Fake News Detection

Shakshi Sharma, William Mukose, Zaure Zhussipbekova

University of Tartu

Introduction

Fake News are widely spread across the Internet especially after the 2016 U.S Presidential Elections. Fake news take advantage of the echo chambers phenomenon; people tend to follow and share mostly information they believe in. Therefore, we study the detection of Fake News through news articles. We utilize two types of analysis while detecting Fake news: first, descriptive analysis and second, predictive analysis.

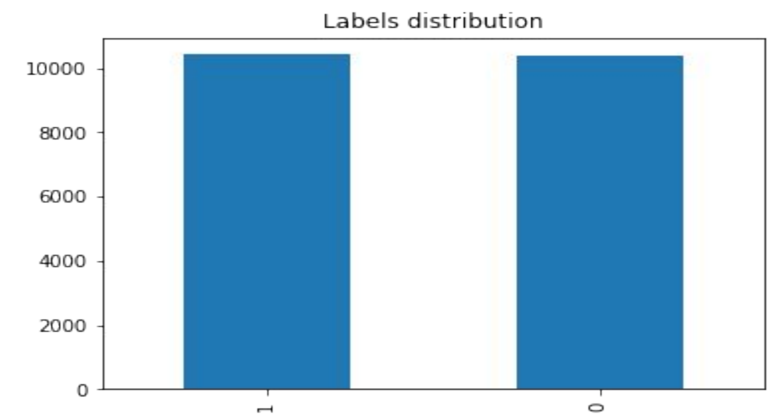
We work on a Kaggle project about fake news detection. We have a labelled dataset with several articles that provide examples of both fake and real news. We build several models to try to learn features of fake vs real news, then apply them to the test dataset, also on Kaggle. Initially, we take the baseline as a logistic regression, Naive model etc, and then we apply some neural networks.

We achieved best results from neural network approach - Convolutional Neural Network, that is, 99.7% test accuracy and 99.1% validation accuracy and validation F1 Score both.

In the previous literature, fake news detection problems have been widely used as classification tasks using traditional as well as neural network techniques. However, the task of looking at the structure of fake and real news is still an unexplored area. Thus, apart from performing classification tasks, we would like to explore descriptive analysis such as Topic Modelling, Sentiment analysis, etc.

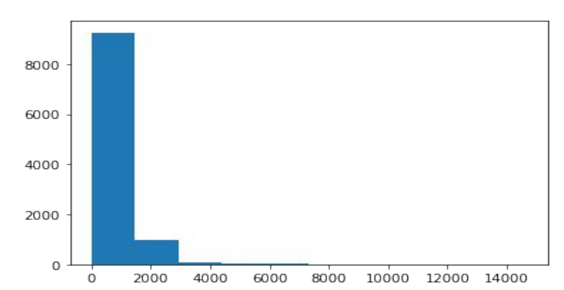
Dataset

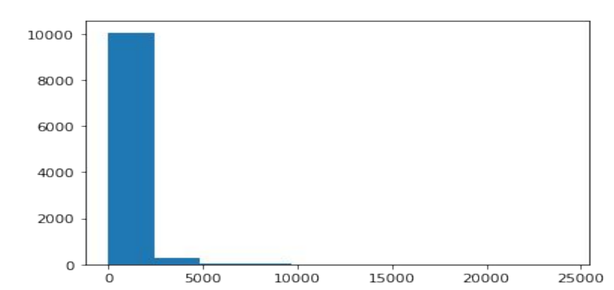
For this project, we used publicly available data set from Kaggle competition- https://www.kaggle.com/c/fake-news The dataset consists of articles with ids, titles, texts, authors and, in the case of train data labels whether the article was true or false. The train set had 20,800 articles, while the test set had 5200 articles. The data was perfectly balanced, with equal numbers of articles falling into the fake news (1) and real news (0) categories.



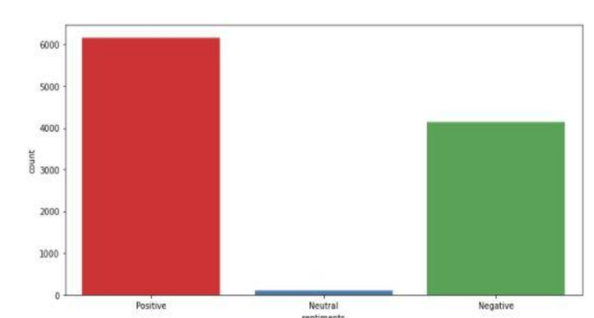
Descriptive Analysis

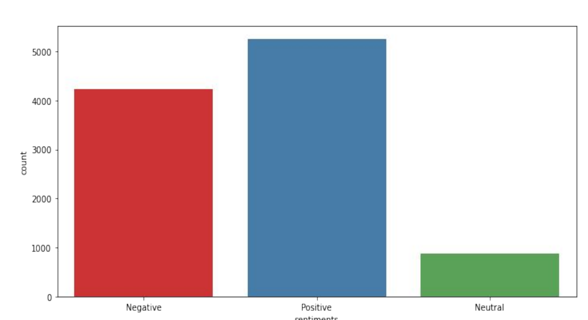
Articles' Length



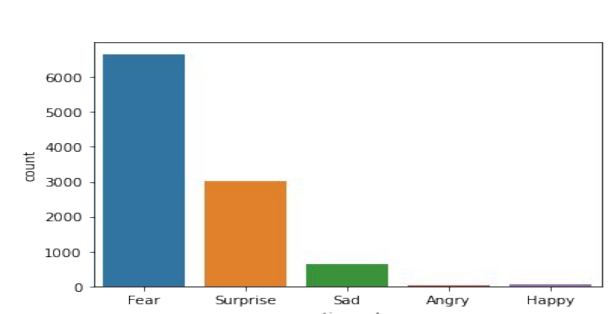


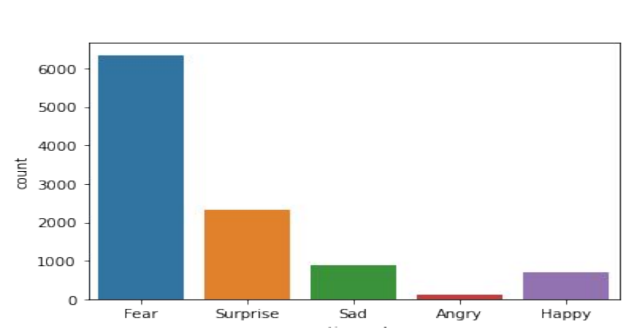
Sentiment



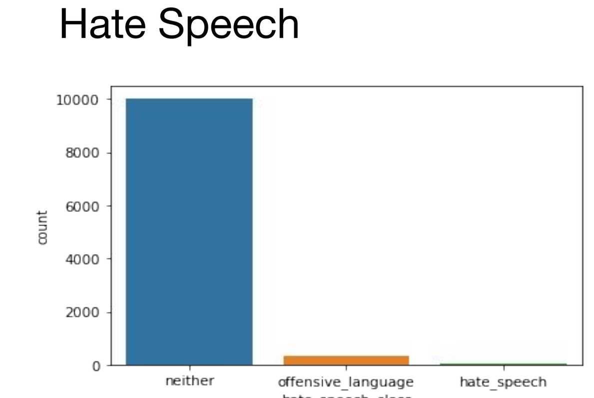


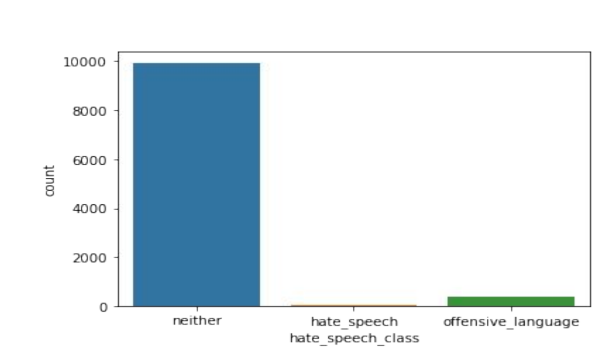
Emotion





Hate Speech





Predictive Analysis

The decision was taken to use only a combination of article title and author (leaving out the text), as this sped up the models appreciably and had minimal effect on performance. Ordinary machine learning techniques with appropriate hyperparameter tuning were first used to classify the articles (Logistic Regression, MultiNomial Naive Bayes and Random Forest). Articles were encoded as one hot vectors of their words (lowercased), using only the 10,000 most common words.

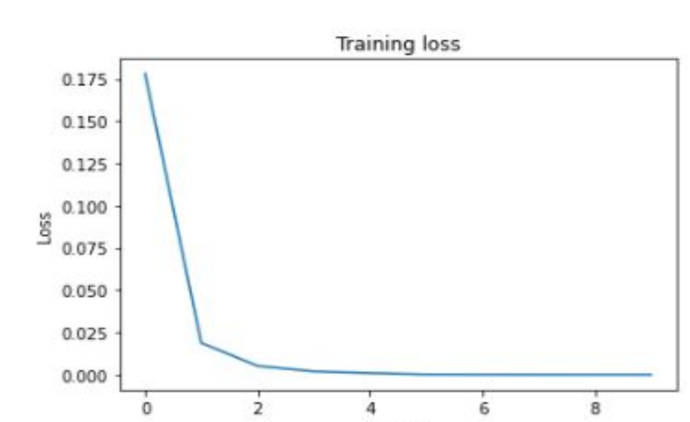
One problem with the machine learning approach was that testing on validation set produced consistently better scores than testing on the test set. This is because the test set did not participate in making the encoded matrix, and some words present in the training set were absent in the test set, impacting performance

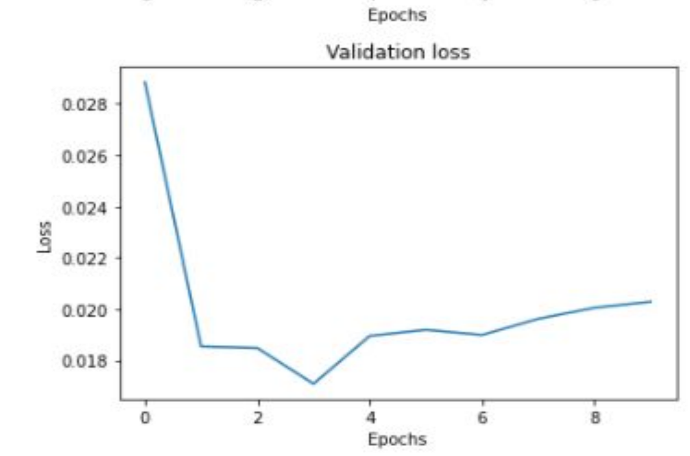
Accuracy values for vanilla machine learning were in the 88%-92.7% range.

Two neural networks were used to predict data classes, a Long Short Term Memory (LSTM) network and a Convolutional Neural Network (CNN). Both neural networks showed superior performance to vanilla machine learning methods, with more than 99% accuracy on the test set on kaggle.

Both neural networks used the same preprocessing architecture: Tokenization and stemming with nltk, turning text into embeddings with keras.preprocessing and padding to maximum sentence length. Both networks had one Convolutional and LSTM layer respectively, as the data was comparatively not very complex.

Train and validation learning curves for our CNN model are shown below. Surprisingly, validation loss actually begins to grow after the 3rd epoch, meaning overfitting is taking place but given that accuracy on the binary classification is 99.7%, this was deemed a minor issue.





Results

Below are results we obtained while submitting to Kaggle

Model	Test Set Accuracy (Graded By Kaggle)
Logistic Regression, l1 regularization	92.6%
Logistic Regression, l2 regularization	91.5%
MultiNomial Naïve Bayes	89.0%
Random Forest	92.2%
Gradient Booster	92.2%
Convolutional Neural Network	99.7%
LSTM Neutral Network	99.0%

Below are results we obtained while training models using a validation set. We are presenting accuracy, as well as the F1 score on the validation set.

Model	Validation Accuracy	Validation F1 score (average over both classes)
Logistic Regression, l1 regularization	99.1%	99.0%
Logistic Regression, l2 regularization	97.5%	97.5%
MultiNomial Naïve Bayes	95.9%	95.9%
Random Forest	98.4%	98.7%
Gradient Booster	98.7%	98.6%
Convolutional Neural Network	99.1%	99.1%
LSTM Neutral Network	99.0%	99.0%

Conclusion

We analyzed the Fake news detection problem in two ways - descriptive and predictive analysis. Descriptive analysis shows satisfactory results in terms of differentiating between the True and Fake news. In case of predictive analysis, we get the best results using Neural Network, that is, Convolutional Neural Network in terms of accuracy and F1 Score both. To summarize, we can say that both descriptive and predictive analysis are essential while detecting Fake news.

References

[1] Mykhailo Granik, Volodymyr Mesyura. Fake News Detection Using Naive Bayes Classifier. 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)

[2] Hadeer Ahmed, Issa Traore, Sherif Saad.Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques. International Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments. 2018

[3] Shu, Kai, et al. "defend: Explainable fake news detection." Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2019.

[4] Wang, Yaqing, et al. "Eann: Event adversarial neural networks for multi-modal fake news detection." Proceedings of the 24th acm signed international conference on knowledge discovery & data mining. 2018.

[5] Oshikawa, Ray, Jing Qian, and William Yang Wang. "A Survey on Natural Language Processing for Fake News Detection." ArXiv:1811.00770 [Cs], March 4, 2020. http://arxiv.org/abs/1811.00770.

[6] Girgis, Sherry, Eslam Amer, and Mahmoud Gadallah. "Deep Learning Algorithms for Detecting Fake News in Online Text," 2018.https://doi.org/10.1109/ICCES.2018.8639198.

[7] https://www.voicemagazine.org/app/uploads/VOICEMAG_2702_2019-01-11.pdf

[8] https://www.kaggle.com/c/fake-news/