Clickbait Exposed: Sentiment Analysis as a Clickbait Detection Method

Name: Vadim Boot

Student number: 4847253

Date: 19 April 2024

Supervisors: Dr. Donya Alinejad, Dr. Fabian Ferrari

Index

Ch. 1: Introduction

As digital news platforms have been rapidly gaining popularity over the past decade, so has the skepticism towards them compared to their traditional media counterparts.[1] There are numerous factors that contribute to this distrust in digital news platforms, from the mishandling of sensitive user data, to the publication of false or misleading information like often associated with sensationalism.[2] [3] The sensationalism of headlines - also known as clickbaiting - by news publications, is a practice that is both observed in new, and traditional media.[4] However, clickbait has become a term synonymous with digital news media as opposed to traditional media due to different reasons. One way in which digital media developed this stereotypical characterization was as a result of the new challenges and opportunities that arose during the transition towards the digital space.

In an era where competition for online attention is cutthroat, digital news platforms are often tempted to over-sensationalize their article headlines in the hopes of getting more clicks, at the cost of the public's trust in digital news media. Previous research on the effects of clickbait headlines in relation to source credibility showed a direct correlation between the use of clickbait headlines and a negative effect on perceived source credibility, making the practice potentially harmful to a news platforms long term success.[5] This makes the development of an accurate and efficient clickbait detection tool extremely valuable as it not only provides a method for detecting the phenomenon for research purposes, but it also provides valuable insight into the motivating forces behind digital news credibility.

This however is no easy feat as clickbait signifies an extremely complex, and multi layered concept with still no clearly defined characteristics. Scholars have been debating over the essence of clickbait ever since the phenomenon first started to appear. Over the years clickbait has been hypothesized to be a lot of things, from "false news" to it "playing a queer

---

[1] Camila Mont'Alverne, "The Trust Gap: How and Why News on Digital Platforms Is Viewed More Sceptically versus News in General," ORA - Oxford University Research Archive, 2022, https://ora.ox.ac.uk/objects/uuid:42cc0bd8-f737-4a79-947f-e528e8116926.
[2] V. Kaushal and K. Vemuri, "Clickbait—Trust and Credibility of Digital News," IEEE Transactions on Technology and Society 2, no. 3 (September 2021): 146-154, doi: 10.1109/TTS.2021.3073464.
[3] Nayla Fawzi et al., "Concepts, Causes and Consequences of Trust in News Media – a Literature Review and Framework," Annals of the International Communication Association 45, no. 2 (April 3, 2021): 154–74, https://doi.org/10.1080/23808985.2021.1960181.
[4] Ángela Bazaco, Marta Redondo, and Pilar Sánchez-García, "Clickbait as a strategy of viral journalism: conceptualisation and methods," Revista Latina de Comunicación Social 74 (2019): 94.
[5] Brett Bowlin, "Investigating the Influence of Clickbait News Headlines - Center for Media Engagement," Center for Media Engagement, July 25, 2022, https://mediaengagement.org/research/clickbait-headlines/.

role in twenty-first-century performances of knowledge".[6] [7] This last phrase specifically refers to the claim that clickbait can have a non-traditional influence on how knowledge is presented and understood in the digital age. According to the Oxford Dictionary, which features the word clickbait since August 2014, clickbait can be defined as "Material put on the internet in order to attract attention and encourage visitors to click on a link to a particular web page".[8] Which would practically qualify 99% of the websites online as clickbait websites, further illustrating the degree to which the exact demarcation of clickbait is still misunderstood.

Over the past decade there have been numerous attempts at utilizing digital text analysis models for the development of a clickbait detection tool. Their main appeal comes from the fact that they function in a methodical and reliable manner that could easily be applied to extensive datasets, providing results significantly faster compared to a manual review for example.[9] One avenue within this sphere of digital text analysis models that has yet to be utilized for the detection of clickbait in academic research is text sentiment analysis. Text sentiment analysis consists of giving a computational sentiment analysis model a group of words as input, and the model in turn giving a score between -1 and 1 as output, this output is also known as a polarity score. The higher the polarity score, the more positive the perceived sentiment of the submitted group of words. The lower the polarity score, the more negative the perceived sentiment, with 0 signifying a neutral sentiment.

This study makes an attempt at contributing to the academic discussion surrounding the relationship between clickbait and digital news credibility by analyzing whether there is a correlation between the characteristics of clickbait and extreme sentimental tones through the hypothesis that clickbait can be detected through a high sentiment discrepancy of headline and corresponding article. Unraveling the role of sentiment misalignment within the detection of clickbait could yield significant results in the detection of clickbait usage, and subsequently digital news credibility. A potential correlation between a sentiment

---

[6] Yimin Chen, Niall J. Conroy, and Victoria L. Rubin, "Misleading Online Content: Recognizing Clickbait as 'False News'," in Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection (Seattle, Washington, USA: Association for Computing Machinery, 2015), 15–19, https://doi.org/10.1145/2823465.2823467.
[7] Christine Hoffmann, "What Is Clickbait? (Check All That Apply)," in Springer eBooks, 2017, 109–28, https://doi.org/10.1007/978-3-319-63751-8_5.
[8] "Clickbait Noun - Definition, Pictures, Pronunciation and Usage Notes | Oxford Advanced Learner's Dictionary at OxfordLearnersDictionaries.com," n.d., https://www.oxfordlearnersdictionaries.com/definition/english/clickbait.
[9] Martin Potthast et al., "Clickbait Detection," in Lecture Notes in Computer Science, 2016, 810–17, https://doi.org/10.1007/978-3-319-30671-1_72.

misalignment of digital news headlines and articles, and the detection of clickbait could open up an entirely new avenue of academic research into digital news credibility as it would expose an integral part of how we perceive trustworthiness. The results will be thoroughly verified from the perspective of tool criticism theory in order to assess not only the potential of sentiment analysis as a clickbait detection tool but also its limitations and ethical considerations.[10]

1.1 Main Research Hypothesis

Comparative sentiment discrepancy analysis between headlines and corresponding articles can serve as an effective method for detecting the use of clickbait within digital news and contribute to our understanding of digital news credibility.

1.2 Sub-Research Questions

- Does a discrepancy between the emotional tone of an article title and its corresponding article text indicate clickbait?
- Are clickbait headlines characterized by extreme sentiment polarity scores?

[10] Karin van Es, Maranke Wieringa, and Mirko Tobias Schäfer. "Tool Criticism: From Digital Methods to Digital Methodology." Proceedings of the 2nd International Conference on Web Studies (WS.2 2018), October 2018, pp. 24–27, https://doi.org/10.1145/3240431.3240436.

Ch. 2: Theoretical Framework

2.1 Defining Clickbait

The term "clickbait" is often referring to some form of digital content crafted to entice readers into clicking a link or button, that is not fully representative of the underlying content.[11] The term itself generally carries a negative connotation within popular culture and has been shown to negatively affect perceived source credibility.[12] [13] However, the characteristics that define clickbait are still widely debated due to the complexity of the concept itself.

For the manual review of the effectivity of the proposed hypothesis, this study requires a demarcation of general characteristics of clickbait as a concept. However, as there is currently no general profile of what defines a headline as clickbait, a combination of different characterizations will be called upon drawn from various studies on the subject.

The list of criteria include:

- Misleading Titles: Headlines that seem to be promising more than the articles actually deliver.
- Withholding Information: Headlines that intentionally withhold information to entice the reader to click.
- Sentiment Discrepancy: A clear discrepancy between the emotional tone of the headline and the tone of the article itself.

The first criterium regarding headlines that seem to be promising more than the articles deliver is likely the most discussed characterization of clickbait as it lends itself extremely well to political discussions, thus garnering more attention. The study "Seeing Is Not Always Believing" An Exploratory Study of Clickbait in WeChat" offers insight into what characterized the headlines that attracted clicks in their specific case. Results highlighted the relation between hyperbolic headlines and user engagement, enforcing the belief that at least

[11] Martin Potthast et al., "Clickbait Detection," in Lecture Notes in Computer Science, 2016, 810, https://doi.org/10.1007/978-3-319-30671-1_72.
[12] "Clickbait Noun - Definition, Pictures, Pronunciation and Usage Notes | Oxford Advanced Learner's Dictionary at OxfordLearnersDictionaries.com," n.d., https://www.oxfordlearnersdictionaries.com/definition/english/clickbait.
[13] Ángela Bazaco, Marta Redondo, and Pilar Sánchez-García, "Clickbait as a strategy of viral journalism: conceptualisation and methods," Revista Latina de Comunicación Social 74 (2019): 104.

on WeChat clickbait is partially characterized as misleading headlines.[14] The second criterium, the withholding of information is one often referred to in literature on the subject as an “artificial curiosity gap”. Scholars elaborate on this term by describing the feeling that this type of clickbait provokes “Users get the feeling of needing to know the information which can be reached only by clicking on the link”.[15] Finally the last criterium that will act as a clickbait determinator is a clear discrepancy between the emotional tone of the headline and the corresponding article as in line with the main hypothesis of this study.

Various digital news publishing organizations have recently made attempts at governing their supposed use of clickbait through redefining the term. Ben Smith, the editor of BuzzFeed, a website featuring articles often used for the testing of clickbait detection models that features numerous headlines like “Killer Robots Are Coming And These People Are Trying To Stop Them” and “12 Amazing Cookies That Are Better Than A Boyfriend” made the claim in 2014 that BuzzFeed “doesn’t do clickbait.” meaning he most likely believes that clickbait consists of content that makes false promises to potential readers.[16] He sees what BuzzFeed does as a way to “persuade you to click because you know you’ll find a story that will satisfy your interest.”, raising an interesting debate on the very nature of clickbait. Translating the claims made by Smith to the context of the hypothesis presented in this study would suggest that his statement supports the presented hypothesis. This would be due to the fact that – if Smith’s claims are true regarding the stories of BuzzFeed matching the over-the-top headlines, and the sentiment analysis models work as predicted – the discrepancies between sentimental scores of the headlines and those of the articles would in theory be minimal, indicating legitimate non-clickbait content. According to Smith, headlines from BuzzFeed competitor Upworthy, e.g. “9 Out Of 10 Americans Are Completely Wrong About This Mind-Blowing Fact” are the real cases of clickbait and fundamentally differ from BuzzFeed’s headlines as they are not “extremely direct” and leave the reader guessing about what content they will find behind the click.

---

[14] Wenping Zhang et al., “Seeing Is Not Always Believing: An Exploratory Study of Clickbait in WeChat,” Internet Research 30, no. 3 (March 13, 2020): 1043–58, https://doi.org/10.1108/intr-09-2019-0373.

[15] Anna-Katharina Jung et al., “Click Me…! The Influence of Clickbait on User Engagement in Social Media and the Role of Digital Nudging,” PLOS ONE 17, no. 6 (2022): e0266743, https://doi.org/10.1371/journal.pone.0266743, 1.

[16] Ben Smith, “Why BuzzFeed Doesn’t Do Clickbait,” *BuzzFeed*, November 6, 2014, https://www.buzzfeed.com/bensmith/why-buzzfeed-doesnt-do-clickbait.

Ben Smith's claims regarding clickbait have since been dissected and debated by cultural studies scholar Christine Hoffmann.[17] Hoffmann believes that digital news publications like BuzzFeed, despite their "extremely direct" headlines that align with the underlying story still do partake in publishing clickbait headlines. Hoffmann's critique stems from the notion that clickbait is not solely defined by misleading headlines, but rather by its potential manipulation of a readers' attention and, or emotions, arguing that even if headlines accurately reflect the underlying content, they still attempt to get readers to click the links through sensational, or extreme language. Hoffmann essentially makes the claim that even if a headline is one hundred percent transparent about the underlying article, it should still be characterized as clickbait if it is designed to exploit a reader's curiosity or emotions. The counter argument that could be made here though would be that "extreme language" has a time and a place. Reports of natural disasters for example would be right to use language like "catastrophic" if the situation is indeed that dire, however that language does – perhaps unintentionally – play into human curiosity and emotion.

Further academic research surrounding clickbait mostly – like Hoffmann - lays an emphasis on the analysis of headlines, often neglecting the connection with the underlying piece of content and the perceived sentiment discrepancy of either. This gap within the literature is remarkable considering the fact that one could present an argument for a sensational headline not being clickbait, if the accompanying article matches the sentimental tone set by the headline like BuzzFeed's Ben Smith claimed. Therefore, a more holistic exploration that accounts for both the headline's sentiment and the underlying piece of content's sentiment is essential for a more nuanced understanding of this phenomenon.

The importance of being able to clearly define, and coincidentally detect clickbait is partially built upon the effect it has on digital news credibility. While traditionally associated with less reputable news sites, recent research from Stanford University highlights the infiltration of clickbait into more reputable journalism platforms. [18] A study carried out by Kaushal and Vemuri explored the correlation between the use of clickbait and the public's trust in digital news. [19] The study underscored the frequently encountered observation that clickbait

[17] Christine Hoffmann, "What Is Clickbait? (Check All That Apply)," in Springer eBooks, 2017, 109–28, https://doi.org/10.1007/978-3-319-63751-8_5.
[18] Angèle Christin, "Counting Clicks: Quantification and Variation in Web Journalism in the United States and France," American Journal of Sociology 123, no. 5 (March 1, 2018): 1382–1415, https://doi.org/10.1086/696137.
[19] V. Kaushal and K. Vemuri, "Clickbait—Trust and Credibility of Digital News," IEEE Transactions on Technology and Society 2, no. 3 (September 2021): 146-154, doi: 10.1109/TTS.2021.3073464.

negatively influences the perception of digital news sources, resulting in decreased public trust.

2.2 Clickbait Detection Models

Determining what is, and what isn't clickbait is currently a difficult task as there is no widely agreed upon model, or method of detection. As previously mentioned, clickbait currently mainly refers to some form of digital content crafted to entice readers into clicking a link or button, that is not fully representative of the underlying content.[20] However this description is clearly subject to interpretation as it is an extremely vague one for such a complex concept.

As technological revolution has accelerated the pace of digital news and event publication, the growing amount of data available to researchers, decision-makers, journalists and consumers alike is becoming extremely valuable.[21] This underscores the importance of an accurate tool for detecting the use of clickbait on digital news platforms, that also accounts for the underlying content of the supposed clickbait headline and their sentimental alignment. Models that fail to account for the content the headline is referring to and their sentimental alignment do not take all relevant factors into consideration. Attempting to calculate whether a headline is clickbait or not purely based off the headline is an outdated practice.

The development of computational clickbait detection models has been ongoing since at least 2016, and seems to have started when M. Potthast, S. Köpsel, B. Stein and M. Hagen became concerned with the adoption of clickbait by news publishers. They presented the first machine learning approach for clickbait detection, tested on the Twitter platform, based on a total of 215 features.[22] The model did not just look at the text of the clickbait title itself, but also at the link underneath the title, the meta information (e.g. is an image or video attached). It is important to note that the model had no way of checking the actual content that was hidden under the link itself, it merely looked at the likeliness of the link being relevant to the headline through metadata.

[20] Martin Potthast et al., "Clickbait Detection," in Lecture Notes in Computer Science, 2016, 810, https://doi.org/10.1007/978-3-319-30671-1_72.
[21] A. C. Rao, A. C. Rao, and Chaitanya Kulkarni, "A Survey on Sentiment Analysis Methods, Applications, and Challenges," Artificial Intelligence Review 55, no. 7 (February 7, 2022): 5731, https://doi.org/10.1007/s10462-022-10144-1.
[22] Martin Potthast et al., "Clickbait Detection," in Lecture Notes in Computer Science, 2016, 810–17, https://doi.org/10.1007/978-3-319-30671-1_72.

Next, there was a study in 2017 that made an attempt at detecting the stance of news headlines in relation to their corresponding articles.[23] The team from Copenhagen, Denmark developed a model that would first look at whether a particular headline/article combination was related or unrelated based on "n-gram matching of the lemmatized input using the CoreNLP Lemmatiser".[24] Only this step would give the model an accuracy of 61.51%, but the following steps that included categorizing the titles and articles as "agree", "disagee" and "discuss" (named the three-class score) and relatively smaller final tweaks raised the perceived accuracy to 89.59%. However, it is important to note that this study was carried out in a controlled environment as part of a competition, did not look at the sentiment of the headlines or articles, and merely looked at how related both factors were.

Finally in February of 2023 a team of three researchers published their study on the effectivity of utilizing a semantic analysis method (a machine learning method for uncovering meaning) for the use of identifying clickbait headlines. The study revealed an incredibly high success rate of 98% but unfortunately failed to take the articles associated with the headlines into account.[25] The dataset that the model was tested on was also dubious as it was crowdsourced and categorized as clickbait or non-clickbait purely by which news publication published the headline.

[23] Peter Bourgonje, Julian Moreno Schneider, and Georg Rehm, "From Clickbait to Fake News Detection: An Approach based on Detecting the Stance of Headlines to Articles," in Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism (Copenhagen, Denmark: Association for Computational Linguistics, 2017), 84–89.
[24] Ibid: 84.
[25] Mark Bronakowski, Mahmood Al-Khassaweneh, and Ali Al Bataineh, "Automatic Detection of Clickbait Headlines Using Semantic Analysis and Machine Learning Techniques," Applied Sciences 13, no. 4 (February 14, 2023): 2456, https://doi.org/10.3390/app13042456.

2.3 Sentiment Analysis

Looking further into digital text analysis models, sentiment analysis appears to not yet have been utilized as a possible clickbait detection tool. Sentiment analysis lends itself perfectly for being able to be used at a large scale, taking both headline and article into account, while effectively being able to determine the sentiment of a given text. The latter point holds value for potential clickbait detection methods if the presented hypothesis is assumed to be true. Sentiment analysis (S.A.) consists of "the computational study of people's opinions, attitudes and emotions towards an entity. Therefore, the target of SA is to find opinions, identify the sentiments they express, and then classify their polarity".[26] Polarity in this statement refers to the sentimental tone of the input text.

Sentiment Analysis could prove valuable as a clickbait detection tool when considering its ability to discern the sentiment of a given headline and its corresponding text. A discrepancy between the polarity score (the score given to texts by sentiment analysis models) of the two could hint at clickbait, which in turn would provide valuable information on the role of headline and article sentiment alignment within the credibility of digital news. The systematic nature of Sentiment Analysis lends itself perfectly for application to news articles like those described in the many different examples given in S. Taj's, B.B. Shaikh's, and A.F. Meghji's "Sentiment Analysis of News Articles: A Lexicon based Approach".[27] They expand onto possible use cases and functions of the sentiment analysis of news articles.

> Sentiment Analysis or Opinion Mining is a way of finding out the polarity or strength of the opinion (positive or negative) that is expressed in written text, in the case of this paper – a news article. Manual labeling of sentiment words is a time consuming process.[28]

The authors continue to detail the various methods in which sentiment analysis can best be applied to news articles, distinguishing between utilizing a lexicon of weighted words and an

[26] Walaa Medhat, Ahmed Hassan, and Hoda Korashy, "Sentiment Analysis Algorithms and Applications: A Survey," Ain Shams Engineering Journal 5, no. 4 (December 1, 2014): 1093, https://doi.org/10.1016/j.asej.2014.04.011.
[27] Soonh Taj, Baby Bakhtawer Shaikh, Areej Fatemah Megji, "Sentiment Analysis of News Articles: A Lexicon-based Approach," in 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET) (Sukkur, Pakistan: 2019), pp. 1-5, doi:10.1109/ICOMET.2019.8673428.
[28] Ibid: 1.

approach based on machine learning. A combination of both will be utilized for this study as further detailed in the methodological section.

Ch. 3: Method

The following section details the utilized approach for the case study showcasing the effectiveness of comparative sentiment analysis as a clickbait detection model. The case study's aim is to offer a glimpse into the extent in which comparative sentiment analysis can potentially serve as an effective method for detecting the use of clickbait within digital news, and the way it could contribute to our understanding of digital news credibility. It is important to note that the study at hand does not propose or claim a definitive solution towards the detection of clickbait. This specific approach merely attempts to unveil one of the many layers that make up the complex concept of clickbait.

The analysis of the datasets relies upon a tool chain consisting of a web scraping tool, a data preprocessing tool, and two sentiment analysis tools, each of these individual tools' documentation resources have been thoroughly assessed for reliability.

3.1 Web Scraping

The web scraping tool's function will concern the gathering of the corpora, consisting of a database of headlines, articles and other metadata of a reputable digital news publication, and a similar database including headlines, articles and other metadata of a less reputable digital news publication. The reputable digital news publication is classified as such by the awards for journalistic excellence they have won throughout the years, while the less reputable digital news publication is classified through its relatively high rates of perceived clickbait in both popular culture as well as previous studies on the subject. It is important to note that the reputable digital news publication differs from the less reputable digital news publication as it uses a subscription service as its main revenue model whereas the less reputable digital news publication focuses on advertisement revenue as its main source of revenue. This study will not dive deeper into the implication of these differences but can potentially instigate new avenues for future research. The reasoning behind the inclusion of these two specific publications in the case study is the idea that results gathered from the less reputable dataset will be amplified compared to the results from the more reputable dataset.

Delving into the gray legal area of the web scraping of intellectual property, first the "User Agreement" of the publications were analyzed. In order to comply with the statement below as well as possible, all mentions of the publications will be redacted or changed to "Reputable Source" and "Non-Reputable Source". The agreement states that the publication prohibits:

> "Absent explicit prior written consent in certain situations, you may not, nor may you allow, enable, authorize, instruct, encourage, assist, suggest, inform, or promote that others, directly or indirectly, do any of the following for any reason: copy, harvest, crawl, index, scrape, spider, mine, gather, extract, compile, obtain, aggregate, capture, access, store, or republish any Content on or through the Service, including by an automated or manual process or otherwise, for any and all purposes other than indexing Content for inclusion in a Search Engine, including but not limited to any purpose related to data mining and/or the training or operation of any software or service to the extent that it incorporates a large language model, foundation model, deep machine learning, generative artificial intelligence, or any other process of a nature commonly referred to as artificial intelligence;"[29]

The web scraping of the digital news publishers was carried out through the "Web Scraper" browser extension.[30] This tool was chosen after an exhaustive review of various different options. The main motivating force behind the use of this specific web scraping tool is its option for "headed" web scraping as opposed to "headless", where the browser runs and collects data in the background without a GUI (graphical user interface) visible.[31] The use of a "headed" web scraping tool allows for direct authentication to the subscription account for the reputable source that will be used to circumvent articles being locked behind the paywall that the "headless" web scraping tool cannot pass.[32]

[29] "REDACTED," REDACTED, n.d., REDACTED/user-agreement.
[30] Web Scraper - The #1 Web Scraping Extension," n.d., https://www.webscraper.io/.
[31] A. S. Bale, N. Ghorpade, R. S, S. Kamalesh, R. R and R. B. S, "Web Scraping Approaches and their Performance on Modern Websites," presented at the 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2022, pp. 956-959, doi: 10.1109/ICESC54411.2022.9885689.
[32] Helle Sjøvaag, "Introducing the Paywall," Journalism Practice 10, no. 3 (March 10, 2015): 304–22, https://doi.org/10.1080/17512786.2015.1017595.

3.2 Data Preprocessing

The initial dataset was subject to digital preprocessing that consisted of removing any links, html code and irregularities. The datasets were each limited to a total of 7400 articles with publication dates ranging between 2014 and 2024 the data preprocessing tool consists of the Visual Code Studio Integrated Development Environment (IDE) running programming language Python, supported by GitHub's coding assistant Copilot, in which the data was cleaned and organized into new CSV files.[33] [34]

3.3 Evaluating Construct Validity Method

In order to validate whether the selected sentiment analysis models are in fact being applied to one dataset containing clickbait headlines and one dataset without, this study applies a manual check involving the scoring of exactly one hundred randomly selected entries per dataset. The manual check emphasizes the importance of construct validity within this study by assessing the rate of clickbait headlines within the reputable, and non-reputable dataset.

3.3.1 Method of Selection and Manual Scoring

To mitigate potential biases and in the pursuit of gathering a representative sample, a Python script was called upon to randomly select exactly one hundred entries of both the reputable, and non-reputable (preprocessed) datasets.[35] These two hundred total entries were then combined into a new, shuffled dataset (CSV) without any columns apart from the ones containing the headline and the article text. One new column was added to each entry, "Clickbait" which was then manually scored "Yes", or "No".

3.3.2 Clickbait Determination Protocol

At the time of this study, there is still no generally agreed upon model, or method for the detection or identification of clickbait, as mentioned in chapter 2.2, Clickbait Detection Methods. Previous studies on the use of computational models for clickbait detection tested their methods' detection accuracy on either human verified clickbait datasets, or simply datasets of publications infamous for their use of clickbait.

Given the subjective nature of defining clickbait, this study utilized a set of predetermined criteria for the manual clickbait validation, further detailed in Ch. 2.1 Defining Clickbait.

---

[33] See Appendix I.
[34] GitHub Copilot · Your AI Pair Programmer," GitHub, 2024, https://github.com/features/copilot.
[35] See Appendix II.

Entries that are aligned with one, or more criteria are flagged as clickbait. The list of criteria include:

- Misleading Titles: Headlines that seem to be promising more than the articles actually deliver.
- Withholding Information: Headlines that intentionally withhold information to entice the reader to click.
- Sentiment Discrepancy: A clear discrepancy between the emotional tone of the headline and the tone of the article itself.

3.4 Sentiment Analysis

3.4.1 Sentiment analysis Models

The sentiment analysis tools involved in this study include the NLP library TextBlob, and VADER.[36] [37] [38] The TextBlob model can return two different scores based on text input. One being the polarity score, further detailed in subchapter 3.4.2 Polarity Score, and one being a subjectivity score that is not taken into consideration for this study, ranging between 0 and 1, meant to help distinguish between subjective and objective text input. TextBlob is known for its straightforward and approach but also tends to encounter difficulties with handling negations and struggles as a result with text input like “not the best” vs. “best”, possibly resulting in similar sentiment in special cases.[39] The specific S.A. approach of TextBlob revolves around a combination of the earlier mentioned lexicon approach, and a machine learning algorithm to determine sentiment.[40] VADER (Valence Aware Dictionary and sEntiment Reasoner) on the other hand is a lexicon and rule-based model designed specifically for web-based media.

> VADER utilizes a mix of lexical highlights (e.g., words) that are, for the most part, marked by their semantic direction as one or the other positive or negative. Thus,

---

[36] Visual Studio Code - Code Editing. Redefined,” November 3, 2021, https://code.visualstudio.com/.
[37] Welcome to Python.Org,” Python.org, December 15, 2023, https://www.python.org/.
[38] C.J. Hutto and E.E. Gilbert, "VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text," presented at the Eighth International Conference on Weblogs and Social Media (ICWSM-14), Ann Arbor, MI, June 2014.
[39] Afaf Athar, “Sentiment Analysis: VADER or TextBlob?,” Analytics Vidhya, August 4, 2022, https://www.analyticsvidhya.com/blog/2021/01/sentiment-analysis-vader-or-textblob/.
[40] Soonh Taj, Baby Bakhtawer Shaikh, Areej Fatemah Megji, "Sentiment Analysis of News Articles: A Lexicon-based Approach," in 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET) (Sukkur, Pakistan: 2019), pp. 1-5, doi:10.1109/ICOMET.2019.8673428.

> VADER not only tells about the Polarity score yet, in addition, it tells us concerning how positive or negative a conclusion is.[41]

VADER finds its strength compared to TextBlob in handling context and negations, or subtleties in sentiment, making it best suited for complex text input. VADER, like TextBlob also offers the option to calculate scores besides the polarity score like "compound score" and "subjectivity score". This study only incorporates the polarity score. The reasoning behind the inclusion of a combination of the forementioned models in this study stem from the need for a robust and dependable analysis. The slightly varying analysis methods that the models follow will offer a more nuanced understanding of any results encountered.

3.4.2 Polarity Score

The dataset consists of a total of 8 columns: id, Starting Link, Headline, Description, Article Link, Article Text, Publication Date, Author and Category. The description column is made up of the short paragraph introducing the article. Each individual dataset entry will receive a polarity score for both the article headline, the description, and the article text. The "Headline Sentiment", "Description Sentiment". and "Article Sentiment" columns will be added to the database and the scores will be added respectively to their dataset entry's corresponding column. The polarity scores can range from a -1 to a 1, where:

- Negative Polarity (-1 to 0): Indicates a negative sentiment.
- Neutral Polarity (0): Indicates a neutral sentiment.
- Positive Polarity (0 to 1): Indicates a positive sentiment.

3.4.3 Sentiment Score Discrepancy

Next, any discrepancies between the headline, and article text polarity scores will be analyzed and compared. The reasoning behind this comparison is to investigate whether there is consistency in sentiment between article headlines and their corresponding articles texts. Inconsistency could indicate headline sensationalism – clickbait - as significant discrepancies in polarity score could indicate a deliberate mismatch to entice readers to click the articles.[42] Discrepancies between description sentiment and article text sentiment will also be tested additionally in case it offers more knowledge in the matter. This data will however be of secondary importance to the study compared to the headline, article text discrepancy scores.

---

[41] Skillcate Ai, "Sentiment Analysis — Using NLTK Vader - Skillcate AI - Medium," Medium, October 7, 2022, https://medium.com/@skillcate/sentiment-analysis-using-nltk-vader-98f67f2e6130.
[42] Martin Potthast et al., "Clickbait Detection," in Lecture Notes in Computer Science, 2016, 810, https://doi.org/10.1007/978-3-319-30671-1_72.

3.4.4 Manual Analysis of Extreme Discrepancy Outliers Method

The essence of sentiment analysis as a clickbait detection tool depends on its ability to distinguish cases of clickbait from cases of non-clickbait. To gain a more nuanced understanding of whether the envisioned approach does in fact give different output in the case of clickbait, a manual analysis was carried out of both the highest, and lowest discrepancy outliers of both datasets.

In order to make sure that the manual analysis was carried out as objectively as possible, a python script was set up to generate a new randomized csv file, containing 400 total entries.[43] These entries consisted of the one hundred highest and lowest scoring sentiment discrepancy elements of each dataset, equally split between sentiment analysis models with only the headline, description and article text column saved to the new CSV file. Again, a column called “Clickbait” was appended to the new file which could hold either a “Yes”, or a “No” value. This file was then manually scored based off of the clickbait detection criteria established in chapter: 3.3.2 Clickbait Determination Protocol.

The following phase of the study involved a comparative analysis where the manually reviewed results were juxtaposed with the discrepancy scores attributed to the dataset entries by the sentiment analysis models. Any potential correlation between extreme cases of sentiment discrepancy and manually identified clickbait was uncovered and further supported using the entries with minimal sentiment discrepancy as a control group presumed to be free of clickbait.

[43] See Appendix III.

Ch. 4: Analysis

The preprocessed CSV files containing both the reputable, and non-reputable dataset were analyzed following the sentiment analysis method detailed in chapter 3.4: Sentiment Analysis. The results are detailed and highlighted in the various sections below and further discussed in chapter 5: Discussion.

4.1 Evaluating Construct Validity

The construct validity of the study's approach was tested by manually scoring a sample of exactly one hundred entries of both the reputable and the non-reputable dataset as clickbait or non-clickbait. Each article was assessed against a set of predetermined criteria to determine whether it would be classified as clickbait. This manual check was instrumental in validating any results sprouting from the sentiment analysis.

The results of the manual scoring process revealed an expected, but significant disparity in encountered clickbait frequency between the two sampled datasets. The reputable dataset sample featured a total of 11% clickbait headlines while the non-reputable dataset featured a total of 80% clickbait headlines. This observation aligns with the literature stating that clickbait is increasingly being adopted by reputable news sources.[44] However, the fact that the reputable news source is protected by a paywall and can only be accessed through a subscription should raise concerns as the entire premise behind the use of a paywall is it being an alternative source of income allowing publications to supply higher quality news without having to utilize clickbait as a competitive tool.[45] These numbers however are still relatively conservative and could also be highly affected by sample size and specific clickbait scoring criteria.

While the presence of clickbait in the reputable dataset initially might seem concerning from a research perspective as its main purpose is drawing a contrast with the non-reputable dataset, it is important to note that clickbait still clearly does not dominate the dataset. For this reason, the reputable dataset analysis should still be able to function as an effective zero measurement next to the analysis results of the non-reputable dataset that does seem to be dominated by clickbait headlines.

---

[44] Angèle Christin, "Counting Clicks: Quantification and Variation in Web Journalism in the United States and France," American Journal of Sociology 123, no. 5 (March 1, 2018): 1382–1415, https://doi.org/10.1086/696137.
[45] Jonas Nygaard Blom and Kenneth Reinecke Hansen, "Click Bait: Forward-reference as Lure in Online News Headlines," Journal of Pragmatics 76 (January 1, 2015): 87–100, https://doi.org/10.1016/j.pragma.2014.11.010.

4.2 Comparative Sentiment Analysis and Discrepancy Metrics

After conducting a detailed sentiment analysis using both the TextBlob and VADER sentiment analysis models, notable findings in the average sentiment scores and discrepancies between reputable and non-reputable news sources were observed. The average sentiment scores given by the TextBlob model were incredibly similar for both the reputable- as the non-reputable dataset, hovering around a 0.08 polarity score.[46] For this reason, the maximum observed discrepancy or this model was only -0.04 for both datasets.

This observation was mimicked by the numbers originating from the analysis through the VADER sentiment analysis model. Although this model returned slightly higher discrepancies maxing out at -0.59 for the headline/article text discrepancy of the reputable dataset, the corresponding discrepancy for the non-reputable dataset was within 0.28.[47] This difference, even though minimal, would've been expected to be the other way around where the non-reputable dataset displayed higher discrepancy numbers.

The following tables present these key metrics, rounded to two decimals for clarity.

TextBlob		
	Non-Reputable Dataset	**Reputable Dataset**
Avg. Title Sentiment	0.07	0.08
Avg. Article Text Sentiment	0.11	0.11
Avg. Description Sentiment	0.08	0.11
Title-Article Text Discrepancy	-0.04	-0.03
Title-Description Discrepancy	-0.01	-0.02

Table 1: TextBlob Analysis Key Metrics

[46] See Table 1: TextBlob Analysis Key Metrics.
[47] See Table 2: VADER Analysis Key Metrics.

VADER		
	Non-Reputable Dataset	**Reputable Dataset**
Avg. Title Sentiment	-0.02	0.03
Avg. Article Text Sentiment	0.29	0.62
Avg. Description Sentiment	0.02	0.09
Title-Article Text Discrepancy	-0.31	-0.59
Title-Description Discrepancy	-0.05	-0.06

Table 2: VADER Analysis Key Metrics

The observed average TextBlob sentiment results seem to barely be impacted by whether the dataset featured clickbait, or non-clickbait headlines. Both the sentiment scores of the reputable and non-reputable dataset are within a range of 0.03 of its counterparts, leading to an average discrepancy of only 0.04 for headline and article text sentiment scores for the non-reputable dataset, and 0.03 for the reputable dataset. The headline and article description average sentiment score discrepancy is even lower coming in at respectively 0.01 and 0.02 for the non-reputable and reputable dataset.

For the VADER sentiment analysis model this contrast is slightly larger but still relatively similar with a maximum difference of 0.33 between the headline, article text, or description sentiment score between the two datasets. This makes the average sentiment score discrepancies of the non-reputable dataset 0.31 for the headline and article text, and 0.05 for the headline and article description. For the reputable dataset these numbers were 0.59 for the headline and article text average discrepancy and 0.06 for the headline and article description average discrepancy. To illustrate the subtle differences in sentiment discrepancies between the headlines and article texts of both reputable and non-reputable sources, a bar chart is visualized below.[48]

This visual representation clarifies the fact that on average, the sentiment discrepancies between headline and article text for both datasets do not exhibit notable divergences. The left side of the figure represents the TextBlob model, while the right side represents the VADER model. The green tinted bars indicate the average sentiment discrepancy for the

[48] See Figure 1: Average Sentiment Discrepancy

reputable dataset while the red tinted bars indicate the average sentiment discrepancy of the non-reputable dataset.

The fact that the TextBlob model scored slightly higher on the headlines, while the VADER model did this for the article texts themselves could be attributed to the previously mentioned strengths and weaknesses of both models. The strength of the TextBlob model is in the analysis of short texts, making it more likely to accurately detect sentiment in the headlines. The opposite should be true for the VADER model, that excels in the analysis of context of larger inputs, making it more suitable for the article texts. However, it is important to note that this assumption is highly speculative as there is no clear indicator that a more accurate

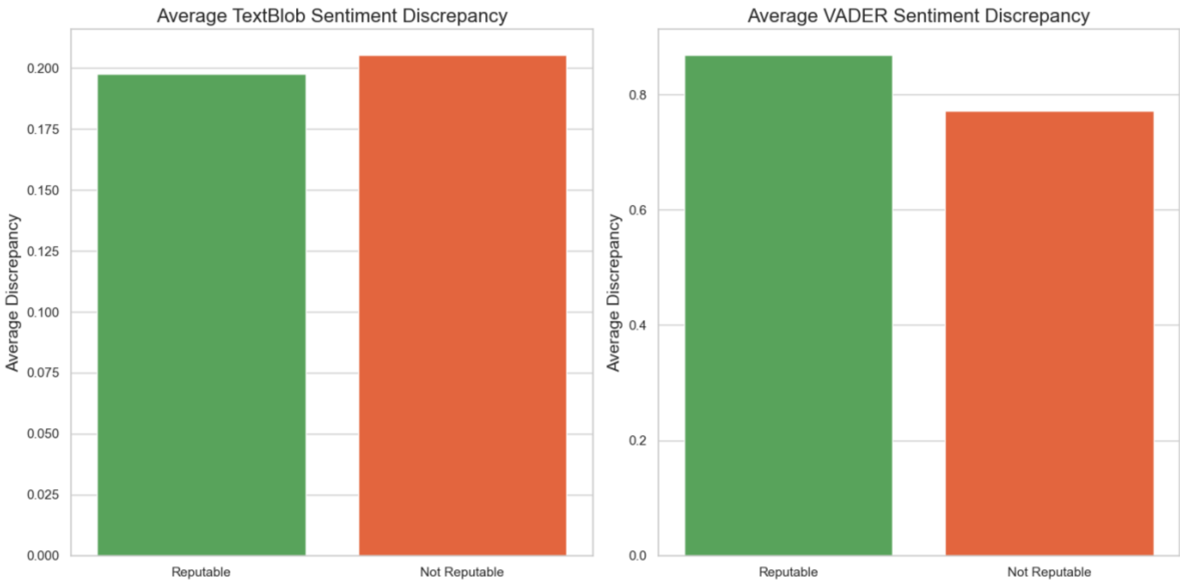

Figure 1: Average Sentiment Discrepancy

4.3 Distribution Range of Discrepancies

Whereas the previously presented average sentiment scores, and discrepancy metrics of the various datasets and sentiment analysis models offer a glimpse into the full analysis, a more detailed breakdown on the full distribution range of the discrepancies between headline and article text, and headline and article description are seen below in the box and whisker plots. The boxes themselves represent the interquartile range (IQR) of the analyzed metrics, meaning the middle 50% of all entries. Whereas Q1 and Q3 represent the 25th, and 75th percentile respectively. This means that the boxes represent where the bulk of the data points lie.

Next, the horizontal line within the boxes represents the median of the data points, indicating the middle value if the data points are arranged in ascending order. The box is divided into two halves by the median line. The top half of the box, between the median and Q3, shows the distribution of the upper 50% of the data within the IQR. The bottom half, between the median and Q1, shows the distribution of the lower 50% of the data within the IQR. If these halves appear to be of different sizes, it indicates skewness in the data distribution; a larger upper half suggests a skew towards lower values, and a larger bottom half suggests a skew towards higher values. The whiskers (lines coming from the box) show the variability outside the upper and lower quartiles, offering a view of the range of the data. They extend to the smallest and largest values within 2 times the IQR from the Q1 and Q3, respectively. Any data points beyond this range are often considered outliers and were not visualized in the figure.

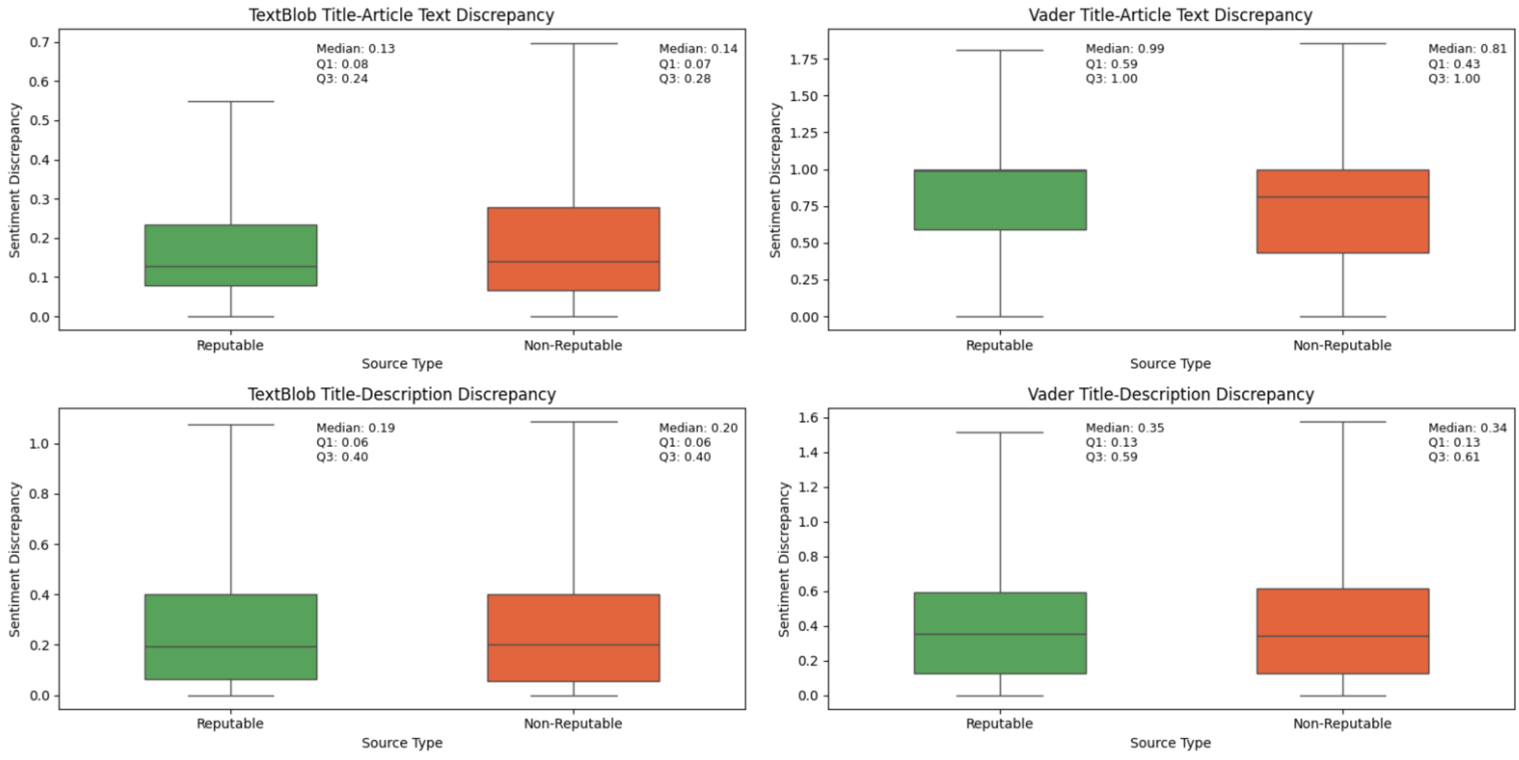

Figure 2: Average Discrepancy Box and Whisker Plot

The data highlights how even though the numbers per utilized sentiment analysis model are relatively similar per dataset, they slightly differ per model as the models use a slightly different approach. Whereas the top of the Q3 IQR of the TextBlob model maxes out at a sentiment discrepancy of about 0.4 for both the headline and article text discrepancy and the headline and description discrepancy, this number for the VADER is about 1.0. The same can be set for the whiskers above the Q3 for both models. The highest observed point for any TextBlob analysis is about 1.0. whereas its 1.75 for the VADER model.

The TextBlob models' headline/description discrepancy range for both datasets are higher compared to the headline/article text discrepancy range for both datasets, while it is the other way around or the VADER model. A proposition that could potentially explain this phenomenon would be the relatively straightforward analysis method of TextBlob being better suited for shorter text formats, while VADER overcomplicates the scoring of shorter text input but more accurately scores longer, and more complex text input like the full articles. This proposition however is pure guesswork based off of what is known of the way the models function and would have to be further analyzed to be accepted as truth.

4.4 Manual Analysis of Extreme Discrepancy Outliers

To gain a more nuanced understanding of whether high discrepancies between the sentiment score of headlines compared to their corresponding article texts were in any way indicative of clickbait, the highest, and lowest 50 discrepancy scored entries of each model per dataset were isolated into a new CSV file that was then manually reviewed for clickbait based off the criteria listed in chapter 3.3.2: Clickbait Determination Protocol.

4.4.1 The Correlation Between Manual Analysis of Clickbait and Discrepancy Score Outliers

The manually reviewed entries were given a value of either "Yes", or "No", indicating the presence or absence of clickbait based on the set of pre-established criteria. These values were then computationally compared to the entries' discrepancy score and whether or not they correlated to the high, or low discrepancy outliers of their respective sentiment model. The percentage of entries that were manually ranked as clickbait that also belonged to the high discrepancy score outliers are detailed in the boxplot below.

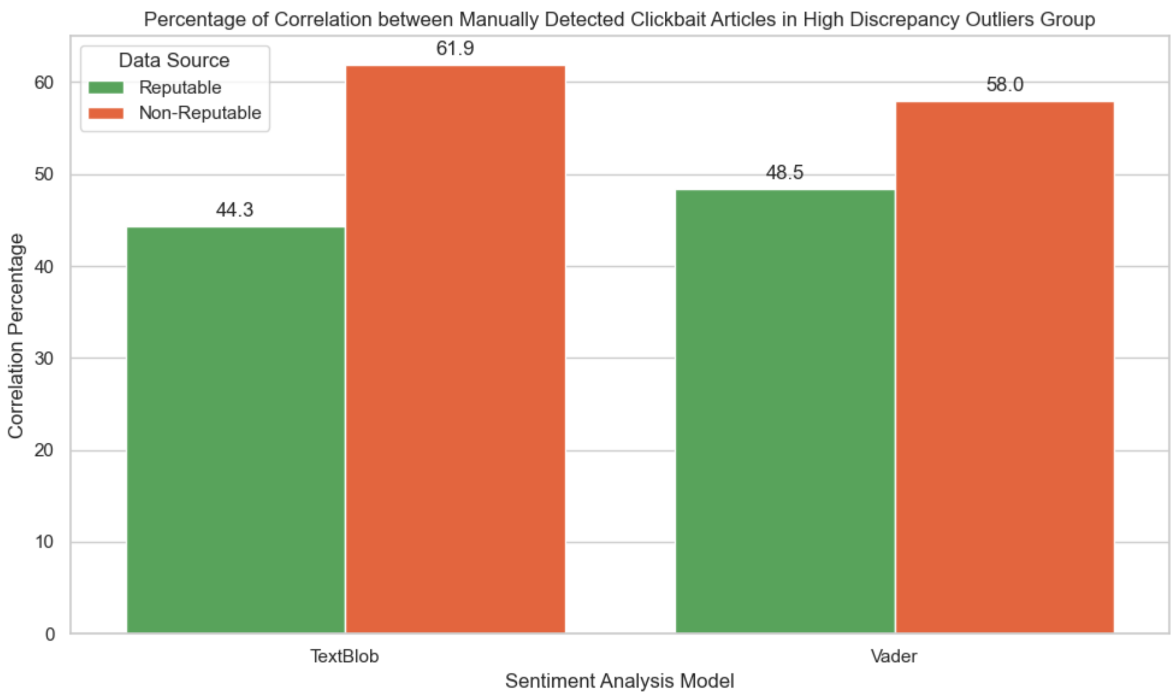

Figure 3: High Discrepancy Outliers vs. Clickbait

4.4.2 Dissecting Characteristics of High Discrepancy Entries

Further manual review of the discrepancy score outlier's dataset revealed some interesting observations regarding the characteristics of high discrepancy entries. For starters, there was an overlap of 79.10% between dataset entries that were in the high discrepancy score group for the TextBlob model, as well as the VADER model. This relatively high rate of overlap suggests that while the models do approach sentiment differently, about 80% of the time they do rate text to a similar degree of sentiment.

Another notable observation that was made after manual review was how both models tended to score overly sensationalist headlines with high sentiment scores, likely due to their usage of extreme language. However, in some of the cases where the headline scored extremely high (0.9 – 1.0 polarity score), the tone of the corresponding article matched the tone of the headline, resulting in incredibly similar sentiment scores, placing the dataset entry in the group with the overall lowest amount of sentiment discrepancy. This observation provided valuable insights into not only clickbait and the importance of a clear way to characterize it, but also the mechanisms behind the sentiment analysis models.

Finally, both models were able to detect clickbait that was also manually classified as clickbait in headline/article combinations that featured headlines with words that could be described as extremely sentiment conveying, like "devastating", "shocking", etc. and accompanying articles that were written in a relatively neutral tone of emotion. The high sentiment score that the headlines would get due to the extremely positive or negative words would create enough of a discrepancy with the articles that would score around a 0.0 sentiment score to stand out from regular headline/article combinations.

Chapter 5: Discussion

This study made an attempt at shedding light at the inner workings of the complex concept of clickbait and its effect on digital news credibility through an original clickbait detection approach involving comparative sentiment analysis. The hypothesis which served as a foundation for the entire study, was based on the idea that digital news clickbait in written form could be detected through a sentiment analysis discrepancy between the headline and the corresponding article. The reasoning behind the envisioned use of this specific computational type of model in the detection of clickbait was the assumption that sentiment analysis models would efficiently and systematically be able to detect an emotional disconnect between headlines and articles. To gather results, the study implemented both the TextBlob and VADER sentiment analysis models as analysis tools for one dataset comprised of headlines, articles and other metadata of a reputable digital news platform, and one dataset comprised of headlines, articles and other metadata of a non-reputable digital news platform.

5.1 Construct Validity and Averages

The results laid bare in Chapter 4: Analysis, paint a nuanced picture of the true complexity of the concept of clickbait. Following the evaluation of the construct validity of the study through manually rating a sample of both datasets for clickbait, the generalization was made that about 80%~ of the non-reputable dataset can be qualified as clickbait, while about 10%~ of the reputable dataset consists of clickbait headlines. Keeping these numbers in mind, the average sentiment and discrepancy scores of both datasets can appear extremely underwhelming with the average sentiment discrepancy results of the TextBlob model coming in at -0.04 for the non-reputable dataset, and -0.03 for the reputable dataset, and the VADER model scoring a -0.31 for the non-reputable dataset, and -0.59 for the reputable dataset. The slightly varying numbers per model were most likely due to the differences in the models' ability to read context and their strengths and weaknesses. VADER should in theory perform better on complex, longer texts (article texts), while TextBlob should excel when given shorter, more straightforward texts (headlines). Looking at these averages, with the assumption that about 80%~ of the non-reputable dataset should set off some indicators in the non-reputable dataset discrepancy numbers compared to the reputable dataset numbers, it became obvious that analyzing sentiment discrepancy surely did not indicate clickbait in all cases of the phenomenon. However, to gain a better understanding of whether there was knowledge to be gained about the inner workings of clickbait in relation to sentiment analysis, a more in-depth approach was necessary.

5.2 The Correlation Between Clickbait Detection and Extreme Sentiment Scores

The sentiment analysis of the manually validated high discrepancy outliers (chapter 4.4 Manual Analysis of Extreme Discrepancy Outliers) provided an avenue for a more in-depth exploration of the potential of comparative sentiment analysis as a clickbait detection method. The results stemming from this part of the analysis highlighted the fact that both models appear to be relatively sensitive to similar factors, substantiated by the 80%~ overlap in their top 50 entries per dataset with the highest discrepancy score.

One notable trend both models followed was the correct assignment of extremely high, or extremely low discrepancy scores to headlines with highly sentimental language that were incidentally also manually, and computationally marked as clickbait. These cases in isolation supported the hypothesis that comparative sentiment discrepancy analysis between headlines and corresponding articles can serve as an effective method for detecting the use of clickbait as their articles in addition did not match the sentiment score of their corresponding headlines, resulting in high discrepancy scores. However, complexities around the hypothesis arose when comparing these specific results to the gross of the extreme sentiment scored headlines, including those that received incredibly low sentiment discrepancy scores.

However, the manual review also uncovered some notable observations that tested the hypothesis on a foundational level as the observed phenomenon called into question the very concept of the definition of clickbait. These entries were similar to the clickbait examples just mentioned with extremely sensational headlines that were picked up by the sentiment analysis models but were different as the corresponding articles matched the overly sensational emotional tone of the headline, resulting in extremely low sentiment discrepancy scores. These types of headline and article combinations would, according to BuzzFeed's previously mentioned Ben Smith not be clickbait, whereas Christine Hoffmann would claim the opposite as the headlines perceived sentiment score signifies the use of extreme language playing into the readers' curiosity and emotions.[49] [50]

Now, in some of these cases these headlines were manually classified as non-clickbait, an example being a news report linguistically structured like "Devastating earthquake hits

[49] Ben Smith, "Why BuzzFeed Doesn't Do Clickbait," BuzzFeed, November 6, 2014, https://www.buzzfeed.com/bensmith/why-buzzfeed-doesnt-do-clickbait.
[50] Christine Hoffmann, "What Is Clickbait? (Check All That Apply)," in Springer eBooks, 2017, 109–28, https://doi.org/10.1007/978-3-319-63751-8_5.

location” where the article continues in the same emotional tone like for example “The damage done is immense”. But in other cases, this phenomenon was observed in very obvious examples of clickbait like “What happened at this event was SHOCKING”, where the article was written in a similar sensational matter, resulting in the analysis not presenting this article as possible clickbait due to a low discrepancy score.

5.3 Comparative Sentiment Discrepancy Analysis as Clickbait Detection Tool

Overall, the verdict on the hypothesis is that it’s an extremely nuanced matter as the results have shown. It appears that in most cases comparative sentiment discrepancy analysis is not able to significantly distinguish clickbait headlines, from legitimate headlines. The models were often able to successfully predict the emotional tone of a given headline or article, the issues however partially arose in the way the study defined clickbait. If, for example clickbait was defined by headlines that are not emotionally indicative of the story, then comparative sentiment analysis would be a relatively efficient indicator for clickbait in extreme scenarios. However, all cases of clickbait simply do not qualify within this limited scope. The results did hint at clickbait often being characterized by extremely high, or low scoring sentiment scores. As was the case for legitimate news that was deserving of headlines with extremely emotional language, e.g. natural disasters, again another factor complicating the computational detection of actual clickbait. These issues highlight how, even though sentiment analysis holds value in the detection of clickbait, any approach involving this computational tool should be highly sensitive to context.

Reflecting upon the sub-research questions of “Does a discrepancy between the emotional tone of an article title and its corresponding article text indicate clickbait?”, and “Are clickbait headlines characterized by extreme sentiment polarity scores?” again highlights the diverse nature of clickbait. The results indicated that both questions could be true, depending on the way in which clickbait is defined. Both questions were incidentally proven to be true within the scope of this study’s definition of clickbait. The occurrences where a discrepancy between the emotional tone of headline and article text indicated clickbait, and the cases where clickbait was characterized by extreme sentiment polarity scores compared to those where it wasn’t were a true testament to the multi-faceted, and complex nature of the concept.

Chapter 6: Conclusion and Reflection

6.1 Did the Hypothesis Hold Up?

Reflecting upon the results encountered in the study, it is clear that while the use of comparative sentiment analysis did show potential in the further understanding of the complex concept of clickbait, the envisioned approach was not able to gather any significant results supporting the presented hypothesis. Although the utilized models were able to quantify, relatively accurately the emotional tone of both headlines and article text's, the entire premise of the study was challenged by what comprised the defining factors of clickbait itself. Had clickbait been strictly defined as headlines emotionally/sentimentally misaligned with their corresponding articles, then the analysis results would have proven more promising. Despite this, the study was however able to shine a light on the intricate relationship between headlines and articles featuring extreme language, and context. Specific cases describing for example natural disasters with extreme language are often legitimate, non-clickbait headlines whether the accompanying article matches the extreme tone of the headline or not. This example suggests that while sentiment analysis can play an important role in the future of clickbait detection and our understanding of digital news credibility, more intelligent tools need to be utilized in order to take context into consideration.

6.2 Reflection

Reflecting upon the entire research process it becomes evident that even though some decisions were made in careful consideration, they may have not necessarily provided the most insights. For instance, the decision to include both the TextBlob and VADER model instead of focusing on a single model was in theory a well thought out one. Multiple models could highlight the results from slightly different angles, providing more potentially valuable data. Retrospectively it might have been wiser to limit the study to one model but include more parameters of this model into the equation like the previously mentioned compound score for example. This combination of polarity score, in relation to one of these other signifiers could have proven more fruitful results.

Another issue that shaped the eventual conclusion of the entire study was the lack of clearly demarcated clickbait characterization. Even though the study did include several characteristics that would indicate clickbait, they were too broadly defined in retrospect. More clear boundaries would have given more detailed results as the datasets could've been more easily navigated and analyzed for those specifics.

6.3 Future Research

The insights gained from this study can give way to multiple avenues for future research. For starters, future studies into the use of computational models for the detection of clickbait could encompass a more holistic approach that incorporates not solely sentiment analysis, but also other NLP tools like the ones mentioned in chapter 2.3 Clickbait Detection Models. Building on the results of not just this, but also other studies on the subject could provide an elevated starting point as researchers build a custom model that incorporates multiple methods.

Next, a similar study could be conducted that focusses more on the fluctuation of sentiment scores over the years. More data on this specific matter could provide extremely valuable based on the results as it would allow for a better understanding of dynamic journalistic practices that could also potentially impact overall clickbait rates. Clickbait detection however should not be the sole vantage point of such a study to not limit its findings.

Finally, the intricacies of what defines clickbait could be further investigated through a study where a test group rate headlines and article combinations as clickbait or non-clickbait. These results could prove extremely valuable for not just our understanding of clickbait and its effect on news credibility, but also the testing of future clickbait detection models.

References

1. Ai, Skillcate. “Sentiment Analysis — Using NLTK Vader - Skillcate AI - Medium.” Medium, October 7, 2022. https://medium.com/@skillcate/sentiment-analysis-using-nltk-vader-98f67f2e6130.

2. Athar, Afaf. “Sentiment Analysis: VADER or TextBlob?” Analytics Vidhya, August 4, 2022. https://www.analyticsvidhya.com/blog/2021/01/sentiment-analysis-vader-or-textblob/.

3. Bale, A. S., Ghorpade, N., S, R., Kamalesh, S., R, R., & B. S, R. "Web Scraping Approaches and their Performance on Modern Websites." In 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2022, pp. 956-959. doi: 10.1109/ICESC54411.2022.9885689.

4. Bazaco, Ángela, Marta Redondo, and Pilar Sánchez-García. "Clickbait as a strategy of viral journalism: conceptualisation and methods." Revista Latina de Comunicación Social 74 (2019): 94.

5. Blom, Jonas Nygaard, and Kenneth Reinecke Hansen. “Click Bait: Forward-reference as Lure in Online News Headlines.” Journal of Pragmatics 76 (January 1, 2015): 87–100. https://doi.org/10.1016/j.pragma.2014.11.010.

6. Bourgonje, Peter, Julian Moreno Schneider, and Georg Rehm. 2017. "From Clickbait to Fake News Detection: An Approach based on Detecting the Stance of Headlines to Articles." In Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism, 84–89. Copenhagen, Denmark: Association for Computational Linguistics.

7. Bowlin, Brett. “Investigating the Influence of Clickbait News Headlines - Center for Media Engagement.” Center for Media Engagement, July 25, 2022. https://mediaengagement.org/research/clickbait-headlines/.

8. Bronakowski, Mark, Mahmood Al-Khassaweneh, and Ali Al Bataineh. "Automatic Detection of Clickbait Headlines Using Semantic Analysis and Machine Learning Techniques." Applied Sciences 13, no. 4 (February 14, 2023): 2456. https://doi.org/10.3390/app13042456.

9. Chen, Yimin, Niall J. Conroy, and Victoria L. Rubin. "Misleading Online Content: Recognizing Clickbait as 'False News'." In Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection, 15–19. Seattle, Washington, USA: Association for Computing Machinery, 2015. https://doi.org/10.1145/2823465.2823467.

10. Chowdhary, K. R. "Natural Language Processing." In Springer eBooks, 603–49, 2020. https://doi.org/10.1007/978-81-322-3972-7_19.

11. Christin, Angèle. "Counting Clicks: Quantification and Variation in Web Journalism in the United States and France." American Journal of Sociology 123, no. 5 (March 1, 2018): 1382–1415. https://doi.org/10.1086/696137.

12. Clickbait Dataset," April 18, 2020. https://www.kaggle.com/datasets/amananandrai/clickbait-dataset/data.

13. "Clickbait Noun - Definition, Pictures, Pronunciation and Usage Notes | Oxford Advanced Learner's Dictionary at OxfordLearnersDictionaries.com," n.d. https://www.oxfordlearnersdictionaries.com/definition/english/clickbait.

14. Condé Nast. "Condé Nast," n.d. https://www.condenast.com/user-agreement.

15. Fawzi, Nayla, Nina Steindl, Magdalena Obermaier, Fabian Prochazka, Dorothee Arlt, Bernd Blöbaum, Marco Dohle, et al. "Concepts, Causes and Consequences of Trust in News Media – a Literature Review and Framework." Annals of the International Communication Association 45, no. 2 (April 3, 2021): 154–74. https://doi.org/10.1080/23808985.2021.1960181.

16. GitHub Copilot · Your AI Pair Programmer," GitHub, 2024, https://github.com/features/copilot.

17. Hoffmann, Christine. “What Is Clickbait? (Check All That Apply).” In Springer eBooks, 109–28, 2017. https://doi.org/10.1007/978-3-319-63751-8_5.

18. Hutto, C.J., and Gilbert, E.E. "VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text." Presented at the Eighth International Conference on Weblogs and Social Media (ICWSM-14), Ann Arbor, MI, June 2014.

19. Jung, Anna-Katharina, Stefan Stieglitz, Tobias Kißmer, Milad Mirbabaie, and Tobias Kroll. “Click Me…! The Influence of Clickbait on User Engagement in Social Media and the Role of Digital Nudging.” PloS One 17, no. 6 (June 29, 2022): e0266743. https://doi.org/10.1371/journal.pone.0266743.

20. Koolen, Marijn, Jasmijn Van Gorp, and Jacco Van Ossenbruggen. “Toward a Model for Digital Tool Criticism: Reflection as Integrative Practice.” Digital Scholarship in the Humanities 34, no. 2 (October 12, 2018): 368–85. https://doi.org/10.1093/llc/fqy048.

21. Kaushal, V., and Vemuri, K. "Clickbait—Trust and Credibility of Digital News." IEEE Transactions on Technology and Society 2, no. 3 (September 2021): 146-154. doi: 10.1109/TTS.2021.3073464.

22. Medhat, Walaa, Ahmed Hassan, and Hoda Korashy. “Sentiment Analysis Algorithms and Applications: A Survey.” Ain Shams Engineering Journal 5, no. 4 (December 1, 2014): 1093–1113. https://doi.org/10.1016/j.asej.2014.04.011.

23. Mont’Alverne, Camila. “The Trust Gap: How and Why News on Digital Platforms Is Viewed More Sceptically versus News in General.” ORA - Oxford University Research Archive, 2022. https://ora.ox.ac.uk/objects/uuid:42cc0bd8-f737-4a79-947f-e528e8116926.

24. Potthast, Martin, Sebastian Köpsel, Benno Stein, and Matthias Hagen. “Clickbait Detection.” In Lecture Notes in Computer Science, 810–17, 2016. https://doi.org/10.1007/978-3-319-30671-1_72.

25. Python.org. “Welcome to Python.Org,” December 15, 2023. https://www.python.org/.

26. Rao, A. C., A. C. Rao, and Chaitanya Kulkarni. “A Survey on Sentiment Analysis Methods, Applications, and Challenges.” Artificial Intelligence Review 55, no. 7 (February 7, 2022): 5731–80. https://doi.org/10.1007/s10462-022-10144-1.

27. Rossetto, Louis. “The Original WIRED Manifesto.” WIRED, September 18, 2018. https://www.wired.com/story/original-wired-manifesto/.

28. Scott, Ben. “A Contemporary History of Digital Journalism.” Television & New Media 6, no. 1 (February 1, 2005): 89–126. https://doi.org/10.1177/1527476403255824.

29. Sjøvaag, Helle. “Introducing the Paywall.” Journalism Practice 10, no. 3 (March 10, 2015): 304–22. https://doi.org/10.1080/17512786.2015.1017595.

30. Smith, Ben. “Why BuzzFeed Doesn’t Do Clickbait.” BuzzFeed, November 6, 2014. https://www.buzzfeed.com/bensmith/why-buzzfeed-doesnt-do-clickbait.

31. Taj, Soonh, Shaikh, Baby Bakhtawer, Megji, Areej Fatemah. “Sentiment Analysis of News Articles: A Lexicon based Approach.” In 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, Pakistan, 2019, pp. 1-5. doi:10.1109/ICOMET.2019.8673428.

32. TextBlob: Simplified Text Processing — TextBlob 0.16.0 Documentation,” n.d. https://textblob.readthedocs.io/en/dev/#.

33. The Editors of Encyclopaedia Britannica. “Wired | Technology, Science, Culture.” Encyclopedia Britannica, December 1, 2023. https://www.britannica.com/topic/Wired.

34. van Es, Karin, Maranke Wieringa, and Mirko Tobias Schäfer. "Tool Criticism: From Digital Methods to Digital Methodology." In Proceedings of the 2nd International

Conference on Web Studies (WS.2 2018), October 2018, 24–27. https://doi.org/10.1145/3240431.3240436.

35. van Es, Karin. 2023. "Unpacking Tool Criticism as Practice, in Practice." DHQ: Digital Humanities Quarterly 17 (2). Utrecht University. https://orcid.org/0000-0001-5184-8640.

36. Visual Studio Code - Code Editing. Redefined," November 3, 2021. https://code.visualstudio.com/.

37. Web Scraper - The #1 Web Scraping Extension," n.d. https://www.webscraper.io/.

38. Zhang, Wenping, Wei Du, Yiyang Bian, Chih-Hung Peng, and Qiqi Jiang. "Seeing Is Not Always Believing: An Exploratory Study of Clickbait in WeChat." Internet Research 30, no. 3 (March 13, 2020): 1043–58. https://doi.org/10.1108/intr-09-2019-0373.

Appendix I: Data Preprocessing

```
import pandas as pd
from textblob import TextBlob
from nltk.sentiment.vader import SentimentIntensityAnalyzer
import re
from datetime import datetime

def clean_reputable_article_text(article_text):
    cleaned_text = re.sub(r'\[{"Article Text":"(.*?)"}\]', r'\1', str(article_text))
    return cleaned_text


def clean_date(date_str):
    try:
        date_obj = datetime.strptime(date_str, "%Y-%m-%dT%H:%M:%S.%fZ")
    except ValueError:
        date_obj = datetime.strptime(date_str, "%Y-%m-%d")
    return date_obj.strftime("%B %d, %Y %I:%M %p")

reputable_df = pd.read_csv("REDACTED.csv")

not_reputable_df = pd.read_csv("REDACTED.csv")

reputable_df["Article Text"] = reputable_df["Article Text"].apply(clean_reputable_article_text)

not_reputable_df["Article Text"] = not_reputable_df["Article Text"].apply(remove_specified_text)

not_reputable_df["Date"] = not_reputable_df["Date"].apply(clean_date)

def calculate_textblob_sentiment(text):
    return TextBlob(str(text)).sentiment.polarity

def calculate_vader_sentiment(text):
    sid = SentimentIntensityAnalyzer()
    scores = sid.polarity_scores(str(text))
    return scores['compound']

reputable_df["Title Sentiment (TextBlob)"] = reputable_df["Title"].apply(calculate_textblob_sentiment)
```

```
reputable_df["Article Text Sentiment (TextBlob)"] = reputable_df["Article
Text"].apply(calculate_textblob_sentiment)
reputable_df["Description Sentiment (TextBlob)"] =
reputable_df["Description"].apply(calculate_textblob_sentiment)

reputable_df["Title Sentiment (Vader)"] = reputable_df["Title"].apply(calculate_vader_sentiment)
reputable_df["Article Text Sentiment (Vader)"] = reputable_df["Article Text"].apply(calculate_vader_sentiment)
reputable_df["Description Sentiment (Vader)"] = reputable_df["Description"].apply(calculate_vader_sentiment)

not_reputable_df["Title Sentiment (TextBlob)"] = not_reputable_df["Title"].apply(calculate_textblob_sentiment)
not_reputable_df["Article Text Sentiment (TextBlob)"] = not_reputable_df["Article
Text"].apply(calculate_textblob_sentiment)
not_reputable_df["Description Sentiment (TextBlob)"] =
not_reputable_df["Description"].apply(calculate_textblob_sentiment)

not_reputable_df["Title Sentiment (Vader)"] = not_reputable_df["Title"].apply(calculate_vader_sentiment)
not_reputable_df["Article Text Sentiment (Vader)"] = not_reputable_df["Article
Text"].apply(calculate_vader_sentiment)
not_reputable_df["Description Sentiment (Vader)"] =
not_reputable_df["Description"].apply(calculate_vader_sentiment)

reputable_df = reputable_df[(reputable_df["Title"].notna()) & (reputable_df["Article Text"].notna()) &
(reputable_df["Title"] != "[]") & (reputable_df["Article Text"] != "[]")]
not_reputable_df = not_reputable_df[(not_reputable_df["Title"].notna()) & (not_reputable_df["Article Text"].notna())
& (not_reputable_df["Title"] != "[]") & (not_reputable_df["Article Text"] != "[]")]

reputable_df.to_csv("cleaned_analyzed_reputable_data.csv", index=False)
not_reputable_df.to_csv("cleaned_analyzed_not_reputable_data.csv", index=False)

print("TextBlob Sentiment Analysis:")
print("Reputable Dataset:")
print("Average Title Sentiment Score:", reputable_df["Title Sentiment (TextBlob)"].mean())
print("Average Article Text Sentiment Score:", reputable_df["Article Text Sentiment (TextBlob)"].mean())
print("Average Description Sentiment Score:", reputable_df["Description Sentiment (TextBlob)"].mean())
print("\nNot Reputable Dataset:")
print("Average Title Sentiment Score:", not_reputable_df["Title Sentiment (TextBlob)"].mean())
print("Average Article Text Sentiment Score:", not_reputable_df["Article Text Sentiment (TextBlob)"].mean())
print("Average Description Sentiment Score:", not_reputable_df["Description Sentiment (TextBlob)"].mean())

print("\nVader Sentiment Analysis:")
```

```
print("Reputable Dataset:")
print("Average Title Sentiment Score:", reputable_df["Title Sentiment (Vader)"].mean())
print("Average Article Text Sentiment Score:", reputable_df["Article Text Sentiment (Vader)"].mean())
print("Average Description Sentiment Score:", reputable_df["Description Sentiment (Vader)"].mean())
print("\nNot Reputable Dataset:")
print("Average Title Sentiment Score:",not_reputable_df["Title Sentiment (Vader)"].mean())
print("Average Article Text Sentiment Score:", not_reputable_df["Article Text Sentiment (Vader)"].mean())
print("Average Description Sentiment Score:", not_reputable_df["Description Sentiment (Vader)"].mean())

print("\nAverage Sentiment Scores (TextBlob and Vader):")
print("Reputable Dataset:")
print("Average Title Sentiment Score (TextBlob):", reputable_df["Title Sentiment (TextBlob)"].mean())
print("Average Article Text Sentiment Score (TextBlob):", reputable_df["Article Text Sentiment
(TextBlob)"].mean())
print("Average Description Sentiment Score (TextBlob):", reputable_df["Description Sentiment
(TextBlob)"].mean())
print("Average Title Sentiment Score (Vader):", reputable_df["Title Sentiment (Vader)"].mean())
print("Average Article Text Sentiment Score (Vader):", reputable_df["Article Text Sentiment (Vader)"].mean())
print("Average Description Sentiment Score (Vader):", reputable_df["Description Sentiment (Vader)"].mean())

print("\nNot Reputable Dataset:")
print("Average Title Sentiment Score (TextBlob):", not_reputable_df["Title Sentiment (TextBlob)"].mean())
print("Average Article Text Sentiment Score (TextBlob):", not_reputable_df["Article Text Sentiment
(TextBlob)"].mean())
print("Average Description Sentiment Score (TextBlob):", not_reputable_df["Description Sentiment
(TextBlob)"].mean())
print("Average Title Sentiment Score (Vader):", not_reputable_df["Title Sentiment (Vader)"].mean())
print("Average Article Text Sentiment Score (Vader):", not_reputable_df["Article Text Sentiment (Vader)"].mean())
print("Average Description Sentiment Score (Vader):", not_reputable_df["Description Sentiment (Vader)"].mean())
```

Appendix II: Construct Validity

```
import pandas as pd
import numpy as np

np.random.seed(0)

reputable_csv_path = 'cleaned_analyzed_not_reputable_data.csv'
non_reputable_csv_path = 'cleaned_analyzed_reputable_data.csv'

reputable_df = pd.read_csv(reputable_csv_path).head(7400)
non_reputable_df = pd.read_csv(non_reputable_csv_path).head(7400)

reputable_sample = reputable_df.sample(100)
non_reputable_sample = non_reputable_df.sample(100)

combined_sample = pd.concat([reputable_sample, non_reputable_sample], ignore_index=True)

combined_sample = combined_sample.sample(frac=1).reset_index(drop=True)

columns_to_keep = ['Title', 'Article Text']
combined_sample = combined_sample[columns_to_keep]

combined_sample['Clickbait'] = np.nan

combined_sample.to_csv('manual_scoring_dataset.csv', index=False)
```

Appendix III: Extreme Discrepancy Outlier Analysis

```
import pandas as pd
import numpy as np

reputable_dataset_path = 'cleaned_analyzed_reputable_data.csv'
non_reputable_dataset_path = 'cleaned_analyzed_not_reputable_data.csv'

reputable_df = pd.read_csv(reputable_dataset_path).head(7400)
non_reputable_df = pd.read_csv(non_reputable_dataset_path).head(7400)

def get_outliers(df, column):
    top_outliers = df.nlargest(50, column)
    bottom_outliers = df.nsmallest(50, column)
    return pd.concat([top_outliers, bottom_outliers])

reputable_textblob_outliers = get_outliers(reputable_df, 'Title Sentiment (TextBlob)')
reputable_vader_outliers = get_outliers(reputable_df, 'Title Sentiment (Vader)')
non_reputable_textblob_outliers = get_outliers(non_reputable_df, 'Title Sentiment (TextBlob)')
non_reputable_vader_outliers = get_outliers(non_reputable_df, 'Title Sentiment (Vader)')

combined_outliers = pd.concat([
    reputable_textblob_outliers,
    reputable_vader_outliers,
    non_reputable_textblob_outliers,
    non_reputable_vader_outliers
])

combined_outliers = combined_outliers.sample(frac=1).reset_index(drop=True)

columns_to_keep = ['Title', 'Article Text']
combined_outliers = combined_outliers[columns_to_keep]

combined_outliers['Clickbait'] = np.nan

combined_outliers.to_csv('manual_scoring_extreme_discrepancy_outliers.csv', index=False)
```