**Development and Application of E2E NLP Pipeline for the
Detection of Fake News and Content Summarization of News Articles**

Due Date: Week 8

Introduction:

Natural Language Processing (NLP) has grown in popularity in recent years as a promising field with many potential applications such classification, summarization, and information retrieval. In particular, applications of NLP to improve user experience when it comes to everyday activities such as reading and writing have increased in recent years to meet the unprecedented influx of information within online platforms. One area in which NLP is currently underrepresented is within the confines of fake news detection given that there very few tools available to help limit or prevent false claims. Given the current popularity of this topic in the news, the focus of this project will be concentrated on that. Within this project, an end-to-end pipeline will be developed which will comprise (1) a classification model by which news articles will be determined to be real or fake (2) a subject classification model by which the category of the topic will be determined, and (3) the content will be summarized in an attempt to generate an appropriate headline. This proof-of-concept will demonstrate the potential utility of a pipeline that can be used my major news organizations to automate the methods by which users add content to websites.

Dataset and Approach:

In order to develop these proposed models, the "*Fake and Real News Article Dataset*" [1] from Kaggle [2] will be used. This rich dataset comprises a title, an article's text, the subject, and the legitimacy classification. Using this single dataset, all of the proposed models will be trained and validated in order to build the end-to-end pipeline.

The two classification models will be developed using a deep learning architecture [3] using the Keras and/or Tensorflow libraries. Several hidden layers will be implemented along with an LSTM layer given the dependence of words on each other [4], allowing us to better capture the context of the text. GLOVE embeddings will be used, and several other variants will be explored to determine the best path forward. In terms of regularization, Adam will be used given its general use capabilities. When it comes to validating the models, the use of accuracy, recall, and F-1 scores will be used to determine the correct path forward. In addition, the loss from the models will be monitored as the model continues its training process. A visual will be prepared showing the training and validation loss and accuracies against the epochs. In addition, an ROC-AUC graph will be prepared.

Finally, the summarization model will likely be implemented using a seq2seq architecture, or, through an open-source hugging-face pretrained summarization model depending on the given timeline. The summarization capabilities will be evaluated using the BLEU and ROUGE scores. In addition, several articles at random will be selected and their summaries measured against ChatGPT. The results will be visualized using the seaborn and matplotlib libraries.

Assessment Methods:

For the two classification models, we will focus on accuracy, recall and precision to assess and monitor the performance of the two models. In addition, as the parameters of the models are changed and the performance is improved, the AUC-ROC curve will be created to track the performance during the ablation process. This of course will be generated using a confusion matrix. For the summarization model, the performance will be assessed using the BLEU and ROUGE scores, which are two metrics commonly used for this task.

Timeline:

- Week 8: Develop and complete the exploration and data preprocessing
- Week 9: Develop the subject classification model
- Week 10: Complete the legitimacy classification model for the text articles
- Week 11: Complete the summarization model for the articles
- Week 12: Write the report and present the results

References & Related Work:

- **[1]** Ahmed H, Traore I, Saad S. "Detecting opinion spams and fake news using text classification", Journal of Security and Privacy, Volume 1, Issue 1, Wiley, January/February 2018.
- **[2]** Ahmed H, Traore I, Saad S. (2017) "Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques. In: Traore I., Woungang I., Awad A. (eds) Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments. ISDDC 2017. Lecture Notes in Computer Science, vol 10618. Springer, Cham (pp. 127-138).
- **[3]** Shervin Minaee, Nal Kalchbrenner, Erik Cambria, Narjes Nikzad, Meysam Chenaghlu, Jianfeng Gao. "Deep Learning Based Text Classification: A Comprehensive Review", Arxiv.org
- **[4]** Shervin Minaee et. al. "Deep Learning Based Text Classification: A Comprehensive Review", Arxiv.org
- **[5]** Nathaniel Hoy, Theodora Koulouri. "A Systematic Review on the Detection of Fake News Articles", Arxiv, **2021**.
- **[6]** Mrinal Rawat, Diptesh Kanojia. "Automated Evidence Collection for Fake News Detection", Arxiv, **2021**.