1. What are some key ideas from this podcast relevant to text sentiment analysis/authorship attribution (1, or 2) or working with large diverse data sets (3)?

* Authorship attribution has been a standard problem within Natural Language Processing (NLP)
* Advanced neural language models (NLMs) are widely used in sequence generation tasks because they are able to produce fluent and meaningful sentences. They can also be used to generate fake reviews, which can then be used to attack online review systems and influence the buying decisions of online shoppers. As Mr. Adelani mentioned, reviews can have great influence on buyers’ decisions and it is not only the content but also the volume of the reviews that plays a crucial role.
* Mr. Adelani discussed the idea of generating sentiment-preserving fake online review by using an example of a positive/negative review and generate more fake reviews similar to it. It is relevant to sentiment analysis/authorship attribution because it preserves the sentiment of the writer and it is like paraphrasing the original review.
* It requires experts to train a tailored LM for a specific topic. A low-skilled threat model can be built just by combining publicly available LMs and show that the produced fake reviews can fool both humans and machines.
* Then the GPT-2 NLM is used to generate a large number of high-quality reviews based on a review with the desired sentiment and then using a BERT based text classifier (with accuracy of 96%) to filter out reviews with undesired sentiments.
* Because none of the words in the review are modified, fluent samples like the training data can be generated from the learned distribution. A subjective evaluation with 80 participants demonstrated that this simple method can produce reviews that are as fluent as those written by people.
* Three countermeasures, Grover, GLTR, and OpenAI GPT-2 detector, were found to be difficult to accurately detect fake review.

2. How do you think the ideas discussed may be relevant in your future work?