Intelligent Systems - NLP

Sentiment Analysis on HuffPost Headlines to Track Polarization in Public Discourse

Presented by: Azeez Abdikarim

January 30th, 2022

Table of Contents

Problem Context

Reflecting on the previous decade, one attribute of society that has received a lot of attention, is a perceived rise in polarization (Boxell, 2020). While often this conversation is rooted within a conversation regarding a widening gap between political ideologies, polarization as a general term is defined as a “division into two sharply contrasting groups or sets of opinions or beliefs“. To investigate this phenomena, this paper seeks to analyze the sentiment of new article headlines published over a six year period of the 2010s (2012-2018).

Reading through a paper’s headlines is a quick way for one to get a sense of the events of a day, as well as the tone surrounding them. Due to their brevity, people often read more headlines than full articles, which therefore makes the sentiment of a headline particularly powerful in establishing society’s opinion of events. With respect to headlines, an increase in polarization may be evident if one can illustrate that the sentiment of headlines have become less ‘neutral’, and actually diverged to stronger ‘positive’ and/or ‘negative’ sentiments.

Dataset

The dataset used to conduct this analysis was found on Kaggle, and represents a set of 200,583 HuffPost (formerly Huffington Post) news headlines published between January 28th, 2012 through May 26th, 2018. HuffPost is a USA based news organization, therefore this dataset includes headlines that precede the 2016 election, as well as those during and succeeding the election cycle. Each row in the dataset is a headline that is associated with 6 columns. These columns are ‘headline’, ‘authors’, a ‘link’ to the article, ‘short_description’, publishing ‘date’, as well as a categorical feature called “category”, which labels the section of the paper in which the news article was published.
An interesting characteristic of this dataset is that the number of news articles published per day is not consistant throughout the time period of the dataset.

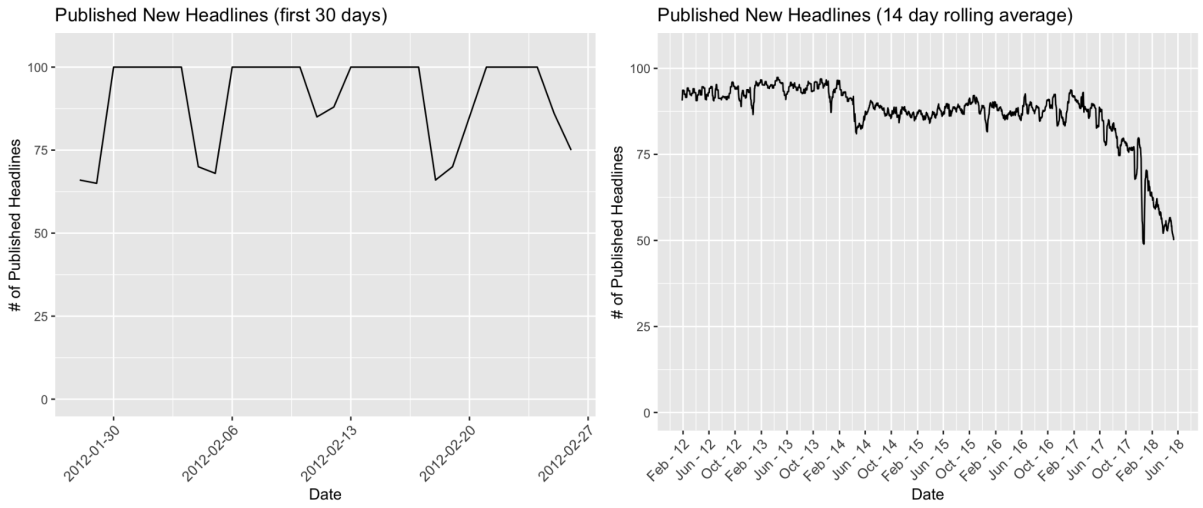

Figure 1: Count of headlines published per day.

The plots above show a raw count of the number of headlines published during the first 30 days of the dataset, along with the 14 day rolling average of the number of headlines published over the six year. These graphs show that the number of headlines published everyday are not consistent. On weekdays, around 100 headlines are published, while on weekends that number drops to around 75. This fact raises a question, how similar is the sentiment of weekday articles to that of articles published over the weekend?

It also appears that over the dataset's six year time window, there is a general downward trend in the number of published headlines, starting after April of 2017. In April of 2017, the formerly known 'Huffington Post' fully rebranded to HuffPost, and with that came changes to the paper's website, as well as the content it reported on. This should be taken into account, since it affected the publishing rate within specific categories.

Experiments

The dataset of HuffPost headlines was analyzed using R. To prepare the headlines for sentiment analysis, each headline was normalized. In this process, each headline was converted to lowercase, punctuations were removed, numbers were removed, and stop words were also removed. The sentiment for each normalized headline was then determined using the R syuzhet package, which is a sentiment analysis tool developed by researchers at Stanford. The package's *get_sentiment()* function has a 'method' parameter that allows you to specify one of four sentiment analysis methods (Syuzhet, Bing, Afinn, CRN) with which to evaluate sentences.

The difference between these methods are the lexicons used to derive the sentiment score, as well as the range of outputs that the methods return (research article). For example, the Syzuhet method produces output on the range of [-1,1], while the method Affin produces output on the range [-5,5] (Naldi, 2019). To account for this, for each method, each negative score was mapped to -1, each positive score mapped to +1, and scores of 0 represent the 'neutral' rating. Because it was uncertain which method was best suited for this context, each normalized sentence was scored by all four methods, therefore each headline received four sentiment scores.

Analysis

The two goals of the analysis conducted with this dataset, are to:

1. How has the sentiment of HuffPost headlines changed from 2012-2018?
2. How does the sentiment of headlines published on weekdays differ from those published on weekends?

To analyze this first goal, a time series graph was constructed to show how the percentage of daily news headlines with 'neutral', 'positive', and 'negative' sentiment changed over time. Due to noise in the day-to-day rates, a 14-day rolling average was applied to the sentiment rates, so that trends would be easier to identify.

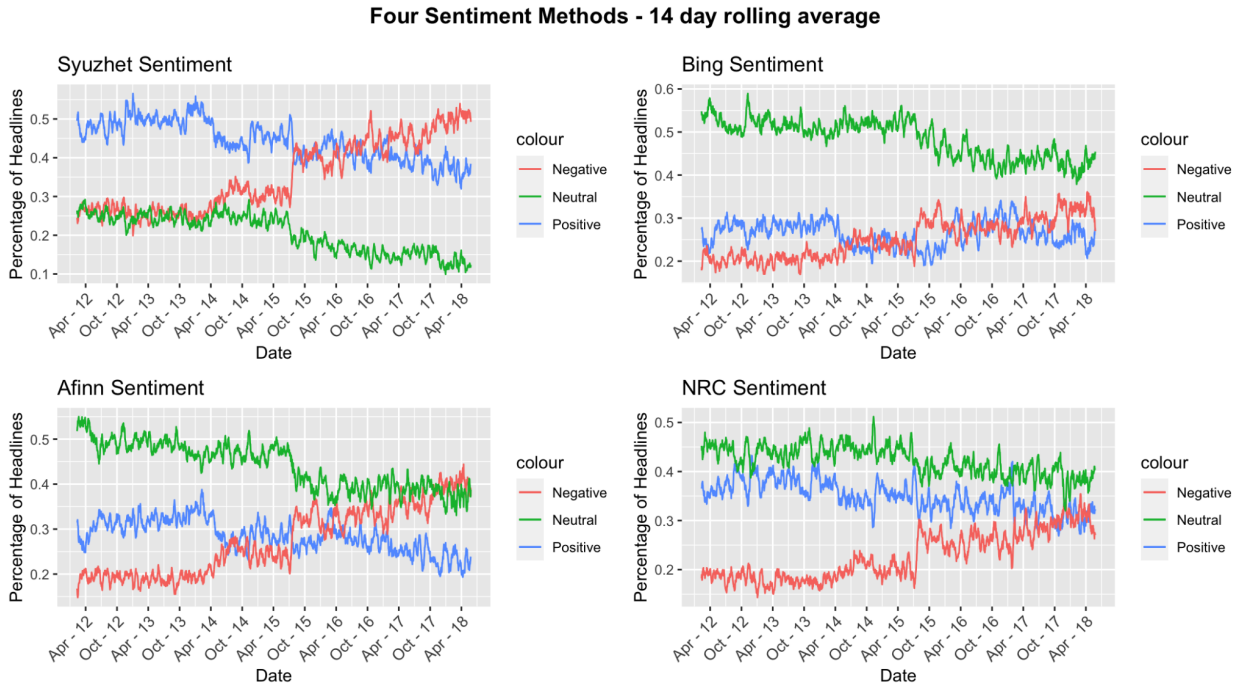

Figure 2: Comparison of 4 sentiment analysis methods on HuffPost headlines 2012-2018.

As seen in the results from all four sentiment analysis methods, there is a slight downtrend in the percentage of ‘neutral’ headlines that seems to start in the summer of 2015. In the results produced by the Syuzhet, Bing, and Afinn sentiment methods, the decrease in ‘neutral’ headlines is compensated for by an increase in ‘negative’ headlines. Although it’s difficult to prescribe this phenomena as polarization since there’s not an equal divergence in the directions of ‘positive’ and ‘negative’, the fact that headlines began to skew from ‘neutral’ towards ‘negative’ could indicate an evolution in HuffPost’s

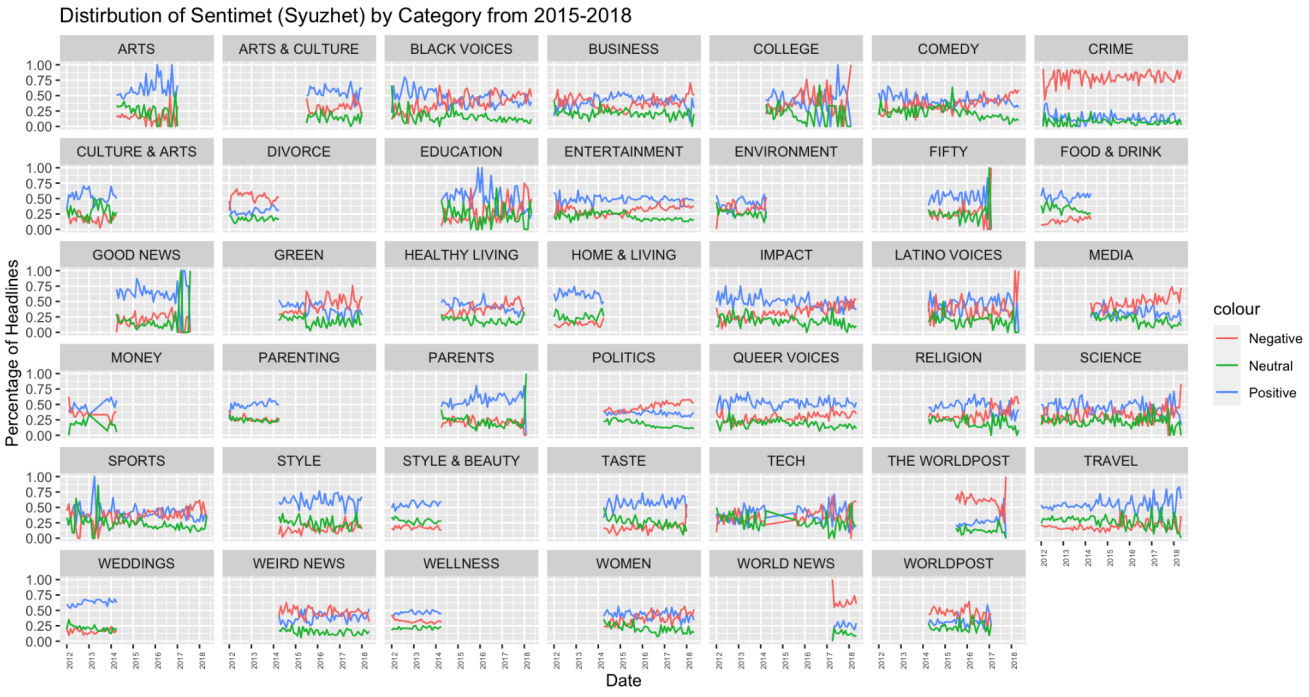

Figure 3: Sentiment evolution within categories.

perspective. For all further analysis, the Syuzhet method results are used to evaluate sentiment. This decision was made due to Syuzhet having the largest lexicon with 10,748 words, compared to the 2,477 word in the Afinn and 6,789 word in Bing.

Figure 3 illustrates how sentiment evolved within individual categories of news at HuffPost over the 2012-2018 time period. By studying the evolution of sentiment on this granular level, one can make inferences as to why the overall sentiment of the dataset became more 'negative' over time. It's interesting to note when certain sections stopped publishing headlines (i.e. Weddings, Wellness, Parenting, Environments…), and others began to. Specifically, categories with a high rate of negative sentiment headlines like The Worldpost, World News, and Politics, were only created after 2015, which could help explain the overall increase in 'negative' sentiment.

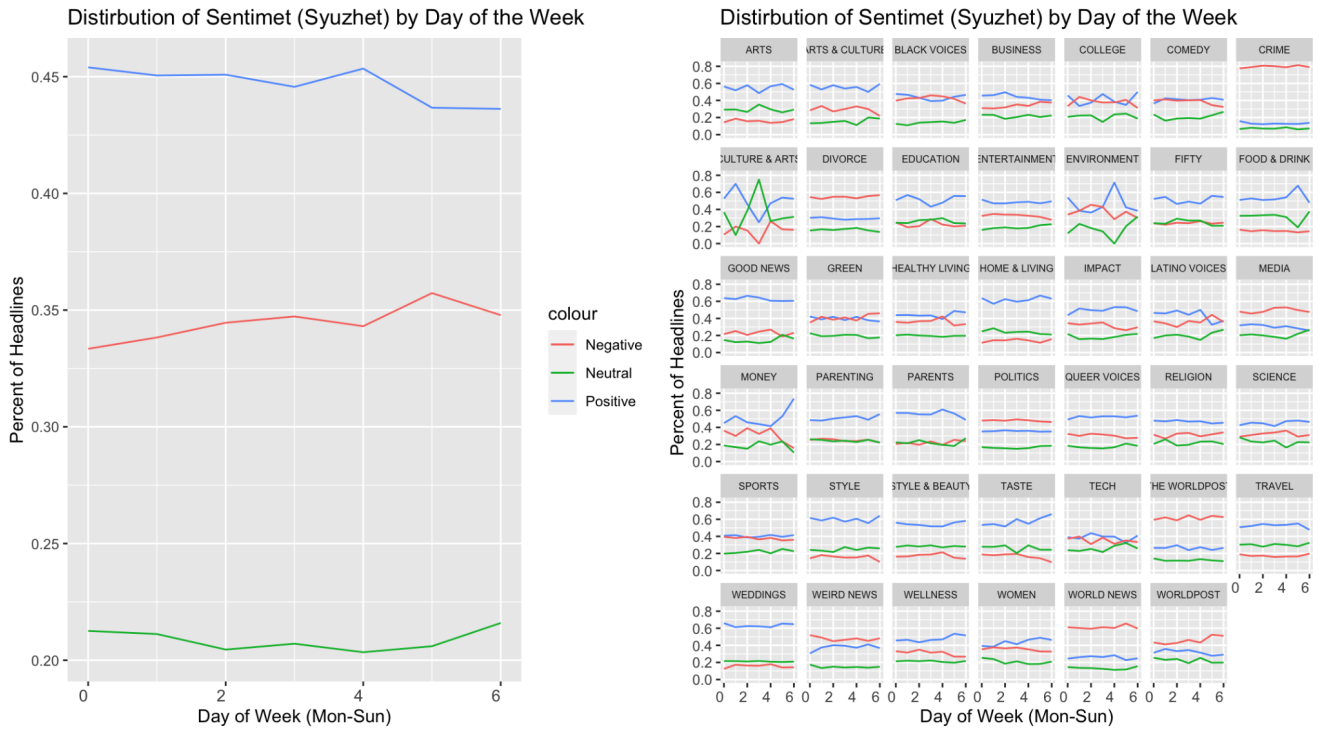

Figure 4: Sentiment evolution within categories.

To analyze how sentiment of published headlines changes throughout the week the graphs in Figure 4 were constructed. The left graph of Figure 4 shows that while sentiment is typically consistent for each weekday, the beginning of the week (Monday) tends to be more positive, however as the week progresses, average sentiment becomes slightly more negative approaching the weekend. The right graph of Figure 4 shows that for most categories, there's no difference in sentiment between articles published on the weekdays vs those published on weekends. Some categories that are notable exceptions to this trend are Environment, Food & Drink, and Mondey. In all of these categories there's a spike in positive sentiment at the tail end of the week. These trends may be indicative of some editorial strategy that prefers to use a positive headline to target weekend readers within specific categories.

References

Boxell, Levi, et al. “Cross-Country Trends in Affective Polarization.” *NBER WORKING PAPER SERIES*, 2020, https://doi.org/10.3386/w26669. Accessed 25 Jan. 2022.

Naldi, Maurizio. “A Review of Sentiment Computation Methods with R Packages.” *ResearchGate*, Jan. 2019, https://www.researchgate.net/publication/330618158_A_review_of_sentiment_computation_methods_with_R_packages.

Syuzhet Package Reference

Bing Liu, Minqing Hu and Junsheng Cheng. "Opinion Observer: Analyzing and Comparing Opinions on the Web." Proceedings of the 14th International World Wide Web conference (WWW-2005), May 10-14, 2005, Chiba, Japan.

Minqing Hu and Bing Liu. "Mining and Summarizing Customer Reviews." Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004), Aug 22-25, 2004, Seattle, Washington, USA. See: http://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html#lexicon

Saif Mohammad and Peter Turney. "Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon." In Proceedings of the NAACL-HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, June 2010, LA, California. See: http://saifmohammad.com/WebPages/lexicons.html

Finn Årup Nielsen. "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs", Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts':Big things come in small packages 718 in CEUR Workshop Proceedings : 93-98. 2011 May. http://arxiv.org/abs/1103.2903. See: http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=6010

Manning, Christopher D., Surdeanu, Mihai, Bauer, John, Finkel, Jenny, Bethard, Steven J., and McClosky, David. 2014. The Stanford CoreNLP Natural Language Processing Toolkit. In Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, pp. 55-60. See: http://nlp.stanford.edu/software/corenlp.shtml

Richard Socher, Alex Perelygin, Jean Wu, Jason Chuang, Christopher Manning, Andrew Ng and Christopher Potts. "Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank Conference on Empirical Methods in Natural Language Processing" (EMNLP 2013). See: http://nlp.stanford.edu/sentiment/

Dataset Reference

@dataset{dataset,
  author = {Misra, Rishabh},
  year = {2018},
  month = {06},
  pages = {},
  title = {News Category Dataset},
  doi = {10.13140/RG.2.2.20331.18729}
}