Computational Assessment of Hyperpartisanship in News Titles

Hanjia Lyu*, Jinsheng Pan*, Zichen Wang*, and Jiebo Luo
Department of Computer Science, University of Rochester
* these authors contributed equally

Introduction

❑ Increasing exposure to media coverage of polarization could lead people to think that not only the political system, but the general public is also polarized [1]
❑ To facilitate a quantitative understanding of the extent of partisanship, we extend previous studies’ definitions of hyperpartisan news to include not only (1) the news that contains one-sided opinions but also (2) the news that describes conflicts and the underlying politically polarized climate because both of them could lead to an increase in the public’s perceived polarization [1, 2, 3]
❑ To that end, we develop a dataset to measure the partisanship in news titles. We focus on titles because it is easier for online users to access them on social platforms and decide whether to read the news content since they often summarize the opinions or events of the full article. Easily available titles that are circulated in like-minded online communities tend to become anchor points for individual opinions [2, 4], especially when individuals do not directly experience the covered events [5].

Method and Material

❑ There are three major issues that prevent us from directly using the existing dataset (SemEval) for hyperpartisan title detection
❑ Class imbalance
❑ Task-label misalignment
❑ Distribution shift
❑ We collect over 1.8 million titles posted by nine representative media outlets of different biases – Left, Central, and Right from January 2014 to September 2022
❑ To label titles more efficiently, we develop our dataset in an active learning manner (until the process converges)

References