Fake News Detection Using Deep Learning

Namratha Shah[1], Sneha Mhaske[1], Nat Hawkins[2]
1. Department of Computer Science and Engineering, Michigan State University, East Lansing, MI, USA
2. Department of Computational Mathematics, Science, and Engineering, Michigan State University, East Lansing, MI, USA

ECE884 SS2021 Final Project

College of Engineering | Computer Science and Engineering

CMSE

Introduction

Online fake news has recently emerged as a major societal problem. Malicious actors spread fallacious viral stories in order to gain advertising revenue, tip influence opinions, and even tip elections. Available fake articles online increase exponentially every day, too much for humans to inspect and filter by hand.

Automated solutions for processing large volumes of information are required. It boils down to determining the legitimacy of information source. Fake News Detection techniques have become the need of the hour.

Existing approaches in the field:

Fully Connected DNN, LSTM, Bi-LSTM, CNN, Graph CNN, RNN (GRU, "Vanilla"), BERT, "Simple Models" (e.g., Log Reg, SVM).

Our approach:

Create a unified and diverse set of tools to tackle Fake News - any of which can be utilized as per the resources available.

Proposed Approaches

LSTM: Only remember the past inputs

- Sequential model
- Embedding layer
- LSTM - 2 layers
- Dense - 2 layers

BiLSTM: Remember both past and the future

- Non-sequential(functional) model
- Embedding layer
- Bidirectional - 2 layers
- Dense layer

Generative Pre-training (GPT-2):

- A semi-supervised approach for language understanding tasks which uses a combination of unsupervised pre-training and supervised fine-tuning
- Employs a two-stage training procedure: 1) uses a language modeling objective on large corpus of unlabeled data to learn the initial parameters of a neural network model; 2) fine-tune these parameters to a target task using the corresponding supervised objective - classification in our case.
- Uses a multi-layer Transformer decoder for the language model

Pre-trained BERT - Bidirectional Encoder Representations from Transformers

- Uses a pre-trained Transformer Bi-directional Encoder then fine-tunes with one additional output layer to create state-of-the-art models for a wide range of tasks without substantial architecture modifications
- BERT model is first initialized with the pre-trained parameters, and all of the parameters are fine-tuned using our labeled data from the classification task

Baselines: Approximately 360 baseline methods evaluated using 10-fold CV

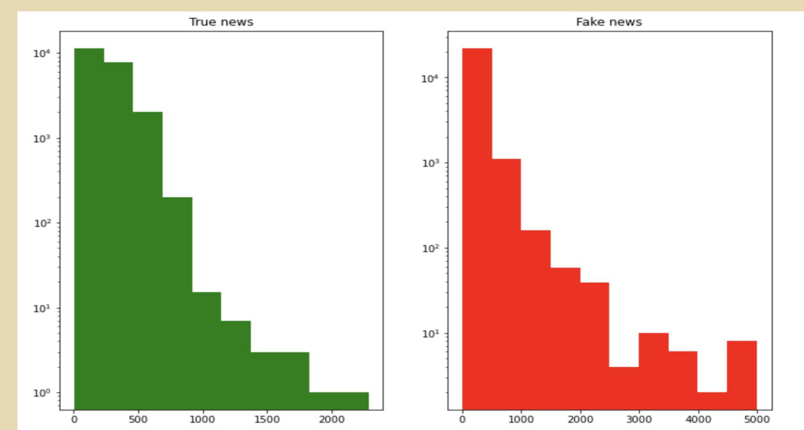
- 9 embedding methods from flairNLP library
- 10 ML classification models from sklearn using default parameters
- 4 feature sets: title, body, concatenation of title and body embeddings, average of title and body embeddings

Dataset

We utilized the [Fake News Dataset (Kaggle)], a publicly available dataset with ~40,000 news articles

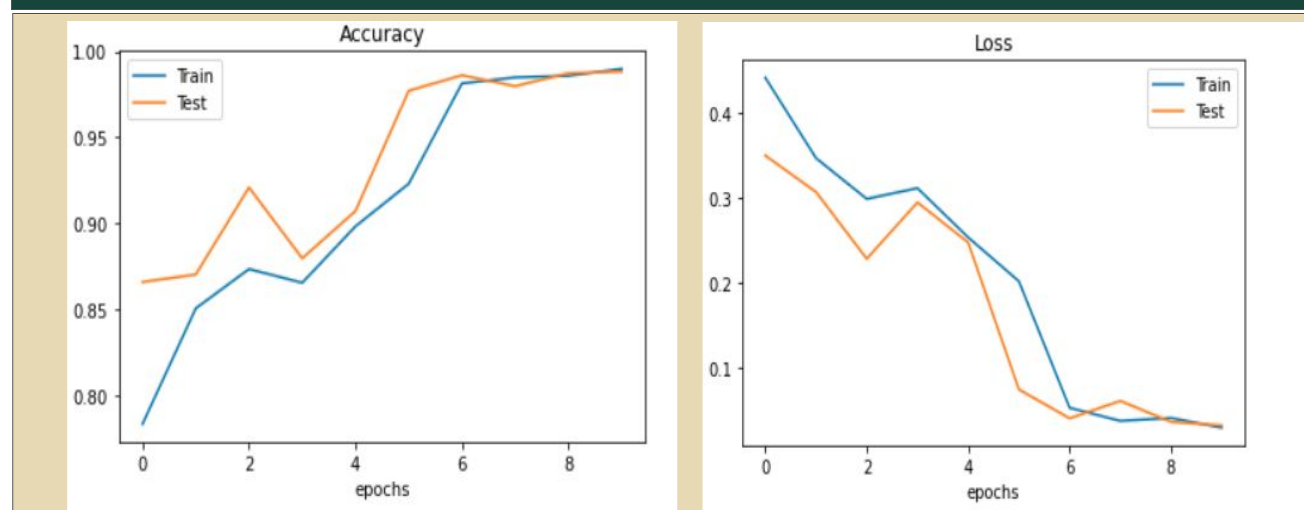
Here we show a breakdown of the number of articles in each class, a histogram of the number of words in each article by class, and a breakdown of number of articles in each topic

Type	Count
True	21,417
Fake	23,481

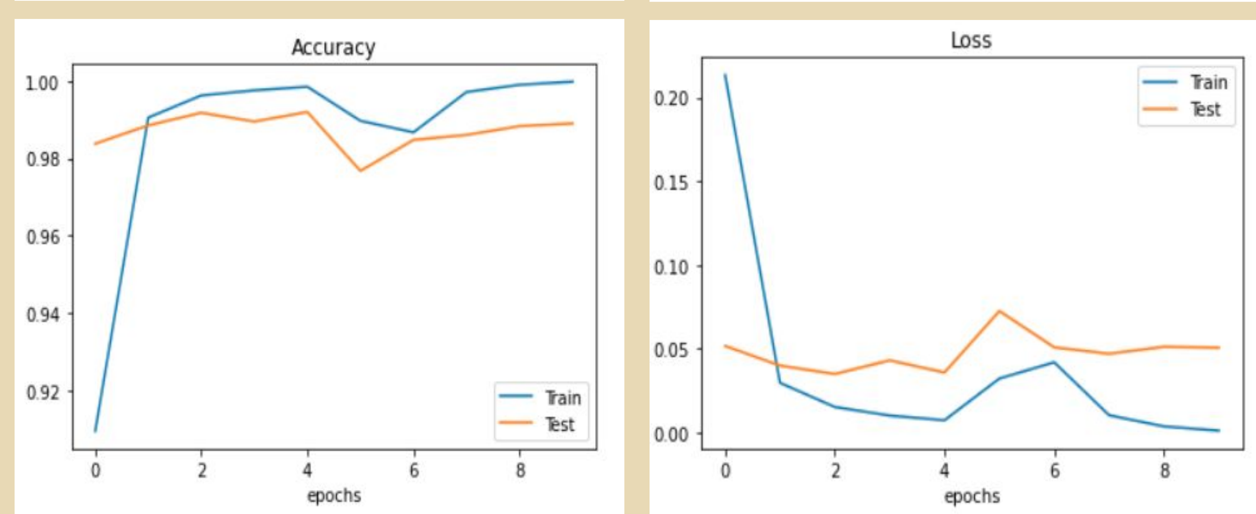


Subject	Count
Political News	11,272
World News	10,145
News	9,050
Politics	6,841
Left-News	4,459
Government News	1,570
US News	783
Middle-East	778

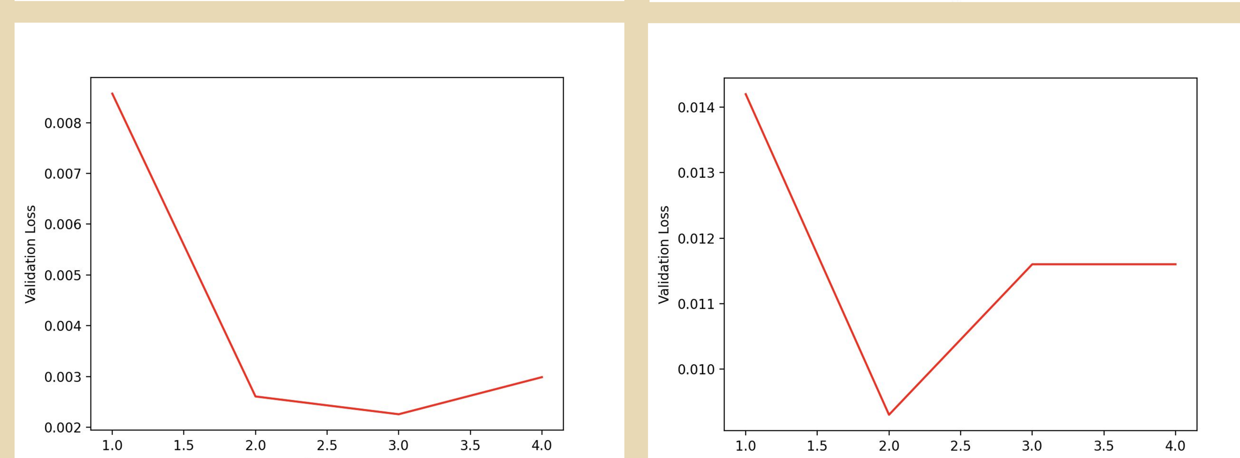
Results



LSTM: 98.99% accuracy on the test set



Bi-LSTM: 98.83% accuracy on the test set



Basic models for both **BERT (left)** and **GPT-2 (right)** were used - 12 layers with 100M+ parameters

100% accuracy on this dataset from epoch 1 for both BERT and GPT-2.

Baseline Model Results:

Title only: MLP trained on elmo embeddings **95.7% accuracy** (MLP bert: 95.1%)
Body only: MLP trained on bert embeddings **95.1% accuracy** (L1 log reg elmo: 94.2%)
Concatenation: MLP trained on bert embeddings **95.1% accuracy** (L1 log reg elmo: 94.3%)
Average: MLP trained on bert embeddings **95.1% accuracy** (L1 log reg elmo: 94.3%)

Discussion and Conclusion

Summary of our Models:

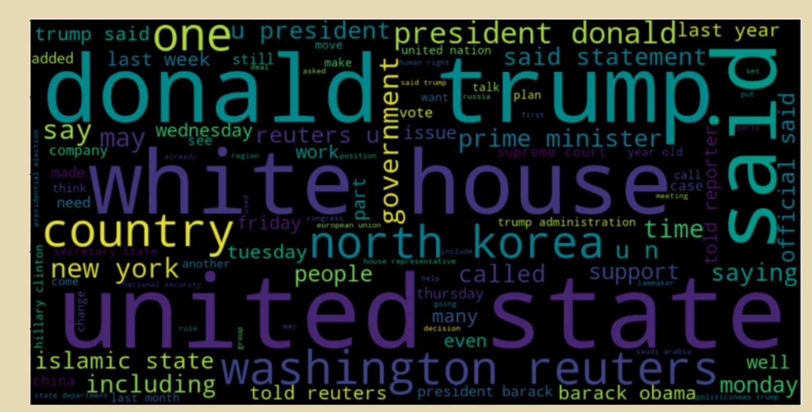
- Best performing baseline models were roughly 95% accuracy - comparable to many SOTA models in the literature
- Both LSTM and Transformer based approaches gave 99 and 100% accuracy. Transformer based models needed to be trained on v100 32G GPU for few hours.
- Best embeddings: BERT (pre-trained, no fine tuning), ELMo from flairNLP library
- Best baseline models: MLP, L1 log reg, L2 log reg. "Good" models: SVM (RBF/Linear), QDA - requires *significant* training time.

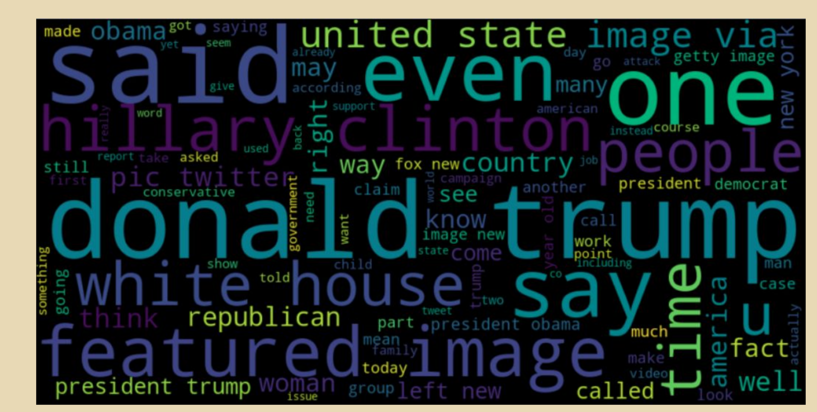
Our Approaches and Contributions:

- Current SOTA: >99% (task depending) in the literature. Our results meet or exceed this
- Transformers are still being explored in this space (not enough to compare to)
  - Our transformer models provide results in a developing area of research
- We have provided a number of computationally lightweight solutions for fake news detection, any of which can be used as per the resources available at hand
  - Example use: use simple logistic regression model to do host-side fake news prediction on a browser plugin
- Comprehensive results covering many different areas for this task
  - Our results are the most thorough benchmarking on this specific task including baselines and our deep learning models

Limitations:

- Limited number of topics and language variety used in our dataset limits generalizability to broader space of articles (see word clouds) - True News (left) and Fake News (right)





References

1. flairNLP: https://github.com/flairNLP/flair
2. R. Oshikawa, J. Qian, and W. Y. Wang, "A survey on natural language processing for fake news detection," in Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020), p. 6086–6093, 2020.
3. S. Girgis, E. Amer, and M. Gadallah, "Deep learning algorithms for detecting fake news in online text," in 2018 13th International Conference on Computer Engineering and Systems (ICCES), pp. 93–97, 2018.
4. K. Popat, S. Mukherjee, A. Yates, and G. Weikum, "Declare: Debunking fake news and false claims using evidence-aware deep learning," 2018.
5. E. Qawasmeh, M. Tawalbeh, and M. Abdullah, "Automatic identification of fake news using deep learning," in 2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS), pp. 383–388, 2019.
6. M. A. Ayat Abedalla, Aisha Al-Sadi, "A closer look at fake news detection: A deep learning perspective," in ICAAI 2019: Proceedings of the 2019 3rd International Conference on Advances in Artificial Intelligence, pp. 24–28, 2019.