

## **Project Proposal: A Tweet Recommendation System**

### **Question/Need:**

Currently, in twitter's UI, 'Relevant Profiles' are recommended in conjunction with a tweet. Generally, they correspond to the author of the tweet, limiting broader reach to new authors and tweets.

The goal of this project is to use NLP to create a tweet recommendation system that would recommend 2-3 similar tweets to a given tweet. This would allow users to explore relevant tweets, expanding reach beyond the tweet and its author.

The practical implementation of this in Twitter's UI would be adding a new button at the bottom of a given tweet that would show the recommended tweets.

### **Data Description:**

The dataset is collected using snsrape and consists of 1000+ rows, with one row of data consisting of a tweet selected at random from three topics determined beforehand.

If time permits, additional columns will be collected/engineered, including but not limited to: sentiment analysis of the comments on a tweet, style, word count, word to image ratio...

### **Tools:**

- Text handling and Topic Modeling with NLTK/ Spacy, Sklearn
- Pandas and Numpy for data exploration
- Sklearn/Scipy for Recommender

### **MVP Goal:**

The goal for the MVP is to get to an initial set of topics from the cleaned tweets using dimensionality reduction techniques.