Twitter Entity Sentiment Analysis: Unveiling Sentiments Towards Entities in Microblogging

Shadman Ahmad Nafee
Computer Science and Engineering
Brac University
Dhaka, Bangladesh
shadman.ahmad.nafee@g.bracu.ac.bd

Jannatus Sakira Khondaker
Computer Science and Engineering
Brac University
Dhaka, Bangladesh
jannatus.sakira.khondaker@g.bracu.ac.bd

***Abstract*—This study analyzes the sentiment towards entities in microblogging, specifically on Twitter. The goal is to uncover the sentiments expressed towards different entities mentioned in tweets. Social media platforms have facilitated the collection of abundant data for sentiment analysis, allowing researchers to obtain valuable insights into public opinions. This study centers on Twitter, a widely used microblogging platform, and utilizes entity-level sentiment analysis to identify sentiments conveyed in tweets regarding particular entities. The dataset consists of messages classified into three sentiment categories: Positive, Negative, and Neutral. Messages that are considered irrelevant are treated as Neutral. We utilize natural language processing methodologies, which encompass text preprocessing, feature extraction through TF-IDF and n-grams, and the application of machine learning algorithms such as Logistic Regression and XGBoost. We aim to precisely forecast sentiment towards entities and investigate the intricacies of sentiment expression on Twitter. The findings illuminate the patterns of emotions within the Twitter community, adding to the wider field of sentiment analysis and offering valuable insights for applications in business intelligence, brand management, and public opinion monitoring.**

I. Introduction

Social media platforms have emerged as valuable sources of data for comprehending public sentiment, with Twitter being particularly notable as a microblogging platform known for its concise nature and swift dissemination of information. Amidst the era of extensive data, it has become crucial to utilize the abundant user-generated content on Twitter to gain valuable insights into public sentiments regarding different entities, such as products, personalities, or events. This study focuses on the analysis of sentiment at the entity level on Twitter. It aims to identify the sentiment expressed in tweets that specifically mention entities. The sentiment classes consist of Positive, Negative, and Neutral. The Neutral class includes messages that are considered irrelevant to the target entity. Using an extensive dataset, we employ natural language processing techniques to preprocess the text data, extracting significant features for further analysis. Our approach involves using TF-IDF to extract features and incorporating n-grams to capture contextual information. In addition, we utilize machine learning models such as Logistic Regression and XGBoost to predict sentiment. Our objective is to reveal the complex patterns of sentiment expression in the Twitterverse using this method. Our analysis of sentiments towards entities on Twitter contributes to the field of sentiment analysis and has practical implications for business intelligence, brand management, and real-time monitoring of public opinions. The following sections elaborate on the methodology, results, and discussions, offering a thorough investigation of our findings.

II. Literature Review

1. Overview of Sentiment Analysis on Social Media: The field of sentiment analysis, also called opinion mining, has become prominent due to the emergence of social media platforms. Twitter, a form of social media, provides a vast reservoir of unedited viewpoints and sentiments shared by users across the globe. Researchers have progressively relied on sentiment analysis to extract valuable insights from this extensive and ever-changing reservoir of data.

2. Sentiment analysis specific to Twitter: Twitter, being a microblogging platform, poses distinct challenges and possibilities for sentiment analysis. Due to the concise nature of tweets, which are typically restricted to 280 characters, specific strategies are necessary to accurately convey subtle emotions. Prior research has examined different methodologies, such as natural language processing (NLP) techniques, machine learning algorithms, and customized deep learning models designed specifically for data collected from Twitter.

3. Entity-Level Sentiment Analysis: Entity-level sentiment analysis is a more specific approach to sentiment analysis that examines the sentiments expressed towards particular entities, such as products, brands, or individuals, rather than focusing on the overall sentiment within a document or sentence. This sophisticated approach enables a more detailed comprehension of how emotions differ across various aspects and entities mentioned in the tweets.

4. Challenges and Opportunities: The existing literature emphasizes the difficulties faced in analyzing sentiment at the entity level on Twitter. These challenges include effectively dealing with short and noisy text, identifying sarcasm, and adapting to the ever-changing nature of language on social media platforms. Nevertheless, the integration of contextual information using n-grams and the utilization of sophisticated machine learning algorithms present encouraging opportunities for enhancement.

5. Twitter Sentiment Analysis Applications: Apart from academic research, sentiment analysis on Twitter has been utilized in various fields. The utilization of Twitter data, ranging from brand management to political analysis and public opinion monitoring, plays a crucial role in facilitating well-informed decision-making and the development of effective strategies.

III. Methodology

A. Dataset:

We used an extensive dataset that consisted of tweets labeled with sentiment at the entity level. The sentiment categories included Positive, Negative, Neutral, and Irrelevant. The Twitter API was utilized to gather tweets, and the dataset underwent preprocessing to address noise and ensure consistency.

B. Text Preprocessing:

To ready the text data for analysis, we executed several preprocessing procedures. These tasks involved transforming the text to lowercase, eliminating special characters, and segmenting it into tokens. We utilized the NLTK library to perform common Natural Language Processing (NLP) tasks and employed regular expressions to cleanse the text data.

C. Feature Extraction:

We utilized two fundamental techniques for feature extraction: TF-IDF (Term Frequency-Inverse Document Frequency) and n-grams. TF-IDF quantified the significance of words within the corpus, whereas 4-grams, a type of n-gram, enabled us to incorporate contextual information in sentiment analysis.

D. Model Selection:

For sentiment prediction, two machine learning models were utilized: Logistic Regression and XGBoost. Logistic Regression was used as a basic model, while XGBoost, renowned for its capability to handle intricate relationships in data, offered a more advanced approach.

E. Evaluation:

We evaluated the performance of our models using conventional metrics such as accuracy, precision, recall, and F1 score. The hyperparameters of the models were fine-tuned using a validation dataset to ensure their generalizability.

F. Ethical Considerations:

To guarantee ethical data usage, we strictly followed privacy standards and guidelines during the collection and handling of Twitter data. Furthermore, measures were taken to reduce biases in the models and interpretations. This approach enabled us to effectively analyze sentiments on Twitter, offering a strong framework for comprehending and forecasting sentiments towards entities in the microblogging domain.

IV. Results

The logistic regression model (model-1) was trained using hyperparameters C=1, solver = "liblinear," and a maximum iteration of 200. The test set yielded an accuracy of 81.51%, suggesting a moderate level of success in sentiment prediction using the bag-of-words (BoW) representation. Furthermore, during the evaluation of the model on the validation set, it demonstrated a superior accuracy of 91.7%, indicating commendable generalization performance.

A second logistic regression model, referred to as model2, was trained using modified hyperparameters (C = 0.9, solver = "liblinear," max_iter = 1500). The model demonstrated enhanced performance compared to the initial model, achieving an accuracy of 90.79% on the test set. When tested on the validation set, model 2 exhibited an exceptionally high accuracy of 98.6%, underscoring its efficacy in forecasting sentiments on unfamiliar data.

Precision	0.8209
Recall	0.8151
F1 score	0.8131

TABLE I
Model 1 Metrics:

PrecisioN	0.9100
Recall	0.9079
F1	0.9079

TABLE II
Model 2 Metrics

V. Discussion

A. Model Comparison:

The comparison between model1 and model2 indicates that making changes to hyperparameters, such as the regularization strength (C), can have a substantial effect on the performance of logistic regression models. Model 2, with a reduced regularization parameter (C=0.9), achieved better performance than model 1 on both the test and validation sets.

B. Generalization:

The superior accuracy observed on the validation set for both models suggests that they have a strong ability to perform well on data that they have not been trained on. Nevertheless, it is crucial to take into account the possibility of overfitting, particularly when attaining nearly flawless accuracy on the validation set, as observed in the instance of the model.

Additional analysis should incorporate precision, recall, and F1 scores to assess the performance of the models across various sentiment categories, in addition to accuracy. Furthermore, investigating alternative methods of representing features and designing model architectures could provide valuable insights for enhancing the overall performance of sentiment analysis.

VI. Limitations

The dataset used in this study may not be representative of the diverse and dynamic nature of Twitter data, as it was collected from a specific time period and focused on a limited number of entities. The sentiment categories used in this study may not capture the nuances and complexities of emotions expressed on Twitter, as they only include positive, negative, and neutral labels. Moreover, the neutral label may be ambiguous, as it encompasses both irrelevant and mixed sentiments. The models used in this study may not be able to handle the challenges posed by Twitter data, such as noise, sarcasm, slang, abbreviations, and emoticons, which may affect the accuracy and interpretability of the results.

VII. Future Work

To gather and examine more up-to-date and varied Twitter data, encompassing a broader spectrum of entities and subjects, and employing various sampling and filtering techniques. The objective is to investigate and assess emotion categories that are more detailed and multifaceted, such as anger, joy, sadness, surprise, etc., and to devise techniques for identifying and managing sentiments that are mixed or contradictory. To utilize and evaluate more sophisticated and tailored machine learning and deep learning models, such as neural networks, convolutional neural networks, recurrent neural networks, transformers, etc., and to integrate additional characteristics and methods, such as word embeddings, sentiment lexicons, attention mechanisms, etc.

VIII. Conclusion

In summary, the logistic regression models, specifically model 2 with optimized hyperparameters, showed impressive accuracy in predicting sentiment on both the test and validation sets. These findings emphasize the significance of hyperparameter tuning in maximizing the effectiveness of logistic regression models for sentiment analysis on Twitter data.

The validation set's high accuracy indicates that these models have the potential to be applied in real-world scenarios, such as brand monitoring or public opinion analysis. Nevertheless, it is crucial to thoroughly examine the risk of overfitting and conduct a comprehensive evaluation of the model's ability to generalize across various datasets. These areas should be prioritized for future research.

These findings add to the ongoing discussion on methodologies for sentiment analysis, highlighting the iterative process of developing models and the necessity for constant improvement to account for the ever-changing nature of language used on social media.