

# CS577 Final Presentation

Zhuo Chen, Yijiang Zhu

# Introduction

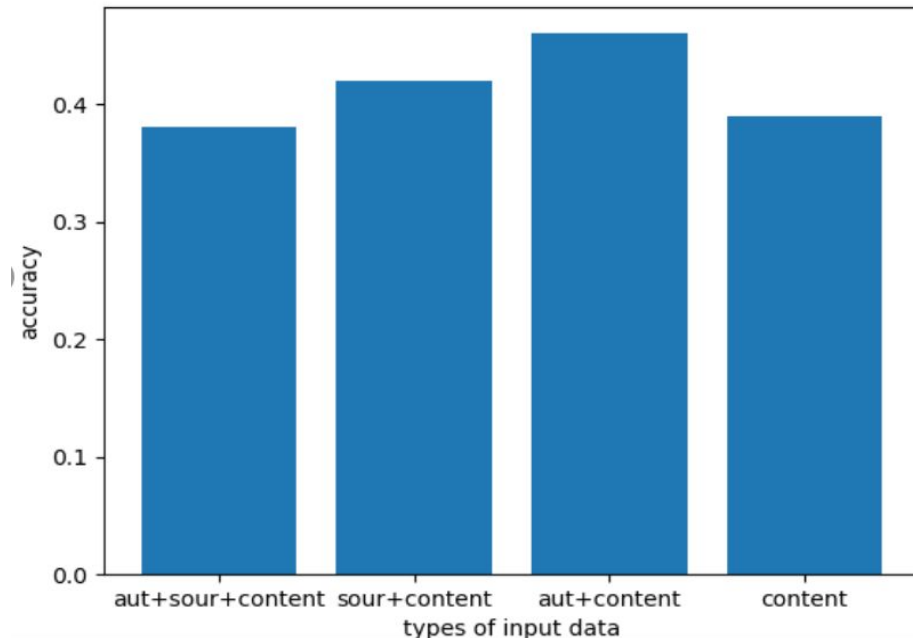
- The primary problem of our project is to identify the various factors that contribute to bias in news articles and determine how to leverage these factors to train the model.
  - By incorporating the priority of authors and sources as novel features in our model, we aim to identify the underlying factors that contribute to bias in news media and assess their relative importance in the detection of bias.
- We aim to provide valuable insights into the mechanisms that underlie bias in news media, and contribute to the development of a more nuanced understanding of the factors that influence the content and presentation of news articles.

# Methods

- Data source—from Kaggle
- Preprocessing:
  - Lower
  - Stopwords removal
  - Tokenization
  - `pad_sequences`
  - `fit_on_text`
- Feature Extraction
- Model Training
  - RNN
  - Adam optimizer

# Experiments

- A controlled variable experiment of four sets of controlled variables including:
  - content, author, and source
  - content, author
  - content, source
  - content



# Experiments

- It is easy to realize that the experimental results do not match the expectations.
- Insights on the result:
  - Simply combined the authors and sources as one input with the content cannot achieve our goal.
  - Padding sequence length is not long enough.  
We used only 8000 lines to train. Using the full set would probably be better.

# Discussion

- Bias detection is a challenging task and requires careful preprocessing of data and selection of appropriate models.
- Future work
  - Use two parallel models instead.
  - Look for other incorporating factors that might contribute towards the biases of the new content