

- Changing project trajectory
  - Starting off with defining which news articles are considered to be “political” or “non political”
  - Referring to t-shikuro github data for inspiration behind our own machine learning model
    - T-Shikuro project analyzed political data from tweets made by political figures
- Highlighted features of t-shikuro project:
  - Defined how political tweets are by assigning percentages between “partisan” and “neutral” categories
  - Focuses on certain tweet features, such as number of characters, number of hashtags, number of unique words, etc.
    - Can fine tune these features to fit with article title features instead
  - Cleans tweets by removing punctuations, stop words to fixate on only the words
- Starting a google colab to have code specifically for our machine learning model
  - Imported kagglehub and created a path to social media post
  - First step: importing files from kaggle to be used for later
    - Figuring out API
  - Filtering out which columns in the inputted dataset for relevancy in our project:
    - Text
    - Bias
    - Bias confidence
    - Message (take note on labels)
      - policy, attack, support, etc.
    - Label (authors and who they are)
  - Things to filter out from individual tweets:
    - Hashtags
    - @
    - Punctuation
    - Make everything lower-case