In Defense of PBS

Gender Bias Classification as a Foundation for Headline Objectivity

Our task is to test machine learning models to predict bias in headlines.

Gender Bias as a Framework

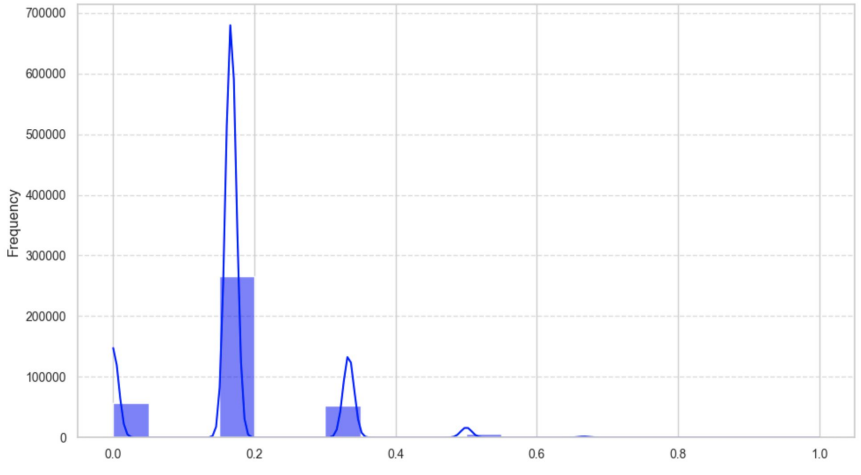
- Headlines are powerful agents that shape our perceptions.
- In the information age, gender bias is pressingly prevalent and relevant
- Models can be applied using other datasets that deal with other forms of bias

Our Data

- Headline text derived from “When Women Make Headlines” - *The Pudding*, June 2022

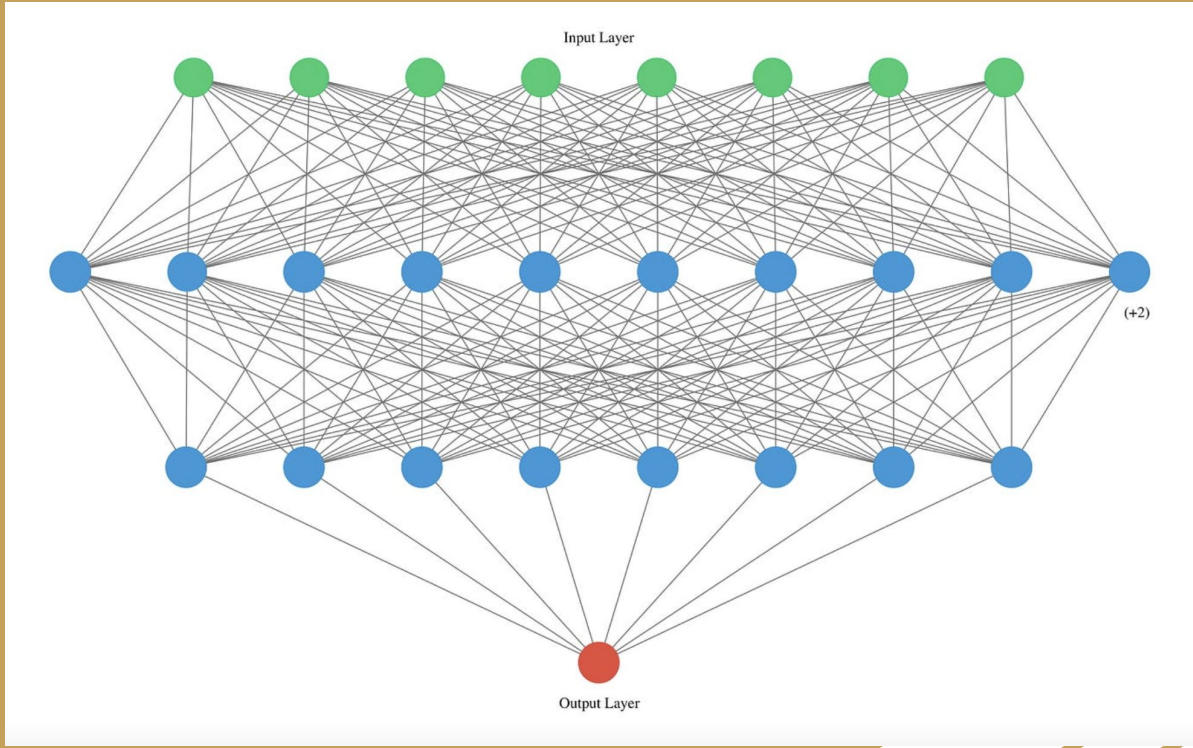
Our Target for Prediction

- Bias Categories:
  - No Bias
  - Low Bias
  - High Bias
- Class imbalance favoring Low Bias



Evaluation Metric: Weighted Average F1 Score

- Used to evaluate all model performances for consistency
- Accounts for a model's false negatives and false positives
- Provides a comprehensive, balanced measure of predictive power


Input Layer
(+2)
Output Layer

Best Performing Model: LSTM
Weighted Average F1 score:
54%

Technical Recommendations

- Continue to test and tune NNs and LSTMs
- Plot loss functions for best performing models to better understand the dynamics of the training process and make informed decisions about model architecture.

Business Recommendation: FairFrame

An app or web extension that takes headlines as input and outputs a predicted bias category, giving guided editing notes for headline writers:

- "Consider neutralizing language in this sentence."
- "Use active voice for greater clarity and objectivity."
- "Adjust the tone to ensure a balanced representation."

Next Steps

- Use engineered features, found in our notebook
- Apply Synthetic Minority Oversampling Technique (SMOTE) to address class imbalance
- Use existing PBS-specific headline text data or data from more reputable news sources

Thank you!

Contact me: mbirchhn@gmail.com | www.github.com/madelinebirch