Journal of Infrastructure, Policy and Development 2024, 8(15), 9003.
https://doi.org/10.24294/jipd9003

Article

Estimation of the LINDA index prediction based on deep learning models

Diego Armando Giral-Ramirez*, Cesar Hernández*, Ernesto Cadena-Muñoz*

Universidad Distrital Francisco José de Caldas, Bogota 111611-111611537, Colombia
*** Corresponding authors:** Diego Armando Giral-Ramírez, dagiralr@udistrital.edu.co; Cesar Augusto Hernández-Suarez, cahernandezs@udistrital.edu.co; Ernesto Cadena-Muñoz, ecadenam@udistrital.edu.co

CITATION

Giral-Ramirez DA, Hernández C, Cadena-Muñoz E. (2024). Estimation of the LINDA index prediction based on deep learning models. Journal of Infrastructure, Policy and Development. 8(15): 9003. https://doi.org/10.24294/jipd9003

ARTICLE INFO

Received: 6 September 2024
Accepted: 9 October 2024
Available online: 16 December 2024

Abstract: Recognizing the importance of competition analysis in telecommunications markets is essential to improve conditions for users and companies. Several indices in the literature assess competition in these markets, mainly through company concentration. Artificial Intelligence (AI) emerges as an effective solution to process large volumes of data and manually detect patterns that are difficult to identify. This article presents an AI model based on the LINDA indicator to predict whether oligopolies exist. The objective is to offer a valuable tool for analysts and professionals in the sector. The model uses the traffic produced, the reported revenues, and the number of users as input variables. As output parameters of the model, the LINDA index is obtained according to the information reported by the operators, the prediction using Long-Short Term Memory (LSTM) for the input variables, and finally, the prediction of the LINDA index according to the prediction obtained by the LSTM model. The obtained Mean Absolute Percentage Error (MAPE) levels indicate that the proposed strategy can be an effective tool for forecasting the dynamic fluctuations of the communications market.

Keywords: artificial intelligence; deep learning; LINDA index; long-short term memory; market prediction

1. Introduction

Recent studies in Colombia reveal a high concentration in the telecommunications market in fixed and mobile services. This situation directly impacts prices, which are higher for users, and competitiveness, which is low (Bardey et al., 2013). Limited competition creates an unfavorable environment for the development of Information and Communication Technologies (ICT), restricting the supply of new-generation services, such as the Internet of Things (IoT), and negatively affecting the growth of companies in the sector (Beckert and Siciliani, 2022; Myers and Tauber, 2011).

The first step is to recognize the importance of competition analysis in Colombian telecommunications markets (Ochuba et al., 2024). Until now, static econometric tools have been used to allow evaluation in each period but make long-term prediction difficult. This underlines the need to establish dynamic algorithms for more accurate prediction and adaptation to the country's realities. In this way, a more competitive environment could be created, and the conditions for users and companies in voice and data services could be improved at the mobile, residential, and corporate levels (Adetunji and Moses, 2022; Beckert and Siciliani, 2022).

Therefore, it is essential to evaluate the algorithms that can perform this task, with one of the most adaptable and dynamic options being the use of AI techniques. These techniques can offer greater accuracy in prediction and adaptability, thus improving the competitiveness and quality of the services provided in the sector (Giral et al., 2021).

In this context, AI is a solution capable of processing large volumes of data and detecting patterns that would be impossible to identify manually (Giral-Ramírez, 2022). The LINDA index, designed to measure telecom markets, is used to measure whether oligopolies exist or whether there are inequalities in the market and benefits significantly from AI (Venegas et al., 2022). Using advanced machine learning and predictive analytics algorithms, AI improves the accuracy and speed of the LINDA Index analyses and provides effective results. Integrating AI using the LINDA Index enhances the ability to respond and adapt to a dynamic and constantly changing telecommunications environment.

1.1. LINDA index

In the current literature, various indices are designed to analyze and evaluate competition in telecommunications markets, mainly through concentration measures of companies in the sector (Venegas et al., 2022). These indices, known as concentration indices, use vital variables such as the revenue generated, the number of users, and the traffic each operator handles. These variables allow a comprehensive assessment of the level of competition, providing a clear view of how resources and benefits are distributed among the different market players.

This index focuses on the distribution of the k companies with the largest market share, as it is designed to assess the degree of inequality in a market. The presence of oligopolies means that, unlike the previous indices, it compares the concentration between two groups of companies, those with the largest market share (leaders) and the others, allowing their joint relative incidence to be calculated about the rest of the companies (Equation (1)) (Venegas et al., 2022).

$$L = \frac{1}{N(N-1)} \sum_{i=1}^{N-1} \frac{\bar{X}_\iota}{\bar{X}_{N-i}} \tag{1}$$

where:

$\bar{X}_i$ is the average market share of the first *i* companies.

$\bar{X}_{N-i}$ is the average market share of the remaining companies.

N is the total number of companies in the market.

The LINDA index's range is between zero and infinity, and its interpretation is given in **Table 1** (Venegas et al., 2022).

Table 1. Interpretation of the LINDA index.

Concentration	Range
Low	< 0.2
Moderate	0.2–0.5
High	0.5–1
Very High	> 1

1.2. Artificial intelligence models

The financial market is random and difficult to predict, and so is the telecommunication services market. Analysts have always sought innovative ways to predict this non-linear and chaotic domain (Adetunji and Moses, 2022; Al-Mhdawi and Qazi, 2024). Analysts used three main techniques to solve this problem. The first one is fundamental analysis. The second one is technical analysis based on historical data and some indicators. Recently, researchers have been trying to get more valuable features like sentiment index and trends from historical data and then feed these data into the input features (Giral-Ramírez, 2022). The third combines the above two techniques: Fundamental analysis and technical analysis. Technical analysis is entirely based on historical data, making applying techniques like AI relevant. Different approaches were adopted to roughly predict the market movement. The most famous ones can be classified into five categories: Statistics, Machine Learning, Pattern Recognition, Sentiment Analysis, and Hybrid (Berradi et al., 2020).

Deep Learning has demonstrated outstanding performance in modeling diverse areas and domains, such as object recognition, image classification, and language recognition (Giral et al., 2021). It has also shown promising results in time series prediction in different disciplines, which underlines the ability of neural networks to process, analyze, and model this type of information. In particular, Long Short-Term Memory (LSTM) neural networks have stood out for their high performance in these tasks (Ding et al., 2022; Lu et al., 2024).

LSTM networks

LSTMs are a particular type of recurrent network (RNNs), which, in turn, are a specific type of Deep Learning (Das et al., 2023). The main characteristic of RNNs is their ability to persist information, allowing them to “remember” previous states and decide on the following state (Li et al., 2022; Shi et al., 2022). Due to the above, they are a suitable candidate for processing time series. However, RNNs can only model short-term dependencies, while LSTMs can learn long dependencies; that is, they have a longer-term “memory” (Rithani et al., 2023).

An LSTM network is an RNN that processes input information and updates the state of the RNN (Prakash et al., 2023). The state of the RNN contains information from previous time units. LSTMs can forecast a time series using information from last time units as input. To do this, the LSTM is trained with sequential output, where the responses (targets) are the training sequences with values shifted by one-time units (Rithani et al., 2023). There are two methods for forecasting: Open-loop and closed-loop forecasting.

1.3. Comparison between prediction techniques

Table 2 summarizes the currently best-performing prediction techniques based on the state of the art performed previously.

Table 2. Comparison between prediction techniques.

AI Technique	Precision Level	Complexity Level	Reference
Time Series	High	High	(Li et al., 2022)
ANN	Medium	Medium	(He and Wu, 2021)
Deep Learning	Medium	Medium	(Li et al., 2022; Shi et al., 2022)
RNN	Medium	Medium	(Ma et al., 2023)
LSTM	Very High	High	(Wang and Zhang, 2020)

1.4. Contribution and objective

The main contribution of this article is developing an AI model designed to predict telecommunications markets. This model is based on the LINDA indicator, an advanced analytical tool that accurately assesses trends and patterns in the telecommunications sector. Through the implementation of machine learning and data analysis techniques, the proposed model offers an innovative methodology that overcomes the limitations of traditional approaches, providing more accurate and relevant predictions for strategic decision-making in this field.

This study is focused on validating the effectiveness of the AI model, which is based on the LINDA indicator, in enhancing the accuracy of predicting telecommunications market dynamics. By achieving this goal, we aim to demonstrate the model's viability and provide a powerful tool for analysts and sector practitioners. This tool has the potential to significantly improve planning and adaptation to market changes in the telecommunications sector.

This article aims not to propose or develop a new prediction strategy but to take advantage of the advantages of the LINDA indicator, which is currently used by government entities, and of LSTM networks, which have demonstrated effective results in prediction. No comparisons are made with other strategies since it is based on validated methodologies. Instead of innovating in the technique, the hybridization of the LINDA indicator and LSTM networks is sought to address the estimation of future oligopoly environments.

2. Materials and methods

The objective of having a definition of test scenarios for the project is to achieve the design of an AI model that allows characterizing and predicting competition in the telecommunications market for mobile service regarding internet traffic, the number of internet accesses, and internet revenues in Colombia. The methodology is summarized in **Figure 1** in the diagram.

The test scenarios for the project are described below.

a) Collect data on the variables Internet traffic, number of accesses, and revenue for the mobile market in Colombia.
b) Clean the available data with the defined variables.
c) Standardize and/or homogenize the variables with which the model is to be generated.
d) The LINDA index is calculated Based on standardization and/or homogenization.

e) Prepare the dataset with 70% of the data for the training process, 15% for testing, and 15% for verification of the prediction levels.
f) Implement the prediction model using LSTM.
g) Perform different iterations to ensure reliability in collecting training and testing data.
h) Check the R-squared value and adjust the variables or the model so that the value exceeds 80%.
i) Suppose the prediction values are below the established limits. In that case, the variables used must be reviewed, or the empirical ranges established for the model must be adjusted until the best possible values are obtained.

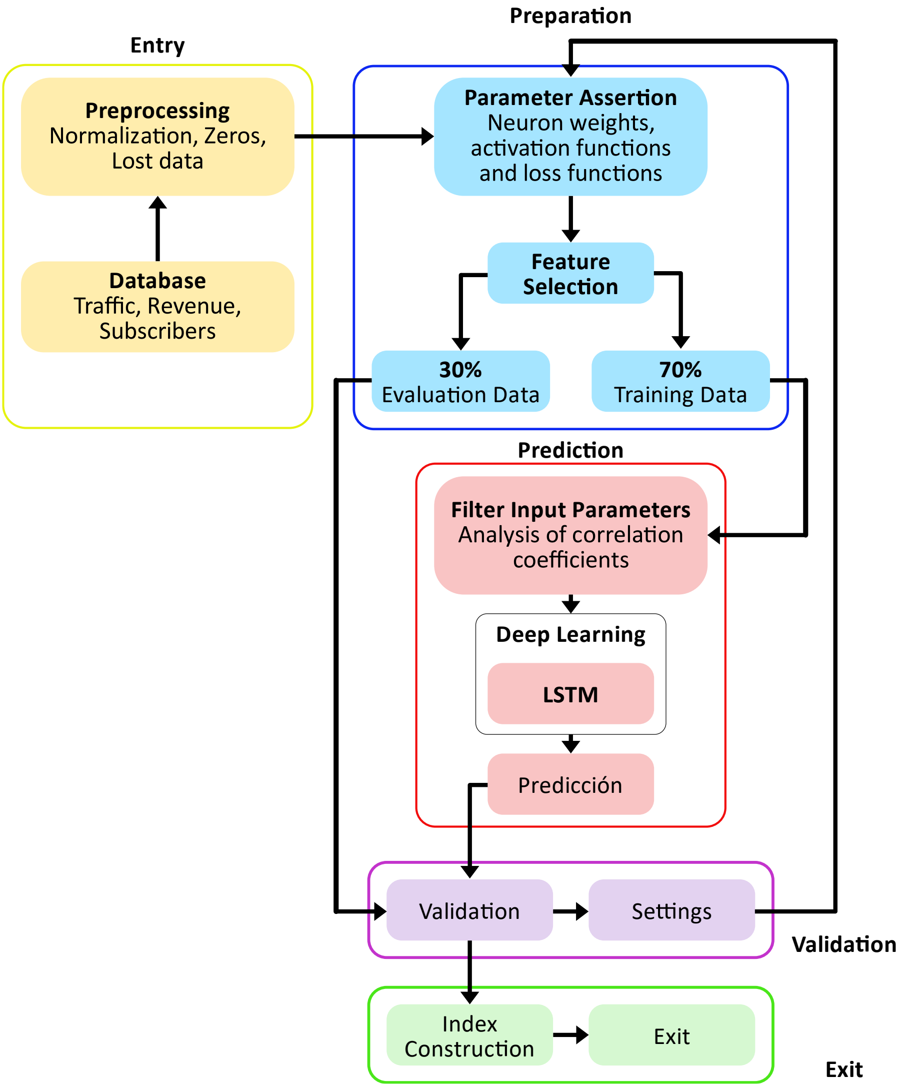

Figure 1. Proposed LSTM model.

2.1. Database

For this activity, the information presented by the Communications Regulation Commission was used (Comisión de Regulación de Comunicaciones-Republica de Colombia, 2024a, 2024b).

The database corresponds to the number of users, revenue, and traffic of 15 network operators. A program was developed in Python to organize and filter the variables of interest (traffic, users, and revenue) for each month and by the operator from 2012 to 2022 (inclusive), cleaning and adapting the information.

Although the data used in this study are accurate and publicly available, it is essential to consider the limitations due to the region and time. The data are limited to a specific region and cover a specific period, which may restrict the generalization of the results. In addition, the region's particularities and the period analyzed may not reflect the conditions of other areas or times, which could affect the applicability of the results.

2.2. LSTM system model

Making predictions with a high degree of accuracy is beneficial for planning and control in many fields of research and development. However, such a degree of accuracy in estimates brings with it a high level of difficulty (Ghaffar Nia et al., 2023); however, there are promising prediction techniques based on AI with the capacity to provide awareness, reasoning, and learning, as is the case of LSTM (He et al., 2010; Hernández et al., 2020).

The future estimation of the time series of interest, in this case, the traffic, revenue, and subscribers of telecommunications companies, or directly the selected market concentration indices such as Stenbacka, IHH, or LINDA, will be carried out using a recurrent neural system based on deep learning such as LSTM. Initially, the theoretical concept of LSTM is defined, how the input time series to the system is modeled is described, and the layered structure of the LSTM network is analyzed; subsequently, the mathematical model that explains the LSTM system is built by describing the interaction between the input neurons, the memory cells and output neurons, during the training or learning process (Hernández et al., 2020).

Traditional artificial neural networks cannot store information, so it is necessary to modify their topology by creating recurrent structures that provide feedback to the neuron and allow information storage; these structures are known as recurrent neurons. The union of a set of these neurons is called Recurrent Neural Networks (RNN). They allow the preservation of subsequent states between different time intervals where their parameters are shared between the model's multiple parts, allowing for better generalization (Veeriah et al., 2015). One of the problems of RNN networks is long-term dependency; this problem raises the need not always to study a whole history to perform a current task, which implies that these neural networks only store the information learned in the past and are not able to store new information in the short term. LSTMs can be explicitly designed to avoid long-term dependency problems, remember information for long periods, and learn new information. LSTM blocks contain memory cells that allow them to remember a value for an arbitrary time and use it when needed; it also has a forgetting layer that can erase the memory content when it is not applicable. All components are built as differentiable functions and trained during backpropagation (Wang et al., 2003).

The structure of an LSTM is shown in **Figure 2**, where the memory cell is symbolized by the letter C, the forgetting layer by the letter O, the input layer by the letter E, and the output layer by the letter S (Hernández et al., 2020).

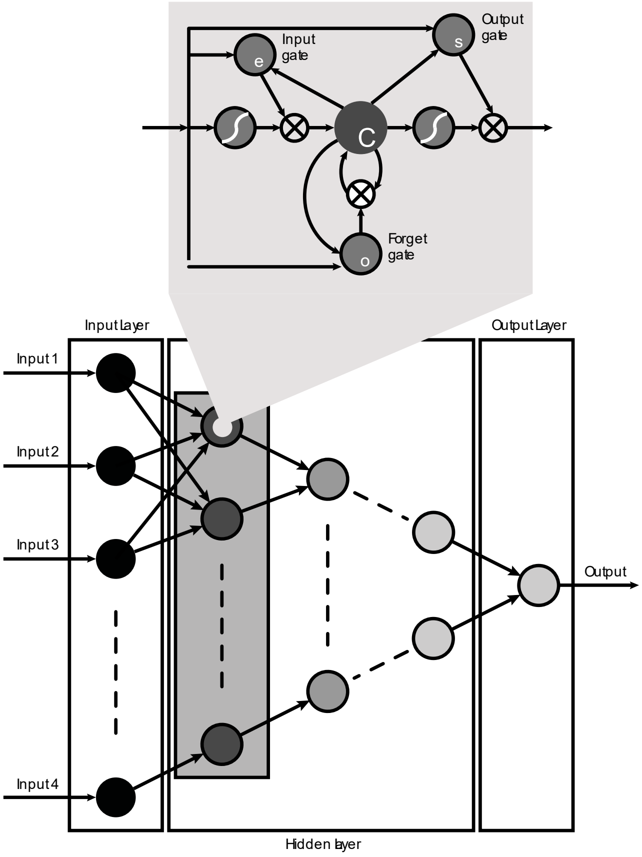

Figure 2. Graphical representation of LSTM-type neural networks (Hernández et al., 2020).

LSTM can be considered a differentiable function approximator, usually trained with gradient descent (Graves, 2012). Although initially, a truncated form of BPTT (Backpropagation Through Time) was used to approximate the error gradient (Hochreiter and Schmidhuber, 1997), in the research, the calculation with BPTT without truncation was used based on what Graves and Schmidhuber (2005) proposed in. The LSTM neural network operation uses the notations established in **Table 3**, consistent with (Graves, 2012).

Table 3. Notations for the development of the mathematical model.

	Memory Block	**Input Gate**	**Forget Gate**	**Output Gate**	**Memory Cell**
Subscript	i	l	$\varnothing$	w	c
Input	x_i	a_l^t	$a_\varnothing^t$	a_w^t	a_c^t, s_c^t
Output	y_i	b_l^t	$b_\varnothing^t$	b_w^t	$b_c^t = b_w^t\ (s_c^t)$
Number of Units	I	N/A	N/A	N/A	C
Trigger Function	N/A	f sigmoide	f sigmoide	f sigmoide	f (in-cell) h (sal-cell)

LSTM system flowchart

The flowchart for training (**Figure 3**) begins its process by randomly initializing each neuron with values ranging from −1 to 1. Each training example is taken, and the output is evaluated against the expected one. Suppose the response delivered does not correspond to the desired one. In that case, the algorithm calculates the error between the output obtained by the system and the expected one, correcting each weight of the gates (input, output, forget) and the cell through the application of weights and using tangential and sigmoid functions until finishing with all the training examples, and in this way approximating the output of the model to the expected one (by decreasing the error) (Hernández et al., 2020).

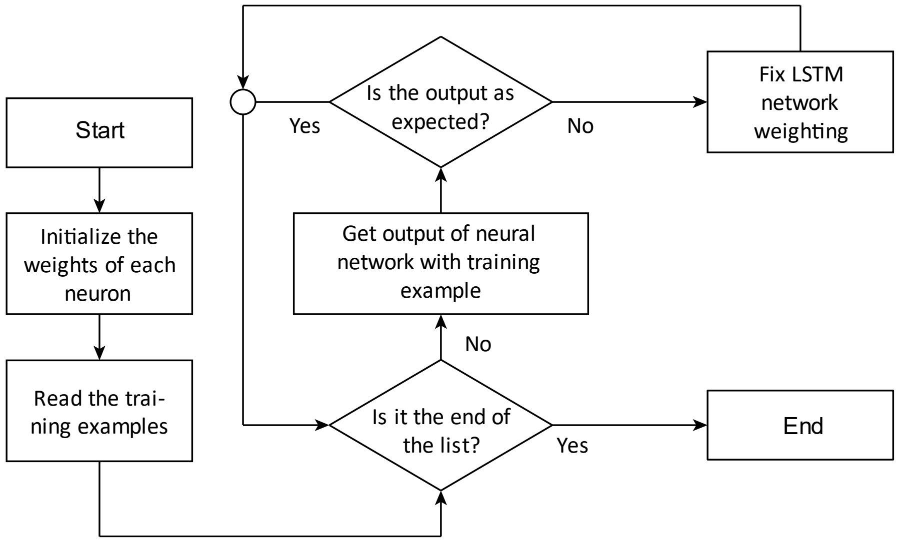

Figure 3. Flowchart for LSTM training.

2.3. Testing protocol

Considering the comprehensive definition of the test scenarios, a detailed description of the Model testing protocol is provided. **Figure 4** depicts this stage meticulously handles the input and output data.

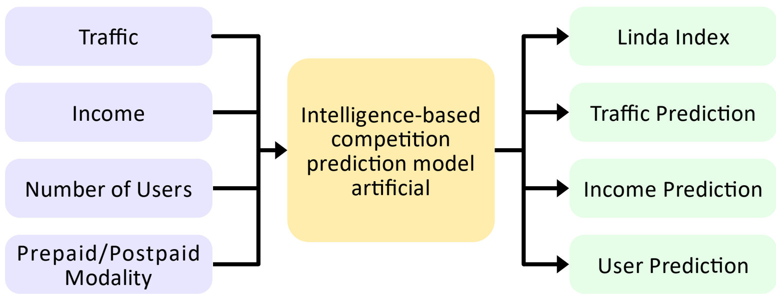

Figure 4. Input and output data of the LSTM model LSTM.

The chosen approach is the hold-out method, a versatile technique for training machine learning models. It involves flexible data division into distinct sets: One for training and another for validation and testing. This method is adept at assessing the performance of a machine-learning model with new data.

Below, the division of the hold-out testing method is presented, which can be worked with 70% for training, 15% for validation, and 15% for testing, or it can also be worked with 80% for training and 20% for testing. Although mobile operators have more than 130 data for this case, it could be trained with 90 or 100 data, validated from

15 to 20 data, and another 15 or 20 data to test the data. A limitation is that recent operators have less than 30 data (Vardhan et al., 2022; Vardhan et al., 2023).

The cross-validation method also obtains at least ten datasets with different data for validation and testing according to the randomness of the selection. This method allows you to verify the model's performance on a dataset it uses for training and testing. With this method, you identify the "prediction error" and not the "training error".

Cross-validation divides your data into parts. Like split validation, you train on one part and then test on the other. However, unlike split validation, cross-validation is not done just once; instead, it takes an iterative approach to ensure that all the data can be used for testing.

With the technique in mind, for the test scenarios, the following steps must be taken:

a) Collect current data on the variables of internet traffic, the number of accesses, and revenue for the postpaid mobile market in Colombia.
b) Clean the available data with the defined variables.
c) Standardize and/or homogenize the variables with which the model is to be generated.
d) After this, the following steps are followed:
e) The prediction uses AI techniques for traffic, revenue, and users, and the LINDA index is calculated.
f) The prediction dataset is prepared with 70% or more of the historical data for training, 15% or less for testing, and 15% for verifying the prediction levels. At this stage, the MAPE is calculated, which allows for estimating the efficiency of each prediction. The objective is for this index to approach zero (Doğan, 2021).

3. Results and discussion

According to the analysis of the variables that affect the competition market and to establish or predict said competition in the mobile services market, the following input variables are defined:

- Traffic: The amount of traffic produced by operators each month.
- Revenue: The amount of revenue reported by operators each month.
- Users or Subscribers: Number of users operators report each month.

The following are the output data of the model:

- LINDA indexes according to the database.
- Prediction using LSTM for traffic over time, revenue over time, and users over time.
- Prediction of the LINDA index.

3.1. LINDA index

This indicator is usually used to measure the possible existence of an oligopoly and the inequality between different market shares. In addition, like the concentration ratio, it is calculated for several leading companies in the market so that their joint relative incidence can be calculated for the rest of the participants at that end of the market (supply or demand).

Initially, the data corresponding to the analysis variables were obtained, such as traffic, revenue, and subscribers, corresponding to each of the telecommunications companies that operated and operated in Colombia from 2012 to September 2022 (inclusive); this information was obtained from the post data database of the Communications Regulation Commission. Subsequently, the data was organized in Excel to create a database with the information of interest organized chronologically. Here, three databases were finally obtained:

a) Traffic from the demand for fixed-charge mobile internet (postpaid).
b) Revenue from fixed-charge mobile internet demand (postpaid).
c) Subscribers from fixed-charge mobile internet demand (postpaid).

Since forecasts of the LINDA index are planned to be made later, it was decided to calculate this index monthly to obtain more data. The procedure required to calculate the LINDA index required that for each period (month), the telecommunications companies be ordered from highest to lowest according to the variable's value to be analyzed (traffic, revenue, or subscribers).

Finally, the LINDA index was calculated for each database mentioned above. It was evident that when some companies had a zero value in the variable of interest, the LINDA index was indeterminate. If it was very close to zero, it increased exponentially. Due to the above, it was decided to eliminate the data equal to zero since the interpretation is that the company did not operate in the said period. Additionally, it was decided to eliminate all data less than 50,000 in the traffic and revenue databases; the amount of data eliminated was 36, which gives approximately a value of less than 0.4% of the total database.

According to the description, **Figure 5** presents the LINDA index for the traffic of fixed-charge mobile Internet demand, **Figure 6** presents the LINDA index for the revenue of fixed-charge mobile Internet demand, and **Figure 7** presents the LINDA index for the number of subscribers of fixed-charge mobile Internet demand.

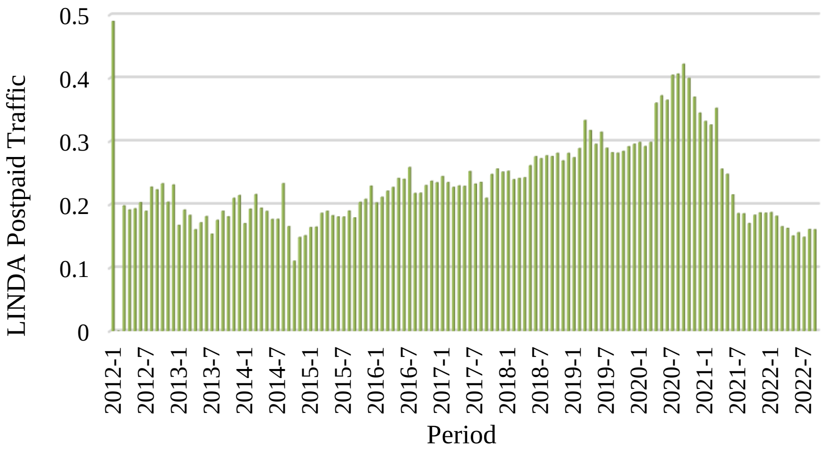

Figure 5. LINDA index for postpaid mobile internet demand traffic.

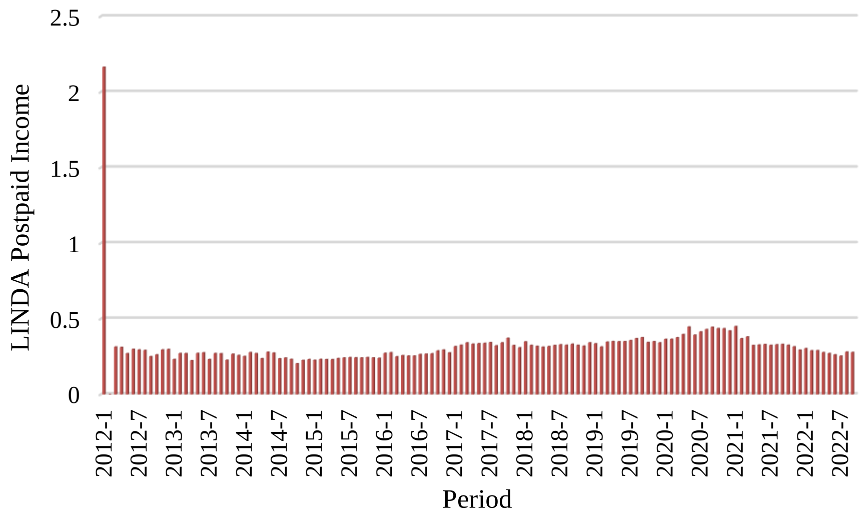

Figure 6. LINDA index for postpaid mobile internet demand revenues.

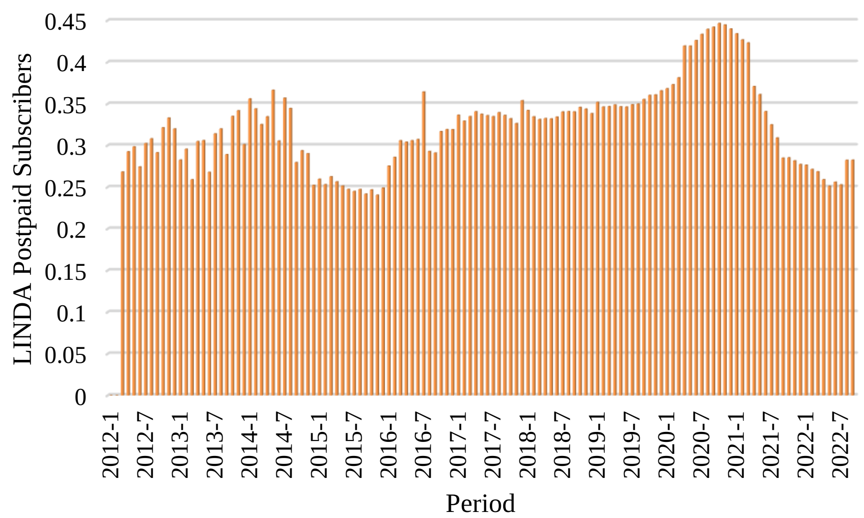

Figure 7. LINDA index for postpaid mobile internet demand subscriber count.

3.2. Prediction using LSTM

This section presents the predictions obtained using the LSTM network to analyze the results for the first half 2025. The results correspond to the total number of users, the total revenue, and the total traffic of one of the operators available in the database (Colombia Telecomunicaciones) for postpaid services. The predictions for the 15 operators need to be presented to avoid duplication and redundancy.

The proposed algorithm's execution generates the following results: **Figure 8** presents the prediction for the number of users, **Figure 9** presents the prediction for total revenue, and **Figure 10** presents the prediction for total traffic.

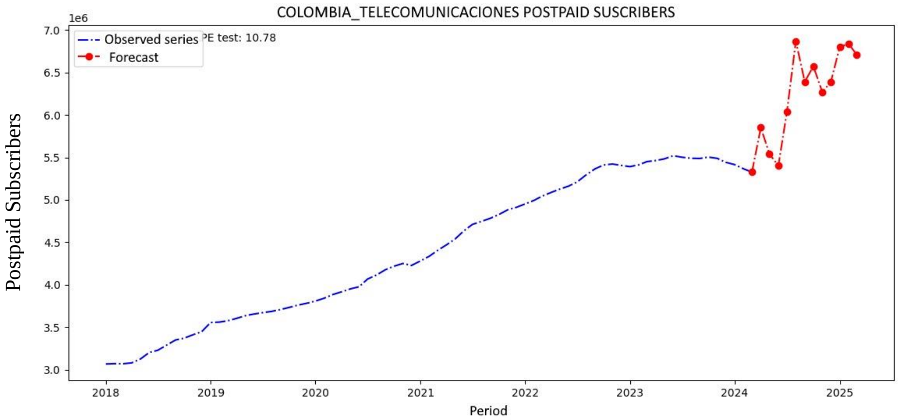

Figure 8. LSTM results for Colombia Telecommunications postpaid subscribers.

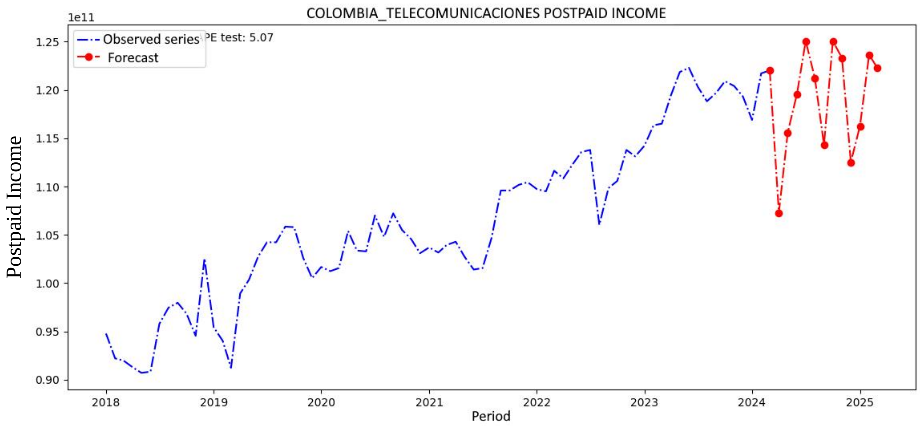

Figure 9. LSTM results for Colombia Telecommunications postpaid income.

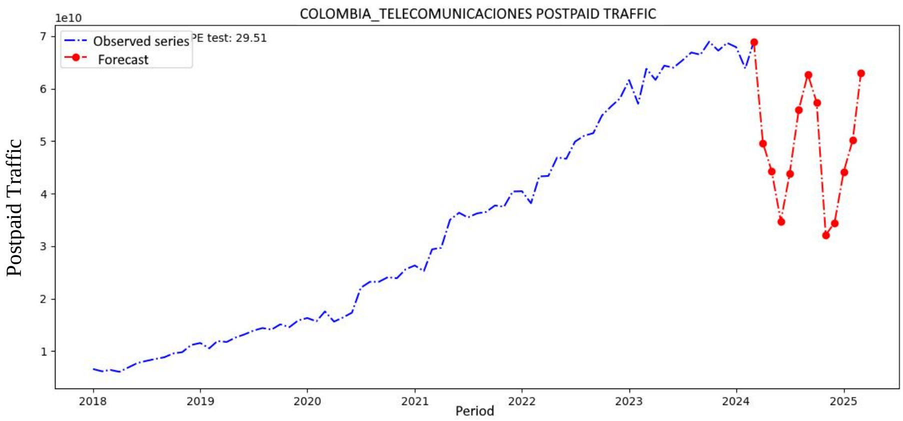

Figure 10. LSTM results for Colombia Telecommunications postpaid traffic.

3.3. LINDA index for prediction

Based on the predictions made, the index for the following year is analyzed using the current data of the different operators and variables. **Figure 11** shows the LINDA index prediction with the LSTM model. According to the prediction for the following 12 months, the index tends to increase for subscribers, and after an increase, it decreases for revenue and traffic. It is clarified that as updated data is entered, the model will generate new predictions based on the new data.

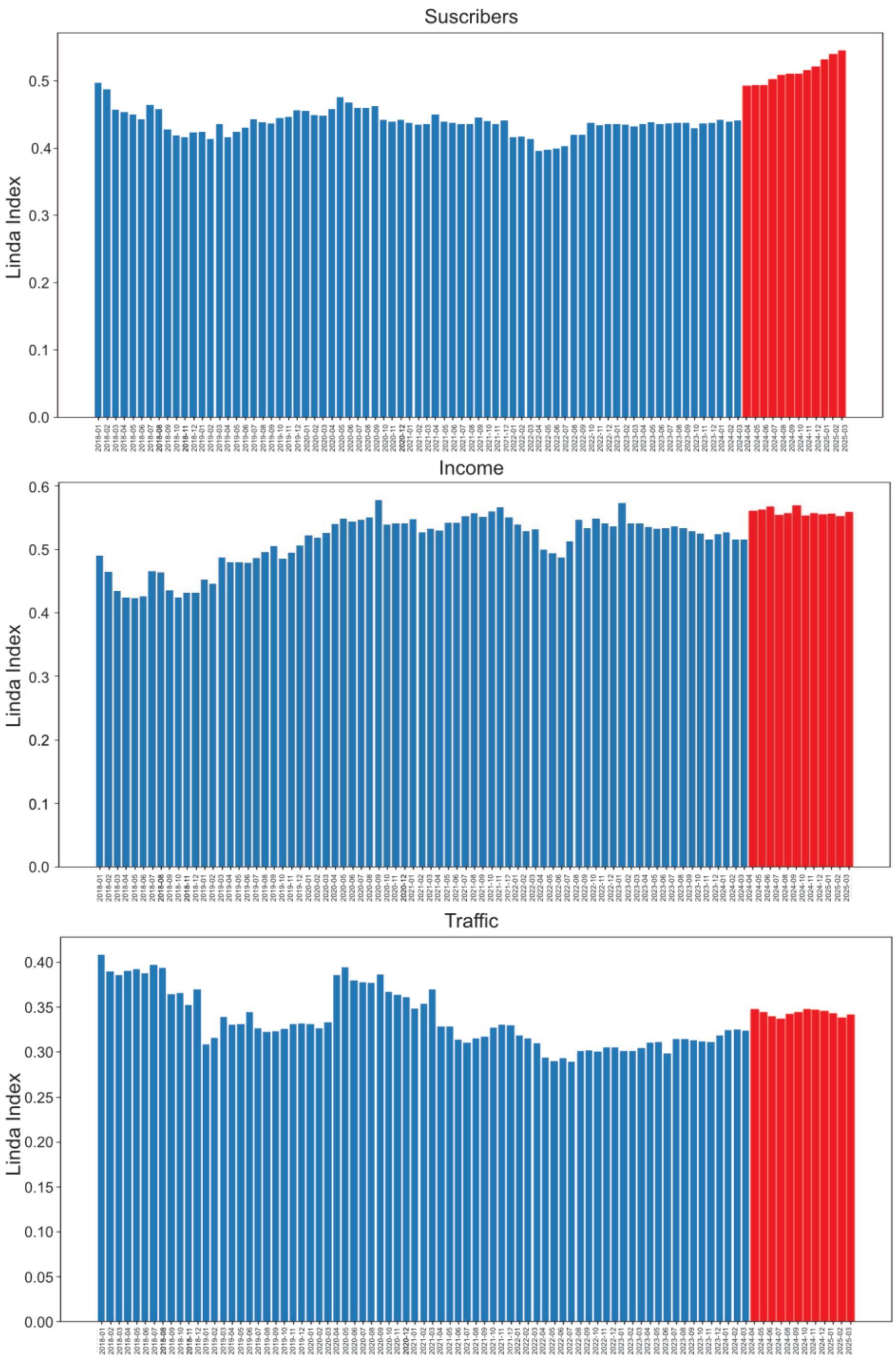

Figure 11. LINDA index prediction with AI model.

3.4. Discussion

During the calculation of the LINDA index, it was observed that when some companies recorded a value of zero in the variable of interest, the index was indeterminate. In addition, this index showed exponential growth when the value was close to zero. Data with a zero value was removed to address this issue, indicating that the company was not operational during the corresponding period. Likewise, it was decided to exclude all data less than 50,000 in the traffic and income databases for postpaid and prepaid, as well as global modalities. In total, 36 data were eliminated, representing less than 0.4% of the database.

The splitting of the data for training and prediction allowed the effectiveness of the models to be assessed. The results showed a low error rate in the traffic and revenue variables, confirming an appropriate model selection. However, the discrepancy in the data related to accesses suggests that improvements could be implemented to adapt the model. This significant divergence points to the need to review and adjust the approach.

It should be noted that the LINDA index compares the groups formed by the two companies with the highest values in the variable of interest (traffic, revenue, or subscribers) with the rest of the companies. In some cases, it is possible to observe a greater concentration in a group composed of the three or four most dominant companies in the market.

In the analysis of the predictions for the next 12 months, a divergent trend is observed in the key indicators. While the number of subscribers shows a continuous increase, both revenues and traffic show a decrease. This dynamic could be indicative of market saturation. It is essential to highlight that these predictions are subject to change as more recent data is introduced into the model, allowing a constant update of the results and better adapting to market fluctuations.

The analysis of the results reinforces the idea that the market is highly concentrated and dominated by a small group of critical providers. This concentration suggests the existence of significant barriers to the entry of new players, which could limit competition and affect innovation in the sector. In addition, the persistent dominance of large players raises concerns about the fairness and long-term sustainability of the telecommunications ecosystem.

The analysis of the results reinforces the idea that the market is highly concentrated and dominated by a small group of critical providers. This concentration suggests the existence of significant barriers to the entry of new players, which could limit competition and affect innovation in the sector.

In this context, regulatory measures and technological advances, such as implementing 5G networks, are emerging as critical factors in reconfiguring the market. Regulatory strategies that promote the entry of new competitors and foster technological innovation could mitigate the risks associated with market concentration. In addition, developing new offerings and diversifying technologies can improve competitiveness and transform the sector’s dynamics towards a more balanced and dynamic environment.

4. Conclusion

This paper presents an AI model based on the LINDA indicator to predict trends in telecommunication markets, using traffic, revenue, and user count data as input parameters. The model provides accurate predictions based on LSTM neural networks and stands out for its excellent performance according to the MAPE levels. These results indicate that the proposed strategy can be an effective tool for forecasting the dynamic fluctuations of the communications market.

Several critical aspects of the telecommunications market structure are highlighted. Adapting the LINDA index and data management has allowed a more precise interpretation of the behavior of dominant companies. Eliminating data equal to zero and those less than 50,000, although minimal in proportion, was crucial to avoid distortions in the analysis. The effectiveness of the applied model is underscored by the low error rate in traffic and revenue predictions, although areas for improvement in access prediction were identified. The 12-month projections reveal potential market saturation, with an increase in the number of subscribers and a decrease in revenue and traffic. This, combined with the market's concentration in a small group of dominant players, underscores the need for regulatory measures that promote competition and mitigate the risks associated with such concentration.

The option of deep learning-based strategies, such as LSTM networks, is an essential tool for anticipating and adapting to the fluctuating dynamics of the communications market. This insight opens new research opportunities and promotes the practical implementation of innovative solutions in the future. By combining advanced predictive capabilities with rapid adaptability, these techniques could significantly revolutionize how companies and researchers face the challenges of a constantly evolving market, providing a robust framework for continuous development and innovation.

Author contributions: Conceptualization, DAGR, CH and ECM; methodology, DAGR, CH and ECM; software, DAGR, CH and ECM; validation, DAGR, CH and ECM; formal analysis, DAGR, CH and ECM; investigation, DAGR, CH and ECM; resources, DAGR, CH and ECM; writing—original draft preparation, DAGR, CH and ECM; writing—review and editing, DAGR, CH and ECM; visualization, DAGR, CH and ECM; supervision, DAGR, CH and ECM; project administration, DAGR, CH and ECM; funding acquisition, DAGR, CH and ECM. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Science, Technology and Innovation (Minciencias) and the Communications Regulatory Commission (CRC)—Colombia. It was conducted under the framework of Contract 80740 082 of 2022, signed between Fiduprevisora and the Universidad Distrital Francisco José de Caldas, which was awarded in Call 908 of 2021 of MinCiencias.

Acknowledgments: The authors thank the Research Office (ODI) of the Universidad Distrital Francisco José de Caldas, Ministry of Science, Technology and Innovation (Minciencias), and the Communications Regulatory Commission (CRC).

Data availability statement: The database corresponds to the number of users, revenue, and traffic of 15 network operators. The data is taken from the Communications Regulation Commission of the Republic of Colombia. The data set is archived and published at https://postdata.gov.co/search/type/dataset.

Conflict of interest: The authors declare no conflict of interest.

References

Adetunji, A. J., and Moses, B. O. (2022). The role of network technologies in the enhancement of the health, education, and energy sectors. Network and Communication Technologies, 7(1), 39.

Bardey, D., Becerra, A., and Cabrera, P. (2013). Análisis económico de la normativa de libre competencia en Colombia.

Beckert, W., and Siciliani, P. (2022). Protecting Sticky Consumers in Essential Markets. Review of Industrial Organization, 61(3), 247–278. https://doi.org/10.1007/s11151-022-09880-z

Berradi, Z., Lazaar, M., Mahboub, O., and Omara, H. (2020). A Comprehensive Review of Artificial Intelligence Techniques in Financial Market. 2020 6th IEEE Congress on Information Science and Technology (CiSt), 367–371. https://doi.org/10.1109/CiSt49399.2021.9357175

Comisión de Regulación de Comunicaciones - Republica de Colombia. (2024a). Postscript - Beyond the data (Spanish). Fixed-Rate Mobile Internet Subscribers, Revenues and Traffic. Available online: https://www.postdata.gov.co/dataset/suscriptores-ingresos-y-tráfico-de-internet-móvil-por-cargo-fijo (accessed on 1 July 2024).

Comisión de Regulación de Comunicaciones - Republica de Colombia. (2024b). Postscript - Beyond the data (Spanish). Mobile Internet Subscribers, Revenues and Traffic on Demand. Available online: https://www.postdata.gov.co/dataset/abonados-ingresos-y-tráfico-de-internet-móvil-por-demanda (accessed on 1 July 2024).

Das, S., Tariq, A., Santos, T., Kantareddy, S. S., and Banerjee, I. (2023). Recurrent neural networks (RNNs): architectures, training tricks, and introduction to influential research. Machine Learning for Brain Disorders, 117–138.

Ding, X., Lv, Q., Zou, Y., and Zhang, G. (2022). Spectrum Prediction for Satellite based Spectrum-Sensing Systems Using Deep Learning. GLOBECOM 2022 - 2022 IEEE Global Communications Conference, 3472–3477. https://doi.org/10.1109/GLOBECOM48099.2022.10000832

Doğan, E. (2021). Performance analysis of LSTM model with multi-step ahead strategies for a short-term traffic flow prediction. Zeszyty Naukowe. Transport/Politechnika Śląska, 111.

Ghaffar Nia, N., Kaplanoglu, E., and Nasab, A. (2023). Evaluation of artificial intelligence techniques in disease diagnosis and prediction. Discover Artificial Intelligence, 3(1), 5.

Giral, D., Hernández, C., and Salgado, C. (2021). Spectral decision in cognitive radio networks based on deep learning. Expert Systems with Applications, 180, 115080. https://doi.org/10.1016/j.eswa.2021.115080

Giral-Ramírez, D. (2022). Intelligent Fault Location Algorithms for Distributed Generation Distribution Networks: A Review. PRZEGLĄD ELEKTROTECHNICZNY, 1(7), 139–146. https://doi.org/10.15199/48.2022.07.23

Graves, A. (2012). Supervised Sequence Labelling with Recurrent Neural Networks (Vol. 385). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-24797-2

Graves, A., and Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural Networks, 18(5–6), 602–610. https://doi.org/10.1016/j.neunet.2005.06.042

He, A., Kyung Kyoon Bae, Newman, T. R., Gaeddert, J., Kyouwoong Kim, Menon, R., Morales-Tirado, L., Neel, J. J., Youping Zhao, Reed, J. H., and Tranter, W. H. (2010). A Survey of Artificial Intelligence for Cognitive Radios. IEEE Transactions on Vehicular Technology, 59(4), 1578–1592. https://doi.org/10.1109/TVT.2010.2043968

Hernández, C., López, D., and Giral, D. (2020). Modelo de decisión espectral colaborativo para mejorar el desempeño de las redes de radio cognitiva (Editorial UD (ed.); Primera Ed).

Hochreiter, S., and Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Li, L., Huang, S., Ouyang, Z., and Li, N. (2022). A Deep Learning Framework for Non-stationary Time Series Prediction. 2022 3rd International Conference on Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications (CVIDL & ICCEA), 339–342. https://doi.org/10.1109/CVIDLICCEA56201.2022.9824863

Lu, X.-Q., Tian, J., Liao, Q., Xu, Z.-W., and Gan, L. (2024). CNN-LSTM based incremental attention mechanism enabled phase-space reconstruction for chaotic time series prediction. Journal of Electronic Science and Technology, 22(2), 100256. https://doi.org/https://doi.org/10.1016/j.jnlest.2024.100256

Ma, Z., Zhang, H., and Liu, J. (2023). MM-RNN: A Multimodal RNN for Precipitation Nowcasting. IEEE Transactions on Geoscience and Remote Sensing, 61, 1–14. https://doi.org/10.1109/TGRS.2023.3264545

Myers, J. H., and Tauber, E. (2011). Market structure analysis. Marketing Classics Press.

Ochuba, N. A., Amoo, O. O., Akinrinola, O., Usman, F. O., and Okafor, E. S. (2024). Market expansion and competitive positioning in satellite telecommunications: a review of analytics-driven strategies within the global landscape. International Journal of Management & Entrepreneurship Research, 6(3), 567–581.

Prakash, S., Jalal, A. S., and Pathak, P. (2023). Forecasting COVID-19 Pandemic using Prophet, LSTM, hybrid GRU-LSTM, CNN-LSTM, Bi-LSTM and Stacked-LSTM for India. 2023 6th International Conference on Information Systems and Computer Networks (ISCON), 1–6. https://doi.org/10.1109/ISCON57294.2023.10112065

Qazi, A., and Al-Mhdawi, M. K. S. (2024). Exploring critical drivers of global innovation: A Bayesian Network perspective. Knowledge-Based Systems, 299, 112127. https://doi.org/https://doi.org/10.1016/j.knosys.2024.112127

Rithani, M., Kumar, R. P., and Doss, S. (2023). A review on big data based on deep neural network approaches. Artificial Intelligence Review, 56(12), 14765–14801.

Shi, J., Jain, M., and Narasimhan, G. (2022). Time series forecasting (tsf) using various deep learning models. ArXiv Preprint ArXiv:2204.11115.

Vardhan, B. V. S., Khedkar, M., and Thakre, P. (2022). A Comparative Analysis of Hold Out, Cross and Re-Substitution Validation in Hyper-Parameter Tuned Stochastic Short Term Load Forecasting. 2022 22nd National Power Systems Conference (NPSC), 448–453. https://doi.org/10.1109/NPSC57038.2022.10069288

Vardhan, B. V. S., Khedkar, M., Srivastava, I., Thakre, P., and Bokde, N. D. (2023). A Comparative Analysis of Hyperparameter Tuned Stochastic Short Term Load Forecasting for Power System Operator. Energies, 16(3), 1243. https://doi.org/10.3390/en16031243

Veeriah, V., Zhuang, N., and Qi, G.-J. (2015). Differential Recurrent Neural Networks for Action Recognition. 2015 IEEE International Conference on Computer Vision (ICCV), 4041–4049. https://doi.org/10.1109/ICCV.2015.460

Venegas, P. B., Porras Quispe, D. K., Bravo Apolinario, Y., and Camacho Gadea, M. J. (2022). Bank concentration, measured by different indicators. The peruvian case. Journal of Globalization, Competitiveness and Governability, 16(1). https://doi.org/10.3232/GCG.2022.V16.N1.05

Wang, G., Liu, Q., Yao, Y., and Skowron, A. (Eds.). (2003). Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing (Vol. 2639). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-39205-X

Wang, X., and Zhang, Y. (2020). Multi-step-ahead time series prediction method with stacking LSTM neural network. 2020 3rd International Conference on Artificial Intelligence and Big Data (ICAIBD), 51–55.

Wu, J., and He, Y. (2021). Prediction of GDP in Time Series Data Based on Neural Network Model. 2021 IEEE International Conference on Artificial Intelligence and Industrial Design (AIID), 20–23. https://doi.org/10.1109/AIID51893.2021.9456509