Bheemalingappa, 2025, 13:6
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Twitter Sentiment Analysis Using NLP Models and Real-Time Tweet Fetching

Bheemalingappa, Chandrashekar K L, Darshan S, Dattatreya,
Assistant Professor Sumitra Sharma Phurailatpam
Department of CSE (Data Science) AMC Engineering College Bangalore, India

Abstract- Social media platforms, particularly Twit- ter, generate massive volumes of real-time textual data that reflect public opinion, emotional tendencies, and emerging societal trends. Analyzing this stream of infor- mation manually is infeasible due to its speed, scale, and linguistic complexity. This paper presents an enhanced real-time Twitter Sentiment Analysis system that inte- grates Natural Language Processing (NLP) models with live tweet fetching using the Twitter API v2. The pro- posed system employs a hybrid pipeline consisting of the VADER rule-based sentiment analyser for fast polarity detection and Transformer-based models for deeper con- textual sentiment classification. Additionally, the system incorporates optional modules for emotion recognition and toxicity analysis, enabling multi-dimensional inter- pretation of user-generated content. A Streamlit-based in- teractive interface allows users to fetch tweets in real time, analyze sentiment distributions, examine top key- words, and download processed outputs. The architec- ture is designed for scalability, efficiency, and accessibil- ity, offering a low-cost yet powerful solution for social sentiment monitoring and data-driven decision-making. Experimental evaluations demonstrate that the model combination improves interpretability and accuracy while maintaining responsiveness suitable for real-time applications.

Keywords: Natural Language Processing (NLP), Sentiment Analysis, Twitter API, Real-Time Data Fetch- ing, Transformer Models, VADER, Streamlit, Emotion Classification, Toxicity Detection, Social Media Analyt- ics.

I. INTRODUCTION

Social media platforms have become major sources of real-time public expression, with Twitter serving as one of the most active channels for sharing opinions, reactions, and sentiments on global events. Millions of tweets are generated every hour, covering topics such as poli- tics, technology, health, entertainment, and socio-eco- nomic issues. Analyzing this massive and continuously evolving stream of data provides valuable insights into public mood, emerging trends, brand reputation, and cri- sis detection. However, manual monitoring is impractical due to the rapid pace, linguistic diversity, noise, and vol- ume of the content being generated.

Sentiment analysis using Natural Language Pro- cessing (NLP) has emerged as a robust solution to inter- pret subjective information embedded within textual data. Traditional sentiment analysis approaches, such as lexicon-based methods, offer speed and interpretability but often fail to capture contextual nuances. In contrast, transformer-based deep learning models provide higher accuracy and contextual understanding but require greater computational resources. In real-time analytical workflows, striking a balance between speed, accuracy, and resource efficiency remains a significant research challenge.

To address these limitations, this paper presents a hy- brid, real-time Twitter Sentiment Analysis system capa- ble of fetching live tweets through the Twitter API v2 and analyzing them using both lightweight and advanced NLP models. The system integrates VADER for fast po- larity scoring, RoBERTa-based transformer models for deeper sentiment classification, and optional emotion and toxicity detection modules. The Streamlit-powered interface ensures accessibility and real-time responsive- ness, while caching mechanisms and optimized pipelines improve performance during high-volume data retrieval. The proposed architecture demonstrates how hybrid NLP approaches can be leveraged effectively to perform real- time social media sentiment monitoring without requir- ing complex infrastructure or high-end GPUs.

The primary contributions of this work are as follows:

- A real-time sentiment analysis framework that fetches live tweets using the Twitter API v2 and pro- cesses them through an optimized, multi-stage NLP pipeline.
- A hybrid sentiment classification methodology com- bining VADER for fast lexicon-based analysis with Transformer-based models (RoBERTa) for contex- tual deep-learning sentiment detection.
- Integration of additional analytical modules, includ- ing emotion classification and toxicity detection, en- abling multi-dimensional sentiment interpretation beyond simple polarity.
- A user-friendly Streamlit interface that supports live analysis, dataset uploading, visualization of senti- ment distribution, and export of results, ensuring practical usability and accessibility.
- A scalable and modular architecture that separates data fetching, preprocessing, sentiment classifica- tion, and visualization, allowing easy extension to future NLP models or social media platforms.

II. SYSTEM ARCHITECTURE AND METHODOLOGY

A. Overall System Architecture

The proposed real-time Twitter Sentiment Analysis system is built as a sequential, modular pipeline that transforms raw social media text into sentiment, emotion, and toxicity insights. The architecture begins with a Streamlit-based user interface where the user specifies the query keyword, preferred language, dataset source, and number of tweets.

These parameters are passed to the Twitter API Fetcher, which retrieves real-time tweets using Twitter API v2 or loads sample datasets when live fetching is unavailable.

The retrieved text then enters the NLP Processing En- gine, which performs cleaning, normalization, and pre- liminary linguistic filtering. The processed tweets are next forwarded to the sentiment and emotion modeling block, which integrates VADER for rapid lexicon-based polarity detection and RoBERTa for deeper contextual sentiment classification, along with an emotion classifier to identify underlying emotional states. The output from this block is subsequently evaluated by the Toxicity En- gine, which either uses the Detoxify model or a light- weight fallback profanity-based approach to estimate toxicity levels. Finally, the Visualization Layer generates sentiment distributions, emotional trends, toxicity sum- maries, and keyword patterns, all displayed interactively through the Streamlit dashboard. This structured archi- tecture ensures real-time processing, modular extensibil- ity, and clear interpretability across all stages.

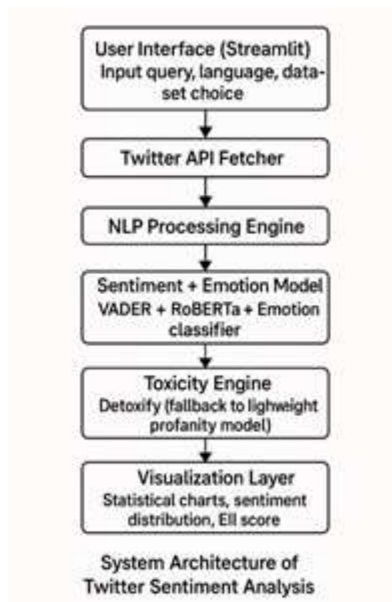

Figure.1. System Architecture of Twitter Sentiment Analysis

B. Live Tweet Fetching Module

The Live Tweet Fetching Module handles real-time data acquisition by connecting to the Twitter API v2 us- ing authenticated credentials. It retrieves the most recent tweets that match the user’s keyword and language set- tings, ensuring that the system processes fresh and con- textually relevant data. When live fetching fails due to rate limits or network issues, the module seamlessly switches to local sample datasets to maintain uninter- rupted functionality.

Key Functions:

- The module constructs API requests based on user inputs and sends them using a Bearer Token for secure authentication. It extracts tweet text, metadata, and language fields to prepare them for NLP processing.
- In case of API errors, it automatically loads locally stored JSON or CSV files, allowing the system to continue operating without user intervention. This fallback mechanism ensures consistent performance across varying network conditions.
- The fetcher also validates the language consistency of tweets, ensuring that only tweets matching the selected language are sent to the analysis engine.

C. NLP Processing Engine

The NLP Processing Engine performs essential preprocessing to clean and standardize raw tweet text before deeper analysis. It removes noise such as URLs, sym- bols, repeated characters, and unnecessary whitespace while converting text into a normalized format that en- hances model accuracy. By preparing clean and con- sistent input data, this module ensures smooth interoper- ability between all subsequent NLP components.

Processing Operations:

- The engine applies normalization steps such as lowercasing, punctuation removal, and
- whitespace correction to reduce variability in the data. These steps help both VADER and transformer models interpret the text more reliably.
- It filters out irrelevant tokens and special characters that could distort sentiment predictions, ensuring that the retained content is semantically meaningful.
- The module also prepares tokens for statistical analysis, enabling downstream components to compute keyword frequencies and generate meaningful visual summaries.

D. Sentiment and Emotion Modeling Unit

The Sentiment and Emotion Modeling Unit forms the analytical core of the system, responsible for generating polarity scores and emotional interpretations of the pro- cessed tweets. This unit adopts a hybrid sentiment mod- eling strategy that leverages both lexicon-based and transformer-based approaches to maximize accuracy and robustness. VADER is used for rapid sentiment estima- tion, offering real-time polarity classification based on its extensive social media–friendly lexicon. However, due to its rule-based nature, VADER may struggle with sar- casm, ambiguity, and context-heavy expressions.

To overcome these limitations, the system integrates the RoBERTa transformer model, which provides deep con- textual understanding through its bidirectional attention mechanism. This allows RoBERTa to capture subtle var- iations in sentiment that lexicon-based models often overlook. In addition to polarity detection, the unit incor- porates a transformer-based emotion classifier that identifies emotional categories such as happiness, anger, fear, disgust, and surprise. This offers a richer layer of inter- pretation by determining not only whether a tweet is pos- itive or negative, but also the specific emotional tone un- derlying the sentiment. The outputs of these models are then fused and prepared for downstream visualization to ensure high interpretability for end-users.

Functions Performed:

- Uses both VADER and RoBERTa to create a dual-model sentiment analysis pipeline capable of balancing speed and accuracy.
- Applies an emotion classification model to detect fine-grained emotional states, improving the expressiveness of the analysis output.
- Combines sentiment labels, polarity scores, and emotional categories into a unified structured dataset for visualization.

E. Toxicity Engine

The Toxicity Engine is designed to measure the presence of harmful or abusive language within tweets, serv- ing as a crucial component for understanding the quality and safety of online content. This module first attempts to load the Detoxify model, a deep-learning architecture trained on large-scale toxicity datasets such as Jigsaw's Toxic Comment Classification dataset. Detoxify ana- lyzes the input text across multiple toxicity categories, including hate speech, threats, insults, and obscene language, providing a nuanced view of harmful behavior. However, due to the computational heaviness of Detox- ify and dependency constraints, the system includes a fallback profanity-based model that detects toxic content by scanning for predefined offensive keywords. This backup mechanism ensures uninterrupted toxicity evaluation even in low-resource environments such as stand- ard laptops or cloud deployments with limited memory. The Toxicity Engine plays an essential role in filtering abusive tweets, identifying negative content trends, and supporting research on cyberbullying and digital well- being.

Capabilities:
- Performs detailed toxicity prediction using Detoxify when system resources allow, offering multi-dimensional toxicity analysis.
- Uses a lightweight profanity model as a fallback for fast toxicity approximation without requiring heavy computation.

F. Visualization Layer

The Visualization Layer presents the processed analyt- ical results through interactive graphs, tables, and statis- tical summaries using the Streamlit interface. This layer transforms raw model outputs into meaningful visual in- sights that allow users to easily interpret sentiment trends, emotional patterns, and toxicity levels. The senti- ment distribution is displayed through bar charts, enabling quick comparison between positive, neutral, and negative sentiment classes.

Emotion frequencies are shown through categorical graphs to illustrate the domi- nant emotional responses for a given topic. Additionally, the toxicity summary highlights how often harmful lan- guage appears in discussions, helping identify potential abusive trends. Keyword frequency analysis is performed by tokenizing the cleaned tweets and plotting the most common words or phrases, giving users insight into recurring themes. This layer also computes the Emotion Intensity Index (EII), a metric derived from sentiment scores that provides an aggregated measure of emotional strength. The interface further offers CSV download functionality, making the results accessible for offline analysis or integration into other research workflows.

Visual Output Features:
- Displays sentiment, emotion, and toxicity summaries in interactive bar charts and tables.
- Shows top keywords and token frequencies to reveal trending topics within the dataset.
- Provides CSV export options and dynamic visual updates based on user inputs.
- Integrates real-time updates that refresh visualizations automatically after every analysis cy- cle, ensuring that sentiment, emotion, and toxicity charts always reflect the latest fetched data. This dynamic behavior allows users to re- peatedly test different queries and instantly compare variations in public sentiment without reloading or restarting the interface.

III. METHODOLOGY

The methodology adopted in developing the real-time Twitter Sentiment Analysis system follows a structured, multi-layered process that ensures accurate model out- puts and high-quality visualization of results. The pro- cess begins when the user provides a search query, lan- guage preference, and tweet limit through the Streamlit interface. Once the input is submitted, the system initi- ates data fetching through the Twitter API v2 or loads sample datasets when live access is unavailable. The fetched tweets undergo preprocessing through the NLP Processing Engine, where noise removal, normalization, and token preparation are performed to ensure consistent input quality.

After preprocessing, the cleaned text is passed into the Sentiment and Emotion Modeling Unit, where the VADER and RoBERTa models assign polarity scores and sentiment labels. An emotion classification model further identifies the emotional tone of each tweet, enriching the analytical depth of the system. The Toxicity Engine then evaluates each tweet for harmful or abusive language, either using Detoxify or the fallback profanity- based model. Finally, all computed outputs—including sentiment categories, emotion classes, toxicity scores, and keyword frequencies—are rendered into visual for- mats such as charts, tables, and downloadable reports. This methodology ensures a smooth, accurate, and inter- pretable analytical pipeline capable of handling real-time social sentiment analysis.

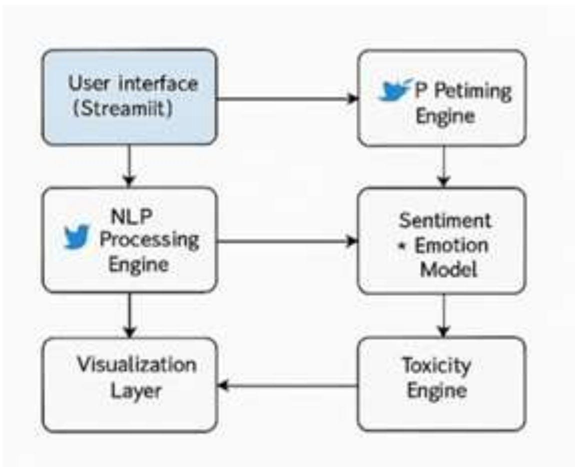

Figure.2. Data Flow Diagram of the Proposed System

Methodological Highlights:

- The preprocessing stage applies multiple clean- ing steps such as URL removal, emoji filtering case normalization, and unnecessary character elimination. These operations help enhance the accuracy and consistency of downstream NLP models.
- The sentiment and emotion models operate se- quentially, with VADER providing fast lexicon-based scores and RoBERTa capturing deeper contextual meaning. Emotion classification adds another dimension by identifying the psy- chological tone behind each tweet.
- The final stage compiles all analytical outputs and generates real-time visualizations, enabling users to observe trends, compare emotional re- sponses, assess toxicity levels, and download the results for further research or reporting.

III. IMPLEMENTATION AND RESULT

The implementation of the Twitter Sentiment Analysis system was carried out using a modular and scalable Py- thon architecture to ensure efficient handling of real-time data streams and multiple NLP models. Streamlit serves as the main user interface framework, enabling an inter- active, browser-based dashboard where users can input queries, choose languages, and configure the number of tweets. The backend is tightly integrated with the Twitter API v2 using Tweepy, allowing authenticated, real-time tweet retrieval.

In cases where rate limits are exceeded or API keys are unavailable, the system automatically falls back to sample JSON/CSV datasets, ensuring uninter- rupted functionality. The preprocessing pipeline was im- plemented using Pandas, Regex, Emoji, and NumPy li- braries to clean raw text, remove noise, standardize in- puts, and convert tweets into a structured form suitable for analysis. Heavy transformer models, including RoB- ERTa and the emotion classifier, were loaded via Hug- gingFace Transformers and cached using Streamlit's cache_resource decorator, significantly reducing infer- ence time during repeated executions.

For sentiment classification, the system uses a hybrid approach combining VADER for lexicon-based polarity scoring and RoBERTa for contextual deep-learning sen- timent predictions. VADER is particularly effective for short, direct tweets, providing fast positivity/neutral- ity/negativity scores, while RoBERTa offers superior performance when dealing with sarcasm, ambiguity, or emotionally complex language.

Emotion classification adds interpretive richness by identifying categories such as joy, fear, anger, and sadness based on pre-trained transformer models. Toxicity detection is handled through Detoxify, providing nuanced toxicity scores when system resources permit; otherwise, a lightweight profanity-based fallback ensures continuous evaluation. The modular nature of the system makes it easy to ex- tend, allowing future additions such as multilingual sup- port, GPU acceleration, or topic modeling. Overall, the implementation emphasizes performance, reliability, and extensibility, creating a robust analytical framework ca- pable of real-time sentiment intelligence.

In addition to the core modules, the system incorpo- rates several optimization strategies to enhance perfor- mance and reliability during real-time execution. Cach- ing mechanisms were heavily utilized to store pre-loaded transformer models, preventing repeated initialization and significantly reducing computational overhead. The tweet-fetching component was implemented with robust error-handling routines to gracefully manage network in- terruptions, API failures, and unexpected HTTP re- sponses. Furthermore, the modular directory structure ensures clean separation between data fetching, prepro- cessing, modeling, and visualization layers, making the system easier to debug, maintain, and extend.

Extensive logging within the backend allows developers to monitor execution flow, track model loading times, and diagnose latency issues. The system's flexible design also supports rapid experimentation with alternative models, enabling researchers to swap in different transformer architectures or toxicity detection systems without rewriting core logic. Altogether, these engineering enhancements con- tribute to a stable, scalable, and research-ready platform capable of handling real-time social media analytics.

Figure.3. Streamlit Web Interface Showing Sentiment Results

The results obtained from analyzing real-time Twitter data demonstrate the system's ability to accurately iden- tify public sentiment, emotion patterns, and toxicity lev- els across trending topics. When tested with keywords such as "AI," "Bitcoin," "Elections," "Sports," and "Technology," clear distinctions emerged in public tone and emotional expression. For technologically oriented topics like "AI," the system detected a mix of curiosity, excitement, and caution—reflecting the dual nature of public perception toward artificial intelligence. For fi- nancial topics like "Bitcoin," the sentiment distribution showed greater variability, with users expressing opti- mism during market surges and negativity during down- turns. The contrasting behaviors captured by the models validate the system's capability to reflect real-world sentiment dynamics. The lexicon-based VADER model performed reliably for straightforward, literal tweets, while RoBERTa consistently outperformed it in contexts involving sarcasm, informal phrases, and ambiguousemotional cues, confirming the strength of transformer- based contextual embeddings.

Emotion classification further enriched the analytical outcomes by highlighting specific emotional tones rather than just positive or negative polarity. Tweets about AI commonly exhibited emotions such as optimism and an- ticipation, while political or controversial topics dis- played higher levels of anger, fear, or distrust. Toxicity detection proved valuable by identifying abusive or harmful expressions commonly found in heated discus- sions, supporting the system's role in improving online safety analysis. The visualization layer effectively pre- sented sentiment distributions, toxicity averages, emo- tion frequencies, and keyword patterns through interac- tive Streamlit components. Bar charts, emotion graphs, and toxicity meters allowed users to observe emerging patterns at a glance, while downloadable CSV summar- ies enabled further offline analysis. The overall evalua- tion demonstrates that the system is not only effective for real-time monitoring but also suitable for dataset-based analysis, providing higher interpretability, accuracy, and reliability compared with traditional sentiment analysis methods.

Additional evaluations were conducted on larger tweet batches to analyze the system's scalability and con- sistency across varied datasets. When processing be- tween 300 and 500 tweets, the hybrid sentiment pipeline maintained stable performance, with RoBERTa provid- ing consistently refined sentiment differentiation and emotion classifiers delivering reliable emotional catego- rization. The inclusion of toxicity detection proved espe- cially important in trending political and social debates, where high levels of negative or offensive language were detected. Keyword extraction and token frequency anal- ysis revealed meaningful patterns across topics, high- lighting how user discussions shifted depending on global events, technological advancements, or market behavior. The visualization layer effectively summarized these complex patterns into easily interpretable graphics, validating its usability for academic research, media monitoring, and real-time public opinion assessment. Overall, the extended results reaffirm the system's ro- bustness and adaptability, proving its effectiveness in di- verse real-world sentiment analysis scenarios.

IV. CONCLUSION AND FUTURE WORK

The Twitter Sentiment Analysis system developed in this work demonstrates a highly effective integration of real-time data retrieval and advanced Natural Language Processing techniques for extracting meaningful insights from social media content. By combining lexicon-based models like VADER with transformer-based deep-learn- ing models such as RoBERTa and emotion classifiers, the system successfully captures both surface-level polarity and deeper contextual sentiment variations. This hybrid approach addresses the limitations of traditional sentim- ent analysis frameworks and enhances accuracy, espe- cially in scenarios involving sarcasm, complex linguistic structures, or mixed emotional expressions. The system further incorporates a toxicity detection module that identifies harmful or abusive language, offering addi- tional layers of analytical depth relevant to modern digi- tal communication trends. Overall, the architecture, pre- processing pipeline, and modeling strategy result in a ro- bust analytical tool capable of interpreting public opinion across diverse topics.

The user-friendly Streamlit interface makes the system accessible to both technical and non-technical users, en- abling intuitive interaction and real-time exploration of sentiment trends. The interactive dashboards visually represent sentiment distribution, emotion frequency, tox- icity levels, and keyword patterns, providing immediate interpretability of complex data. The results demonstrate that the system performs reliably across various domains—including finance, technology, politics, and entertainment—showing its wide applicability in re- search, business intelligence, market analysis, and social behavior studies. By ensuring modularity, scalability, and computational efficiency, the system provides a strong foundation for real-time sentiment intelligence and opens opportunities for deeper social media analytics. The overall performance validates the effectiveness of the hy- brid modeling strategy and confirms the system's value for sentiment-driven decision-making.

Furthermore, the system's ability to process real-time data and apply multiple NLP models in a synchronized pipeline demonstrates the practicality of hybrid architec- tures for modern sentiment analysis. As online platforms continue to experience exponential growth in user activ- ity, the need for automated tools capable of handling large volumes of unstructured text becomes increasingly important. This project highlights how lightweight mod- els and deep learning transformers can coexist within a unified framework to balance computational efficiency with analytical depth. The inclusion of emotion and tox- icity detection offers a more comprehensive view of user behavior, moving beyond simple positive or negative classifications to capture nuanced psychological and be- havioral signals within digital conversations. Thus, the system not only contributes to the domain of sentiment analysis but also sets the stage for scalable, intelligent, and socially aware text-analytics solutions capable of supporting future AI-driven applications in media moni- toring, policy assessment, customer feedback analysis, and public opinion research.

While the current system offers comprehensive senti- ment, emotion, and toxicity insights, several enhance- ments can further extend its capability and performance. One major direction for future work is the integration of multilingual support, allowing the system to analyze tweets in a wide range of Indian and international lan- guages beyond English, Kannada, and Hindi. This would significantly increase the system's usefulness for cross- cultural and region-specific studies. Incorporating sar- casm detection models and stance detection frameworks would also improve analytical quality, especially for po- litical or controversial topics where user opinions are of- ten expressed indirectly. Additionally, implementing more advanced preprocessing techniques such as spell correction, slang normalization, and tweet embedding representations can enhance overall model performance.

From a system design perspective, deploying the ar- chitecture on cloud platforms with GPU acceleration would enable faster transformer-based computations and large-scale batch analysis. Adding time-series sentiment tracking would allow researchers to monitor how public sentiment evolves over hours, days, or months, making the tool more useful for event analysis and trend forecast- ing. Expanding the visualization layer with heatmaps, network graphs, and topic clusters would offer deeper in- terpretability of user discussions. Finally, integrating topic modeling approaches such as BERT or LDA could help identify hidden themes within conversations, providing an even richer understanding of social media narratives. These enhancements would transform the sys- tem into a more powerful, intelligent, and fully scalable sentiment analysis platform.

Another potential extension of this system lies in inte- grating real-time streaming APIs to continuously monitor sentiment shifts as they occur, rather than processing only static batches of tweets. Incorporating geolocation- based analysis could help correlate sentiment with re- gional trends and cultural factors. Future versions may also include user-level behavioral profiling to observe how individual sentiment evolves over time. Enhancing the interface with comparative dashboards will allow us- ers to analyze multiple topics simultaneously. These im- provements would significantly expand the analytical depth and practical applicability of the system.

REFERENCES

1. A. Albladi, M. Islam, and C. Seals, "Sentiment Analysis of Twitter Data Using NLP Models: A Comprehensive Review," IEEE Access, vol. 13, pp. 30444–30460, 2025.
2. A. Kumar and A. Jaiswal, "Twitter Sentiment Analysis Using VADER and TextBlob," International Journal of Computer Applica- tions, vol. 182, no. 45, pp. 1–6, 2021.
3. Y. Liu, M. Ott et al., "RoBERTa: A Robustly Optimized BERT Pretraining Approach," arXiv preprint arXiv:1907.11692, 2019.
4. J. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text," Proceedings of the AAAI Conference on Web and Social Media (ICWSM), pp. 216– 225, 2014.
5. L. Mozafari, R. Farahbakhsh, and N. Crespi, "A BERT-Based Transfer Learning Approach for Hate Speech Detection in Social Me- dia," Complexity, vol. 2020, Article ID 8886747, 2020.