Project Proposal

Group 47

Clickbait Detector

**Introduction and Background**

- Topic

- Literature

- Datasets

**Problem Definition**

The advent of monetization for online media created a massive market for consumer attention. Everyone who posts content on the internet is incentivised to garner as many clicks as possible and maximize the amount of time consumers spend watching. Users decide what piece of media to consume based on short titles, headlines, or thumbnails so all creators must optimize these short snippets of information to be as clickable as possible. This has led to the prevalent problem of clickbait, where creators will make misleading or greatly exaggerated titles in order to “bait” users into looking at the content.

In an effort to solve this problem, we aim to create an AI model that will detect whether or not a given title is clickbait, which could later be used to automatically block clickbait content. There has been prior work to create models that detect clickbait in specific forms of media, but we hope to create a model that will find more general insights into common clickbait techniques and work across different types of media. To that end we will explore methods such as training the model on the larger datasets we found for clickbait articles, then testing its efficacy in accurately predicting the smaller YouTube title dataset.

**Methods**

- Preprocessing methods (3+)

- ML algorithms (3+, unsupervised and supervised)

**Potential Results and Discussion**

One of the metrics that we could use to measure the performance of our models is accuracy, which is simply the fraction of correct predictions on the test set. As our task is a binary classification task, we can perform more detailed analysis by categorizing the model predictions as true or false positives and negatives. For example, if an article is clickbait but our model predicts that it is not clickbait, the prediction would be a false negative. Commonly-used metrics that are based on this categorization include precision, recall, and F1 score. Using these metrics would help us determine what exactly our model excels and struggles at. We could also visualize this information using a confusion matrix.

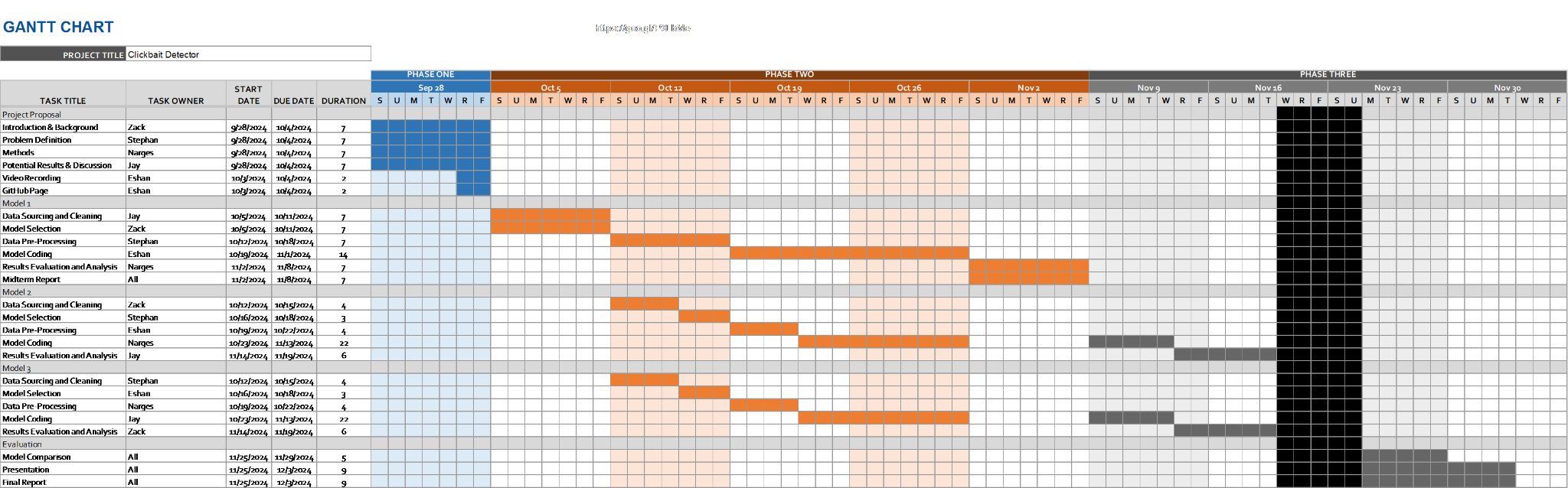
If all goes well, we expect that our final model will be able to successfully identify clickbait titles across various media. This could then be deployed in the real world, allowing users to make informed decisions about the content they engage with and discouraging the spread of deceptive and harmful practices on the internet.

**References (3+)**

- IEEE format

* **Gantt Chart**

**Link to chart:** [**Gantt Chart.xlsx**](https://gtvault-my.sharepoint.com/:x:/g/personal/stserovski3_gatech_edu/EauHMjPS8zZApkkLDNeOTaUBeoK81qgKnSyrHNyyek6beQ?e=Fddn7J)

****

* **Contribution Table**

| Name | Proposal Contributions |
| --- | --- |
| Stephan Tserovski | Problem Definition and Gantt Chart |
| Jay Cha | Potential Results and Discussion |
|  |  |
| Narges Moeini | Methods |
|  |  |