research is made, where the different methods proposed by the authors are analyzed, the information on the data used and the time period of the study are detailed, in addition to the metrics evaluated to validate the models.

II. CONTRIBUTION

The application of deep learning in sales prediction has been the topic of analysis on several research studies, as described in the related works in the previous section. The difference and contribution we make with our research is based on the evaluation of various architectures of recurrent neural networks, such as GRU, LSTM and RNN, which allows us to determine more accurately which architecture will provide us with the best metrics for predicting sales in a commercial company.

Based on the review of related works and the objective of our research, a broad conceptual framework is proposed to understand and define the fundamental concepts of Deep Learning, the architectures of neural networks and the company's need for an efficient sales management. In this context, it is important to clearly establish the key terms and concepts that will be discussed in this work, below we will find the essential definitions that will lay the foundation for solid understanding and deeper analysis in our research.

A. Sales Management

Sales management in [8] refers to the comprehensive process of planning, coordinating, and controlling activities related to a company's sales force, with the objective of achieving optimal sales performance and reaching established business objectives. It involves developing effective strategies and tactics, as well as constantly monitoring and evaluating the performance of sales teams.

TABLE I. COMPARISON BETWEEN PROPOSED MODEL VS. RELATED WORKS

Year	Author's References	Description of method / technique	Dataset Size	Performance Metric/s (value/s)
2020	Paper 1[3]	The proposed method is based on a state-of-the-art sequential deep learning method – long-short-term-memory networks (LSTM) and a machine learning method – random forest (RF).	Sales data of 16 products from online and offline stores over a period of 30 months.	The error measures employed were ARME, ARMAE, and ARMSE as proxies for the bias, accuracy, and variance of the evaluated forecasts.
2020	Paper 2[4]	Cross-Temporal Forecasting Framework (CTFF) to generate coherent forecasts at all levels of a retail supply chain, using a deep learning method, the long-short-term-memory network (LSTM).	Weekly sales data for 10 products (SKU) in a period of more than two years.	ARMAE, ARMSE, ARMAPE were used for the comparisons of forecasts errors across different approaches.
2021	Paper 3[5]	Meta-learning framework with automatic feature learning for retail product sales forecasting, based on convolution neural networks.	Sales data for 6 categories of products sold on 100 stores over a 153-week period.	Three error measures were used to compare the forecasting performance of the models: sMAPE, AvgRelMAE, MPE.
2023	Paper 4[6]	This study proposes a machine learning approach based on import and export economics statistics, the input features selection according to existing classical economic structure models under economics theories.	Data is collected from the of the export and import values of goods in USD records of 10 countries. All indices used in this study were adjusted using 2015 as the base year.	The methods used to evaluate the forecasts are the root mean square error (RMSE) and mean average percentage error (MAPE).
2023	Paper 5 [7]	The study aims to develop a prediction model for the electricity market of Turkey with more comprehensive Deep Learning models applying the TEDSE model.	In this study, 5-year of hourly historical data between 2017 and 2021 in the Turkish electricity market is used.	Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) statistics along with Mean Standard Error (MSE) were used to measure the model's performance.
2023	The model proposed	In this study, a Deep Learning based sales prediction system is proposed using three neural network architectures provided by the Brain.js library in JavaScript: GRU, LSTM and RNN.	The data for this study corresponds to historical sales records on 10 products in a period of 12 months.	To compare the performance of the models, MAE (Mean Absolute Error) and MAPE (Mean Absolute Percentage Error) are used.

Source: Own elaboration

The importance of sales management lies in its ability to accelerate and improve the achievement of business objectives, as well as in the creation of competitive advantages and greater value in the marketplace. In addition, sales management facilitates the follow-up and monitoring of processes, identifying problems early and establishing training guidelines for the growth of the sales team. The three main objectives of sales management are to increase sales volume, contribute to company profits and increase ROI, and achieve long-term sustainable growth. The system is dedicated to helping meet each of these objectives. Moreover, these objectives imply additional responsibilities for sales managers, who not only focus on sales, but must also manage and train a team of sales representatives to align them with organizational objectives. Effective sales management involves being informed about the market and making tactical decisions based on data and trend analysis.

In summary, sales management is a strategic process that seeks to maximize the performance of the sales force and achieve the company's business objectives by planning, coordinating, and controlling sales-related activities. Its importance lies in improving efficiency, generating competitive advantages, and ensuring the long-term growth of the organization.

B. Deep Learning

Deep learning [9], is a subset of machine learning that relies on multi-layered neural networks to learn and make predictions from large volumes of data. Unlike conventional machine learning, deep learning can work with unstructured data, such as images or text, without requiring pre-processing to organize it into a specific format. Deep neural networks take advantage of algorithms such as gradient slope and back propagation to adjust their parameters and improve their accuracy over time. These algorithms automate feature extraction, which reduces the reliance on human intervention in manually defining relevant features. By eliminating the need for preprocessing and enabling learning directly from unstructured data, deep learning has driven the development of artificial intelligence applications in a variety of areas, such as speech recognition, fraud detection, autonomous driving and many more. Its ability to process complex information and learn hidden patterns has enabled significant advances in solving problems that were previously considered difficult or even impossible to tackle.

In short, deep learning is a branch of machine learning that uses multilayered neural networks to learn and make predictions from unstructured data. Its ability to automatically extract relevant features and its flexibility to work with various types of data have contributed to important advances in artificial intelligence.

C. Neural Networks

In study [10], a neural network is a computational model inspired by the biological nervous system and consists of a collection of processing units called artificial neurons. These neurons are linked together by weighted connections that transmit and process information. Neural networks are used in many areas, including pattern recognition, computer vision, natural language processing and artificial intelligence in general. Their architecture consists of layers, including an input layer, one or more hidden layers and an output layer. Neural networks learn automatically through a training process in which connection weights are adjusted to improve their ability to make predictions and classifications. This process is based on learning algorithms such as supervised and unsupervised learning.

In short, a neural network is a computational model that imitates the nervous system, made up of artificial neurons connected to each other.

D. JavaScript

In study [11], JavaScript is a programming language used to create interactive web pages. It enhances the user experience by enabling features such as social media updates, animations, and interactive maps. JavaScript was born as a technology intended for the browser, with the purpose of giving greater dynamism to web pages. Browsers were able to respond to user interactions and modify the layout of website content.

E. RNN (Recurrent Neural Network)

The structure of an RNN in [12] consists of a series of repeating units connected in loops, also called "cells". Each repetitive unit receives input and produces output. There are also connections that can receive information from previous iterations. This allows information to flow recursively through the network and considers temporal dependencies.

F. GRU (Gated Recurrent Unit)

GRU [13] is an efficient variant of SRRNs (standard recurrent neural network) used in natural language processing with its simple structure, it is easy to train and computationally efficient. It has proven its effectiveness in machine translation and speech recognition, being popular in machine learning and artificial intelligence.

G. LSTM (Long Short-Term Memory)

In research [14], LSTMs are a special type of RNN whose main characteristic is that information can persist by introducing loops in the network diagram, which means that it can remember previous states and use this information to decide which will be the next one. LSTMs have a longer-term memory.

A great asset of LSTM [15], is its capacity to unravel intricate temporal structures and seize the fluid dynamics of systems that undergo time-driven fluctuations. By dissecting the inherent motifs and inclinations embedded within the data, LSTM architectures can reveal interconnections that may not be easily distinguished using established statistical or machine learning methodologies.

In comparison to LSTM, GRU [16] has a simplified cell structure that also operates based on a Gating system, but only has an update and reset gate. The main difference to LSTM is the circumstance that the cell state can be completely revised at each iteration and updated with short-term information via the reset gate. Instead, LSTM provides a mechanism that limits the change gradient that can be done at each iteration. Therefore, information is not completely gone with LSTM in contrast with GRU.

H. MAE (Mean Absolute Error)

MAE [17] is a popular metric because the error value units match the predicted target value units. In MAE, different errors are not weighted more or less, but the scores increase linearly with the increase in errors. The MAE score is measured as the average of the absolute error values. The Absolute is a mathematical function that makes a number positive.

Therefore, the difference between an expected value and a predicted value can be positive or negative and will necessarily be positive when calculating the MAE.

I. MAPE (Mean Absolute Percentage Error)

MAPE [18] is a relative error measure that uses absolute values to prevent positive and negative errors from canceling each other and uses relative errors that allows to compare forecast accuracy between time series methods.

Mean Absolute Percentage Error [19] is often used due to its very intuitive interpretation in terms of relative error. MAPE is relevant in finance, considering that profits and losses are often measured in relative values. It is also useful for calibrating product prices, customers are sometimes more sensitive to relative variations than to absolute ones. MAPE is frequently used when it is known that the quantity to be predicted remains well above zero. More generally, MAPE is well suited for forecasting applications, especially in situations where sufficient data is available.

J. Inference Time

In research [20], inference time is the time it takes for a machine learning model to make predictions after training. It is important to optimize the model and hardware to meet the inference time requirements.

III. METHOD

In this study, a sales management system was implemented using the Brain.js library in JavaScript. Fig. 1 presents the steps for the method used.

A. Recording Historical Sales Data

The data for this study corresponds to historical sales records compiled from records provided by the company owner. We have focused on 10 products specifically selected because of their relevance. These products were accessed in the period between August 2022 and July 2023.

B. Use of Neural Networks in JavaScript

To carry out sales analysis and prediction effectively, the Brain.js library was used in the JavaScript environment. This library offers several neural network architectures, such as GRU, LSTM and RNN, which were used to explore various approaches to prediction. This choice made it possible to investigate and select the most appropriate architecture to obtain accurate and relevant forecasting based on the patterns present in the historical data.

C. Loading Data for Neural Network Training

At this stage, the training data for the neural network was prepared. An approach was followed in which the sales history was used as the input data, and the expected number of sales for the following month was used as the expected output value. In specific terms, the input consists of the 12 sales records that occurred in the past months (see Table II).

This structuring of the data allows the neural network to gain insight into the patterns and relationships underlying the sales history and future projections. This approach is essential to achieve highly accurate predictions of future sales.

D. Modeling and Training the Neural Network

At this stage, neural network training was carried out using the architectures provided by the Brain.js library: GRU, LSTM and RNN. Each architecture offers a different approach to modeling and learning relationships in the training data. Neural networks were modeled with the appropriate parameters and settings for each architecture, and then trained using the previously prepared data. During training, the neural networks were exposed to historical sales data to learn patterns and trends. As training progressed, the networks adjusted their weights and parameters to improve prediction accuracy. At the end of this stage, each of the three neural networks was ready to be used for prediction. Each architecture offers a unique perspective for making decisions related to future sales.

E. Model Prediction

At this stage, we proceeded to use each of the trained models for the purpose of making predictions. The same procedure was followed for the three architectures available in Brain.js: GRU, LSTM and RNN. The prediction was carried out by supplying the sales value corresponding to each period in the dataset as input to the respective model. In this way, each architecture generated a sales prediction for the period in question. During this process, the inference time was recorded, reflecting the time required to complete each prediction, which made it possible to evaluate the performance of each model in terms of speed and efficiency. At the conclusion of each prediction, the sales estimate for each product was obtained and presented in the form of rounded integer values, which were displayed on the console along with the corresponding dates. This procedure was repeated for each of the three architectures, allowing a thorough comparison between the predictions generated by each model, as well as an evaluation of their accuracy and performance based on the training data.

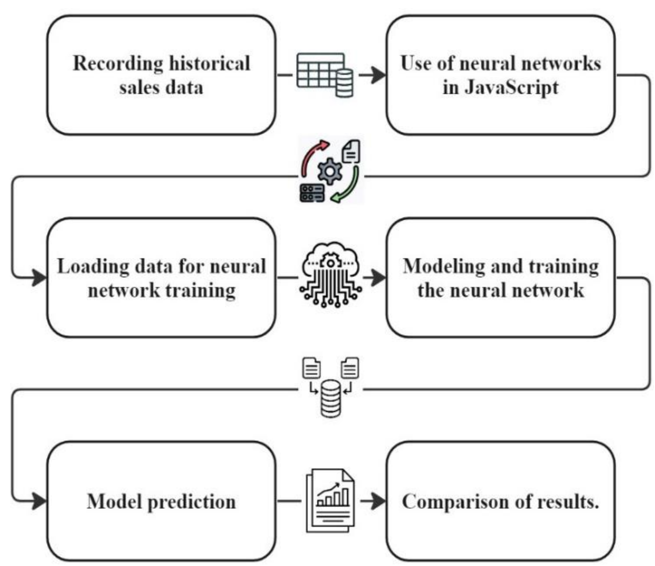

Fig. 1. Flow of steps for the proposed system. Source: Own elaboration.

TABLE II. MONTHLY SALES OF PRODUCTS (MONTHS 1-12)

Products	Month											
	1	2	3	4	5	6	7	8	9	10	11	12
P-01	50	50	50	100	100	150	150	300	150	200	100	150
P-02	50	50	100	100	100	150	100	300	150	150	100	100
P-03	50	50	80	80	100	100	150	300	150	150	120	120
P-04	200	250	250	300	250	300	300	200	150	100	100	100
P-05	30	30	50	50	50	100	100	300	100	150	120	100
P-06	30	20	20	50	50	50	80	200	100	100	100	150
P-07	30	50	50	50	100	100	150	300	100	150	100	150
P-08	50	50	40	40	80	100	100	300	100	150	150	100
P-09	20	30	30	50	50	80	100	200	100	150	150	150
P-10	50	50	100	50	50	100	100	300	100	150	100	200

Source: Own elaboration

F. Comparison of Results

Once the prediction for the next month has been completed using each trained model, the corresponding predicted sale is displayed. In addition, as part of the comparative analysis of the models, MAE (Mean Absolute Error) and MAPE (Mean Absolute Percentage Error) performance indicators are calculated. These indicators are obtained by comparing the sales predicted by each model with the corresponding actual values of the data set, which provides a quantitative assessment of the accuracy of each model in its predictions and allows identifying and comparing its performance against the training data.

IV. RESULTS AND DISCUSSION

This section provides the research result and discussion of the proposed models.

TABLE III. RESULTS OF GRU (GATED RECURRENT UNIT)

Products	GRU			
	Predicated Sale	MAE	MAPE	Inference Time (ms)
P-01	139	129	7,64%	0,74
P-02	61	41	38,53%	0,80
P-03	80	66	52,24%	0,82
P-04	88	38	12,30%	0,90
P-05	99	89	0,93%	1,31
P-06	50	20	75,17%	1,06
P-07	132	122	11,90%	1,15
P-08	99	89	0,54%	1,06
P-09	104	94	30,55%	1,00
P-10	117	97	41,57%	0,86
Average		79	27,14%	0,97

Source: Own elaboration

To provide a clearer understanding of the results in Tables II, III and IV, data obtained from the three neural network architectures: GRU, LSTM and RNN, used to predict sales based on 10 products with a 12-month sales history as a dataset. The results focus on the sales forecast for the following month, analyzing the Mean Absolute Error (MAE), the Mean Absolute Percentage Error (MAPE) and the inference time for each product.

TABLE IV. RESULTS OF LSTM (LONG SHORT-TERM MEMORY)

Products	LSTM			
	Predicated Sale	MAE	MAPE	Inference Time (ms)
P-01	137	117	8,34%	0,81
P-02	101	81	0,60%	0,85
P-03	121	96	0,66%	0,81
P-04	83	33	17,01%	0,72
P-05	99	80	1,08%	0,72
P-06	136	96	9,21%	0,82
P-07	149	120	0,93%	0,86
P-08	100	90	0,40%	0,75
P-09	151	141	0,74%	0,82
P-10	180	90	31,70%	0,88
Average		94	7,07%	0,80

Source: Own elaboration

TABLE V. RESULTS OF RNN (RECURRENT NEURAL NETWORK)

Products	RNN			
	Predicated Sale	MAE	MAPE	Inference Time (ms)
P-01	91	81	39,58%	0,87
P-02	93	73	7,48%	0,78
P-03	33	18	72,68%	0,59
P-04	115	65	15,30%	0,80
P-05	84	74	15,64%	0,64
P-06	77	67	48,73%	0,60
P-07	99	89	34,12%	0,59
P-08	66	56	34,42%	0,61
P-09	89	79	40,49%	0,63
P-10	94	74	53,27%	0,60
Average		68	36,17%	0,67

Source: Own elaboration

Tables detailing the results for each architecture, GRU (see Table III), LSTM (see Table IV) and RNN (see Table V) provide clear and concise overview of the differences between them.

After analyzing the numerical results of the three evaluated architectures, the following results can be observed:

1) Mean Absolute Percentage Error (MAPE): The results in Table IV reveal that LSTM exhibits significantly lower MAPE values compared to GRU (see Table III) and RNN (see Table V). With an average MAPE of 7.07% for the 10 products, LSTM significantly outperforms GRU (27.14%) and RNN (36.17%). These findings highlight the remarkable precision of LSTM in predictions, demonstrating its ability to reduce the absolute error between the predicted mean value and the actual value.

Additionally, the incorporation of MAPE in this study provides a more comprehensive metric, as it not only considers the magnitude of the error but also the relative proportion to the actual value, offering a more detailed assessment of the predictive performance of each architecture.

2) Mean Absolute Error (MAE): Table V shows that the RNN architecture exhibits a lower average MAE (68) for the 10 products, surpassing GRU (see Table III) and LSTM (see Table IV), which have MAE values of 79 and 94, respectively. This suggests that, although RNN does not always predict with the same precision as other architectures, when it makes errors, the predictions do not differ significantly from the actual sales figures.

The inclusion of MAE in the analysis provides a valuable perspective on the average magnitude of errors, allowing for a more direct comparison between architectures in terms of absolute deviation of predictions from actual values.

3) Inference Time: The average inference times, presented in Tables III, IV and V are 0.97 milliseconds for GRU, 0.80 milliseconds for LSTM, and 0.67 milliseconds for RNN, respectively. While there are marginal differences between architectures in terms of inference speed, these results suggest that the choice of architecture should not be primarily based on inference time for this application. It is important to note that the effectiveness of each architecture should be evaluated considering both predictive accuracy and computational efficiency.

TABLE VI. ARCHITECTURE PERFORMANCE COMPARISON

	GRU	LSTM	RNN
MAE	79	94	68
MAPE	27,14%	7,07%	36,17%
Inference Time (ms)	0,97	0,80	0,67

Source: Own elaboration

A key finding of the thorough analysis of the tables, is that LSTM stands out as the most favorable architecture for predicting product sales, as shown in Table VI. It not only achieves the lowest MAPE, indicating high prediction accuracy, but also maintains an acceptable MAE and a competitive average inference time, balancing speed with performance.

Our research has been characterized by working with three different architectures to enrich the literature and analyze different forms of data processing. This contributes to future research and obtaining a more efficient and timelier sales prediction, considering, that it is an important factor in companies knowing how much is going to be sold, which leads us to assess the quantities of items to be produced or sold, eliminating storage costs or overproduction, this being a success factor in any business.

LSTM has been successfully used in a variety of forecasting systems, ranging from retail sales prediction [3, 4, 5] to global trade forecasting in several major countries [6] and electricity price prediction in energy markets [7], highlighting its effectiveness and versatility in various forecast contexts. However, with technological advances in neural networks, it should have finer precision, noting that while the current minimum values of Mean Absolute Percentage Error and Mean Absolute Error are acceptable, they could be optimized.

V. CONCLUSION

The paper has analyzed several algorithms with the objective of improving the accuracy of future sales predictions, which supports the acquisition of products without the need to maintain a high warehouse stock and reduce unnecessary purchase costs per month. The results show that the LSTM architecture excels in terms of accuracy, implying a potential positive impact on the Lima trading company's operational efficiency and its ability to make informed strategic decisions.

Of the three architectures evaluated, the strongest result is LSTM, for its outstanding prediction accuracy that significantly outperforms GRU and RNN in terms of Mean Absolute Percentage Error (MAPE), achieving an average value of the 10 products at 7.07%, in contrast to 27.14% for GRU and 36.17% for RNN. In addition, an acceptable value of the Mean Absolute Error (MAE) with 94 and an average time of 0.80 milliseconds validated among the three architectures evaluated.

Finally, when considering prediction accuracy, Mean Absolute Percentage Error (MAPE) and inference time, the preferred choice is clearly LSTM. This is due to the importance of MAPE in this context. Although RNNs have advantages over MAE, LSTM remains the strongest choice for applications where prediction accuracy is crucial. LSTM focuses on measuring the relative accuracy of forecasts, which is essential for business decision making, effective comparisons, accurate inventory management, informed investments, efficient resource allocation and continuous improvement of business strategies. This underscores LSTM's ability to make highly accurate forecasts, which can be critical in applications where estimation accuracy is essential.

ACKNOWLEDGMENT

Thanks to the Research Directorate of the Peruvian University of Applied Sciences for their support to carry out this research work through the UPC-EXPOST-2023-2 incentive.

REFERENCES

[1] S. Cheriyan, S. Ibrahim, S. Mohanan and S. Treesa, "Intelligent Sales Prediction Using Machine Learning Techniques," International Conference on Computing, Electronics & Communications Engineering (iCCECE), Southend, UK, pp. 53-58, 2018, https://doi.org/10.1109/iCCECOME.2018.8659115.

[2] A.L.D. Loureiro, V.L. Miguéis, Lucas F.M. da Silva, "Exploring the use of deep neural networks for sales forecasting in fashion retail", Decision Support Systems, vol 114, pp. 81-93, 2018, https://doi.org/10.1016/j.dss.2018.08.010.

[3] Sushil Punia, Konstantinos Nikolopoulos, Surya Prakash Singh, Jitendra K. Madaan & Konstantia Litsiou, "Deep learning with long short-term memory networks and random forests for demand forecasting in multi-channel retail", International Journal of Production Research, vol. 58 no. 16, pp. 4964-4979, 2020, https://doi.org/10.1080/00207543.2020.1735666.

[4] Sushil Punia, Surya P. Singh, Jitendra K. Madaan, "A cross-temporal hierarchical framework and deep learning for supply chain forecasting". Computers & Industrial Engineering. vol. 149, 2020, https://doi.org/10.1016/j.cie.2020.106796.

[5] Shaohui Ma, Robert Fildes,- "Retail sales forecasting with meta-learning", European Journal of Operational Research, vol. 288, no. 1, 2021, pp. 111-128, https://doi.org/10.1016/j.ejor.2020.05.038.

[6] Cheng-Hong Yang, Cheng-Feng Lee, Po-Yin Chang, "Export- and import-based economic models for predicting global trade using deep learning", Expert Systems with Applications, vol 218, 2023, https://doi.org/10.1016/j.eswa.2023.119590.

[7] Mustafa Kaya, Mehmet Baha Karan, Erdinç Telatar, "Electricity price estimation using deep learning approaches: An empirical study on Turkish markets in normal and Covid-19 periods", Expert Systems with Applications, vol 224, 2023, https://doi.org/10.1016/j.eswa.2023.120026.

[8] Palacios D. " Sales management: what it is, why it matters and how to implement it". HubSpot, 2021, https://blog.hubspot.es/sales/gestion-de-ventas.

[9] "What is Deep Learning? | IBM", https://www.ibm.com/es-es/topics/deep-learning (accessed Jun. 25, 2023).

[10] "The neural network model | IBM Documentation", https://www.ibm.com/docs/es/spss-modeler/saas?topic=networks-neural-model (accessed Jun. 25, 2023).

[11] "What is JavaScript (JS)? | AWS", https://aws.amazon.com/es/what-is/javascript/#:~:text=AWS%20SDK%20para%20JavaScript%20es,Node (accessed Jun. 25, 2023).

[12] Zachary C. Lipton, John Berkowitz, Charles Elkan. "A Critical Review of Recurrent Neural Networks for Sequence Learning". arXiv preprint arXiv:1506.00019v4, 2015, https://doi.org/10.48550/arXiv.1506.00019.

[13] "Gated Recurrent Units – Understanding the Fundamentals | Data Science", https://datascience.eu/machine-learning/gated-recurrent-units-understanding-the-fundamentals/ (accessed Jul. 10, 2023).

[14] Muñoz González A., Salazar Guillen F. "LSTM neural networks application for the prediction of the daily movement direction problem for Bitcoin". In XI Congress of Applied Mathematics, Computational and Industrial - MACI 2021, Argentine Association of Applied Mathematics, 2021.

[15] Dongliang Li, Youyou Li and Zhigang ZHANG, "Automatic Extractive Summarization using GAN Boosted by DistilBERT Word Embedding and Transductive Learning" International Journal of Advanced Computer Science and Applications (IJACSA), vol. 14 no. 11, 2023. http://dx.doi.org/10.14569/IJACSA.2023.0141107.

[16] Benjamin Lindemann, Timo Müller, Hannes Vietz, Nasser Jazdi, Michael Weyrich, "A survey on long short-term memory networks for time series prediction", Procedia CIRP, vol. 99, 2021, pp. 650-655, https://doi.org/10.1016/j.procir.2021.03.088.

[17] Patrick Schneider, Fatos Xhafa, "Anomaly Detection and Complex Event Processing over IoT Data Streams, in Chapter 3 - Anomaly detection: Concepts and methods", 2022, pp. 49-66, https://doi.org/10.1016/B978-0-12-823818-9.00013-4.

[18] "MAPE Mean Absolute Percentage Error | Oracle® Fusion Cloud EPM Working with Planning", https://docs.oracle.com/cloud/help/es/pbcs_common/PFUSU/insights_metrics_MAPE.htm (accessed Jul. 10, 2023).

[19] Arnaud de Myttenaere, Boris Golden, Bénédicte Le Grand, Fabrice Rossi, "Mean Absolute Percentage Error for regression models", Neurocomputing, vol. 192, 2016, pp. 38-48, https://doi.org/10.1016/j.neucom.2015.12.114.

[20] Vergara J., "What are training and inference in artificial intelligence?" Neuroons, 2021, https://neuroons.com/es/que-son-entrenamiento-e-inferencia-en-inteligencia-artificial/ (accessed Jul. 10, 2023).