FERNUNIVERSITÄT IN HAGEN

APPLIED STATISTICS

Bachelor Thesis

Advanced Architectures in LSTM Networks for Effective Management of Temporal Data

Luca Zangari

" Science knows no country, because knowledge belongs to humanity, and is the torch which illuminates the world. Science is the highest personification of the nation because that nation will remain the first which carries the furthest the works of thought and intelligence."

Louis Pasteur

Acknowledgments

Abstract

Contents

Artificial intelligence is not a substitute for human intelligence; it is a tool to amplify human creativity and ingenuity.
— Fei-Fei Li

1. Introduction

Accurate temporal data forecasting, a task with broad applications, is a complex and multidimensional problem. From energy to finance, healthcare to advertisement, the effective management of this data is essential for optimizing operations[1]. In reality, predictions of long sequential time-series data present challenges due to inherent complexities and the dynamic nature of its information[1, 2, 3]. Dependencies can span from short-term daily fluctuations to long-term seasonal trends, complicating the forecasting process as capturing all relevant information becomes difficult. Non-stationary, where statistical properties such as mean and variance change over time, is an additional hurdle, necessitating models adaptable to temporal changes and must address irregular events and noise to maintain accuracy[4]. The interdependencies among variables further compound the complexity, creating a polymorph and multidimensional problem. With advancements in machine learning (ML), the ability to forecast temporal data has seen notable improvements despite the aforementioned hindrances[5].

Historically, the first applicable research efforts of time-series forecasting have been AutoRegressive Integrated Moving Average (ARIMA) models[6, 7], which assumed linear relationships within the data that simplified the prediction process. However, ARIMA and its variants, such as Seasonal ARIMA (SARIMA)[8, 9], often fell short when dealing with intricate, nonlinear relationships naturally found in many real-world temporal datasets. To this front, Aritificial Neural Networks (ANNs) have become increasingly applicable[10], chiefly because the underlying assumption here is the one of non-linearity. These networks have proven particularly useful in extracting information when handed large amounts of data in fields such as image and speech recognition, natural language processing, and financial forecasting. Among neural networks, Recurrent

Neural Networks (RNNs) stands out due to their ability to handle sequential data, where the output at a given time step relies not just on the current input but also on previous information. Using backpropagation through time for training, RNNs utilize their memory capability derived from the feedback loops; this makes them well-suited for applications in sequential time series prediction tasks[11], where order of output and context are of non-negligible importance. However, RNNs tend to struggle with learning dependencies in the long term, an issue which has been mitigated by the introductions of extension models known as long short term memory (LSTM)[12] to address the issue of gradient flow behavior in[rnn]. LSTM outperforms traditional methods like ARIMA in handling complex, nonlinear relationships within data, as evidenced by the work of Kong et al.[13] on the forecasting of electrical grid load. The ability of LSTMto remember information over extended periods makes them heterogeneously applicable: they enhance text classification, language modeling, and environmental data forecasting[14], improve prediction of production efficiency[15] and support cybersecurity for attack detection in IoT environments[16], to name a few.

Integration of advanced architectures such as attention mechanisms and sequence-to-sequence (Seq-to-Seq) models with LSTM-based networks enhances their performance in handling sequential data even further in a variety of tasks, for instance, machine translation and time series forecasting, while the attention mechanisms additionally enable the model to focus on the pertinent parts of the input sequence during decoding. These qualities are an opportunity for applications in accurate energy usage prediction. This is especially important considering the increasing global population and urbanization (4.5 Billion of urban citizens in 2023 compared to the 2.28 billion in 1990[17]), which surged the demand for reliable electricity supply has surged[10, 18, 9], necessitating accurate load forecasting to maintain system stability amidst the integration of renewable energy sources. The selective focus of attention mechanisms in LSTM leads to more accurate predictions and allows for better resource allocation and load balancing in power grids[19].

This research endeavors to illustrate the superiority of LSTM networks in managing long-range temporal dependencies and their practical application in predicting imbalances in German electricity generation and consumption. The study considers the LSTM model as its baseline for forecasting these imbalances with a horizon of one day ahead, including 24 hours, one hour, and 15 minutes. By utilizing an advanced architecture, namely an attention-based Seq-to-Seq model, as a benchmark, this research aims to advance the understanding and applications of LSTM networks in addressing real-world energy sector challenges.

A. Appendix*

* All Figures were created with BioRender.com

List of Figures

Abbreviations

ANNs Aritificial Neural Networks.

ARIMA AutoRegressive Integrated Moving Average.

LSTM long short term memory.

ML machine learning.

RNNs Recurrent Neural Networks.

SARIMA Seasonal ARIMA.

Seq-to-Seq sequence-to-sequence.

Bibliography

[1] B. Lim and S. Zohren. "Time-series forecasting with deep learning: a survey". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 379.2194 (Feb. 2021), p. 20200209. ISSN: 1471-2962. DOI: 10.1098/rsta.2020.0209.

[2] J. F. Torres, D. Hadjout, A. Sebaa, F. Martínez-Álvarez, and A. Troncoso. "Deep Learning for Time Series Forecasting: A Survey". In: *Big Data* 9.1 (Feb. 2021), pp. 3–21. ISSN: 2167-647X. DOI: 10.1089/big.2020.0159.

[3] J. G. De Gooijer and R. J. Hyndman. "25 years of time series forecasting". In: *International Journal of Forecasting* 22.3 (Jan. 2006), pp. 443–473. ISSN: 0169-2070. DOI: 10.1016/j.ijforecast.2006.01.001.

[4] C. Cheng, A. Sa-Ngasoongsong, O. Beyca, T. Le, H. Yang, Z. (Kong, and S. T. Bukkapatnam. "Time series forecasting for nonlinear and non-stationary processes: a review and comparative study". In: *IIE Transactions* 47.10 (Jan. 2015), pp. 1053–1071. ISSN: 1545-8830. DOI: 10.1080/0740817x.2014.999180.

[5] I. Goodfellow, Y. Bengio, and A. Courville. *Deep learning*. Ed. by A. Courville and Y. Bengio. Adaptive computation and machine learning. Includes bibliographical references and index. Cambridge, Massachusetts: The MIT Press, 2016. 1775 pp. ISBN: 9780262337373.

[6] J. Contreras, R. Espinola, F. Nogales, and A. Conejo. "ARIMA models to predict next-day electricity prices". In: *IEEE Transactions on Power Systems* 18.3 (Aug. 2003), pp. 1014–1020. ISSN: 0885-8950. DOI: 10.1109/tpwrs.2002.804943.

[7] G. Zhang. "Time series forecasting using a hybrid ARIMA and neural network model". In: *Neurocomputing* 50 (Jan. 2003), pp. 159–175. ISSN: 0925-2312. DOI: 10.1016/s0925-2312(01)00702-0.

[8] B. M. Williams and L. A. Hoel. "Modeling and Forecasting Vehicular Traffic Flow as a Seasonal ARIMA Process: Theoretical Basis and Empirical Results". In: *Journal of transportation engineering* 129.6 (Nov. 2003), pp. 664–672. ISSN: 1943-5436. DOI: 10.1061/(asce)0733-947x(2003)129:6(664).

[9] M. Gürtler and T. Paulsen. "Forecasting performance of time series models on electricity spot markets: a quasi-meta-analysis". In: *International Journal of Energy Sector Management* 12.1 (Dec. 2017), pp. 103–129. ISSN: 1750-6220. DOI: 10.1108/ijesm-06-2017-0004.

[10] T. Ahmad and D. Zhang. "A critical review of comparative global historical energy consumption and future demand: The story told so far". In: *Energy Reports* 6 (Nov. 2020), pp. 1973–1991. ISSN: 2352-4847. DOI: 10.1016/j.egyr.2020.07.020.

[11] W. Yin, K. Kann, M. Yu, and H. Schütze. *Comparative Study of CNN and RNN for Natural Language Processing*. 2017. DOI: 10.48550/ARXIV.1702.01923.

[12] S. Hochreiter and J. Schmidhuber. "Long Short-Term Memory". In: *Neural Computation* 9.8 (Nov. 1997), pp. 1735–1780. ISSN: 1530-888X. DOI: 10.1162/neco.1997.9.8.1735.

[13] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang. "Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network". In: *IEEE Transactions on Smart Grid* 10.1 (2019), pp. 841–851. DOI: 10.1109/TSG.2017.2753802.

[14] K. Yokoo, K. Ishida, A. Ercan, T. Tu, T. Nagasato, M. Kiyama, and M. Amagasaki. "Capabilities of deep learning models on learning physical relationships: Case of rainfall-runoff modeling with LSTM". In: *Science of The Total Environment* 802 (Jan. 2022), p. 149876. ISSN: 0048-9697. DOI: 10.1016/j.scitotenv.2021.149876.

[15] Z. Li, H. Guo, A. V. Barenji, W. M. Wang, Y. Guan, and G. Q. Huang. "A sustainable production capability evaluation mechanism based on blockchain, LSTM, analytic hierarchy process for supply chain network". In: *International Journal of Production Research* 58.24 (2020), pp. 7399–7419.

DOI: 10.1080/00207543.2020.1740342. eprint: https://doi.org/10.1080/00207543.2020.1740342. URL: https://doi.org/10.1080/00207543.2020.1740342.

[16] A. Diro and N. Chilamkurti. "Leveraging LSTM Networks for Attack Detection in Fog-to-Things Communications". In: *IEEE Communications Magazine* 56.9 (2018), pp. 124–130. DOI: 10.1109/MCOM.2018.1701270.

[17] W. Bank. "World Development Indicators: Urban population". In: (2023). URL: https://databank.worldbank.org/reports.aspx?dsid=2%5C&series=SP.URB.TOTL.

[18] G. Plessmann, M. Erdmann, M. Hlusiak, and C. Breyer. "Global Energy Storage Demand for a 100% Renewable Electricity Supply". In: *Energy Procedia* 46 (2014), pp. 22–31. ISSN: 1876-6102. DOI: 10.1016/j.egypro.2014.01.154.

[19] G. Zhang, X. Bai, and Y. Wang. "Short-time multi-energy load forecasting method based on CNN-Seq2Seq model with attention mechanism". In: *Machine Learning with Applications* 5 (Sept. 2021), p. 100064. ISSN: 2666-8270. DOI: 10.1016/j.mlwa.2021.100064.