International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 11 | Nov - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Real-Time Sentiment Analysis

Shikhar Gupta
Department of Computer Science, Chandigarh University
guptashikhar7305@gmail.com

Sidhant Singh
Department of Computer Science, Chandigarh University
ankitsingh88590@gmail.com

Er. Neetu Bala
Department of Computer Science, Chandigarh University
neetu.e15382@cumail.in

Divyanshu Gupta
Department of Computer Science, Chandigarh University
divyanshugupta3044@gmail.com

Vaideeka Sharma
Department of Computer Science, Chandigarh University
sharmavaideeka59@gmail.com

Devyansh Sharma
Department of Computer Science, Chandigarh University
sharmadevyansh77@gmail.com

Abstract **- Real-time sentiment analysis is an emerging field that focuses on evaluating and interpreting the emotions, opinions, and sentiments expressed in textual data as it is being generated. With the growing volume of user-generated content on social media, news platforms, and online forums, the need for real-time insights has become crucial for businesses, governments, and organizations. This paper presents a comprehensive framework for real-time sentiment analysis using a combination of natural language processing (NLP) techniques, machine learning models, and streaming data technologies. The system is designed to process continuous data streams, classify sentiments into categories such as positive, negative, and neutral, and provide insights instantly. Key challenges, such as handling noisy data, ensuring low-latency processing, and achieving high accuracy across different domains, are addressed. The framework demonstrates the potential to support real-time decision-making processes in fields such as market analysis, customer feedback, and crisis management. Experimental results show the effectiveness of the proposed system in providing timely and sentiment insights.**

Keywords – Natural Language Processing, Logistic Regression, Support Vector Machines (SVM), Random Forest, Recurrent Neural Networks(RNN), Convolutional Neural Networks(CNN).

I. INTRODUCTION

Real-time sentiment analysis is an advanced technology that allows businesses, organizations, and individuals to monitor and assess public sentiment as it is expressed across various digital platforms, such as social media, blogs, forums, and online reviews. By leveraging natural language processing (NLP) and machine learning (ML) algorithms, this tool analyzes large volumes of text data, automatically identifying and categorizing sentiments—positive, negative, or neutral—almost instantaneously. This enables decision-makers to gain immediate insights into how customers, stakeholders, or the general public perceive a brand, product, event, or topic. The ability to assess sentiment in real time is particularly valuable in today's fast-moving digital landscape, where public opinion can shift quickly andunpredictably.

Real-time sentiment analysis is widely used in a range of industries and applications. For example, in social media monitoring, companies track conversations about their brand, products, or competitors to understand public perception and identify emerging trends. In customer service, sentiment analysis helps prioritize responses by flagging negative sentiments, allowing businesses to quickly address complaints and improve customer satisfaction. Market researchers use it to gauge public opinion on new products, marketing campaigns, or emerging trends, making it easier to adjust strategies in response to real-time feedback. Additionally, sentiment analysis plays a critical role in crisis management, enabling organizations to detect spikes in negative sentiment and take immediate action to mitigate potential PR issues.

Real-time sentiment analysis empowers businesses with the ability to make proactive, data-driven decisions. It transforms raw, unstructured data into actionable insights, helping organizations stay agile and responsive to public opinion in a world where information flows rapidly and customer expectations are constantly evolving.

II. LITERATURE REVIEW

Some notable researches in the field of Real-Time SentimentAnalysis are mentioned in this section.

In 2002, Pang and Lee proposed a technique called "subjectivity summarization" to identify the most important subjective phrases in movie reviews and classify them as positive or negative using a minimum cut algorithm. The approach uses syntactic and semantic information to identify the most salient subjective phrases in a review and assigns a score to each phrase based on its polarity. The scores of the phrases are then used to compute the overall sentiment ofthe review.

Pointwise Mutual Information (PMI): Turney proposed the use of PMI to extract sentiment from movie reviews by identifying words that were strongly associated with positive or negative sentiment. The approach calculates the PMI between each word in a review and a set of positive and negative sentiment words to determine its sentiment polarity. The approach is computationally efficient and can be easily applied to large datasets of movie reviews.

Hybrid approach: Akhtar, Zhang, and Nichele proposed a hybrid approach to sentiment analysis that combined a Naive Bayes classifier with a lexicon-based approach to identify the sentiment of movie reviews. The Naive Bayes classifier uses word frequency and conditional probabilities to classify reviews as positive or negative. The lexicon-based approach uses a sentiment lexicon to assign a polarity

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 11 | Nov - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

score to each word in a review and aggregates the scores to determine the overall sentiment of the review.

Recurrent Neural Networks (RNNs): Ruder, Ghaffari, and Breslin proposed the use of RNNs, specifically LSTM networks, for sentiment analysis. RNNs are well-suited for sequential data such as text, and LSTMs are particularly effective at capturing long-term dependencies. The approach uses word embeddings to represent each word in a review and trains an LSTM network to predict the sentiment of the review.

Word embeddings: Maas et al. introduced a new method for training word embeddings specifically for sentiment analysis using a technique called "paragraph vector." The approach uses a neural network to learn a distributed representation of words that captures the semantic meaning of words in the context of the movie review. The word embeddings are then used to train a classifier to predict the sentiment of thereview.

Hierarchical Attention Networks (HAN): Zhang, Zhang, and Wang proposed a new architecture for sentiment analysis called HAN, which uses a hierarchical structure to model the relationships between words, sentences, and documents. HAN can identify the most important words and sentences in the reviews for predicting their sentiment. The approach uses word embeddings to represent each word in a review and trains an HAN network to predict the sentiment of the review. HAN has shown to outperform other state-of-the-art models on several benchmark datasets.

III. OBJECTIVE

The main goals of Real-Time Sentiment Analysis are:

1. Monitor Brand Reputation: Continuously track public opinion across social media and online platforms to understand how a brand, product, or service is perceived by consumers in real time.

2. Enhance Customer Experience: Identify and respond to customer feedback—especially negative comments—promptly, allowing businesses to resolve issues quickly and improve overall customer satisfaction.

3. Detect Emerging Trends: Recognize patterns, trends, and shifts in consumer sentiment as they happen, enabling companies to stay ahead of market dynamics and make timely adjustments to their strategies.

4. Support Crisis Management: Identify potential public relations crises early by monitoring spikes in negative sentiment, allowing organizations to take swift corrective actions before problems escalate.

5. Optimize Marketing Campaigns: Assess the real-time impact of marketing efforts, advertisements, or product launches by analyzing consumer sentiment, leading to data-driven decisions that enhance campaign effectiveness.

6. Improve Decision-Making: Enable businesses to make informed, data-driven decisions quickly by providing actionable insights into public perception, helping to align business strategies with customer expectations.

IV. PROBLEM STATEMENT

In today's digital era, vast amounts of textual data are generated every second across various online platforms such as social media, news outlets, forums, and customer feedback channels. This surge of unstructured data contains valuable insights into public opinions, emotions, and trends that are crucial for businesses, organizations, and policymakers. However, extracting and interpreting these insights in a timely manner presents significant challenges. Traditional sentiment analysis methods often operate in batch processing modes, leading to delays that render the insights less effective for immediate decision-making needs. The problem is to develop an efficient real-time sentiment analysis system capable of processing high-velocity, large-volume textual data streams. This system must accurately detect and classify sentiments—positive, negative, or neutral—as they are expressed, and present the findings in an actionable format. Addressing this problem involves overcoming obstacles related to natural language processing complexities, computational resource limitations, and the need for scalable algorithms that maintain high accuracy and speed. Solving this issue is essential for enabling organizations to respond promptly to public sentiment, manage brand reputation, engage with customers effectively, and make informed decisions in a rapidly changing information landscape.

V. METHODOLOGY

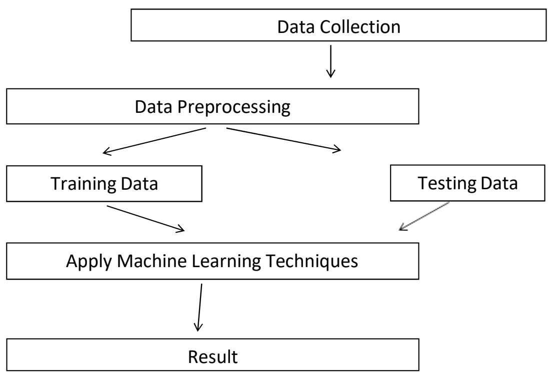

Fig 5.1. Design Flow

1. Data Collection: The first step involves gathering real-time data from various sources such as social media platforms (Twitter, Facebook, Instagram), online reviews, news articles, blogs, and customer support interactions. This data is often collected via APIs or web scraping tools to ensure continuous and real-time input.

2. Data Pre-processing: Raw text data is often unstructured and noisy, containing irrelevant information, such as hashtags, special characters, or URLs. This step is used to clean and pre-process the data. Common pre-processing tasks include:

i. Tokenization: Splitting text into individual words or tokens.
ii. Lowercasing: Converting all text to lowercase for uniformity.

iii. Stop-word removal: Removing commonly used words (e.g., "the", "is") that don't contribute to sentiment.
iv. Stemming/Lemmatization: Reducing words to their base form (e.g., "running" to "run").
v. Handling misspellings and slang: Correcting common misspellings and interpreting slang terms (important for social media data).

3. Feature Extraction: After preprocessing, the relevant features are extracted from the text data to prepare it for sentiment analysis. Popular techniques include:

i. Bag of Words (BoW): Converts text into a set of words and their frequencies.
ii. TF-IDF (Term Frequency-Inverse Document Frequency): Highlights important words in the text based on their frequency and importance relative toother documents.
iii. Word Embeddings (Word2Vec, GloVe): Represent words in a continuous vector space where words with similar meanings are closer together.

4. Sentiment Classification: In this step, machine learning or natural language processing (NLP) models are applied to classify the sentiment expressed in the text as positive, negative, or neutral. This can be done using:

i. Rule-based methods: Using predefined lexicons of words associated with positive or negative sentiment (e.g., VADER for social media).
ii. Machine Learning models: Training classifiers like Logistic Regression, Support Vector Machines (SVM), or Random Forest on labeled sentimentdatasets.
iii. Deep Learning models: Leveraging advanced models like Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM), or Transformers (BERT, GPT) for more nuanced sentiment analysis.

5. Real-Time Processing: The pre-trained models are deployed in a real-time environment where incoming data is analyzed as soon as it is available. Data pipelines are set up to process the data continuously, ensuring that sentiment analysis is performed on each new batch of incoming text with minimal delay. Stream processing platforms like Apache Kafka or Apache Flink may be used to handle the real-time data flow.

6. Sentiment Scoring and Aggregation: Each piece of analyzed text is assigned a sentiment score (positive, negative, or neutral) or a probability value. The individual sentiment scores can then be aggregated over a certain time period or by topic, product, or user group to understand broader sentiment trends.

7. Visualization and Reporting: The results of the real-time sentiment analysis are visualized using dashboards or reports, often in tools like Tableau, Power BI, or custom dashboards. Key performance indicators (KPIs) such as overall sentiment, changes in sentiment over time, and sentiment distribution across platforms or topics are displayed, allowing decision-makers to track public perception and respond accordingly.

8. Actionable Insights and Response: Based on the real-time sentiment analysis, organizations can take immediate action, such as:

i. Addressing negative feedback promptly in customer service.
ii. Amplifying positive sentiment through marketing or engagement efforts.
iii. Adjusting strategies or campaigns to respond to emerging trends.
iv. Notifying crisis management teams of any spikes in negative sentiment for immediate intervention.

9. Continuous Model Improvement: To maintain accuracy and effectiveness, the sentiment analysis models should be regularly updated and retrained on new data. This ensures that they remain relevant in handling emerging language patterns, new slang, or shifts in public sentiment. Continuous evaluation and fine-tuning are necessary to keep the system efficient over time.

These steps can be followed to build a robust real-time sentiment analysis system that processes large amounts of data continuously, delivering actionable insights for decision-making.

Real-time sentiment analysis requires specialized algorithms and techniques that can handle large amounts of data at high speed, with the ability to classify sentiment accurately. Below are some key algorithms and techniques commonly used for real-time sentiment analysis, ranging from traditional machine learning approaches to more advanced deep learning and hybrid methods:

1. Lexicon-Based Methods: Lexicon-based methods use predefined dictionaries of words, where each word is associated with a sentiment score (positive, negative, or neutral). These methods are relatively simple and fast, making them suitable for real-time analysis. However, they might not capture the context or nuances of language effectively.

i. VADER (Valence Aware Dictionary and sEntiment Reasoner): Sentiment analysis for social media can be done by using VADER which is a rule-based sentiment analysis tool. It can handle a variety of language patterns including emojis, slangs, and emoticons. VADER assigns sentiment polarity scores to individual words based on a predefined lexicon. It then uses heuristics to account for intensifiers, negations, punctuation, and capitalization to refine the sentiment score. Ideal for social media sentiment analysis, especially for real-time applications due to its lightweight processing needs.
ii. SentiWordNet: SentiWordNet is an extension of WordNet that provides sentiment scores (positive, negative, and neutral) for each synset (a set of synonyms that represent a single concept). Sentiment is derived from the scores assigned to the words in a sentence, and an aggregate sentiment score is produced based on these word-level scores. Useful for applications that rely on a large and general-purpose lexical resource.

2. Machine Learning-Based Methods: Machine learning algorithms are trained on labelled sentiment data to classify new, unseen text. These models can be very effective but

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 11 | Nov - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

require significant training data to perform well. They are capable of learning contextual and linguistic nuances, improving their performance over time.

i. Naive Bayes Classifier: A simple probabilistic classifier based on Bayes' Theorem, often used for text classification tasks like sentiment analysis. The classifier assumes that the presence of a particular feature (word) in a class is independent of the presence of any other feature. Despite this "naive" assumption, Naive Bayes works surprisingly well for text classification. Suitable for applications where fast classification is needed, and it can be used in real-time systems when combined with online learning techniques.

ii. Support Vector Machines (SVM): SVM is a supervised learning model that classifies data by finding the hyperplane that best separates classes (positive, negative, neutral). Text is transformed into numerical feature vectors (e.g., using TF-IDF), and SVM finds the optimal boundary (hyperplane) that separates the different sentiment classes. Useful for high-dimensional text data, though real- time performance can be slower compared to simpler models like Naive Bayes.

iii. Logistic Regression: Logistic regression is a linear model that predicts the probability that a given input belongs to a particular class (sentiment). After text is vectorized (BoW or TF-IDF), logistic regression computes probabilities for each sentiment class and classifies the text based on the highest probability. Commonly used in sentiment analysis due to its simplicity and effectiveness for binary or multiclass classification problems.

3. Deep Learning-Based Methods: Deep learning models, particularly those using neural networks, have significantly improved sentiment analysis by capturing complex relationships in text data, including context, sarcasm, and multi-layered meanings. They generally require more computational resources but excel in accuracy for more nuanced tasks.

i. Recurrent Neural Networks (RNNs): RNNs are a class of neural networks designed to handle sequential data, making them well-suited for sentiment analysis where the order of words matters. RNNs have memory cells that allow them to "remember" previous words when processing a sentence, which helps in capturing context. Effective for longer texts where context and word order play a crucial role in determining sentiment. However, traditional RNNs may struggle with long-term dependencies.

ii. Long Short-Term Memory (LSTM) Networks: LSTMs are an improvement over RNNs, designed to better capture long-term dependencies by selectively remembering important information and forgetting irrelevant parts of the sequence. LSTMs use "gates" to decide what to remember and what to forget from the input sequence, making them ideal for capturing context and meaning in long sentences or paragraphs. Widely used for real-time sentiment analysis, especially when processing longer and more complex text data, such as reviews, articles, or social media threads.

iii. Convolutional Neural Networks (CNNs) for Text: Though CNNs are commonly used in image processing, they can also be applied to text data. CNNs excel at capturing local patterns in text, such as phrases or combinations of words that determine sentiment. CNNs use convolutional filters to extract features from word embeddings, focusing on n-grams or word combinations that provide cues to sentiment. Often used in combination with other models like LSTM for sentiment analysis in real- time systems.

iv. Transformers (BERT, GPT): Transformer models like BERT (Bidirectional Encoder Representations from Transformers) and GPT (Generative Pre- trained Transformer) have revolutionized NLP tasks by capturing complex dependencies and contextual information at scale. BERT uses bidirectional context to understand the relationship between words in a sentence, while GPT uses a unidirectional context but is powerful in text generation. These models can be fine-tuned for real-time sentiment analysis and are highly effective at understanding the nuances of sentiment in a text, including sarcasm, ambiguity, and context. However, they require significant computational resources.

4. Hybrid Approaches: In real-time sentiment analysis, a combination of different techniques may be used to balance speed, accuracy, and scalability. For example:

i. Rule-based + Machine Learning: Lexicon-based approaches can be combined with machine learning models to handle out-of-vocabulary words or capture domain-specific sentiments.

ii. Deep Learning + Traditional ML: Deep learning models like LSTM or CNN can be used for feature extraction, while traditional machine learning classifiers (like Logistic Regression or SVM) can handle the classification task.

5. Feature Extraction Techniques: To feed data into these algorithms, you need to transform raw text into numerical representations. Some commonly used techniques include:

i. Bag of Words (BoW): Transforms text into a set of words, counting the frequency of each word in the text. BoW ignores grammar and word order, focusing on individual word occurrences.

ii. TF-IDF (Term Frequency-Inverse Document Frequency): Assigns weights to words based on how frequently they appear in a document compared to how frequently they appear across all documents. This helps highlight words that are more important for distinguishing sentiment.

iii. Word Embeddings (Word2Vec, GloVe, FastText): Transforms words into continuous vector representations, capturing semantic relationships between words (e.g., "king" and "queen" being closely related in vector space).

6. Real-Time Data Processing Techniques:

i. Apache Kafka: A distributed event streaming platform used for real-time data collection and processing.

ii. Apache Flink: A stream processing framework that can handle real-time analytics and process data at scale.
iii. Spark Streaming: Part of the Apache Spark ecosystem, this tool allows for real-time data streaming and fast processing.

Implementation plan:

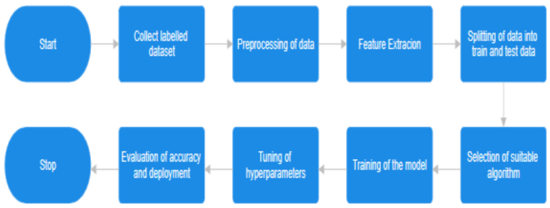

Fig 5.2. Algorithm

VII. CONCLUSION

Real-time sentiment analysis offers invaluable insights by instantly gauging public opinion, emotions, and reactions across various platforms such as social media, reviews, or customer feedback. Its applications span across industries, including marketing, finance, and customer service, where timely insights can drive strategic decisions and enhance customer engagement. Challenges like accuracy, handling sarcasm, and managing large datasets still remain. With advancements in natural language processing (NLP) and machine learning, future developments will likely overcome these hurdles, making sentiment analysis even more powerful and precise. In essence, real-time sentiment analysis is a key tool for businesses aiming to remain competitive in an increasingly data-driven world. As real-time sentiment analysis evolves, it will play an even more critical role in shaping the future of personalized services and automated decision-making. By understanding customer moods and reactions almost instantly, businesses can tailor their responses, enhance user experiences, and pre- emptively address concerns, leading to increased customer loyalty and satisfaction. In finance, real-time sentiment analysis can be used to predict market trends or shifts in public perception, providing investors and analysts with the ability to make more informed decisions. In the political and social arenas, it can serve as a tool to monitor public sentiment on key issues, offering policymakers a clearer understanding of voter concerns or reactions to policies. The growing reliance on this technology also raises important ethical considerations. The potential misuse of real-time sentiment data—such as for manipulating public opinion or invading individual privacy—must be addressed through clear regulations and ethical guidelines. Additionally, the challenge of processing and interpreting sentiment across different cultures and languages remains, requiring more sophisticated algorithms capable of understanding nuanced human emotions and context. In conclusion, while real-time sentiment analysis has already proven to be a game-changer in various fields, its future impact will depend on continued advancements in AI, data processing capabilities, and the responsible application of the technology. By addressing current limitations and challenges, sentiment analysis will continue to offer businesses and individuals alike a powerful tool for navigating the complexities of human emotions in the digital age.

VIII. FUTURE WORK

Future work in real-time sentiment analysis must address several important challenges to improve accuracy, scalability, and ethical considerations. One key area is the integration of multimodal data, incorporating not just text but also audio, video, and visual content from social media and other sources to provide a richer, more holistic understanding of sentiment. This would enable systems to analyze voice tone, facial expressions, and visual elements alongside written words, improving sentiment detection in complex and dynamic scenarios. Additionally, there is a pressing need to develop more advanced models that can handle multilingual and cross-cultural data, as existing sentiment analysis tools often fail to accurately capture the nuances of language, emotion, and expression across different regions and cultures. Future models should be able to better understand slang, idioms, and cultural context in real-time, possibly through improved translation algorithms and training datasets that represent a wider variety of languages. Another critical area for future work involves improving the detection of more complex emotions, such as sarcasm, irony, or mixed sentiments, which current models often struggle to interpret. Advanced natural language processing (NLP) techniques, including deep learning approaches like transformer-based models (e.g., BERT, GPT), offer promising paths for better handling these subtleties. However, this progress also needs to be balanced with concerns around transparency and model interpretability, as real-time sentiment analysis systems become increasingly complex. Ethical issues, such as the potential misuse of sentiment data for manipulation or invasion of privacy, will require on-going attention. Researchers must explore ways to ensure that sentiment analysis is conducted responsibly, with safeguards for user data and transparency in how insights are generated and used. Addressing these areas will be crucial for advancing the field of real-time sentiment analysis and ensuring its applications continue to serve users in an ethical, effective, and meaningful way.

REFERENCES

[1] Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013). "New avenues in opinion mining and sentiment analysis." IEEE Intelligent Systems, 28(2), 15-21.

[2] Pang, B., & Lee, L. (2008). "Opinion mining and sentiment analysis." Foundations and Trends in Information Retrieval, 2(1-2), 1-135.

[3] Liu, B. (2012). "Sentiment analysis and opinion mining." Synthesis Lectures on Human Language Technologies, 5(1), 1-167.

[4] Vinodhini, G., & Chandrasekaran, R. M. (2012). "Sentiment analysis and opinion mining: A survey." International Journal of Advanced Research in Computer Science and Software Engineering, 2(6), 282-292.

[5] Sharma, S., & Dey, S. (2012). "A comparative study of feature selection and machine learning techniques for sentiment analysis." International Journal of Computer Applications, 50(15), 7-11.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 11 | Nov - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

[6] Zimbra, D., Abbasi, A., Zeng, D., & Chen, H. (2018). "The state-of-the-art in Twitter sentiment analysis: A review and benchmark evaluation." ACM Transactions on Management Information Systems (TMIS), 9(2), 1-29.

[7] Medhat, W., Hassan, A., & Korashy, H. (2014). "Sentiment analysis algorithms and applications: A survey." Ain Shams Engineering Journal, 5(4), 1093-1113.

[8] Araque, O., Zhu, G., & Iglesias, C. A. (2019). "A deep learning-based approach for real-time sentiment analysis of tweets." Future Generation Computer Systems, 92, 651-662.

[9] Zhao, J., Dong, L., Wu, J., & Xu, K. (2012). "Moodlens: An emoticon-based sentiment analysis system for Chinese tweets." Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1528-1531.

[10] Mohammad, S. M., Kiritchenko, S., & Zhu, X. (2013). "NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets." Proceedings of the 7th International Workshop on Semantic Evaluation (SemEval 2013), 321-327.

[11] A., Bar, F., & Narayanan, S. (2012). "A system for real-time Twitter sentiment analysis of 2012 US presidential election cycle." Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: System Demonstrations, 115-120.

[12] Saif, H., He, Y., & Alani, H. (2012). "Semantic sentiment analysis of Twitter." International Semantic Web Conference, 508-524.

[13] Ren, F., & Wu, Y. (2013). "Predicting user-topic opinions in Twitter with social and topical context." IEEE Transactions on Affective Computing, 4(4), 412-424.

[14] Rosenthal, S., Farra, N., & Nakov, P. (2017). "SemEval-2017 task 4: Sentiment analysis in Twitter." Proceedings of the 11th InternationalWorkshop on Semantic Evaluation (SemEval-2017), 502-518.

[15] Smailović, J., Grčar, M., Lavrač, N., & Žnidaršič, M. (2014). "Stream-based active learning for sentiment analysis in the financial domain." Information Sciences, 285, 181-203.

[16] Dos Santos, C. N., & Gatti, M. (2014). "Deep convolutional neural networks for sentiment analysis of short texts." Proceedings of the 25th International Conference on Computational Linguistics: Technical Papers, 69-78.

[17] Zhang, L., Wang, S., & Liu, B. (2018). "Deep learning for sentiment analysis: A survey." Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4), e1253.

[18] Bollen, J., Mao, H., & Zeng, X. (2011). "Twitter mood predicts the stock market." Journal of Computational Science, 2(1), 1-8.

[19] Kumar, A., & Sebastian, T. M. (2012). "Sentiment analysis: A perspective on its past, present, and future." International Journal of Intelligent Systems and Applications, 4(10), 1-14.

[20] Jain, A., & Dandannavar, M. (2016). "Application of machine learning techniques to sentiment analysis." 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), 628-632.