Exploring LSTMs and RNNs in the Context of Machine Learning Theory and Their Application in Weather Prediction

Sarwan Shah

12th March, 2024

Team Member Details

Name	Program	ASU ID
Sarwan Shah	Master of Electrical Engineering	1230055790

Table 1: Member Details Table

Abstract

Long Short-Term Memory (LSTM) networks and Recurrent Neural Networks (RNNs) are fundamental architectures in machine learning, particularly when it comes to sequential data analysis. Through this project I aim to explore the mathematical theory and modelling of LSTMs and RNNs, focusing on their theoretical foundations and practical implications, while also exploring their applications in context of weather prediction and meteorology *(research interests of mine)*.

I believe understanding the mathematics and modelling behind these architectures is crucial for gaining insights into their functioning, optimizing their performance, and working them in the future. The significance of this project is that it enables me to expand my understanding of this topic so that I may apply it to weather and climate research, which is of critical importance to our collective future.

Primary Papers to Read

- Aggarwal, C. C. (2023). Neural Networks and deep learning a textbook. Springer International Publishing AG.
- Goodfellow, Ian, et al. "Deep Learning." Deeplearningbook.org, 2016, www.deeplearningbook.org/.
- Hochreiter, Sepp, and Jürgen Schmidhuber. "Long Short-Term Memory." Neural Computation, vol. 9, no. 8, Nov. 1997, pp. 1735–1780.
- Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-kin Wong, and Wang-chun Woo. 2015. Convolutional LSTM Network: a machine learning approach for precipitation nowcasting. In Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1 (NIPS'15). MIT Press.
- Xingjian Shi, Zhihan Gao, Leonard Lausen, Hao Wang, Dit-Yan Yeung, Wai-kin Wong, & Wang-chun Woo. (2017). Deep Learning for Precipitation Nowcasting: A Benchmark and A New Model.
- Hou, L., Zhu, J., Kwok, J., Gao, F., Qin, T., & Liu, T.Y. (2019). Normalization Helps Training of Quantized LSTM. In Advances in Neural Information Processing Systems. Curran Associates, Inc..
- Jiahao Su, Wonmin Byeon, Furong Huang, Jan Kautz, & Animashree Anandkumar (2020). Convolutional Tensor-Train LSTM for Spatio-temporal Learning. CoRR, abs/2002.09131.

Expected Outcome

Through this project I expect to develop an understanding of the mathematical principles underlying LSTMs and RNNs, including their architectures, training algorithms, and optimization techniques. Furthermore, I anticipate gaining insights into recent advancements in the field and their applications in weather prediction. The outcome of which will be a review paper that summarizes existing literature, identifies key challenges, and proposes potential avenues for future research in this domain.