

ABSTRACT

The credibility of news cannot be isolated from that of its source. The concept of credibility is mainly associated with a news source's trustworthiness and expertise. In an effort to measure the trustworthiness of a news source, several factors related to news content must be addressed. Among these; “is biased or unbiased”, “does or does not separate fact and opinion”, and “is factual or opinionated” take prominence.

In this work, we propose an unsupervised probabilistic lexicon-based opinion mining approach to describe a news source as “being factual or opinionated”. Sentence-level measurements are combined to make the corpus-level decision. The idea behind the statistical approach is inspired from the relativism that each word is evaluated with its difference from the average word through the use of cumulative distribution function. In order to test the effectiveness of the approach, three different news sources are chosen. These sources exhibit important distinction in their characteristic of being opinionated. They are editorials, New York Times articles, and Reuters articles respectively. Thus, the experimental validation is done by the analysis of variance on these different groups of news. The results prove that our technique can distinguish the news articles from these groups with respect to “being factual or opinionated” in a statistically significant way.