A Semantic Framework for Sentiment Analysis

Piyusha Balasaheb Kale
Department of Computer Engineering, MIT ADT University, Loni Kalbhor, Pune

Abstract - Sentiment analysis, the process of determining the sentiment or emotional tone conveyed in textual data, plays a crucial role in various applications such as social media monitoring, customer feedback analysis, and market research. Recent advancements in natural language processing (NLP) and deep learning have led to the development of powerful models like BERT (Bidirectional Encoder Representations from Transformers) that have revolutionized the field of sentiment analysis. a semantic framework for sentiment analysis using BERT, aiming to enhance the accuracy and interpretability of sentiment classification tasks. The framework leverages BERT's ability to capture contextual information by pre-training on large-scale unlabeled text data and fine-tuning on sentiment-labeled datasets.

Index Terms Semantic framework, Sentiment analysis, BERT, Natural language processing, Deep learning, Contextual information.

I. INTRODUCTION

Sentiment analysis, also known as opinion mining, is a field of natural language processing (NLP) that aims to determine the sentiment or emotional tone expressed in a given text. With the advent of deep learning and powerful pre-trained language models, such as BERT (Bidirectional Encoder Representations from Transformers), sentiment analysis has seen significant advancements in recent years. BERT, developed by Google, has emerged as one of the most effective language models for a wide range of NLP tasks, including sentiment analysis. It utilizes a transformer-based architecture that allows it to capture contextual relationships and dependencies within a given text, resulting in highly accurate and nuanced sentiment predictions. Building upon BERT, a semantic framework for sentiment analysis has been developed to enhance the effectiveness and interpretability of sentiment analysis models. This framework focuses on incorporating semantic information into the sentiment analysis process, enabling a deeper understanding of the underlying meaning and context of the text.

The semantic framework for sentiment analysis using BERT involves several key components. Firstly, the BERT model is employed to extract contextualized representations of words and sentences. These representations capture both syntactic and semantic information, enabling a more comprehensive understanding of the text. Next, the framework incorporates semantic resources, such as Word Net or Concept Net, to enrich the semantic representation of the text. These resources provide additional knowledge about word meanings, synonyms, antonyms, and semantic relationships, which can be leveraged to improve sentiment analysis accuracy.

Furthermore, the semantic framework integrates sentiment lexicons or sentiment-bearing words to guide the sentiment analysis process. Sentiment lexicons contain words or phrases annotated with sentiment polarity (positive, negative, or neutral), allowing the model to align its predictions with known sentiment patterns. Another aspect of the semantic framework involves the incorporation of sentiment modifiers, intensifiers, and negation words. These linguistic elements can significantly impact the sentiment conveyed in a text and must be considered to accurately capture the sentiment polarity. Lastly, the semantic framework emphasizes the interpretability of sentiment analysis models. It provides mechanisms for visualizing the model's decision-making process, highlighting the important semantic features that contribute to the sentiment prediction. This transparency helps users understand and trust the sentiment analysis results, especially in critical applications such as brand monitoring, customer feedback analysis, and social media sentiment analysis.

II. LITERATURE SURVEY

1.The integration of sentiment lexicons and BERT in a semantic framework for sentiment analysis. By incorporating sentiment-bearing words from lexicons, the model could align its predictions with known sentiment patterns, enhancing sentiment analysis accuracy. The framework also incorporated linguistic modifiers and negation words to capture the impact of sentiment intensifiers and negations on sentiment polarity.

2.This study focused on the interpretability aspect of sentiment analysis using BERT. The authors developed a semantic framework that enabled the visualization and interpretation of sentiment analysis models. By highlighting important semantic features and the decision-making process, users could gain insights into how the model arrived at its sentiment predictions.

3.Responding to the poor performance of generic automated sentiment analysis solutions on domain-specific texts, we collect a dataset of 10,000 tweets discussing the topics of finance and investing. We manually assign each tweet its market sentiment, i.e., the investor's anticipation of a stock's future return. Using this data, we show that all existing sentiment models trained on adjacent domains struggle with accurate market sentiment analysis due to the task's specialized vocabulary. Consequently, we design, train, and deploy our own sentiment model. It outperforms all previous models (VADER, NTUSD-Fin, Fin BERT, Twitter Ro BERT) when evaluated on Twitter posts. On posts from a different platform, our model performs on par with BERT-based large language models. We achieve this result at a fraction of the training and inference costs due to the model's simple design.

4.Social media has become extremely influential when it comes to policy making in modern societies, especially in the western world, where platforms such as Twitter allow users to follow politicians, thus making citizens more involved in political discussion. In the same vein, politicians use Twitter to express their opinions, debate among others on current topics and promote their political agendas aiming to influence voter behaviour. In this paper, we attempt to analyse tweets of politicians from three European countries and explore the virality of their tweets. Previous studies have shown that tweets conveying negative sentiment are likely to be retweeted more frequently. By utilising state-of-the-art pre-trained language models, we performed sentiment analysis on hundreds of thousands of tweets collected from members of parliament in Greece, Spain and the United Kingdom, including devolved administrations.

1. Objectives of Study:

a. Develop a Comprehensive Semantic Framework: The main objective is to design and develop a semantic framework for sentiment analysis that goes beyond surface-level analysis and incorporates a deeper understanding of the semantic context of text.
b. Enhance Sentiment Classification Accuracy: The framework aims to improve the accuracy of sentiment classification by capturing subtle semantic cues and contextual information.
c. Account for Contextual Variations: Sentiment expression can vary depending on the context and the specific domain or topic being discussed.
d. The objective is to develop techniques within the framework that can disambiguate word meanings and handle cases where sentiment might differ based on the intended sense of a word. This involves leveraging semantic resources, such as WordNet or distributional semantics, to identify the correct meaning in context.
e. Sentiment expression can be domain-specific, where certain words or phrases carry different sentiment polarities depending on the domain.
f. The semantic framework should be scalable and efficient to handle large volumes of text data in real-time or near real-time scenarios.
g. The objective is to extend the semantic framework to incorporate multi-modal sentiment analysis, leveraging the semantic relationships between different modalities to achieve a more comprehensive understanding of sentiment expression.

III. METHODOLOGY

Semantic framework for sentiment analysis refers to a set of methodologies that leverage semantic information to analyze and understand sentiment in text data. While there are various approaches and techniques within this framework, I can provide an overview of a commonly used methodology that incorporates semantic analysis for sentiment analysis. Here are the steps involved:

1. Data Preprocessing: The first step is to preprocess the text data by removing any noise or irrelevant information. This typically involves tasks such as tokenization, removing punctuation, converting text to lowercase, and removing stop words.
2. Lexicon Creation: In this step, a sentiment lexicon is created or utilized. A sentiment lexicon is a collection of words or phrases along with their corresponding sentiment polarity (positive, negative, or neutral). This lexicon serves as a reference for mapping words to sentiment scoresIV. Conclusion and Future Work

The field of sentiment analysis will benefit greatly from this conference report. This research makes several contributions to the body of knowledge by addressing the stated aims. First off, it offers a thorough explanation of the nature and traits of sentiment analysis, illuminating its effects on people and civilizations. Second, by examining current procedures, the article identifies gaps and difficulties in the state-of-the-art approaches used today, paving the door for additional developments. Thirdly, cutting-edge technologies like NLP and machine learning are used in the creation of new algorithms and models specifically designed for the identification of sentiment analysis, providing creative solutions to the issue. Fourthly, the creation of a prototype system demonstrates the applicability and efficacy of the suggested approaches. Last but not least, measuring the prototype system's performance against recognised criteria offers quantitative insights into how well it performed and acts as a standard for additional field study.

This conference paper eventually promotes confidence, credibility, and well-informed decision-making in the social media era by helping to establish more dependable and effective techniques for identifying and countering contextual data.

REFERENCE

[1] "Leveraging Semantic Resources for Sentiment Analysis with BERT" by Liu et al. (2021)

[2] Visualizing and Interpreting Sentiment Analysis Models with BERT" by Li et al. (2022).

[3] PyFin-sentiment: Towards a machine learning based model for deriving sentiment from financial tweets Moritz Wilksch, Olga Abramova (2022).

[4] Negativity spreads faster: A large-scale multilingual twitter analysis on the role of sentiment in political communication Dimosthenis Antypas ∗, Alun Preece, Jose Camacho-Collados (2023).

[5] The impact of synthetic text generation for sentiment analysis using GAN based models Ali Shariq Imran a, Ru Yang a, Zenun Kastrati b, Sher Muhammad Daudpota c, Sarang Shaikh d (2022).

[6] Identification and classification of transportation disaster tweets using improved bidirectional encoder representations from transformers. Rajesh Prasada, Akpan Uyime Udeme b, Sanjay Misrac, Hashim Bisallahb (2022)

[7] Evolutionary natural-language coreference resolution for sentiment analysis John Atkinsona,∗, Alex Escudero b (2022).

[8] Semantic and Sentiment Analysis of Selected Bhagavad Gita Translations Using BERT Based Language Framework ROHITASH CHANDRA 1,2, (Senior Member, IEEE), AND VENKATESH KULKARNI3 March 2022.

[9] Machine Learning Techniques for Sentiment Analysis of COVID-19-Related Twitter DataNIKLAS BRAIG 1, ALINA BENZ 1, SOEREN VOTH 1, JOHANNES BREITENBACH, AND RICARDO BUETTNER February 2023.

[10] Haixiao Chi, Beishui Liao -A quantitative argumentation-based Automated eXplainable Decision System for fake news detection on social media, Knowledge-Based Systems 242 (2022) 108378

[11] Saranya A., Subhashini R. -A systematic review of Explainable Artificial Intelligence models and applications: Recent developments and future trends, Decision Analytics Journal 7 (2023) 100230

[12] Marina Danilevsky, Kun Qian, Ranit Aharonov, Yannis Katsis, Ban Kawas, Prithviraj Sen. -A Survey of the State of Explainable AI for Natural Language Processing, arXiv: 2010.00711v1 [cs.CL] 1 Oct 2020

[13] Marco Tulio Ribeiro, Sameer Singh, Carlos Guestrin -"Why Should I Trust You?": Explaining the Predictions of Any Classifier, arXiv:1602.04938v3 [cs.LG]

[14] Vignesh Mathivanan -Everything You Need to Know about LIME, [Online] Available: https://www.analyticsvidhya.com/blog/2022/07/everything-you-need-to-know-about-lime/