Application of LSTM Neural Networks and Complex Network Analysis for Earthquake Forecasting in Colombia

Santiago Giraldo[1*], Daniel León[2†] and Victor Bucheli[3†]

[1,2,3*]Escuela de Ingeniería de Sistemas, Universidad del Valle, 13th st., Cali, 760033, Valle del Cauca, Colombia.

*Corresponding author(s). E-mail(s):
giraldo.santiago@correounivalle.edu.co.com;
Contributing authors: daniel.leon@correounivalle.edu.co;
victor.bucheli@correounivalle.edu.co;
[†]These authors contributed equally to this work.

Abstract

Earthquakes are unexpected and high risk events that may cause human casualties and damage to infrastructures. Over the last years, there's been an increasing amount of earthquake forecasting research in machine learning and neural networks with astonishing results. This work approaches the earthquake forecasting problem with an application of Long Short-Term Memory (LSTM) Neural Networks to historical seismic events in Colombia's pacific shore retrieved from the United States Geological Survey service. We analyzed the seismic catalog over the time, included complex networks features for each set, and trained multiple LSTM neural networks with hyper-parameter tuning. We also used features like latitude, longitude, magnitude, days elapsed between events and complex networks properties. We found models with R^2 score greater than 98% when targeting multiple features at once, and greater than 99% when targeting one feature.

Keywords: Earthquake Forecasting, LSTM Neural Networks, Machine Learning, Complex Networks

1 Introduction

Earthquakes are classified a high risk event due to their unexpected nature and the limited time available to take preventive actions. Furthermore, the earthquakes may cause the human casualties and infrastructure damage. Over the past few years, numerous papers have explored the use of machine learning algorithms and neural networks for earthquake forecasting [1–6]. The growing interest in the use of machine learning has been documented, as well as the use of artificial intelligence to forecast seismic events, indicating future work and various challenges to be addressed.

One approach involves the application of deep neural networks to learn from data directly to identify complex patterns, though, the forecasting of features such as magnitude, is limited by the lack of events that exceed certain threshold [REFERENCIAS]. The most interesting approach, uses a set of different characteristics and regions as input to the neural network in a moving window of time [7].

Some of the documented features in literature are magnitude, frequency of events [8–12], occurrence of aftershocks [13], probability of detection and rate of false alarms [14] , geomagnetic field variation, soil temperature and relative humidity [15], seismic indicators [16], location [17], Gutenberg-Richter b-value [18–20], gravity, georesistivity and water level [21], elapsed days, mean magnitude, mean time between characteristic events, coefficient of variation of the mean time [22], and electronic signals captured on seismographs [23–28].

Artificial intelligence and machine learning algorithms are powerful tools to address a wide set of classification and forecasting problems. They have the potential to learn from hidden patterns within the provided data [6]. As seen in [6, 23] their application in earthquake forecasting is useful for risk assessment and early prevention.

In [4] is stated that studies in seismology are carried out for the following objectives: actions to reduce long-term disasters generated by earthquakes, preparation and adjustment to disasters, disaster response strategies and post-disaster recovery planning. To Satriano et al [29], information systems focused on the early warning of seismological events must quickly and reliably provide information on different parameters such as location, time, and size in order to improve the emergency response, and Espinosa et al [30] mention that the Mexican SAS (seismic alert system), an early warning system, implemented in Mexico has allowed seismic waves to be detected up to 60 seconds in advance, a time frame that allows security procedures to be executed to protect equipment susceptible to damage such as power plants, computer systems, and communication networks.

Additionally, there has been an increasing interest to represent data through graphs, as complex relationships can be modeled using nodes and edges [31], we can model earthquakes as a complex network representing the location as nodes and the sequence as edges [32]. Learning from these structures has been successfully applied in particle interaction, recommendation systems, molecular properties, natural language processing, and others [33].

As stated by León et al [32] seismological events as complex networks allows to study the seismicity in a region and unravel the relationship among these events. The earthquake networks are based on the temporal sequence of events and their specific location, different network features like path length, degree distribution, and

clustering lead to define high relation between events, hubs and small-worldness network properties. Thus, integrating the analysis of complex networks to structure the relationship between seismic events seems to be helpful to improve forecast results.

In this document we processed Colombian earthquake data and trained various neural network models with LSTM layers. We found a LSTM-NN with X layers and activation functions, this model displays a R2 MSE MAE scores of ZZZ/WWW, proposed models compares training with earthquake features alone, and the addition of complex network features. Finding out the addition of complex network features IMPROVE/WORSEN the forecast, results are promising and offer XXXXX at seismic events in the region

2 Methods

In this section, we take a look at the training and test data, and the steps to generate the earthquake features used later for training.

The earthquake events used in this study were obtained from the U.S. Geological Survey (USGS) for the Colombian region, with latitudes ranging from -0.132° to 9.796° and longitudes from -80.343° to -72.466°. The events are plotted by magnitude and depth in **Figure 1**. The retrieved dataset contained 5462 events since 1975, which was reduced to 5025 events after processing. The data cleaning process involved removing empty values, converting string values to numerical values for certain features, and filtering out events that were not validated by multiple sources, had missing data points, or had zero values. The selected events were of the mb type, representing the body wave magnitude, and included the depth of the event.
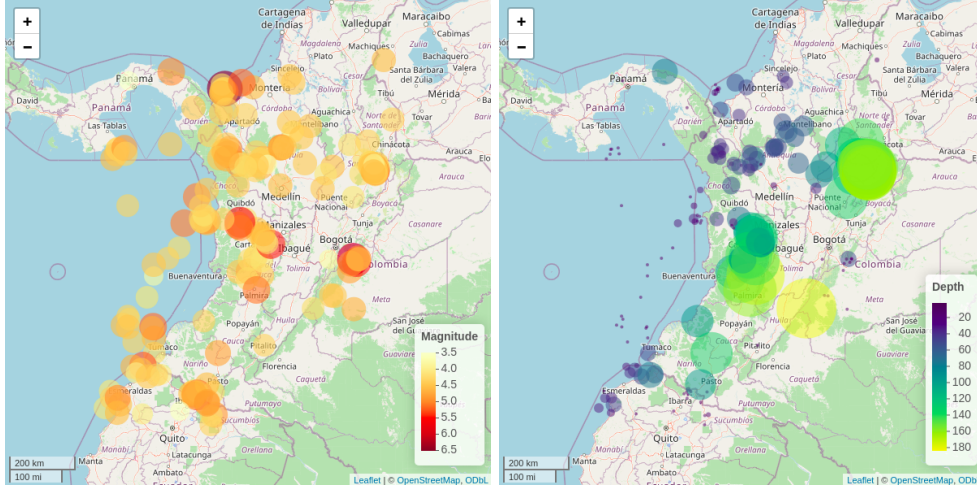

Fig. 1 Earthquake events by Magnitude (left) and Depth (right)

Table 1 Caption text

Feature	Min	Max	Median	Mean	SD	Variance	Range
Latitude	-0.13	9.79	6.73	5.73	2.08	4.31	9.92
Longitude	-80.34	-72.47	-75.68	-75.55	2.49	6.20	7.88
Depth	0.00	436.50	68.10	89.68	65.50	4290.78	436.50
Magnitude	0.00	7.80	4.40	4.19	0.99	0.97	7.80

Source: This is an example of table footnote. This is an example of table footnote.

[1]Example for a first table footnote. This is an example of table footnote.

[2]Example for a second table footnote. This is an example of table footnote.

2.1 Earthquake Networks

Following the methodology proposed by Leon et al. [32], constructing an Earthquake Complex Network involves dividing the selected geographical area into a grid of cells, each representing a specific distance in kilometers. Sequential earthquake events are then used to establish connections between these cells, forming the network.
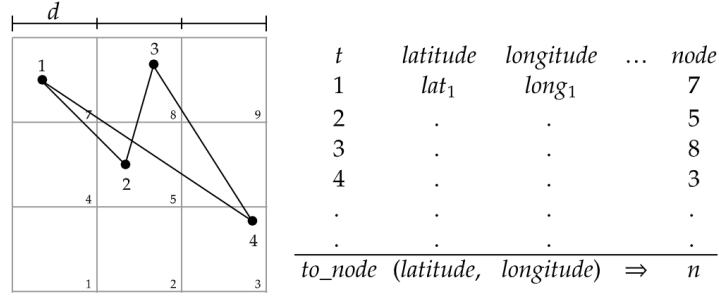

Fig. 2 Illustration of the creation of an earthquake network. **a)** A geographical region is divided into cells of equal size, with earthquakes represented as black dots. Each number identifies an event by its order in the time sequence, and a cell is selected if an earthquake occurs within it. **b)** The selected cells are used as nodes in the network, connected according to the time sequence. The brackets near each link indicate the events that caused the connection.

2.2 Dataset

After we had the processed and cleansed the data, we proceeded to tag each one of the events with their corresponding node. After this, we normalized all the features using scikit-learn StandardScaler.

Following the methodology proposed by Lin et al. [34], multivariate time-series forecasting involves using a rolling window approach. In this context, the input to the model is a sequence of values $\{y_1, y_2, \ldots, y_l\}$, and the output is the value y_{l+l}, where l

represents the lookback number of events. We applied this approach to the earthquake data, using the features within each window as inputs to the neural network and the selected features of the subsequent event as the target.

The sequences were explored according to the lookback parameter during, we generated different matrices of size (S, L) resulting in a dataset of size (N, S, L). The following figure shows the structure of the dataset for the LSTM network.

$$
W_{t:\,1 \to N}
\begin{bmatrix}
F_{1,t} & F_{1,t+1} & . & . & F_{1,t+L} \\
. & . & . & . & . \\
. & . & . & . & . \\
F_{S,t} & F_{S,t+1} & . & . & F_{S,t+L}
\end{bmatrix}
\Rrightarrow
\begin{bmatrix} C_{t+L+1} \end{bmatrix}
$$

$$\qquad\qquad\qquad \textit{input} \qquad\qquad\qquad\qquad \textit{target}$$

Fig. 3 Dataset for the LSTM neural network.

To calculate the complex network properties, we took each moving window and built the respective graph, taking cell distance values of 10, 50, and 100 kilometers, per each event sequence of 50 and 100 events. The networks features we included in the moving windows were degree centrality, clustering, betweenness centrality, closeness centrality, and page rank.

3 Results

Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text. Sample body text.

4 Discussion

Discussions should be brief and focused. In some disciplines use of Discussion or 'Conclusion' is interchangeable. It is not mandatory to use both. Some journals prefer a section 'Results and Discussion' followed by a section 'Conclusion'. Please refer to Journal-level guidance for any specific requirements.

5 Conclusion

Conclusions may be used to restate your hypothesis or research question, restate your major findings, explain the relevance and the added value of your work, highlight any limitations of your study, describe future directions for research and recommendations.

In some disciplines use of Discussion or 'Conclusion' is interchangeable. It is not mandatory to use both. Please refer to Journal-level guidance for any specific requirements.

Supplementary information. If your article has accompanying supplementary file/s please state so here.

Authors reporting data from electrophoretic gels and blots should supply the full unprocessed scans for key as part of their Supplementary information. This may be requested by the editorial team/s if it is missing.

Please refer to Journal-level guidance for any specific requirements.

Acknowledgements. Acknowledgements are not compulsory. Where included they should be brief. Grant or contribution numbers may be acknowledged.

Please refer to Journal-level guidance for any specific requirements.

Declarations

Some journals require declarations to be submitted in a standardised format. Please check the Instructions for Authors of the journal to which you are submitting to see if you need to complete this section. If yes, your manuscript must contain the following sections under the heading 'Declarations':

- Funding
- Conflict of interest/Competing interests (check journal-specific guidelines for which heading to use)
- Ethics approval and consent to participate
- Consent for publication
- Data availability
- Materials availability
- Code availability
- Author contribution

If any of the sections are not relevant to your manuscript, please include the heading and write 'Not applicable' for that section.

Editorial Policies for:

Springer journals and proceedings: https://www.springer.com/gp/editorial-policies

Nature Portfolio journals:
https://www.nature.com/nature-research/editorial-policies

Scientific Reports: https://www.nature.com/srep/journal-policies/editorial-policies

BMC journals: https://www.biomedcentral.com/getpublished/editorial-policies

Appendix A   Section title of first appendix

An appendix contains supplementary information that is not an essential part of the text itself but which may be helpful in providing a more comprehensive understanding of the research problem or it is information that is too cumbersome to be included in the body of the paper.

References

[1] Banna, M.H.A., Taher, K.A., Kaiser, M.S., Mahmud, M., Rahman, M.S., Hosen, A.S.M.S., Cho, G.H.: Application of artificial intelligence in predicting earthquakes: State-of-the-art and future challenges **8**, 192880–192923 https://doi.org/10.1109/ACCESS.2020.3029859 . Conference Name: IEEE Access

[2] Florido, E., Aznarte, J.L., Morales-Esteban, A., Martínez-Álvarez, F.: Earthquake magnitude prediction based on artificial neural networks: A survey, 159–169. Accessed 2020-11-11

[3] Mignan, A., Broccardo, M.: Neural network applications in earthquake prediction (1994–2019): Meta-analytic and statistical insights on their limitations **91**(4), 2330–2342 https://doi.org/10.1785/0220200021 . Accessed 2020-10-28

[4] Jiao, P., Alavi, A.H.: Artificial intelligence in seismology: Advent, performance and future trends **11**(3), 739–744 https://doi.org/10.1016/j.gsf.2019.10.004 . Accessed 2020-11-25

[5] Galkina, A., Grafeeva, N.: Machine learning methods for earthquake prediction: a survey, 8

[6] Kong, Q., Trugman, D.T., Ross, Z.E., Bianco, M.J., Meade, B.J., Gerstoft, P.: Machine learning in seismology: Turning data into insights **90**(1), 3–14 https://doi.org/10.1785/0220180259 . Accessed 2020-11-11

[7] Wang, Q., Guo, Y., Yu, L., Li, P.: Earthquake prediction based on spatiotemporal data mining: An LSTM network approach **8**(1), 148–158 https://doi.org/10.1109/TETC.2017.2699169 . Conference Name: IEEE Transactions on Emerging Topics in Computing

[8] Han, Z.-Q., Wang, B.-Q.: The medium- and short-term prediction methods of strong earthquakes based on neural network **10**(4), 459–467 https://doi.org/10.1007/s11589-997-0054-7

[9] De Falco, I., Iazzetta, A., Luongo, G., Mazzarella, A., Tarantino, E.: The seismicity in the southern tyrrhenian area and its neural forecasting https://doi.org/10.1007/s000240050003

[10] Bodri, B.: A neural-network model for earthquake occurrence https://doi.org/

10.1016/S0264-3707(01)00039-4

[11] Wang, W., Liu, Y., Li, G., Wu, G., Ma, Q., Zhao, L., Lin, M.: Support vector machine method for forecasting future strong earth-quakes in chinese mainland

[12] Sri Lakshmi, S., Tiwari, R.K.: Model dissection from earthquake time series: A comparative analysis using modern non-linear forecasting and artificial neural network approaches **35**(2), 191–204 https://doi.org/10.1016/j.cageo.2007.11.011

[13] Lin, F.C., Elhassan, N., Hassan, A., Yousif, A.: Forecast of seismic aftershocks using a neural network. In: ICONIP 2002 - Proceedings of the 9th International Conference on Neural Information Processing: Computational Intelligence for the E-Age, vol. 4, pp. 1796–1799. https://doi.org/10.1109/ICONIP.2002.1198983 . https://www.scopus.com/inward/record.uri?eid=2-s2.0-84966264019&doi=10.1109%2fICONIP.2002.1198983&partnerID=40&md5=a712499b22454157916039bab9fa0499

[14] Panakkat, A., Adeli, H.: Neural network models for earthquake magnitude prediction using multiple seismicity indicators **17**(1), 13–33 https://doi.org/10.1142/S0129065707000890

[15] Suratgar, A.A., Setoudeh, F., Salemi, A.H., Negarestani, A.: Magnitude of earthquake prediction using neural network. In: Proceedings - 4th International Conference on Natural Computation, ICNC 2008, vol. 2, pp. 448–452. https://doi.org/10.1109/ICNC.2008.781 . https://www.scopus.com/inward/record.uri?eid=2-s2.0-57649133274&doi=10.1109%2fICNC.2008.781&partnerID=40&md5=cbdf23cc8c8534627d064cfa48163706

[16] Panakkat, A., Adeli, H.: Recurrent neural network for approximate earthquake time and location prediction using multiple seismicity indicators **24**(4), 280–292 https://doi.org/10.1111/j.1467-8667.2009.00595.x

[17] Veri, J., Wah, T.Y.: Earthquake prediction based on the pattern of points seismic motion. In: 2012 International Conference on Advanced Computer Science Applications and Technologies (ACSAT), pp. 209–212. https://doi.org/10.1109/ACSAT.2012.67

[18] Morales-Esteban, A., Martínez-Álvarez, F., Reyes, J.: Earthquake prediction in seismogenic areas of the iberian peninsula based on computational intelligence https://doi.org/10.1016/j.tecto.2013.02.036

[19] Reyes, J., Morales-Esteban, A., Martínez-Álvarez, F.: Neural networks to predict earthquakes in chile **13**(2), 1314–1328 https://doi.org/10.1016/j.asoc.2012.10.014 . Accessed 2020-09-23

[20] Asencio–Cortés, G., Morales–Esteban, A., Shang, X., Martínez–Álvarez, F.: Earthquake prediction in california using regression algorithms and cloud-based

big data infrastructure **115**, 198–210 https://doi.org/10.1016/j.cageo.2017.10.011 . Accessed 2020-09-30

[21] Cai, Y., Shyu, M.-L., Tu, Y.-X., Teng, Y.-T., Hu, X.-X.: Anomaly detection of earthquake precursor data using long short-term memory networks https://doi.org/10.1007/s11770-019-0774-1

[22] Banna, M.H.A., Ghosh, T., Nahian, M.J.A., Taher, K.A., Kaiser, M.S., Mahmud, M., Hossain, M.S., Andersson, K.: Attention-based bi-directional long-short term memory network for earthquake prediction **9**, 56589–56603 https://doi.org/10.1109/ACCESS.2021.3071400

[23] Böse, M., Wenzel, F., Erdik, M.: PreSEIS: A neural network-based approach to earthquake early warning for finite faults https://doi.org/10.1785/0120070002

[24] Lakkos, S., Hadjiprocopis, A., Comley, R., Smith, P.: Neural network scheme for earthquake prediction based on the seismic electric signals. In: Neural Networks for Signal Processing - Proceedings of the IEEE Workshop, pp. 681–689. https://www.scopus.com/inward/record.uri?eid=2-s2.0-0028699912&partnerID=40&md5=46b2b75c6754a02f06e4362b9f98b247

[25] Rovithakis, G.A., Vallianatos, F.: A neural network approach to the identification of electric earthquake precursors https://doi.org/10.1016/S1464-1895(00)00049-1

[26] Ifantis, A., Papadimitriou, S.: Support vector identification of seismic electric signals https://doi.org/10.1142/S0218001403002484

[27] Radeva, S., Radev, D.: Real-time model for earthquake prediction. In: Improvement of Buildings' Structural Quality by New Technologies - Proceedings of the Final Conference of COST Action C12. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84857007188&partnerID=40&md5=92165af4a61858a05156172de7122e29

[28] Bernardo-Torres, A., Gómez-Gil, P.: One-step forecasting of seismograms using multi-layer perceptrons. In: 2009 6th International Conference on Electrical Engineering, Computing Science and Automatic Control, CCE 2009. https://doi.org/10.1109/ICEEE.2009.5393349 . https://www.scopus.com/inward/record.uri?eid=2-s2.0-77949822759&doi=10.1109%2fICEEE.2009.5393349&partnerID=40&md5=93eab0868d763b15666643ea38f123ef

[29] Satriano, C., Wu, Y.-M., Zollo, A., Kanamori, H.: Earthquake early warning: Concepts, methods and physical grounds **31**(2), 106–118 https://doi.org/10.1016/j.soildyn.2010.07.007 . Accessed 2021-11-04

[30] Espinosa-Aranda, J.M., Cuellar, A., Garcia, A., Ibarrola, G., Islas, R., Maldonado, S., Rodriguez, F.H.: Evolution of the mexican seismic alert system

(SASMEX) **80**(5), 694–706 https://doi.org/10.1785/gssrl.80.5.694 . Accessed 2021-11-04

[31] Chami, I., Abu-El-Haija, S., Perozzi, B., Ré, C., Murphy, K.: Machine learning on graphs: A model and comprehensive taxonomy 2005.03675. Accessed 2022-01-23

[32] León, D.A., Valdivia, J.A., Bucheli, V.A.: Modeling of colombian seismicity as small-world networks **89**(5), 1807–1816 https://doi.org/10.1785/0220180076 . Publisher: GeoScienceWorld. Accessed 2020-11-11

[33] Grattarola, D., Alippi, C.: Graph neural networks in TensorFlow and keras with spektral 2006.12138. Accessed 2020-11-12

[34] Lai, G., Chang, W.-C., Yang, Y., Liu, H.: Modeling long- and short-term temporal patterns with deep neural networks 1703.07015. Accessed 2021-05-17