



Fake News Challenge Stage 1 (Fnc-1): Stance Detection

Sushmitha Suresh, Department of Management Sciences,
University Of Waterloo



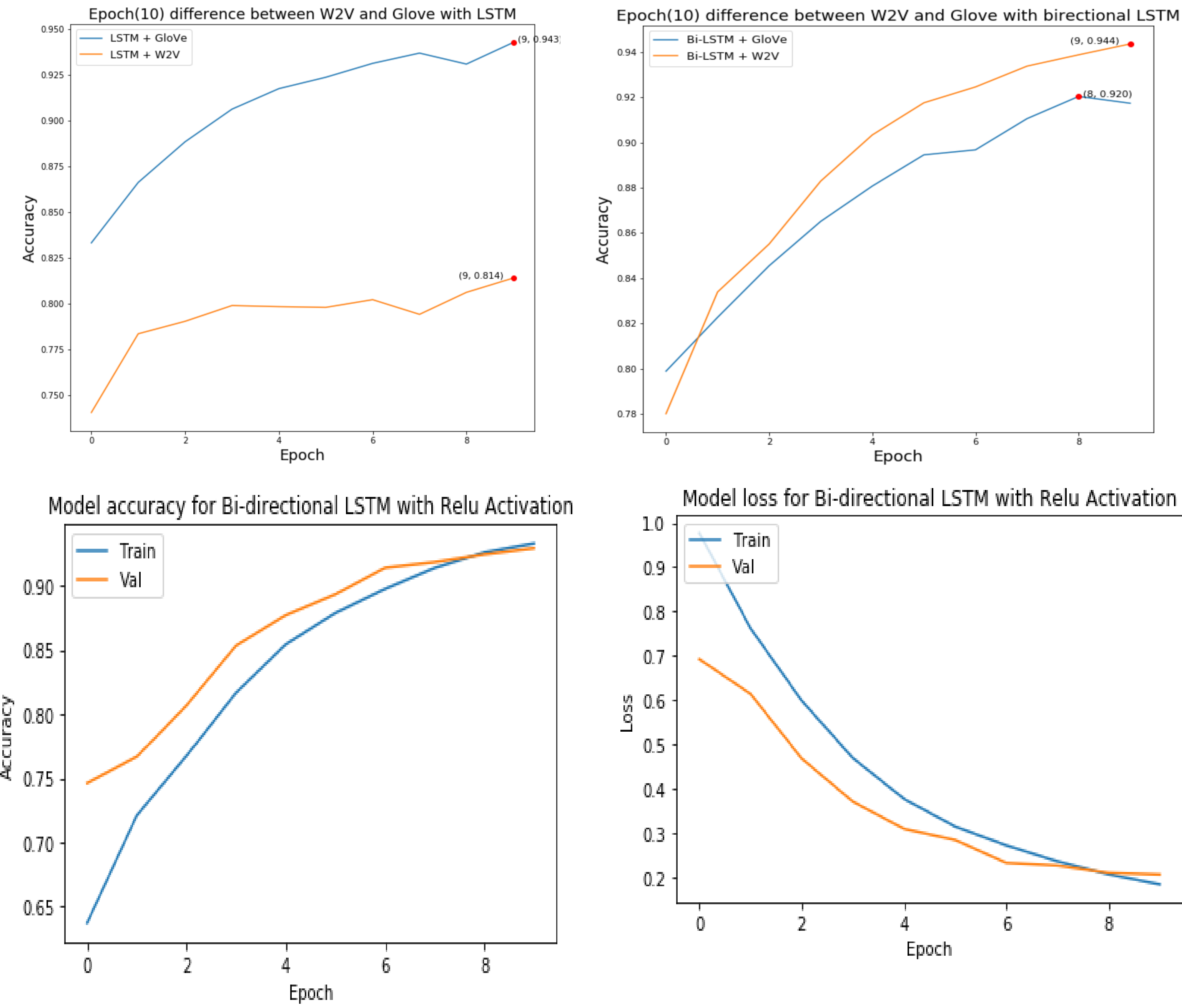
Abstract:

The goal of the Fake News Challenge is to explore how artificial intelligence technologies, particularly machine learning and natural language processing, might be leveraged to combat the fake news problem. Assessing the veracity of a news story is a complex and cumbersome task, even for trained experts. Fortunately, the process can be broken down into steps or stages. A helpful first step towards identifying fake news is to understand what other news organizations are saying about the topic. We believe automating this process, called Stance Detection, could serve as a useful building block in an AI-assisted fact-checking pipeline. So stage #1 of the Fake News Challenge (FNC-1) focuses on the task of Stance Detection.

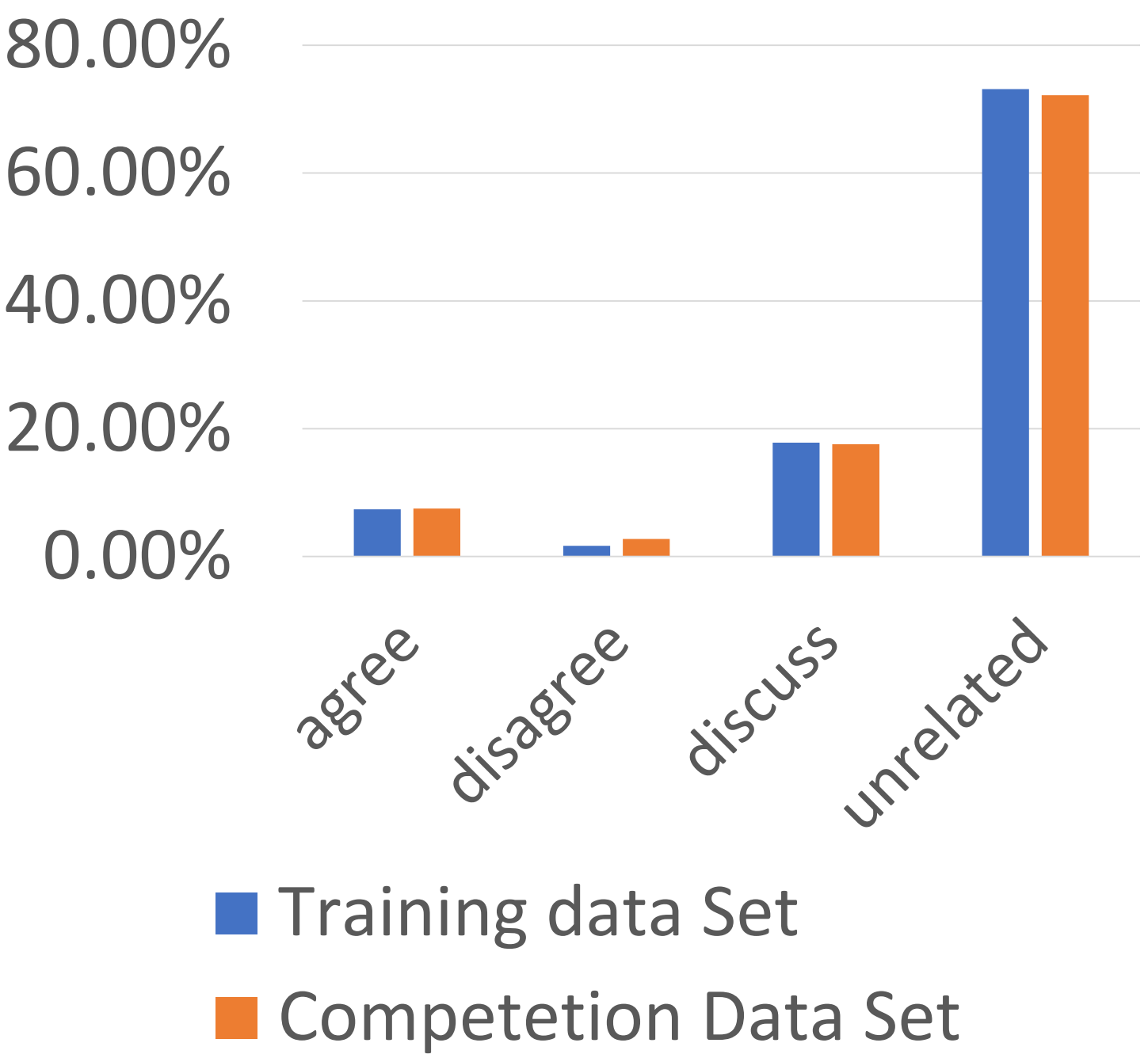
Architecture:



Performance:



Stance Distribution:



Qualitative Analysis:

Classifiers	Accuracy in %
LogisticRegression	85
RandomForestClassifier	77
MultinomialNB	56
W2V LSTM	85
GloVe LSTM	85
W2V Bi-LSTM	86
GloVe Bi-LSTM	96

Results:

Using bidirectional will run the inputs in two ways, one from past to future and one from future to past. The LSTM that runs backwards preserve information from the future and using the hidden states combined we can preserve both past and future. Hence BiLSTMs show very good results as they can understand context better. The graph shows that there is a constant increase and decrease in accuracy and loss throughout the epochs. The validation accuracy of LSTM and BiLSTM is almost the same hitting around 90%, however, BiLSTM performed better on the competition dataset.

Future Work:

In addition to LSTM and BiLSTM, there are two other techniques that can possibly perform better on the unknown data. While using Attention is believed to increase the accuracy, it is very unlikely to incorporate the mechanism in a classification task. Selecting states to attend to is more appropriate for translation tasks. Hence the Attention mechanism is only stated as a suggestion and not implemented as a part of this project. Second approach is to use conditional encoding, one LSTM to encode the headlines and another LSTM to encode the bodies. Finally, the last output vector of the LSTM of bodies is used to predict the stance of the headline-body pairs. The second technique is yet to be implemented in this project and is expected to be done by the end of the final report

References:

- Isabelle Augenstein, Tim Rocktaschel, Andreas Vlachos, and Kalina Bontcheva. 2016. Stance Detection with Bidirectional Conditional Encoding. In Proceedings of EMLNP.
- Benjamin Riedel, Isabelle Augenstein, Georgios P Spithourakis, and Sebastian Riedel. 2017. A simple but tough-to-beat baseline for the Fake News Challenge stance detection task. ArXiv:1707.03264.
- Andreas Hanselowski, Avinesh PVS, Benjamin Schiller, Felix Caspelherr, Debanjan Chaudhuri, Christian M. Meyer, and Iryna Gurevych. 2018. A retrospective analysis of the fake news challenge stance-detection task. In Proceedings of the 27th International Conference on Computational Linguistics, COLING '18, pages 1859–1874, Santa Fe, NM, USA