

Uncovering Media Bias in Eviction Reporting: A Comprehensive Analysis Utilising Sentiment Analysis Framework and Social Media Data

Mabokela Koena Ronny
Applied Information Systems, University of
Johannesburg
krmabokela@gmail.com

Primus Mpho
Institute for Intelligent Systems, University of
Johannesburg
mraborife@gmail.com

Mahlobo Linda
Applied Information Systems, University of
Johannesburg

Abstract

This study investigates the prevalence of evictions in South Africa and examines potential disparities between traditional media reporting and social media discourse. Employing a sentiment analysis framework, we extend its application to compare the reporting of evictions in newspaper articles (i.e. conventional media) and Twitter data (i.e. social media). Statistical machine-learning methods are utilized to predict sentiment scores for both types of content, and a chi-square test is employed to evaluate bias between news articles and tweets. The test results reveal a significant bias in the sentiment distribution, suggesting that the dissimilarities observed between articles and tweets are not merely coincidental.

Keywords: evictions, sentiment analysis, social media, machine learning, bias, discrepancies

1 Introduction

The occurrence of evictions in South Africa has become a subject of concern, as both legal and illegal evictions persist despite the constitutional requirement for court orders. While South Africa is recognized globally for its housing and eviction laws, instances of illegal displacements are prevalent (Muller 2013; Muller et al., 2019, du Plessis, 2005). Reports indicate that many evictions are not court-ordered, leading to vulnerable populations, such as farm workers being left homeless and stranded.

In the study by Bosch & Mutsvairo (2017), the way traditional media covers social issues like illegal evictions has been questioned for possible bias. They found that during the “fees must fall” movement, traditional media often depicted students as violent, while students’ own images and Twitter data showed a more positive portrayal (Bosch & Mutsvairo, 2017; Beukes, 2017). This new research aims to explore how traditional media and social media report on evictions, highlighting potential differences and biases in covering this important social problem. Additionally, Huchzermeyer (2003) highlighted the growing influence of micro-blogging sites like Twitter in raising awareness and engaging the public in social issues. Bosch & Mutsvairo’s study on the Fees Must Fall Campaign also demonstrated differences in reporting between social platforms and traditional media through Twitter images (Bosch & Mutsvairo, 2017). In this study, we aim to investigate and compare the reporting of evictions in traditional media and social media, shedding light on potential discrepancies and biases in the coverage of these critical social issues.

Numerous models have been developed in the past decade to analyse media (Grefenstette et al., 2004). Still, they need to cover developing a model that analyses discrepancies between social media and traditional media (Grefenstette et al., 2004). The above raises the question: *How can an existing sentiment analysis framework for social media be applied to accurately analyse the sentiment discrepancies in reporting social issues between social media content and mainstream news data?* To answer this question, we explore the current sentiment analysis framework for social media proposed by Afia et al. (2018).

This study aims to address the research question by leveraging the sentiment analysis framework proposed by Afia et al. (2018) for social media analysis. As a major contribution, the framework is extended and applied to compare datasets from Twitter and online news articles, specifically focusing on sentiments related to evictions in South Africa. By utilizing this approach, the study seeks to achieve two main objectives: firstly, to identify the most effective machine-learning model for predicting sentiments related to



evictions, and secondly, to uncover any potential discrepancies that may exist between the sentiment distributions in the two datasets. By accomplishing these objectives, this research contributes to a deeper understanding of the sentiment dynamics surrounding evictions in South Africa and sheds light on the variations in sentiment representation between traditional media reporting and social media platforms.

This paper is structured as follows: Section 2: provides a literature review and related works. In section 3, we describe our data collection strategies and the machine learning methods. The experimental setup is presented in Section 4. Section 5 describes the analyses of the results and then discusses the results. Finally, we provide concluding remarks and possible future work.

2 Literature Review

2.1 Bias

According to Hamborg et al. (2019), bias means favouring one opinion in an article intentionally and repeatedly. Discrimination can happen at different stages of making news: when collecting information, writing, and editing. Bias during information gathering can include choosing which events to cover, picking sources, leaving out or adding details. When writing, bias might be shown through labelling and word use. In the editing phase, bias can be seen in where things are placed, how much space they get, the pictures chosen, and how those pictures are explained (Hamborg et al., 2019).

Grefenstette et al. (2004) demonstrated that journalists tend to choose and shape the information they publish, leading to bias in their articles. They use specific language and select particular credible sources to influence how readers perceive the news. This bias can stem from motivations like potential profits, media ownership, or personal beliefs (University of Michigan, 2014). Efforts have been made to detect media bias, with social science methods like content analysis, explanatory models, and frame analysis (Hamborg et al., 2019). However, these methods are manual, challenging to follow and hinder the study of media bias in the social

sciences. Manual approaches can introduce subjectivity, as shown by Enevoldsen and Hansen (2017) who found that researchers' subjectivity using questionnaires or surveys influenced bias.

In the field of computational sciences, the exploration of media bias is relatively new compared to social sciences, which have been studying this since the 1960s (Hamborg et al., 2019). Park et al. (2009) created "NewsCube" to reduce discrimination effects, offering multiple perspectives on a topic for a more balanced view. However, this approach assumes bias in the text and does not detect it. Choy et al. (2011) improved "NewsCube" with version 2.0, allowing users to estimate bias in news articles. Nevertheless, the drawback is that these users might not be experts in detecting bias.

Saez-Trumper et al. (2013) explored various models that use NLP, machine learning, and deep learning to detect biases in statements in an unsupervised way. These methods show promise for being faster, scalable, and more automated compared to traditional social science methods. In Hamborg's work (2020), the detection of bias in language shifted from people to computational linguistics like NLP with deep learning models, enhancing the NewsCube model for automatically recognizing bias in articles. Despite its success, the model's inability to be widely applied is due to data imbalances. Nadeem & Raza (2022) used various NLP methods to detect bias in political news sources. A deep neural network with a bag of words to represent the input sentences was used. Then employ the Term Frequency - Inverse Document Frequency (TF-IDF) as a weighting factor to the DNN input data. To add an unsupervised K-Means clustering algorithm was used to analyse patterns discovered. The main goal was to identify whether a news article is biased or unbiased.

Several studies were conducted to understand media bias in both traditional and social media sources. Wang & Mark (2013) surveyed Chinese citizens to determine their trust in different news sources but did not find bias. Diehl et al., (2013) suggested bias between social media and online data and explored how interaction between journalists and people on Twitter affects



perceptions of bias. Younus et al. (2012) researched how social media can identify bias in traditional media, introducing the concept of citizen journalism and creating a model using latent Dirichlet allocation and Jaccard similarity to measure bias.

Thomsen (2018) used sentiment analysis to study media bias in newspaper tweets and discovered that while some media companies showed bias, most did not. The study required a detailed understanding to include sentiment analysis. Choy et al. (2011) justified selecting commission and omission bias to study differences in reporting social issues between social media and mainstream news data because it helps understand how information is manipulated and its impact on public sentiment. Commission and omission bias involve purposefully adding or removing information that can influence how readers perceive an event. Hamborg et al. (2019) explain that this bias can significantly shape the story told to the audience, affecting their understanding and feelings about a topic. Analyzing this bias is crucial to understanding how media and social platforms selectively present information to influence public opinion. No studies have explored how social and traditional media perceive evictions through sentiment analysis and compared these sources to find bias.

2.2 Sentiment Analysis

Sentiment analysis, also known as opinion mining, involves extracting sentiments from text, which can be positive, negative, or neutral (Devika, 2016). It encompasses three levels of analysis: document, sentence, and aspect. Document sentiment analysis focuses on determining the overall opinion of an entire text (Mabokela et al, 2023). Sentence-level analysis extracts sentiments from individual sentences. Lastly, aspect-level analysis delves into the most granular level, examining opinions themselves (Devika, 2016). Sentiment analysis methods can be categorized into four main types: Machine learning, Rule-Based, Lexical-Based, and hybrid methods (Mulvenna, 2015; Al-Shabi, 2020).

The understanding of text by machines requires feature extraction techniques to convert text into

a machine-readable format. Various methods have been introduced in NLP. N-grams and TF-IDF are notable techniques used in current trends (Barnaghi et al., 2016; Shikomba et al., 2021). N-grams involve extracting a specified number of sequential words simultaneously, like Unigrams (1 word) and bigrams (2 words). On the other hand, TF-IDF focuses on selecting top features using statistical methods to reduce feature count (Barnaghi et al., 2016). This study aims to compare the effectiveness and superiority of these approaches in sentiment detection.

In the field of sentiment analysis, previous research has outlined a standard process involving data collection, preparation, model training, and evaluation. This framework has led researchers to create specialised frameworks tailored to specific issues (Devika, 2016; Hadjidj et al., 2017). This study evaluates various sentiment analysis frameworks to choose the most suitable one for the researcher's objectives. Marques-Lucena et al. (2014) introduced a framework for analyzing hotel reviews to enhance service delivery based on customer feedback. Olugbara & Zvarevashe (2018) expanded this framework by creating an automatic labelling system for training datasets, reducing labelling efforts and errors. However, these frameworks are unsuitable for the current study as they only cater to data from single sources, such as websites.

In their work, Karpurapu & Jololian (2017) presented a streaming data framework for sentiment analysis using Twitter data. However, the framework's limitation lies in its inability to integrate different data sets without substantial modifications. Moreover, the exclusive reliance on Naïve Bayes (NB) for sentiment analysis might lead to inefficiencies. Comparison studies, like Diaz-Martins (2015), have consistently demonstrated that NB is outperformed by models such as support vector machines (SVM) and Random Forest (RF), which are shown to be superior choices for sentiment analysis.

Hadjidj et al. (2017) propose a framework treating each processing step as a web service component. The framework consists of four sequential features: data collection, noise removal, geo-location, and sentiment analysis. Each element is



loosely coupled, enabling changes in one component with minimal impact on the framework’s overall features. It is particularly suitable for scenarios involving both social media datasets and geo-location information which is not suitable for our study.

The field of sentiment analysis has witnessed extensive research, but sentiment analysis frameworks have received limited attention (Afia et al., 2018). Afia et al. (2018) address this gap by presenting a comprehensive framework that covers various essential components, including data acquisition, processing, analysis, and visualization (see Figure 1). The framework includes extracting, storing, and distributing data in data acquisition, pre-processing, plugin management in data processing, and extracting emotions and polarity in data analysis. Data visualization provides a platform for quick data visualization.

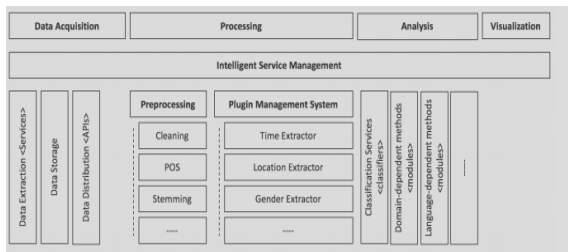


Figure 1: Sentiment Analysis Framework (Afia et al. 2018)

The orchestration layer contains a rule-based semantic engine responsible for assigning optimal resources for each sentiment analysis task (at the components level) (Afia et al., 2018). Unlike other frameworks, which are generic, this framework is specifically built to handle social media content. However, it has the flexibility to handle additional data sources as well.

3 Methodology

This study aims to apply an existing sentiment analysis framework for social media to analyze sentiment discrepancies between social media content and mainstream news data concerning evictions. To achieve this, the study selects the framework developed by Afia et al. (2018) which facilitates all necessary steps for sentiment analysis using social media data. However, the current state of the framework is insufficient to fully address

the research question. Therefore, we extended the framework to include data extraction services for website data or online news articles and plugin management for analysing discrepancies between social media data and online news articles. The extensions are depicted in yellow in Figure 2.

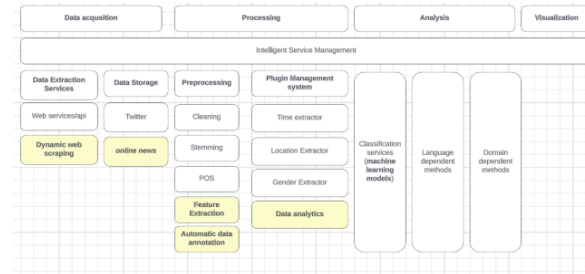


Figure 2: The Extended Sentiment Analysis Framework

In addition, we explore the following research objectives to ultimately answer the research question:

- Conduct a comparative analysis of feature extraction methods to determine the most effective approach and select the optimal machine learning model for sentiment prediction.
- Expand the existing framework to accommodate multiple datasets beyond social media data and apply the framework to predict sentiments regarding South African land and housing eviction topics.
- Utilize statistical methods on predicted sentiments to identify discrepancies between news articles and tweets about the same eviction topics.

3.1 Data Acquisition

The extended framework was proposed to work in the following manner: First, news articles are collected using dynamic web scraping on online news websites. For Twitter data, we used Twitter API to collect tweets using the geo-location-based data collection feature with special keywords and hashtags (Mabokela & Schlippe, 2022). The date range of the collected news articles and tweets ranged from 2014 to 2020. We also used a language detection process to ensure that are collected data was in English. The tweets and articles were divided into their respective topics



(i.e. the hastags). We only search use hastags and keywords to collect housing and land eviction topics in South Africa.

3.2 Data pre-processing

Data pre-processing is a crucial step in preparing extracted data for classification models (Afia et al., 2018). For this study's extended framework component, for pre-processing, we involved several tasks, including converting text to lowercase, removing URLs and HTML tags, tokenization, correcting misspelt words using a spell checker, eliminating stop words and numbers, lemmatizing words, removing question marks and special characters (including emojis), and discarding duplicate tweets and articles, while also performing parts of speech tagging. We also translated local languages into English were required. The distribution of the datasets collected is presented in Table 1.

Table 1: Distributions of the collected datasets: Tweets vs News Articles.

	Negative	Neutral	Positive
Tweets	4052	1888	355
Article	46	67	3

In total, we collected 6295 tweets and 116 news articles. Next, the collected data was stored in a Microsoft SQL server database for easy access and management.

3.3 Automatic Data Annotation

To automatically label the dataset, a combination of three (3) existing English sentiment lexicons such as AFFIN, Vader, and Senti-Strength were used (Ribeiro et al., 2016). To determine the final sentiment label for each tweet, we applied the following carefully reviewed rules presented in Table 2. Each sentiment lexicon is applied to the two datasets to determine the sentiment labels and then the rules are used to determine the final sentiment label for each tweet. In cases where both negative and positive labels are present (**Negative + Positive + Neutral**), it indicates conflicting sentiment. For this, the rule assigns a neutral label since there is no clear dominant sentiment (Ribeiro et al., 2016).

Table 2: Rules for Determining the Final Sentiment for Tweets and News Articles

Rules (AFFIN+Vader+Senti-Strength)	Final Label
Negative + Negative + Positive	Negative
Negative + Negative + Neutral	Negative
Negative + Negative + Negative	Negative
Positive + Positive + Negative	Positive
Positive + Positive + Neutral	Positive
Positive + Positive + Positive	Positive
Neutral + Neutral + Positive	Neutral
Neutral + Neutral + Negative	Neutral
Neutral + Neutral + Neutral	Neutral
Negative + Positive + Neutral	Neutral

Overall, the justification behind these rules aims to provide consistent and accurate sentiment labels for sentences based on collective evidence from multiple lexicons. Also ensuring that the assigned sentiment label aligns with the strength of the expressed sentiment.

3.4 Machine Learning Models

Next, Subsequently, we employed the scikit-learn toolkit to construct our sentiment classification models. The algorithms employed encompassed NB RF, Logistic Regression (LR), and SVM. NB, a simple classification model rooted in Bayes' theorem, was realized through the following formula.

$$f(x) = \operatorname{argmax}_y P(y) \prod_{i=1}^n P(x_i|y) \dots (1)$$

where $f(x)$ = [Negative, Positive, or Neutral] given a set of articles/tweet features (x_i).

SVM creates a linear decision boundary for different classes (i.e., negative, positive, and neutral) of sentiment (Kirchner & Signorino, 2018). This linear decision boundary is designed to maximise the margin of the types and the hyperplane (Kirchner, A., and Signorino, 2018). The margin ensures that the newly placed object/element is classified correctly. When the



data cannot be linearly separable, it is transformed into a higher dimension using the kernel trick.

RF is a supervised classification and regression model that uses more than one decision tree to make a prediction (Song and Lu, 2015). Each decision tree is given a random training set sample (Mulvenna, 2013). The decision trees of the model have supervised sub-classifiers that contribute to the final decision, using the majority vote approach for categorical labels (see Figure 3). The majority vote means that when there are ten decision trees and eight of the 10 classify a tweet as positive, this classification is taken as the final classification.

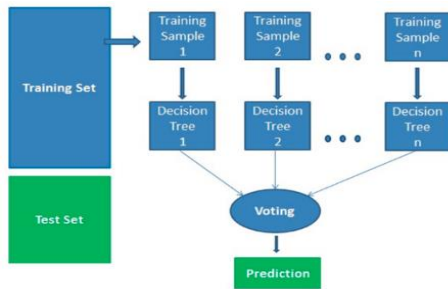


Figure 3: Random Forest Model

For decision trees, we apply the Gini Index (GI) to measure the probability an element (when selected at random) is incorrectly classified (Zakariah, 2014). A grouping (dataset) is considered pure if all group features are of the same class. Since it is a probability measurement, the values range from 0 to 1. A dataset with a gini index of zero means that all elements within the dataset are of the same class (pure), and all parts belong to various categories (impure). **LR** aims to convert continuous dependent variables (features) of a linear function to equivalent categorical values (sentiment labels) based on the independent variables (tweets and news articles).

3.5 Data Analysis

In this study, we utilised the well-established chi-square test, a statistical method with a strong theoretical foundation, to determine discrepancies within our dataset. The chi-square test has found extensive application in various fields, including social sciences, market research, and epidemiology. Its wide acceptance makes it suitable for detecting bias between two datasets,

making it valuable for hypothesis testing and constructing confidence intervals for categorical data. The chi-square statistic involves comparing observed frequencies with expected frequencies under a null hypothesis, providing valuable insights into the differences and associations within the data. (Franke et al., 2012). The chi-square statistic applies to this study for the following reasons:

- Hypothesis testing: The chi-square test allows us to test the null hypothesis, which assumes no bias or association between the datasets.

By comparing the calculated chi-square to critical values or calculating the associated p-value, we can make an informed decision about rejecting or accepting the null hypothesis, thereby determining the presence or absence of bias.

- Flexibility in categories: The chi-square test applies to situations where the categories being compared can have multiple classes (Franke et al., 2012). It allows for simultaneously reaching proportions or frequencies across various types, enabling a comprehensive bias analysis.

In our case, we have two datasets with multi-class labels.

4 Experimental Setup

After training the algorithms, we evaluated accuracy using the test dataset. Pre-processing techniques were applied to ensure data quality, and feature extraction used N-grams and TF-IDF methods to capture relevant patterns and information from the texts. Subsequently, the data were divided into training (70% - 4406 tweets and 81 news articles) and testing datasets (30% - 1889 tweets and 35 news articles) to prepare for the supervised techniques' application. The selected models were evaluated using both standard testing and stratified K-Fold cross-validation with 10-fold to assess their performance.

To validate the performance of the algorithms (SVM, LR, RF and NB), the following evaluation criteria with the confusion matrix:

- TP = True Positive (Correct Classification)
- FP = False Positive (Incorrect Classification)
- TN = True Negative (Correct Classification)
- FN = False Negative (Incorrect Classification)

Using the defined measures above, we also measure the accuracy of the classification models as follows:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \dots (2)$$

5 Results and Discussion

5.1 Results

A total combination of 16 experiments was performed. The experiments were performed concerning implementing the following objectives: *To compare and determine the best feature extraction method and sentiment classification.* We used N-grams (average accuracy of 95%) to perform better than the TF-IDF (average accuracy of 73%) for news articles and tweets. Outliers were excluded using the 1.5 Interquartile Range Rule (IQR). The overall results are in Table 3 in Appendix A.

Secondly, *we needed to determine the best machine-learning model for sentiment classification.* The experiment below shows the results of the four compared machine learning models. SVM is the best-performing classifier regarding stratified average accuracy (96%) although the RF also performed slightly better in all the experiments (See Figure 4 and Table 6).



Figure 4: Machine Models Comparison

Lastly, *we use statistical methods on predicted sentiments to determine if there are discrepancies between news articles and tweets on the same eviction topics.* To quantitatively measure the bias between the two sets (news

articles and tweets) of sentiments (Negative, Positive, and Neutral), the statistical method test, the chi-square test, is applied to determine if there is a significant difference in the proportions of sentiments between the two datasets. The below dataset (Table 1) is used.

To perform a chi-square test for independence, we need to set up the null and alternative hypotheses:

- Null Hypothesis (H_0): There is no bias regarding sentiment distribution between the article and tweets.
- Alternative Hypothesis (H_1): There is a bias between the news article and tweets regarding sentiment distributions.

To calculate the chi-square test, we follow these steps: **Step 1:** Set up the contingency table.

Table 3: Contingency values

	Negative	Neutral	Positive	Total
Tweets	4052	1888	355	6295
Articles	46	67	3	116
Total	4098	1995	358	6411

Step 2: We calculate the expected frequencies for each cell. To calculate the expected frequencies, the following formula is used: Expected Frequency = (row_total * col_total) / grand_total.

Table 4: Expected frequencies

	Negative	Neutral	Positive	Total
Tweets	4043,68	1931,85	354,47	6295
Articles	54,32	25,95	4,76	116
Total	4098	1955	358	6411

Step 3: Calculate the chi-square test statistic. The chi-square test statistic can be calculated using the formula: Chi-square = $\sum ((\text{Observed Frequency} - \text{Expected Frequency})^2 / \text{Expected Frequency})$

Table 5: Chi-square formula values

	Negative	Neutral	Positive
Tweets	1,04	1,18	0,13



Articles	7,78	8,76	0,845
-----------------	------	------	-------

Then, we sum up the calculated values for each category we get the Chi-square value equal to 19,74.

Step 4: Now, we determine the degrees of freedom (df).

$$df = (2-1) * (3-1) = 2$$

Step 5: Determining the p-value.

To calculate the p-value for the chi-square test, we determine the probability of obtaining a chi-square test statistic as extreme as, or more extreme than, the calculated value under the assumption of the null hypothesis. In this case, the calculated chi-square test statistic is 19.74, and the degrees of freedom are $df = 2$.

To find the p-value, we consult a chi-square distribution table with the given degrees of freedom. The cumulative probability associated with the calculated chi-square test statistic (19.74) and $df = 2$ is found to be 0.99. Therefore, the p-value is $1 - 0.99 = 0.01$.

Since the p-value (0.01) is less than the chosen significance level (0.05), we reject the null hypothesis. This indicates that the observed data has a statistically significant association, and we can conclude that the variables being tested are not independent in the population.

5.2 Discussion

In this section, the results are described, analysed, and interpreted. The sentiment framework by Afia et al. (2018) was initially proposed to handle social media datasets. It was successfully extended to handle an additional source dataset: online news articles. Furthermore, the data analysis component carried out 16 machine learning models to predict sentiments on eviction topics as shown in Table 6.

When comparing the feature extraction methods across different dimensions (datasets and machine learning methods), there seems to be no single model that performs consistently. The results, however, clearly indicate that the N-grams feature extraction method performs better than the TF-

IDF method. The results suggest that N-grams feature extraction performed consistently well across both datasets (articles and tweets), achieving high accuracies for most machine learning models.

Table 6: Results of the sentiment classifications with N-grams and TF-IDF for news articles and tweets.

Model	Dataset	Features	Stratified accuracy
SVM	Articles	N-grams	96%
LR	Articles	N-grams	92%
NB	Articles	N-grams	39%
RF	Articles	N-grams	94%
SVM	Tweets	N-grams	97%
RF	Tweets	N-grams	96%
LR	Tweets	N-grams	97%
NB	Tweets	N-grams	95%
LR	Articles	TF-IDF	92%
RF	Articles	TF-IDF	94%
NB	Articles	TF-IDF	39%
SVM	Articles	TF-IDF	39%
SVM	Tweets	TF-IDF	96%
RF	Tweets	TF-IDF	96%
LR	Tweets	TF-IDF	92%
NB	Tweets	TF-IDF	39%

For the N-grams representation, the SVM performs consistently well across both Articles and Tweets, achieving the highest accuracy, with values of 96-97%. RF and LR follow closely behind, with accuracies ranging from 92-96%. However, NB performs notably lower, achieving accuracies of 39-95%. On the other hand, for the TF-IDF representation, the results vary. Again, SVM and RF perform well across both data types, achieving high accuracies (96-96% for SVM and 96-96% for RF), whereas LR maintains moderate accuracy (92-92%). NB, on the other hand,



performs significantly lower, around 39% accuracy, irrespective of the data type. SVM seems to be a consistent performer across both feature representations and data types, while LR and RF are also strong candidates. However, NB struggles to achieve high accuracies in most scenarios. These results suggest that the choice of feature representation (N-grams or TF-IDF) and the selection of the machine learning model significantly impact the classification accuracy.

Similarly, when contrasting machine learning models across different dimensions (datasets and feature extraction method), there seems to be no single model that performs consistently; however, from an overall perspective, SVM is significantly better than other models. SVM's superior performance was similarly reported by Malhar & Guddeti (2014) in election result validation and by Al-Sarraj & Lubbad (2018) in detecting bias in Western media.

Although sufficient data was collected for the Twitter dataset, the dataset collected for online news articles was insufficient. Despite this limitation, adequate data was organised for the Twitter dataset, but the sentiments were not evenly distributed. The results are, however, still acceptable as the models did not underfit significantly, i.e., the model consistently performed well across 10 stratified K-folds. Furthermore, the chosen sentiment framework by Afia et al. (2018) can also handle imbalanced datasets.

Finally, the Chi-Squared test was applied as a statistical method to detect discrepancies between news articles and tweets on the same eviction topics. The null hypothesis (H_0) was rejected because the calculated p-value (0,01) was less than the significance level of 0,05. This p-value indicates a bias between the news articles and tweets regarding sentiment distribution, and this bias is not due to chance. The research question addresses how an existing sentiment analysis framework by Afia et al. (2018) can be effectively applied to analyse sentiment discrepancies in reporting social issues (eviction topics) between social media data and mainstream news data. Our objectives entail assessing various feature extraction methods and machine learning models

in sentiment predictions. The results answer the research question in the following manner: (1) the N-grams feature extraction method's consistency in outperforming TF-IDF, showing higher accuracy across datasets. Notably, (2) the SVM stands out as a consistently effective model, confirming the findings from other studies. Despite dataset limitations, the models perform well, and the chosen sentiment framework can accommodate an imbalanced dataset. Moreover, (3) a Chi-squared test unveils significant sentiment distribution disparities between news articles and tweets on the same eviction topics, highlighting the importance of considering how different sources can affect the reporting of social issues.

Finally, we have effectively addressed the research question and objectives, shedding light on sentiment discrepancies in reporting social issues and validating the sentiment framework's applicability to diverse datasets. The research also effectively bridges computational analysis and social science expertise, yielding a comprehensive understanding of sentiment discrepancies in reporting social issues across diverse platforms.

6 Conclusion

This study constructed two distinct sentiment analysis datasets using Twitter tweets and online news articles, extending the social media sentiment analysis framework proposed by Afia et al. (2018). The SVM model outperformed other methods, with N-grams being the most effective feature extraction technique. The analysis revealed a significant sentiment distribution discrepancy between tweets and news articles on eviction topics, indicating the presence of bias. The framework covered various aspects of data acquisition, pre-processing, processing, and analysis, providing a robust platform for sentiment analysis tasks. In future, we could explore language-independent models to address the challenge of multilingual content in sentiment analysis. As pre-trained language models demand substantial data and computational resources, exploring ways to leverage their potential while optimising resource usage could lead to improved sentiment analysis accuracy and effectiveness for well-researched models.



Acknowledgements

We extend our heartfelt gratitude to all those who contributed to the successful completion of this research. First and foremost, we express our sincere appreciation to reviewers and supervisors for their guidance, support, and invaluable insights throughout the research project. Lastly, we acknowledge the support of NRF BAAP for funding this research project.

References

- Afia, A. E., Chiheb, R., Faizi, R., & Haddaoui, B. E. (2018), "Toward a Sentiment Analysis Framework for Social Media", *Association for Computing Machinery*, pp.1–5.
- Al-Shabi, M. A. (2020), "Evaluating the Performance of the most important lexicons used to sentiment analysis and opinions", *International Journal of Computer Science and Network Security*, pp. 51-57.
- Barnaghi, P., Breslin, J. G., & Ghaffari, P. (2016), Opinion mining and sentiment polarity on Twitter and correlation between events and sentiment. *IEEE Second International Conference on Big Data Computing Service and Applications (Big data service)*, pp.1-6.
- Beukes, S. (2017), "An Exploration of the Role of Twitter in the Discourse Around Race in South Africa: Using the #Feesmustfall Movement as a Pivot for Discussion". *Mediarep.org*, pp.195-210.
- Bosch, T. (2016). "Twitter activism and youth in South Africa: the case of #rhodesmustfall". *Information, Communication & Society*, pp.1-12.
- Bosch, T. (2016), "Twitter and Participatory Citizenship: #feesmustfall in South Africa", *Digital Activism in the Social Media Era*, pp. 159-173.
- Bosch, T., & Mutsvairo, B. (2017), "Pictures, Protests and Politics: Mapping Twitter Images during South Africa's Fees Must Fall Campaign", *African Journalism Studies*, pp.71-89.
- Devika, M. D., Sunitha, C., & Ganesh, A. (2016), "Sentiment analysis: a comparative study on different approaches," *Procedia Computer Science*, Vol. 87, pp. 44-49.
- Choy, M., Cheong, M. L., Laik, M. N., & Shung, K. P. (2011). "A sentiment analysis of Singapore Presidential Election 2011 using Twitter data with census correction". *Arxiv.org*, pp.1-11.
- Díaz-Martins, A., Menéndez, H., & Suplet, M. R. (2015). "Using Twitter to Engage with Customers: A Data Mining Approach". *Internet Research*, pp.416-434.
- du Plessis, J. (2005). "The growing problem of forced evictions and the crucial importance of community-based, locally appropriate alternatives", *Environment and Urbanisation*, pp. 123-133.
- Franke, T. and Ho, T. and Christie, C. (2012). "The Chi-Square Test Often Used and More Often Misinterpreted", *American Journal of Evaluation*, Vol. 33, pp.448-458.
- Grefenstette, G., Qu, Y., Shanahan, J. G., & Evans, D. A. (2004), "Coupling niche browsers and affect analysis for an opinion mining application", *Proceedings of Recherche d'Information Assistée par Ordinateur (RIA/O)*, pp. 186–194.
- Huchzermeyer, M. (2003). "Housing rights in South Africa: Invasions, evictions, the media, and the courts in the cases of Grootboom, Alexandra, and Bredell", *In Urban Forum*, Vol. 14, No. 1, pp. 80-107.
- Hamborg, F., Donnay, K., & Gipp, B. (2019). Automated identification of media bias in news articles: an interdisciplinary literature review. *International Journal on Digital Libraries*, 20(4), pp. 391-415.
- Hamborg, F. (2020). "Media Bias, the Social Sciences, and NLP: Automating Frame Analyses to Identify Bias by Word Choice and Labeling". *Annual Meeting of the Association for Computational Linguistics.*, pp. 79-87
- Karpurapu, B. S., & Jololian, L. (2017). "A Framework for Social Network Sentiment Analysis Using Big Data Analytics", *Big Data and Visual Analytics*, pp. 203-217.
- Kirchner, A., and Signorino C. S., (2018), "Using Support Vector Machines for Survey Research." *Survey Practice*, Vol. 11, No. 1, pp. 1-10



- Mabokela K. R., Celik T. and Raborife M, (2023), "Multilingual Sentiment Analysis for Under-Resourced Languages: A Systematic Review of the Landscape", *In IEEE Access*, vol. 11, pp. 15996-16020
- Mabokela, K.R., Schlippe, T. (2022). "AI for Social Good: Sentiment Analysis to Detect Social Challenges in South Africa". In: Pillay, A., Jembere, E., Gerber, A. (eds), *Artificial Intelligence Research. SACAIR 2022. Communications in Computer and Information Science*, vol.1734. Springer. pp. 309-322.
- Malhar, A., & Guddeti, R. M, (2014), "Influence Factor Based Opinion Mining of Twitter Data Using Supervised Learning, *Sixth International Conference on Communication Systems and Networks (COMSNETS)*, pp.1-8.
- Marques-Lucena, C., Fonseca, J., Grilo, A., Jardim-Gonçalves, R., & Sarraipa, J. (2014). Framework for Customers' Sentiment Analysis. *Intelligent Systems*, pp. 849-860.
- Muller, G. (2013). "The Legal-Historical Context of Urban. *Fundamental*, pp. 367-396.
- Muller, G., & Alimohammadi, E. (2019). "The illegal eviction of undocumented foreigners from South Africa", *African Human Rights Law Journal*, pp. 793-818.
- Mulvenna, M. D., Baumgarten, M., & Rooney, N. (2013). "Keyword-based classifier sentiment mining using Twitter, *International Journal of Ambient Computing and Intelligence*, pp.56-69.
- Nadeem, M. U., & Raza S. (2021), "Detecting Bias in News Articles using NLP Models", *Stanford CS224N Natural Language Processing with Deep Learning*, pp. 1-7
- Olugbara, O. O., & Zvarevashe, K. (2018). A "Framework for Sentiment Analysis with Opinion Mining of Hotel Reviews", *Conference on Information Communications Technology and Society*, pp. 1-4.
- Park, S., Kang, S., Chung, S., & Song, J. (2009). "NewsCube: delivering multiple aspects of news to mitigate media bias", *In Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 443-452.
- Rawat, S., & Vadivu, G. (2022). "Media Bias Detection Using Sentimental Analysis and Clustering Algorithms", *In Proceedings of International Conference on Deep Learning, Computing and Intelligence (ICDCI)*, pp. 485-494.
- Ribeiro, F. N., Araújo, M., Gonçalves, P. et al. (2016). "SentiBench - a benchmark comparison of state-of-the-practice sentiment analysis methods", *EPJ Data Science*, Vol. 5, No. 23, pp.1-23
- Saez-Trumper, D., Castillo, C., & Lalmas, M. (2013). "Social media news communities: gatekeeping, coverage, and statement bias". *In Proceedings of the 22nd ACM International Conference on Information & Knowledge Management*, (pp. 1679-1684).
- Shikomba, M. T., Suresh, N., Hashiyana, V., Nelulu, J., & Sverdlik, W. (2021), "Correlation of Sentiment Analysis between Tweets and Newspaper Headlines Using NLP". *In Proceedings of the International Conference on Data Science, Machine Learning and Artificial Intelligence*, pp. 13-19.
- Song Y. Y. and Lu Y. (2015). "Decision tree methods: applications for classification and prediction", *Shanghai Arch Psychiatry*", Vol. 27, No. 2, pp.130-135.
- Thomsen, T. (2018). Do media companies drive bias? using sentiment analysis to measure media bias in newspaper tweets. Masters Dissertation, Johns Hopkins University.
- Wang, Y., & Mark, G. (2013), "Trust in online news: Comparing social media and official media use by Chinese citizens", *In Proceedings of the 2013 conference on Computer supported cooperative work*, pp. 599-610.
- Younus, A., Qureshi, M. A., Kingrani, S. K., Saeed, M., Touheed, N., O'Riordan, C., & Gabriella, P. (2012, April). "Investigating bias in traditional media through social media", *In Proceedings of the 21st International Conference on World Wide Web*, pp. 643-644.

