Adversarial Attacks on Fake News Detection Methods

Varsha Meghana Kanmuri

UNI: vk2497

Email: vk2497@columbia.edu

1. Synopsis

Overview: Machine Learning (ML) and Natural Language Processing (NLP) techniques are gaining popularity in fake news detection as it is challenging and time-consuming for humans to go through a large number of news articles manually. As part of the project, I have implemented Machine learning (ML) algorithms like LSTM Neural Networks, XGBoost, and Support Vector Machines (with RBF Kernel), along with employing Natural Language Processing techniques (NLP) such as Word2Vec, RoBERTa for detecting fake news on the ISOT dataset. Furthermore, due to the growing popularity of Large Language Models (LLMs) such as ChatGPT for generating fake data [3], I tried to inject adversarial attacks by using GPT-J Language model to generate synthetic fake news stories. I then evaluated the performance of the above ML/NLP models when they were tested on fake news generated by these LLMs.

The code is available on GitHub, making it publicly accessible for anyone to reproduce the experimental results. In addition, I have included the instructions for setting up the environment and running the code for the three different AI pipelines, hyper-parameter tuning, generating synthetic fake data, and testing demo news articles. Hence, anyone in the research community can use the above AI systems combining NLP and machine learning / DL models for fake news detection.

Novelty: Along with the traditional NLP techniques such as Word2Vec to perform feature engineering to generate the word embeddings, I also used state-of-the-art Large Language transformer models like RoBerta to generate text embeddings as input to my machine learning models and compare performance.

I performed a literature review in NLP about using LLMs for adversarial attacks (generating fake synthetic data), and I found that it is an emerging and novel research area; Existing approaches are more manual and rules-driven to perturb text examples such as perturbing numerical values [6], etc. Furthermore, I experimented and realized that ChatGPT could not generate fake news directly. So, I used a Large-Language Model, GPT-J (6 Billion parameter model), to create synthetic fake news examples from factual news titles as input prompts. The temperature hyper-parameter controls how much randomness is in the output generated by the model.

I evaluated the impact of the fake news adversarial attacks on these trained models and further tested the performance of ChatGPT in detecting some of these samples. I created 500 synthetic fake news examples for an initial performance evaluation as a novel contribution. I also provided the code to generate additional examples so the research community can have a benchmark to test the robustness of their fake news detection models to adversarial attacks. The appendix section also includes a figure of a generated fake news article. Can you spot the real one? (Page 14)

Value to user community: A recent analysis by Statista [4] reveals that people worldwide find it increasingly challenging to distinguish fake news from real news and lose their trust in the information they are exposed to. Fake news detection is crucial today as the fake news generation has greatly influenced decisions on political and socio-economic issues in every society and country. There will be real consequences of not having mechanisms and AI systems to detect fake news, such as election interference, less trust in public health information, increased risk of user safety and privacy, etc.

These statistics have made it critical to address and develop effective techniques to detect fake news. We can measure the performance of the ML/NLP models (accuracy, F-1 score, precision, recall) to address Research Question (RQ) (1). In addition, we can compare different machine learning / deep learning models (using different NLP embeddings) to determine the best-performing model for fake news detection. The performance metrics will help address RQ (2) to replicate the findings and utilize these models in real time for fake news detection. With the popularity of 'ChatGPT' and other Large-Language Models (LLMs), it's essential to evaluate the performance of these models when the actual news is perturbed (adversarial attacks) RQ (3). This will demonstrate the robustness and identify future-research opportunities to improve ML models for fake-news detection RQ (4).

2. Research Questions

Q1. Can we use Machine Learning algorithms with NLP embeddings to detect fake news articles successfully on the ISOT dataset?

Yes, AI models can detect fake news articles on the ISOT dataset with good performance. This is achieved by using an AI pipeline employing NLP to generate embeddings with language models such as RoBERTa / BERT / Word2Vec feeding into ML models such as SVM, XGBoost or DL Models such as LSTM neural networks.

Q2. What percentage of real-news articles are tagged as 'fake' by these ML/DL algorithms?

A very small percentage (<1-2%) of real articles in the ISOT dataset are tagged as 'fake' by the three different AI model pipelines. This provides evidence that AI models trained on appropriate datasets will be significantly helpful to detect fake news.

Q3. Can LLM models like GPT-J successfully generate human-like fake news?

Yes, Although some safety measures exist in directly using ChatGPT for generating fake news, we can easily trick ChatGPT to overcome the guard-rails in some scenarios. In this project, we show that Large Language Models (LLMs) like GPT-J, GPT-4, etc. can be easily used to generate synthetic human-like fake news examples. As a novel contribution, we also generate 500 synthetic fake news examples from real-world news titles as input prompts. We use the temperature hyper-parameter to make the model be more creative while generating synthetic data.

Q4. Can adversarial attacks (generating fake news articles from real news articles) reduce the performance of Machine Learning models for fake news detection ?

Yes, We observe from our experiments that the performance of some ML/DL models are significantly affected by adversarial attacks (generating fake news articles from real news articles). The Word2Vec model does not capture the contextual knowledge of the sentences well when compared to the RoBERTa model. This is an essential factor for why the Word2Vec + SVM model has the most significant drop in performance for successfully detecting fake news articles. The RoBERTa + XGBoost model performs very well in detecting the synthetic fake news as the RoBERTa model is able to capture the context meaning of words in the articles.

3. Related Work

There has been a lot of research done on the topic of Fake News detection using the ISOT dataset [1], [2]. The performance of AI models varies depending on the NLP Language Model used for feature engineering, and the ML / DL model employed for classification along with the dataset used to measure test performance. These approaches have also been extended to detect fake news on social media such as Twitter, Facebook etc. A key aspect which has not been explored in the research community, and a novel contribution of this project is measuring the impact of these AI pipeline models in detecting fake news articles synthetically generated by LLMs [3], [5] based on real-world news titles.

4. Experimental Results and Discussion

Dataset

I have used the ISOT Fake News dataset, which is made publicly available for research by the University of Victoria [1], [2]. The ISOT Fake News dataset [1], [2] is a compilation of several thousands of fake news and truthful articles. Truth news articles were obtained by crawling articles (21,417 examples) from Reuters.com (News website) and Wikipedia. The fake news articles were collected from unreliable websites that were flagged by Politifact (a fact-checking organization in the USA). The dataset contains articles on different topics, however, the majority of articles focus on political and World news topics.

Methodology

Data Pre-processing: We use NLP techniques such as tokenization, stemming, lemmatization, stop-word removal, etc. Additionally, we also remove certain hyperlinks, image captions and punctuations to further clean the data and identify common topics/themes amongst real and fake news examples. We observe that the key topics and common words / word distributions are similar; This highlights the difficulty of the problem in building an AI system which understands the complete context of the article. We also add the Word Clouds in the appendix section of the report.

Model Training: We first split the ISOT dataset using a 80-20 train/test split and then perform k-fold validation on the training set. The dataset is valuable as they have multiple editors to fact-check the news. I then evaluated the performance of fake news detection using ML/NLP approaches on the ISOT 'test' dataset.

We experiment and compare 3 different types of NLP and ML/DL Architectures for fake news classification commonly used in literature for fake news detection. We do not re-train the models after generating the adversarial fake news samples, assuming a Black Box attack approach - typically assumed for measuring performance of the 3 different models.

Figure:- Machine Learning/NLP models used for training on the ISOT dataset.

Feature Engineering - Word Embeddings (NLP Model)	**Classification Model (Machine Learning / Deep Learning)**
Word2Vec (NLP Language Model)	Support Vector Machines (SVM with RBF Kernel)
RoBERTa (NLP Language Model)	Neural Networks (LSTM based architecture)
RoBERTa (NLP Language Model)	XGBoost (Ensemble Methods)

The Word2vec model uses a shallow neural network model to create word embeddings, where as the RoBERTa [7] model uses the same bidirectional transformer architecture as the BERT LLM, but with more training data and a longer training time. The BERT based models are also more suitable for tasks such as classification.

While other NLP models exist, they have lower performance on the same data set and other publicly available fake news datasets; I have focused the experiments on state-of-the-art models to measure impact of adversarial attacks (for fake news classification).

I trained the models on an Apple Macbook Pro laptop with 32 GB RAM and M1 chip processor. The overall computation time for training the models is around 147 hours (~ 6 days) for the experiments including 5/10 fold cross-validation for hyper-parameter tuning in the machine learning and deep learning models models. The most expensive computation time is for GridSearchCV, and tuning the XGBoost and Neural Network architecture when tuning the hyper-parameters over a large search space. Since my laptop does not have a GPU, I could not train Large Neural Networks with longer epochs, to get higher performance measures.

Figure:- Accuracy metrics of the train and test set using the NLP + ML/DL models.

NLP + ML/DL architecture used	Accuracy Train data	Accuracy Test data
Word2Vec + SVM with RBF Kernel	99.96 %	99.93 %
RoBERTa + LSTM Neural Network	96.35 %	96.34 %
RoBERTa + XGBoost	100.00 %	98.1 %

We observe that all 3 models perform well, with the Word2Vec + SVM pipeline having the best Test Accuracy, followed by the RoBERTa model + XGBoost pipeline. Does this mean that the Word2Vec + SVM pipeline model is indeed capturing the meaning of what is real vs fake ? or is it just memorizing certain words or phrases seen in the ISOT dataset?

Generating fake news using LLM (GPT-J):

We experiment with the GPT-J language model to generate fake news stories from real news titles from the ISOT dataset as input prompts. We generate synthetic fake news examples using the GPT-J (6 Billion parameter model) setting the 'temperature' = 0.9. The GPT models are better for natural language generation / summarization.

The model is ~24 GB in size, and consumes nearly ~ 24 GB in RAM for processing. As I generated the fake news examples on my Apple Macbook Pro laptop with 32 GB RAM and M1 chip processor. The overall computation time for generating around 500 synthetic fake news examples (each of ~150 words) is ~ 48 hours (~ 5 to 10 min per sample) .

Figure:- A sample prompt with generated fake news and real news text.

```
[333… prompt1 = shuffled_df['title'].iloc[3]
      prompt1

[333… 'California AG pledges to defend birth control insurance coverage'

[334… input_ids1 = tokenizer(prompt1, return_tensors="tf").input_ids
      •••

[ ]:  # Fake Generated Text

[336… gen_text1

[336… 'California AG pledges to defend birth control insurance coverage\n\nCalifornia Attorney General Kamala Harris has pledged to file briefs i
      n support of the contraceptive mandate in the Hobby Lobby case before the U.S. Supreme Court.\n\nIn a press release announcing this pledge,
      Harris said:\n\n"This case impacts millions of Americans, and it is crucial that their voices are heard. I will file briefs to make sure th
      ey are included in the legal proceedings. It will be my highest priority to defend the constitutionality of the Affordable Care Act, includ
      ing the mandate that employers provide coverage for contraceptives. As the state's chief legal officer, my duty is to defend the laws of th
      e state and the will of its people."\n\nLast'

[ ]:  # Real News Text

[291… shuffled_df['text'].iloc[3]

[291… 'California AG pledges to defend birth control insurance coverage SAN FRANCISCO (Reuters) - California Attorney General Xavier Becerra said
      on Friday he was "prepared to take whatever action it takes" to defend the Obamacare mandate that health insurers provide birth control, no
      w that the Trump administration has moved to circumvent it. The administration's new contraception exemptions "are another example of the T
      rump administration trampling on people's rights, but in this case only women," Becerra told Reuters.  Becerra and other Democratic attorne
      ys general have filed courtroom challenges to other Trump administration policies involving healthcare, immigration and the environment.  '
```

Another figure of a generated fake news article is also added in the appendix section - Page 14.
Check if you can guess it right ?

Experiment: We now evaluate the performance of the three NLP + ML/DL models to test their robustness on the synthetic fake-news data (500 samples) generated using GPT-J.

Figure:- Accuracy Metrics of the models on the synthetic fake samples generated using GPT-J:

NLP + ML/DL architecture used	Accuracy on detecting the generated synthetic fake Test data
Word2Vec + SVM with RBF Kernel	24.75 %
RoBERTa + LSTM Neural Network	50.69 %
RoBERTa + XGBoost	100%
ChatGPT (tested manually on 10 samples)	10 %

Test Demo Result: Using the 3 models on sample data points.

Actual news: *California AG pledges to defend birth control insurance coverage SAN FRANCISCO (Reuters) - California Attorney General Xavier Becerra said on Friday he was "prepared to take whatever action it takes" to defend the Obamacare mandate that health insurers provide birth control, now that the Trump administration has moved to circumvent it. The administration's new contraception exemptions "are another example of the Trump administration trampling on people's rights, but in this case only women," Becerra told Reuters. Becerra and other Democratic attorneys general have filed courtroom challenges to other Trump administration policies involving healthcare, immigration and the environment.*

Figure:- Predictions made by the three ML/DL + NLP models for the real news article

```
Running the different NLP models + ML/DL classifiers for example in "text" column in file -  sample_true2.csv

converting words to features: 100%|██████████| 1/1 [00:00<00:00, 309.59it/s]
Test Demo output - Prediction of rbf_svm + word2vec model:  real news
2023-05-01 22:47:41.832052: W tensorflow/tsl/platform/profile_utils/cpu_utils.cc:128] Failed to get CPU frequency: 0 Hz
1/1 [==============================] - 2s 2s/step
Test Demo output - Prediction of RoBERTa + LSTM Neural Networks model:  real news
Test Demo output - Prediction of RoBERTa + XGBoost model:  real news
```

Input prompt title for GPT-J : California AG pledges to defend birth control insurance coverage

Fake news generated: *California AG pledges to defend birth control insurance coverage California Attorney General Kamala Harris has pledged to file briefs in support of the contraceptive mandate in the Hobby Lobby case before the U.S. Supreme Court. In a press release announcing this pledge, Harris said:"This case impacts millions of Americans, and it is crucial that their voices are heard. I will file briefs to make sure they are included in the legal proceedings. It will be my highest priority to defend the constitutionality of the Affordable Care Act, including the mandate that employers provide coverage for contraceptives. As the state's chief legal officer, my duty is to defend the laws of the state and the will of its people."*

Figure:- Predictions made by the three ML/DL + NLP models for the synthetic fake news article

```
Running the different NLP models + ML/DL classifiers for example in "text" column in file -  sample_fake_syn_2.csv

converting words to features: 100%|██████████| 1/1 [00:00<00:00, 318.06it/s]
Test Demo output - Prediction of rbf_svm + word2vec model:  fake news
2023-05-01 00:48:31.000120: W tensorflow/tsl/platform/profile_utils/cpu_utils.cc:128] Failed to get CPU frequency: 0 Hz
1/1 [==============================] - 1s 908ms/step
Test Demo output - Prediction of RoBERTa + LSTM Neural Networks model:  real news
Test Demo output - Prediction of RoBERTa + XGBoost model:  fake news
```

Discussion:

Although the Word2Vec + SVM model architecture had the best performance metrics on the ISOT dataset - its robustness to adversarial inputs is very poor. The accuracy performance drops significantly 99% -> 24.75%. This is likely due to the word embeddings not capturing the context of the words in a sentence when they appear in the news article.

The RoBERTa + XGBoost model is more robust to adversarial attacks (generating fake news from real-world news titles) because it is able to better capture the word context much better than the Word2Vec + SVM model. This results in the model having very good performance on synthetic fake news samples generated by GPT-J. Finally, The RoBERTa + LSTM NN model has performance between the other two models.

We also tested 10 samples of the synthetic fake news data with 'ChatGPT', and observed that 'ChatGPT' had poor accuracy in detecting fake news examples.

5. Deliverables

The code is open-sourced, and publicly available through GitHub for access with no cost for using the models. https://github.com/varsha-mkanmuri/fake_news_detection_6156.git

I have included all the code, configurations, models, data files, documentation, presentations, demo links within the repository. The project and experiments are completely reproducible and I have given complete instructions in the ReadMe file. Including some high-level instructions and key-points below:-

Installation - Setup the python environment with the packages required - given in the requirement.txt file. You can also use the '.yaml' file to set up the anaconda environment as well.
Data: The dataset files are included in the data folder.
Trained_models: This folder contains all the model parameters, values, weights of the trained NLP, ML/DL models for evaluation of performance.
Syn_fake_data: This folder contains the synthetic fake news data generated using the GPT-J model, with the script 'generate_fake_syn_data.py'. It's currently set up to generate 500 examples, but you can increase this limit to any value you want.
Sample_demo: This folder contains test input samples for demo purposes and testing of real data from the ISOT dataset, and synthetically generated fake data. Refer to GitHub link for additional information.
main.py - The main script runs the pipeline for the 3 different approaches, training the models, saving them, and evaluating performance - comparing test performance of the 3 models on the ISOT dataset.
Test_syn_fake_data.py - This script compares performance of the 3 models on the 500 synthetic fake news examples generated by GPT-J model in the script 'generate_fake_syn_data.py'.
Other scripts.py - Other scripts implement the individual pipelines, hyper-parameter tuning, etc.

6. Conclusion

From the experiments conducted above, it is clear that it is becoming all the more challenging to detect fake news with LLMs making it easier to generate 'human-like' simulated data using prompts and just a few keywords. Most LLMs today can easily be manipulated to break their hard-coded ethical rules by rewording the input prompts. While some Machine Learning models perform exceptionally well while testing on unseen data from the test dataset, not all of them retain their performance when they are exposed to adversarial attacks from the simulated data. This calls for the immediate need to understand how these Machine Learning models work internally and the tasks each model is better suited for. Additionally, this also points out the limitations of relying entirely on metrics like accuracy or precision for measuring a model's performance. This observation stresses the need to utilize Explainability techniques like SHAP (Shapley Additive exPlanations) to trust these AI models better. Knowing why a model classifies a specific news article as fake or real is essential, rather than just knowing whether it classifies it correctly to ensure they work well when deployed in the real world.

7. Self-Evaluation

7.1 What I learned working on this project: I learned the end-to-end process of building a research project from scratch while working on this project. I started with identifying the most crucial topics where machine learning in security could play a significant role in helping the user community to make my research worthwhile. While I had many issues in mind, the recent upsurge in the spread of fake news during the covid-19 pandemic, its role in financial abuse, and political outrage concerned me the most. To make things worse, recognizing fake news has gotten highly challenging, with Large Language Models like ChatGPT being accessible to the general public. Parallely, I also learned the importance of shortlisting an appropriate dataset (ISOT Fakenews dataset) with accurate information from credible sources to ensure I trained my models on valuable data samples. Moreover, I learned to understand the insights and limitations of other related/baseline work and tried to investigate what was missing and how I could improve on them. Additionally, I learned how to think more like a researcher to see how I can introduce novelty through my work and address some of the most worrying concerns in the current world. Furthermore, I also gained broader technical knowledge on shortlisting the best ML/NLP models and the importance of hyperparameter tuning and preprocessing of the text input to optimize for the most accurate results. I learned the importance of starting early to have sufficient time and resource bandwidth to train models on large datasets. I also managed to train over the whole dataset with limited access to GPUs by starting early and achieving parallelization. Finally, I learned how to make the project reusable and open source so the research community can collaborate or share their valuable thoughts.

7.2 What I learned working on this course: Overall, this course was one of the most memorable courses I took because it was the first time I learned techniques to read state-of-the-art research papers and the first time I worked on a research project. As a result, my interest in ML in the Security field and Ethics has broadened widely, and I hope to contribute more by publishing papers at AI conferences.

7.3 Planned but did not do: I also proposed to re-train the machine learning and deep learning models on the augmented dataset (synthetically generated fake news + original dataset) to measure the performance.
However, since my laptop does not have a GPU or more significant RAM, I could generate up to 500 synthetic fake news samples (which took close to 48 hours) and could not create samples with longer words/tokens (~500-1000 words each) - as the generation time is exponential. Since the synthetic sample size is small, instead of retraining the models with augmented data, I used these unseen samples to see how the models behave in a black box adversarial attack setting environment. It is however possible with more compute power and time to generate ~10,000 adversarial samples and combine them with the original ~40,000 data points for re-training the 3 different models to evaluate performance.

7.4 Limitations and Future Work: Further experimentation with Neural Networks to get better results will require more compute time and resources. Even models like RoBERTa + XGBoost need continuous training on latest news articles with sufficient labeled examples to have good performance for fake news detection. Another possible extension of the work would be to train the model on additional data with labels that are not related to 'news' so that AI models can predict other categories of output when an input data sample is not a 'news' article or if the model is not confident about the probability of the news article being real or fake. Furthermore, it would be important to use SHAP and other explainable AI techniques to better understand the output predictions of these models and identify research opportunities for improvement.

8. References

[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. https://onlinelibrary.wiley.com/doi/epdf/10.1002/spy2.9

[2] Ahmed H, Traore I, Saad S. (2017) "Detection of Online Fake News Using NGram 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).
https://link.springer.com/chapter/10.1007/978-3-319-69155-8_9

[3] Abstracts written by ChatGPT fool scientists
https://www.nature.com/articles/d41586-023-00056-7

[4] Level of difficulty differentiating between true and false information online among consumers in selected countries worldwide as of June 2020
https://www.statista.com/statistics/1227193/identifying-misinformation-difficulty-worldwide/

[5] Harshit Sharma, July 15, 2022. "Softmax Temperature"
https://medium.com/mlearning-ai/softmax-temperature-5492e4007f71#:~:text=Temperature%20is%20a%20hyperparameter%20of%20LSTMs%20(and%20neural%20networks%20generally,utilize%20the%20Softmax%20decision%20layer.

[6] Zhou, Z., Guan, H., Bhat, M.M. and Hsu, J., 2019. Fake news detection via NLP is vulnerable to adversarial attacks. arXiv preprint arXiv:1901.09657.
https://arxiv.org/pdf/1901.09657

[7] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L. and Stoyanov, V., 2019. Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692.
https://arxiv.org/pdf/1907.11692.pdf

Appendix

Figure:- The graph shows the class distribution of Real News(Label 0) and Fake News (Label 1) classes. The classes are quite well distributed, indicating there is not a lot of class imbalance to account for.

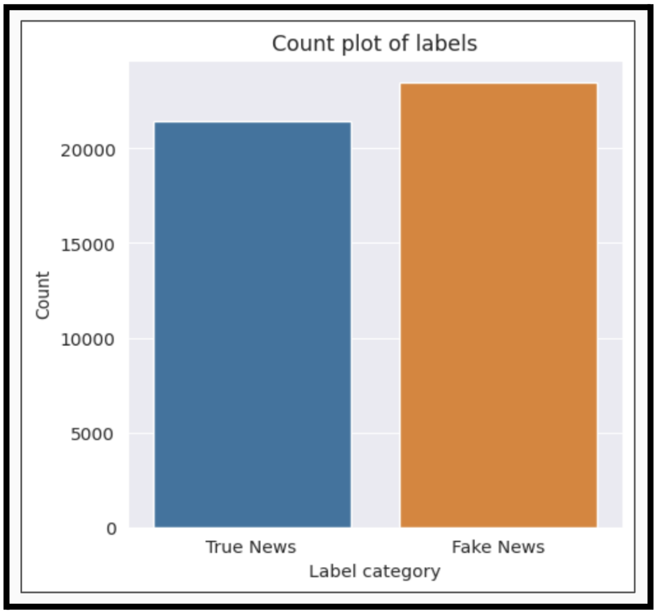


Figure:- The graph below shows the distribution of various topics to which the real and fake news articles belong.

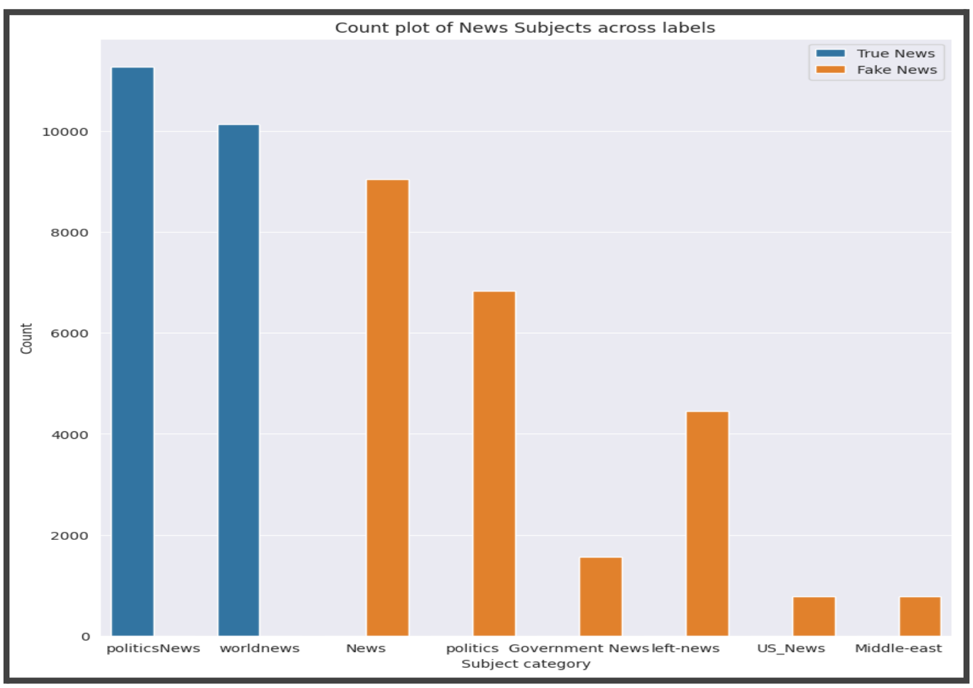


Figure:- Word Cloud generated from Real News articles in ISOT Fake News Dataset after initial preprocessing

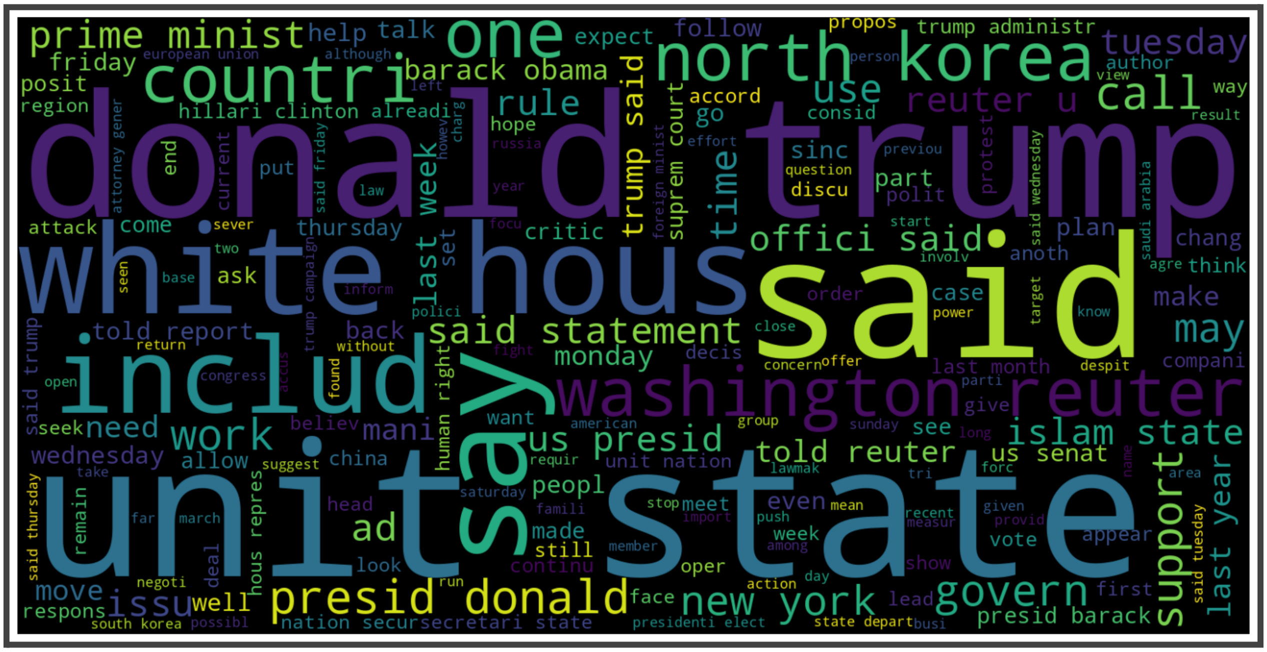


Figure:- Word Cloud generated from Fake News articles in ISOT Fake News Dataset after initial preprocessing

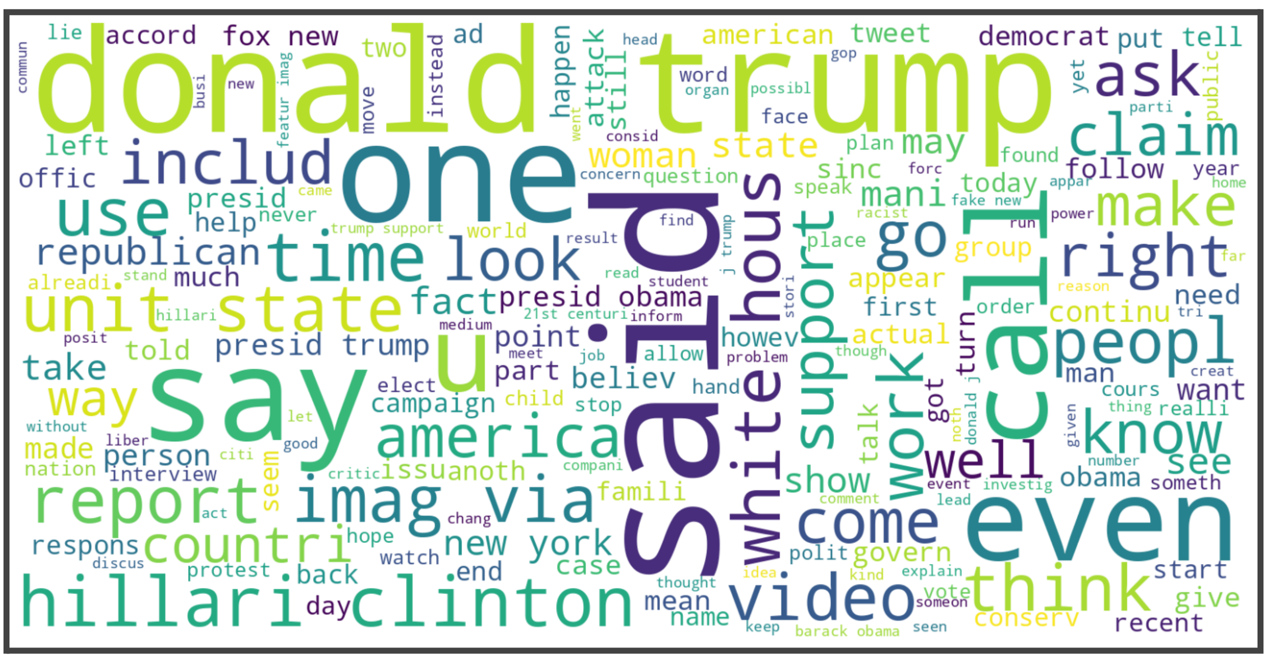


Figure:- Confusion matrix generated upon testing on the ISOT Fake News test dataset using the trained XGBoost Ensemble model (ML model) with RoBERTa (NLP Language Model).

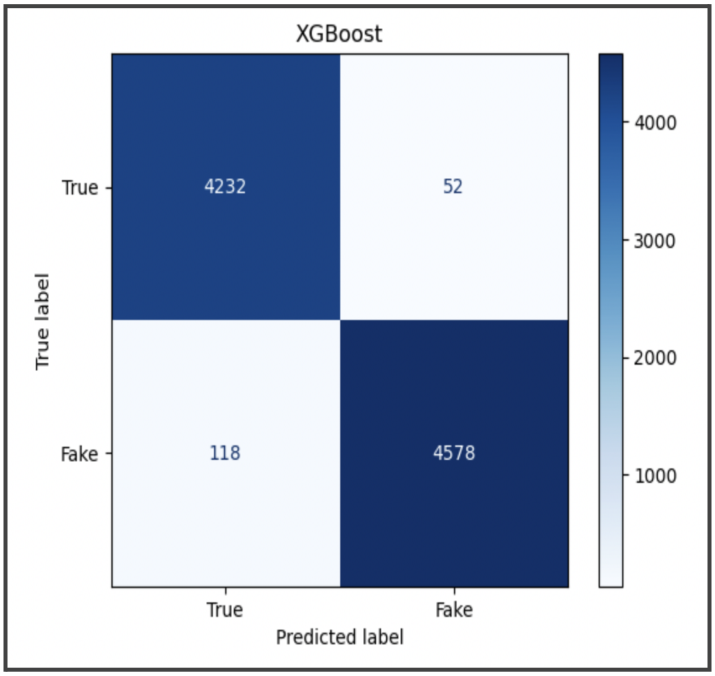


Figure:- Confusion matrix generated upon testing on the ISOT Fake News test dataset using the trained Long Short Term Memory Neural Net model (ML model) with RoBERTa (NLP Language Model).

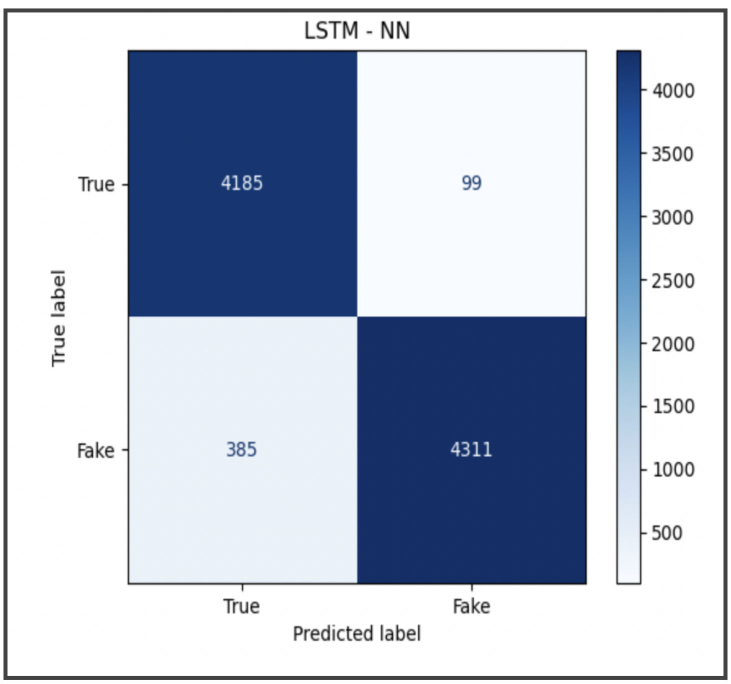


Figure:- Confusion matrix generated upon testing on the ISOT Fake News test dataset using the trained Radial Basis Function (RBF) kernel SVM (ML model) with RoBERTa (NLP Language Model).

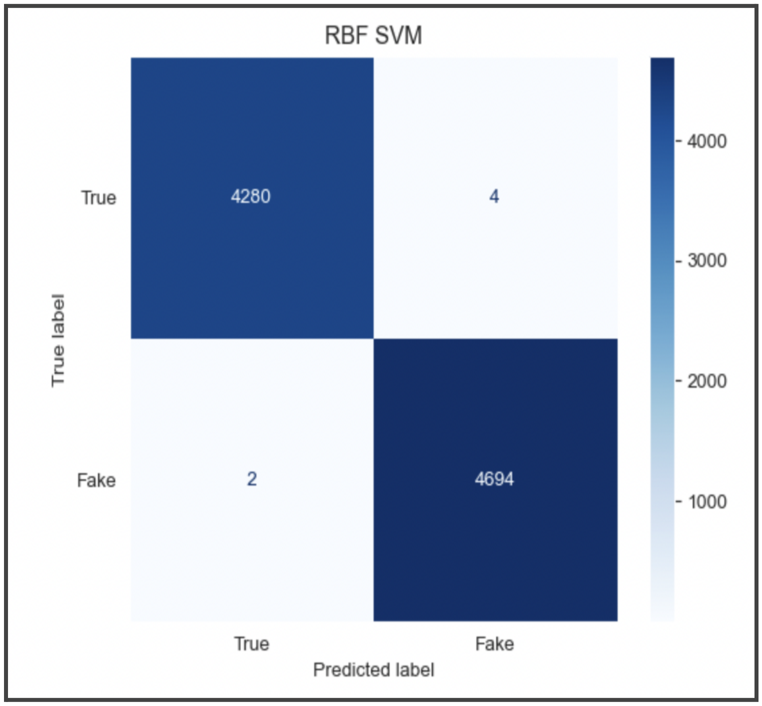
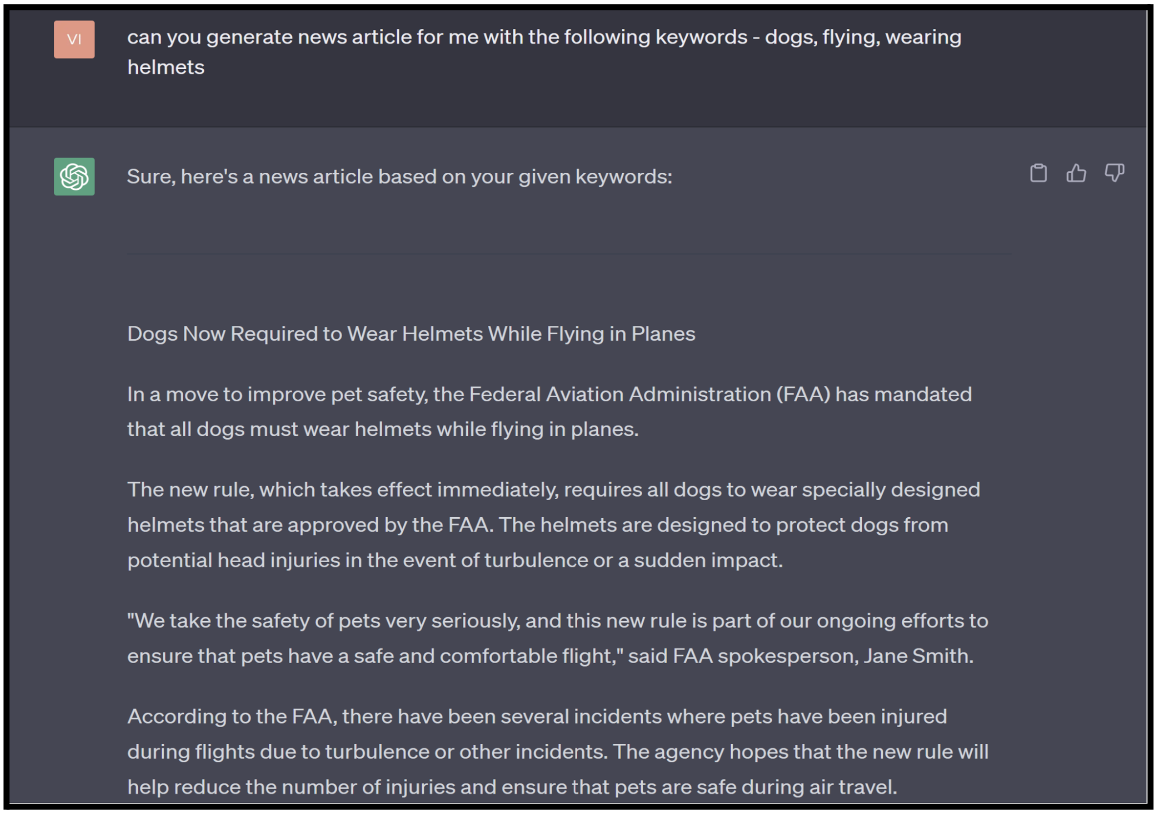


Figure:- Fake news generated by ChatGPT - “Dogs Now required to Wear Helmets While Flying in Planes”



VI

can you generate news article for me with the following keywords - dogs, flying, wearing helmets

Sure, here's a news article based on your given keywords:

Dogs Now Required to Wear Helmets While Flying in Planes

In a move to improve pet safety, the Federal Aviation Administration (FAA) has mandated that all dogs must wear helmets while flying in planes.

The new rule, which takes effect immediately, requires all dogs to wear specially designed helmets that are approved by the FAA. The helmets are designed to protect dogs from potential head injuries in the event of turbulence or a sudden impact.

"We take the safety of pets very seriously, and this new rule is part of our ongoing efforts to ensure that pets have a safe and comfortable flight," said FAA spokesperson, Jane Smith.

According to the FAA, there have been several incidents where pets have been injured during flights due to turbulence or other incidents. The agency hopes that the new rule will help reduce the number of injuries and ensure that pets are safe during air travel.

Figure:- A sample Generated by the “title” prompt I created using GPT-J.

Can you spot the synthetically generated fake news article?

Brazil seeks to revoke asylum of Italian ex-guerrilla convicted of murder

A former Italian leftist guerrilla and former member of the Italian Red Brigades who escaped prison while serving a 15-year sentence for a 1978 murder is facing expulsion from Brazil after prosecutors there launched an investigation into his claim for asylum.

Fabio Tullo, a former member of the Italian Communist Party (PCI) as well as of the Italian People’s Liberation Movement (ML), was sentenced in March 1981 to 15 years behind bars after a military court in Rome found him guilty of murdering in 1978, three years after escaping Italy during a police raid in which six people were arrested.He escaped after being transferred to San Paolo prison in Milan.

Brazil seeks to revoke asylum of Italian ex-guerrilla convicted of murder

The Brazilian government told the Supreme Court on Monday that President Michel Temer has the authority to revoke the asylum status of a former left-wing guerrilla convicted of murder in Italy and extradite him at his country’s request.

Cesare Battisti committed four murders in the 1970s when he belonged to a guerilla group called Armed Proletarians for Communism, according to the Italian government. He escaped from prison in 1981 and lived in France before fleeing to Brazil to avoid being extradited to Italy.
Brazil’s Supreme Court authorized Battisti’s extradition in 2009, but he was not sent back to Italy because former leftist President Luiz Inacio Lula da Silva granted him refugee status on his last day in office in 2010.

Check if you guess it right ? - https://www.reuters.com/article/uk-brazil-italy-idUKKBN1CT022