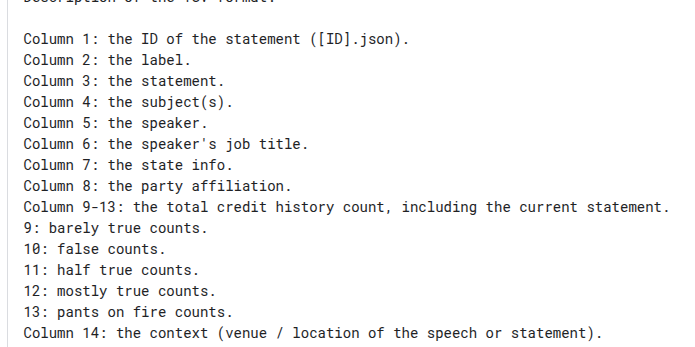
# **Final Project Interim Report**

### **1. Project Summary**

The NLP problem I'm focusing on for this project is to analyze how effective LLMs like Google's BERT can be at detecting and classifying false information. As stated before, as more and more people have access to the internet and use it as their main source of news, there is more opportunity for the injection of false information to wider audiences specifically through social media platforms. By using LLMs to detect the validity of news, I hope to find out how useful LLMs are at mitigating the effects of 'fake news'.

### **2. Data**

For this project I will be attempting to train my model using both the FakeNewsNet dataset in tandem with the LIAR dataset, both found on Kaggle.com. The FakeNewsNet dataset consists of 23,502 fake news articles with 21,417 valid articles. Each row of data consists of the title of the article, article body, subject of the article, and the date the article was published. The LIAR dataset is still a work in progress as I'm trying to find the total true counts and false counts of the data provided. For a more detailed view of the data structuring of the dataset please see the below screenshot.



### **3. Approach**

As for my approach, I'm still trying to finalize it. My blueprint so far is to utilize the pre-trained model, BERT, and add new layers on top of it so when it comes time to train the model, I can freeze the pretrained layers (or maybe use them as I still haven't decided) and use the new layers to train my model on fake news detection. For the tools I will be using on my project: BERT (via Hugging Face) as my core model, PyCaret machine learning library, Pandas data analysis library, and PyTorch library. This is not 100% finalized but will be used as my initial toolkit when working on my fake news classifier.

### **4. Progress So Far**

If I'm being honest, none from an implementation stage yet. I feel that more time should be spent researching approaches to fake news classification through LLMs like BERT. I plan on spending most of the next day (just 1 more) finalizing my research before I start working on the classifier. Given the large time demands of one specific class I'm taking (not this one), I need to work more efficiently in the coming days as I've reached a resting point from that class in question for the next 7 days.

### **5. Next Steps**

Finish reading up on different research papers. I've already started but I do feel like there's a more optimal approach to training BERT and it's bugging me. By tomorrow, begin working on the preprocessing stage of my model after downloading BERT and all the libraries stated above. I do plan on training this model locally so there may be some unforeseen roadblocks getting the model to run initially.

### **6. Questions or Concerns**

Of all the research papers I've looked at, none of them really went into good enough detail for how to train a pre-trained model to classify new data (hence why I'm still looking through research papers). However, I did find a step-by-step blueprint for how to create a fake news classifier from an online article. While I want to utilize this blueprint from a high level (i.e., approach to take, libraries to use, etc.) given I want this classifier to work and I'm in a crunch for time, is this is frowned upon and/or am I cutting myself short? I will likely bring this up in the office hours on Thursday morning but in the meantime, I'm asking it here.