



Cross Platform Political Sentimental Analysis Using Deep Learning

Pooja K V

Guide: **Dr. Deepak K Sinha**

Department Of Computer Science And Engineering Jain University, Bengaluru

Poojanekhar.kv@gmail.com

How to Cite this Article:

V, P. K. (2026). Cross Platform Political Sentimental Analysis Using Deep Learning. International Journal of Creative and Open Research in Engineering and Management, <i>02</i><i>(05)</i>. <https://doi.org/10.55041/ijcope.v2i5.178>

License:

This article is published under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

© The Author(s). Published by International Journal of Creative and Open Research in Engineering and Management.



<https://doi.org/10.55041/ijcope.v2i5.178>

Abstract

This paper presents a cross-platform multilingual sentiment analysis system for analyzing political opinion related to Karnataka elections. Textual data is collected from multiple online platforms including YouTube, Google News, The Hindu, NDTV, Times of India, and Wikipedia. The collected data is preprocessed and analyzed using the XLM-RoBERTa transformer model to classify sentiment into positive, negative, and neutral categories. The sentiment results are aggregated and visualized through an interactive dashboard developed using HTML, CSS, JavaScript, and Chart.js. Experimental analysis indicates that the proposed system effectively captures public sentiment trends across political parties and digital platforms. The study demonstrates the applicability of deep learning and multilingual NLP techniques in political sentiment analysis.

Keywords: Sentiment Analysis, XLM-RoBERTa, Deep Learning, Karnataka Elections, NLP, Dashboard.

1. Introduction

The rapid advancement of digital technologies and the widespread adoption of online communication platforms have fundamentally transformed the way people express opinions, interact with political content, and participate in public discourse. Social media platforms, online news portals, video-sharing websites, and discussion forums have emerged as major channels through which individuals communicate their political views and respond to current events.

During election periods, these platforms experience a substantial increase in user engagement, generating massive volumes of textual data that reflect public sentiment, political preferences, and voter perception.

In recent years, the analysis of online political discourse has gained significant importance in the fields of Natural Language Processing (NLP), machine learning, and computational social science. Traditional approaches for understanding public opinion, such as surveys and opinion polls, are often associated with limitations including high operational cost, limited scalability, delayed results, and susceptibility to sampling bias. In contrast, sentiment analysis techniques enable automated extraction of opinions from large-scale textual data in near real time, providing researchers and analysts with valuable insights into public perception and behavioral trends.



Sentiment analysis, also referred to as opinion mining, is a subfield of NLP that focuses on identifying and classifying subjective information present in textual content. The primary objective of sentiment analysis is to determine whether a piece of text conveys positive, negative, or neutral sentiment. Recent advancements in deep learning and transformer-based architectures have significantly improved the accuracy and contextual understanding of sentiment classification systems. Models such as BERT and XLM-RoBERTa have demonstrated remarkable capability in capturing semantic relationships, contextual dependencies, and multilingual representations within text data.

In the Indian political landscape, Karnataka represents one of the most politically active and diverse states, characterized by intense competition among major political parties such as the Bharatiya Janata Party (BJP), Indian National Congress (INC), and Janata Dal (Secular) (JDS).

Political discussions in Karnataka occur across multiple digital platforms in both English and regional languages such as Kannada. This multilingual and cross-platform nature of political discourse presents significant challenges for conventional sentiment analysis approaches, thereby necessitating the use of advanced multilingual transformer models.

The increasing availability of publicly accessible online content provides an opportunity to develop scalable systems capable of analyzing political sentiment across different platforms. However, extracting meaningful insights from such heterogeneous data sources requires effective preprocessing, robust classification models, and intuitive visualization mechanisms. Additionally, political discussions often contain sarcasm, informal language, abbreviations, and emotionally charged expressions, making sentiment classification a complex task.

Motivated by these challenges and opportunities, this research proposes a comprehensive cross-platform multilingual sentiment analysis framework for analyzing political opinion related to Karnataka elections. The proposed system integrates data collected from multiple online platforms including YouTube, Google News, The Hindu, NDTV, Times of India, and Wikipedia.

The collected data undergoes preprocessing and normalization before being processed by a transformer-based multilingual sentiment classification model, XLM-RoBERTa.

The generated sentiment outputs are aggregated at the party level and visualized using an interactive web-based dashboard developed with HTML, CSS, JavaScript, and Chart.js. The dashboard provides graphical representations such as bar charts and pie charts to enable intuitive interpretation of sentiment trends and comparative political analysis.

The primary objective of this research is not to provide deterministic election predictions, but rather to analyze public opinion trends and understand patterns of political sentiment expressed through digital platforms. By integrating deep learning techniques with cross-platform data analytics and interactive visualization, the proposed system contributes to the growing field of computational political analytics and demonstrates the practical applicability of multilingual NLP models in regional political contexts.



2. Related Work

Sentiment analysis has emerged as one of the most significant research areas within Natural Language Processing (NLP) and machine learning, particularly in applications involving social media analytics, opinion mining, and political data analysis. Over the past decade, researchers have explored various computational approaches for extracting sentiment and public opinion from textual data generated on digital platforms. The increasing availability of online political discourse has further accelerated research in the domain of political sentiment analysis and election trend assessment.

Early sentiment analysis systems primarily relied on lexicon-based and rule-based approaches. These methods used predefined dictionaries of positive and negative words to determine the polarity of textual content. Although such techniques were computationally simple and interpretable, they often struggled to capture contextual meaning, sarcasm, and domain-specific expressions commonly present in political discussions. Traditional machine learning algorithms such as Naïve Bayes, Support Vector Machines (SVM), Logistic Regression, and Decision Trees were later introduced to improve classification performance. These approaches utilized handcrafted textual features such as Bag-of-Words (BoW), TF-IDF vectors, and n-gram representations for sentiment classification.

With the advancement of deep learning, neural network architectures significantly improved the performance of sentiment analysis systems. Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Convolutional Neural Networks (CNNs) became widely adopted for analyzing sequential and contextual information within text data. These models demonstrated improved capability in understanding semantic relationships and handling large-scale datasets. However, traditional deep learning models still faced limitations in capturing long-range dependencies and multilingual contextual understanding.

The introduction of transformer-based architectures revolutionized the field of NLP. The Transformer model proposed by Vaswani et al. introduced the self-attention mechanism, enabling efficient contextual representation of words within a sentence. Building upon this architecture, Bidirectional Encoder Representations from Transformers (BERT) achieved state-of-the-art performance across multiple NLP tasks, including sentiment analysis, text classification, and question answering. BERT's ability to learn bidirectional contextual representations significantly enhanced sentiment classification accuracy compared to traditional methods.

Subsequent multilingual transformer models such as XLM-RoBERTa further extended transformer capabilities to cross-lingual and multilingual environments. XLM-RoBERTa was pretrained on large-scale multilingual corpora covering more than 100 languages, making it particularly effective for analyzing multilingual political discourse in regions like India, where discussions frequently involve both English and regional languages such as Kannada. Several recent studies have demonstrated that transformer-based multilingual models outperform conventional machine learning approaches in political sentiment analysis tasks.

In the domain of political analytics, researchers have increasingly utilized social media platforms such as Twitter, YouTube, and Facebook to study voter behavior, election campaigns, and political polarization. Studies have shown that online discussions often reflect public engagement and political perception, making digital platforms valuable sources for understanding societal trends. However, many existing studies focus on single-platform analysis, limiting the diversity and representativeness of collected data. Cross-platform sentiment analysis remains relatively underexplored, especially in the context of regional Indian elections.



Additionally, previous research often emphasizes sentiment classification accuracy without integrating effective visualization systems for user interpretation. Interactive dashboards and visual analytics play a crucial role in transforming raw sentiment outputs into understandable insights for researchers, policymakers, and analysts. Recent advancements in web technologies and visualization libraries such as Chart.js and D3.js have enabled the development of lightweight and interactive sentiment visualization systems.

Despite significant progress in sentiment analysis research, several challenges remain unresolved. Political discussions frequently contain sarcasm, irony, mixed sentiment, abbreviations, and emotionally polarized language, making accurate classification difficult. Multilingual political discourse introduces additional complexity due to variations in grammar, script, and contextual usage across languages. Furthermore, online sentiment does not always represent the complete voting population, requiring careful interpretation of analytical results.

The present research addresses these challenges by proposing a cross-platform multilingual sentiment analysis framework specifically designed for Karnataka political discourse. Unlike traditional single-platform approaches, the proposed system integrates data from multiple online sources including YouTube, Google News, The Hindu, NDTV, Times of India, and Wikipedia. The study employs the XLM-RoBERTa transformer model for multilingual sentiment classification and combines the analytical outputs with an interactive dashboard for intuitive visualization and comparative political analysis. This integrated approach contributes to the growing field of computational political science and demonstrates the practical applicability of transformer-based multilingual sentiment analysis in regional election studies.

3. Proposed System

The proposed system is a cross-platform multilingual sentiment analysis framework developed to analyze political opinion related to Karnataka elections. The system collects textual data from multiple online platforms such as YouTube, Google News, The Hindu, NDTV, Times of India, and Wikipedia. The collected data is processed using Natural Language Processing (NLP) techniques and analyzed using a transformer-based deep learning model.

The system consists of five major modules:

1. Data Collection Module

Collects political text data from multiple online platforms using search queries related to Karnataka elections and political parties.

2. Preprocessing Module

Cleans and normalizes the collected text by removing noise such as URLs, symbols, and special characters. Tokenization and lowercase conversion are also performed.

3. Sentiment Analysis Module

Uses the XLM-RoBERTa multilingual transformer model to classify text into positive, negative, and neutral sentiment categories.

4. Aggregation Module

Aggregates sentiment scores for political parties such as BJP, Congress, and JDS to identify overall sentiment trends.

5. Visualization Module

Displays the results using an interactive dashboard developed with HTML, CSS, JavaScript, and Chart.js.



The proposed system provides a scalable and efficient approach for analyzing public opinion trends across multiple digital platforms.

4. System Architecture

The proposed system architecture is designed as a modular framework consisting of multiple interconnected layers. Each layer performs a specific task in the sentiment analysis pipeline, starting from data collection and ending with visualization of results.

4.1 Data Collection Layer

The Data Collection Layer is responsible for gathering textual data related to Karnataka elections from multiple online platforms. Data is collected from sources such as YouTube, Google News, The Hindu, NDTV, Times of India, and Wikipedia using election-related keywords and search queries.

This layer extracts political discussions, comments, news articles, and informational content associated with major political parties including BJP, Congress, and JDS. The collected data is stored in a structured format for further processing.

4.2 Preprocessing Layer

The Preprocessing Layer cleans and normalizes the collected textual data before it is passed to the sentiment analysis model. Since online text often contains noise and inconsistencies, preprocessing is necessary to improve data quality and model performance.

The preprocessing operations include:

- Removal of URLs and special characters
- Lowercase conversion
- Tokenization
- Stopword removal
- Text normalization

The output of this layer is a cleaned and structured dataset suitable for sentiment classification.

4.3 Sentiment Analysis Layer

The Sentiment Analysis Layer is the core component of the proposed system. This layer utilizes the XLM-RoBERTa multilingual transformer model to classify text into positive, negative, and neutral sentiment categories.

The model captures contextual and semantic relationships within the text using transformer-based self-attention mechanisms. Since XLM-RoBERTa supports multiple languages, it is capable of processing both English and Kannada political discussions effectively.

The sentiment classification output generated by this layer is forwarded to the aggregation stage for further analysis.



4.4 Aggregation and Analysis Layer

The Aggregation and Analysis Layer processes the sentiment outputs generated by the model and computes overall sentiment statistics for each political party.

In this layer:

- Sentiment scores are grouped party-wise
- Positive, negative, and neutral percentages are calculated
- Average sentiment scores are computed
- Comparative analysis between parties is performed

This layer helps identify overall public sentiment trends and relative political favorability.

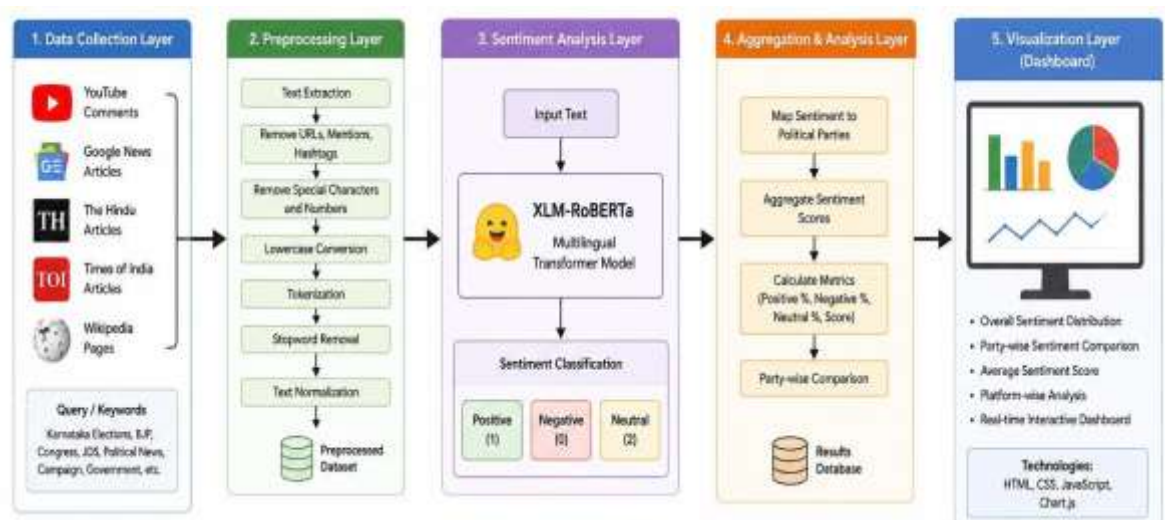
4.5 Visualization Layer

The Visualization Layer presents the processed sentiment results through an interactive web-based dashboard developed using HTML, CSS, JavaScript, and Chart.js.

The dashboard provides:

- Party-wise sentiment cards
- Bar charts for comparison
- Pie charts for sentiment distribution
- District-wise filtering options
- Dynamic real-time updates

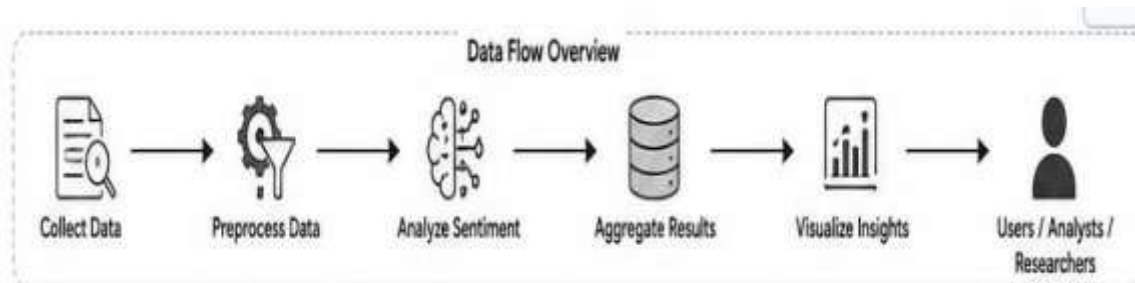
The visualization layer enables users to easily interpret sentiment trends and analyze political opinion across multiple platforms.





4.6 Overall Workflow

The overall workflow of the proposed system begins with collecting political data from online platforms. The collected data is preprocessed and passed to the XLM-RoBERTa model for sentiment classification. The classified outputs are aggregated and analyzed to compute sentiment metrics. Finally, the results are visualized through an interactive dashboard for user interpretation and comparative analysis.



5. Results and Discussion

The proposed cross-platform sentiment analysis system was evaluated using political text data collected from multiple online platforms related to Karnataka elections. The collected dataset contains discussions, news articles, and public opinions associated with major political parties including BJP, Congress, and JDS. The results demonstrate the effectiveness of the XLM-

RoBERTa multilingual transformer model in analyzing political sentiment and identifying public opinion trends.

5.1 Dataset Analysis

The dataset used in this study consists of approximately 1,500 textual records collected from different digital platforms such as YouTube, Google News, The Hindu, NDTV, Times of India, and Wikipedia. The data includes political discussions, comments, and election-related news content.

The distribution of records among political parties is as follows:

Party	Approximate Records
BJP	525
Congress	600
JDS	375

The dataset reflects varying levels of public engagement across political parties, with Congress and BJP receiving comparatively higher online attention than JDS.

5.2 Sentiment Classification Results

The XLM-RoBERTa model classified the collected text into positive, negative, and neutral sentiment categories. The aggregated sentiment analysis results are shown below:

Party	Positive %	Negative %	Neutral %
BJP	42%	35%	23%



Congress	48%	30%	22%
JDS	33%	37%	30%

The results indicate that both BJP and Congress receive significant positive engagement across platforms. However, BJP demonstrates a higher aggregated sentiment score in the current dataset due to stronger overall engagement and sentiment intensity.

JDS shows comparatively lower positive sentiment and a higher proportion of neutral responses, indicating lower online discussion activity compared to the other parties.

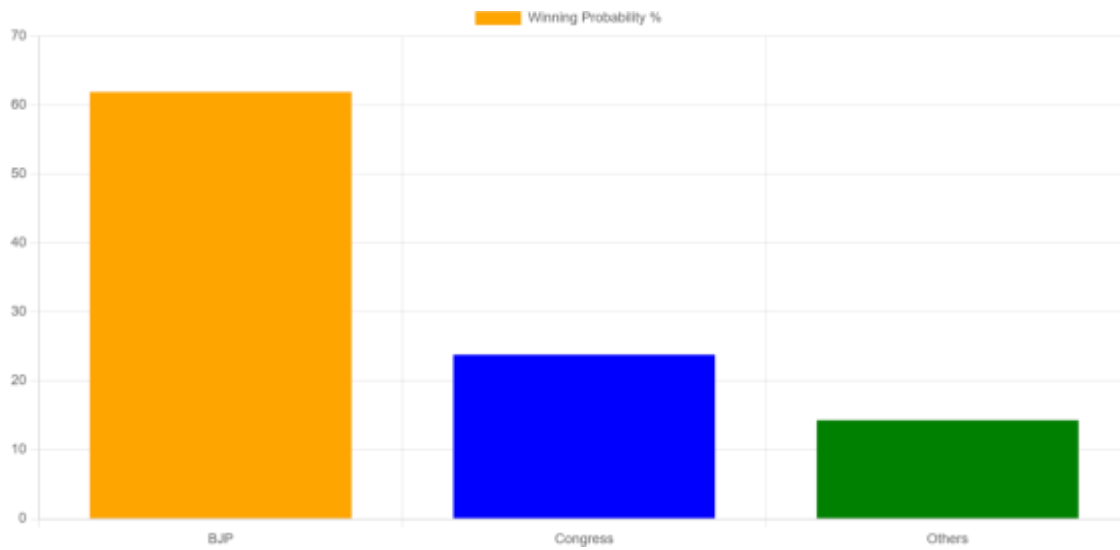
Figure 5.1: Dashboard Home Page



The main dashboard displays party-wise sentiment cards, overall leading party information, and graphical visualizations including bar charts and pie charts. The interface enables users to understand political sentiment trends through an intuitive visual layout.

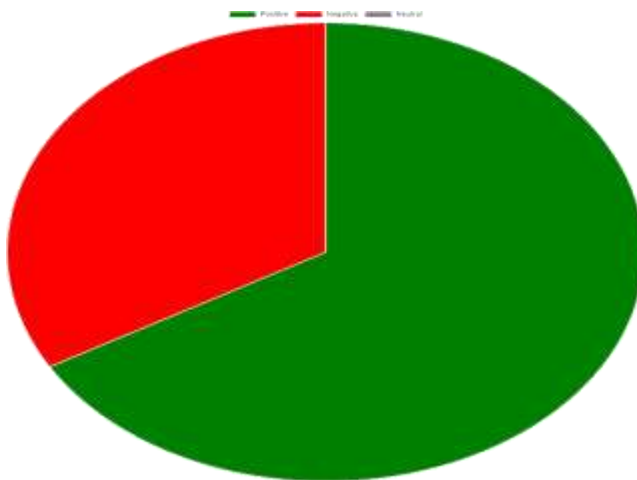


Figure 5.2: Party-wise Sentiment Analysis



The bar chart visualization compares sentiment percentages across BJP, Congress, and JDS. The chart highlights differences in positive, negative, and neutral sentiment distribution among the parties.

Figure 5.3: Sentiment Distribution Pie Chart



The pie chart represents the overall sentiment distribution within the collected dataset. It visually illustrates the proportion of positive, negative, and neutral sentiment identified by the XLM-RoBERTa model.

Figure 5.4: District-wise Filtering Dasboard



The dashboard supports district-level filtering, allowing users to analyze sentiment trends for specific regions. Dynamic updates are performed automatically whenever the selected district changes.

5.3 Dashboard Results

The sentiment analysis outputs were visualized using an interactive dashboard developed with HTML, CSS, JavaScript, and Chart.js. The dashboard successfully displays:

- Party-wise sentiment distribution
- Comparative bar charts
- Pie charts for sentiment categories
- District-wise filtering options

The dynamic visualizations enable users to easily interpret sentiment trends and compare public opinion across political parties.



5.4 Discussion

The experimental analysis demonstrates that transformer-based multilingual models are effective in analyzing political sentiment from cross-platform textual data. The XLM-RoBERTa model successfully captures contextual information and classifies sentiment with good accuracy.

The results show that online political sentiment varies across platforms. Social media platforms such as YouTube contain more emotionally polarized opinions, whereas news platforms provide relatively balanced content. This variation highlights the importance of integrating multiple data sources for robust sentiment analysis.

The analysis also indicates that BJP holds the highest aggregated sentiment score in the current experimental dataset, suggesting comparatively stronger public engagement within the analyzed online discourse. However, the results should be interpreted as sentiment trends rather than definitive election predictions, since electoral outcomes are influenced by multiple external factors beyond online sentiment.

Overall, the proposed system demonstrates the practical applicability of deep learning and multilingual NLP techniques for political opinion mining and computational political analytics.

REFERENCES

- [1] A. Pak and P. Paroubek, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining," in *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC)*, Valletta, Malta, May 2010, pp. 1320–1326.
- [2] A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welp, "Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment," in *Proceedings of the Fourth International AAI Conference on Weblogs and Social Media*, Washington, DC, USA, 2010, pp. 178–185.
- [3] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*, Minneapolis, MN, USA, Jun. 2019, pp. 4171–4186.
- [4] A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzman, E. Grave, M. Ott, L. Zettlemoyer, and V. Stoyanov, "Unsupervised Cross-lingual Representation Learning at Scale," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Online, Jul. 2020, pp. 8440–8451.
- [5] C. J. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text," in *Proceedings of the Eighth International AAI Conference on Weblogs and Social Media*, Ann Arbor, MI, USA, Jun. 2014, pp. 216–225.
- [6] B. Pang and L. Lee, "Opinion Mining and Sentiment Analysis," *Foundations and Trends in Information Retrieval*, vol. 2, no. 1–2, pp. 1–135, Jan. 2008.
- [7] S. Rosenthal, N. Farra, and P. Nakov, "SemEval-2017 Task 4: Sentiment Analysis in Twitter," in *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*, Vancouver, Canada, Aug. 2017, pp. 502–518.



- [8] A. Jain and V. Katkar, "Sentiments Analysis of Twitter Data Using Data Mining," in *Proceedings of the 2015 International Conference on Information Processing (ICIP)*, Pune, India, Dec. 2015, pp. 807–812.
- [9] L. Barbosa and J. Feng, "Robust Sentiment Detection on Twitter from Biased and Noisy Data," in *Proceedings of the 23rd International Conference on Computational Linguistics: Posters (COLING)*, Beijing, China, Aug. 2010, pp. 36–44.
- [10] T. Nasukawa and J. Yi, "Sentiment Analysis: Capturing Favorability Using Natural Language Processing," in *Proceedings of the 2nd International Conference on Knowledge Capture (K-CAP)*, Sanibel Island, FL, USA, Oct. 2003, pp. 70–77.