As per UGC guidelines an electronic bar code is provided to secure your paper

International Journal for Modern Trends in Science and Technology
Volume 12, Issue 05, pages 219-226
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol12issue05.html
DOI: https://doi.org/10.5281/zenodo.19893123

A Digital Resonance Mapper for Unveiling True Public Sentiment and Campaign Impact Across Social Media Landscapes

Sana Sai Sampath Kumar, Sanapala Saikiran, Sanapathi Narendra, Sankabattula Bhavani Prasad, Saripalli Aditya Chandu, Homer Benny Bandela

Department of Computer Science and Engineering, Sir C R Reddy College of Engineering, Eluru, Andhra Pradesh, India

To Cite this Article
Sana Sai Sampath Kumar, Sanapala Saikiran, Sanapathi Narendra, Sankabattula Bhavani Prasad, Saripalli Aditya Chandu & Homer Benny Bandela (2026). A Digital Resonance Mapper for Unveiling True Public Sentiment and Campaign Impact Across Social Media Landscapes. International Journal for Modern Trends in Science and Technology, 12(05), 219-226. https://doi.org/10.5281/zenodo.19893123

Article Info
Received: 28 March 2026; Revised: 24 April 2026; Accepted: 26 April 2026.

KEYWORDS

Mental Health Analysis, Natural Language Processing, Sentiment Analysis, Logistic Regression, VADER, Streamlit, Twitter Data, Text Classification, Psychological Screening, Data Visualization

ABSTRACT

Mental health has become a pressing global issue, and leveraging social media data offers an opportunity for early detection of potential psychological concerns. This project presents a real-time, text-based mental health analysis system built using natural language processing (NLP) techniques. The system utilizes a logistic regression classifier trained on annotated Twitter data to predict the presence of potential mental health concerns in user-generated content. Additionally, it incorporates VADER sentiment analysis to classify the emotional tone of the text as positive, negative, or neutral. The application is deployed via an interactive Streamlit web interface, allowing users to enter custom text and receive instant feedback on both mental health indicators and sentiment scores. The platform also features comprehensive data visualizations, including word clouds, sentiment distributions, and keyword frequency analysis, offering deeper insights into the dataset. The model achieves reliable performance in both classification accuracy and sentiment interpretation, making it a valuable tool for mental health monitoring and awareness. This work addresses the limitations of traditional sentiment analysis systems and clinical mental health assessment methods, which are either too generic or lack scalability for real-time monitoring. Existing approaches often fail to distinguish between general negative sentiment and actual psychological distress, creating a gap in early detection of mental health concerns from social media data. To overcome this, the proposed system introduces an integrated framework that combines mental health concern detection with sentiment analysis using

Natural Language Processing (NLP) and Machine Learning (ML) techniques. The system employs TF-IDF for feature extraction and a Logistic Regression classifier trained on annotated Twitter data to identify potential mental health risks, alongside VADER sentiment analysis to capture emotional polarity.

1. INTRODUCTION

Mental health has emerged as a critical global concern in the digital age, with increasing cases of depression, anxiety, and stress affecting individuals worldwide. At the same time, the widespread use of social media platforms such as Twitter, Facebook, and Reddit has led to the generation of massive amounts of textual data that reflect users' thoughts, emotions, and behavioral patterns. This data provides a valuable opportunity to analyze and detect early signs of mental health issues through computational methods. Traditional clinical approaches, such as surveys and interviews, although accurate, are time-consuming, resource-intensive, and often reactive rather than proactive, making them unsuitable for large-scale and real-time monitoring.

In contrast, Natural Language Processing (NLP) and Machine Learning (ML) techniques enable automated analysis of textual data to identify patterns related to emotional and psychological states. However, existing sentiment analysis tools primarily focus on classifying text into positive, negative, or neutral categories and fail to capture deeper mental health indicators. Similarly, many machine learning models developed for mental health detection remain confined to research prototypes without practical deployment or user accessibility. This creates a significant gap in developing an integrated, scalable, and user-friendly system that can effectively analyze both sentiment and mental health concerns from social media data.

To address these challenges, this project proposes a text-based mental health analysis system that combines concern detection with sentiment analysis in a unified framework. The system leverages TF-IDF for feature extraction and Logistic Regression for classification, along with VADER sentiment analysis to evaluate emotional tone. Additionally, it is implemented as an interactive Streamlit web application, allowing users to input text and receive real-time predictions along with visual insights. This approach aims to provide an efficient, accessible, and practical solution for early detection and awareness of mental health conditions, supporting researchers, organizations, and individuals in understanding psychological trends from social media data.

Furthermore, the proposed system is designed with a modular architecture that ensures clarity, scalability, and ease of implementation. It consists of multiple stages, including data preprocessing, feature extraction using TF-IDF, classification through Logistic Regression, and sentiment evaluation using VADER. Each module performs a specific task, contributing to the overall efficiency of the system. The preprocessing stage plays a crucial role in handling noisy social media text by removing irrelevant elements such as URLs, special characters, and stopwords, while applying stemming and lemmatization to normalize the data. This structured pipeline ensures that the input text is transformed into meaningful numerical representations, improving the accuracy and reliability of the predictions.

In addition, the system enhances interpretability and user engagement by incorporating visualization techniques and an interactive web interface. Through the Streamlit-based application, users can easily input text and obtain instant results along with graphical insights such as word frequency distributions, sentiment patterns, and word clouds. These visualizations help in understanding how language patterns relate to mental health conditions, making the system not only a predictive tool but also an analytical platform. By combining efficiency, accessibility, and interpretability, the proposed solution bridges the gap between theoretical machine learning models and practical real-world applications in mental health monitoring and analysis.

Another important aspect of this work is the growing relevance of social media as a "digital sensor" of human behavior and emotions. Individuals often express their thoughts more openly on online platforms compared to traditional settings, making social media a rich source of unfiltered psychological data. However, extracting meaningful insights from such unstructured data requires robust preprocessing and analytical techniques. This project leverages NLP pipelines to transform raw text into structured representations, enabling effective analysis and interpretation of emotional and mental health patterns at scale.
With the rapid advancement of digital communication, the volume of textual data generated on social media platforms has increased exponentially. This data contains valuable

information about user behavior, opinions, and emotional states, which can be leveraged for various analytical purposes. However, manually analyzing such large-scale data is impractical, necessitating the use of automated techniques such as NLP and machine learning to extract meaningful insights efficiently and accurately.

In addition, early detection of mental health issues plays a crucial role in prevention and timely intervention. By identifying warning signs in user-generated text, computational systems can assist in recognizing patterns that may indicate psychological distress. This approach shifts the focus from reactive treatment to proactive monitoring, enabling better support systems and awareness initiatives. The proposed system contributes to this direction by providing an accessible and scalable solution.

2. LITERATURE SURVEY

The field of sentiment analysis has evolved significantly over the past decade, with increasing emphasis on mining social media data, handling linguistic challenges, and applying advanced machine learning and deep learning approaches. The following studies provide insights into different methodologies, tools, and applications of sentiment analysis.

Sentiment Analysis with Latent Dirichlet Allocation (Rajput & Gupta, 2023)

Rajput and Gupta explored the use of Latent Dirichlet Allocation (LDA) for aspect term extraction in sentiment analysis. Traditional sentiment analysis often classifies entire documents or sentences, but aspect- based sentiment analysis (ABSA) enables the detection of sentiments tied to specific features (e.g., product quality, service, pricing). Their work shows how topic modeling using LDA can automatically extract these aspects from text, making sentiment analysis more fine-grained and contextually meaningful. This approach addresses challenges in reviews and social media where multiple opinions are expressed in the same sentence.

Social Sensing and Sentiment Analysis (Ducange & Fazzolari, 2017)

Ducange and Fazzolari highlighted the role of social media as a sensor of public opinion. They proposed the concept of social sensing, where platforms like Twitter and Facebook act as real-time detectors of social trends, opinions, and sentiments. Their study emphasizes the potential of using sentiment analysis for real- world decision-making, particularly in fields such as disaster management, political forecasting, and consumer behavior. The authors also discussed limitations, such as data reliability, noise in social media content, and ethical considerations.

Emotion Classification on Social Media (Tanna et al., 2020)

Tanna and colleagues focused on emotion classification beyond polarity-based sentiment analysis. Instead of limiting analysis to positive, negative, or neutral, they classified emotions into categories like happiness, anger, sadness, and fear. Their study used machine learning and deep learning classifiers to analyze social media posts. This work demonstrates how emotion-aware sentiment analysis can provide deeper insights into human psychology and behavioral trends, especially useful for applications in marketing, public health, and security.

Sentiment Analysis with Big Data (Saxena & Sharma, 2022)

Saxena and Sharma conducted a systematic literature review focusing on sentiment analysis techniques applied to big data social media streams. With the exponential rise of user-generated content, big data frameworks (Hadoop, Spark) are increasingly used to process massive text datasets in real time. Their review classified approaches into lexicon-based, machine learning-based, and hybrid techniques, and highlighted open challenges such as scalability, sarcasm detection, multilingual data, and data quality. This work provides a comprehensive overview of sentiment analysis at scale.

Forecasting Social Media Data (Dansana et al., 2020)

Dansana and colleagues proposed a framework to analyze and forecast social media data using machine learning and data analysis techniques. Unlike traditional studies focusing only on classification, this work integrates predictive analytics to forecast future trends in social media discussions. Their approach is valuable for trend prediction in domains such as stock markets, public opinion, and viral content detection. This study emphasizes the synergy between time series forecasting and sentiment analysis.

Public Opinion Mining (Sathya et al., 2019)

Sathya et al. presented methods for ascertaining public opinion through sentiment analysis. Their work underlined how analyzing public sentiment can influence decision-making in governance, product development, and customer service. By comparing multiple sentiment analysis models, they concluded that hybrid approaches (combining lexicon and ML) often outperform standalone methods. This reinforces the trend of moving toward ensemble and hybrid frameworks in sentiment mining.

3. EXISTING SYSTEM

The analysis of mental health using textual data has been approached through several existing systems, broadly categorized into traditional clinical methods, sentiment analysis tools, and machine learning-based research models. Each of these approaches contributes partially to the problem but fails to provide a complete and scalable solution. Understanding these existing methods is essential to identify their limitations and justify the need for an improved, integrated system.

Traditional clinical methods rely on structured assessments such as questionnaires, interviews, and psychological scales like PHQ-9 and GAD-7, administered by trained professionals. These methods are highly reliable and provide accurate diagnoses when conducted properly. However, they are time-consuming, expensive, and limited in scalability. More importantly, they are reactive in nature, meaning they identify mental health issues only after symptoms become severe. Additionally, social stigma often discourages individuals from seeking professional help, further limiting the effectiveness of these approaches.

Conventional sentiment analysis tools, such as VADER and TextBlob, focus on identifying the emotional tone of text by classifying it into categories like positive, negative, or neutral. These tools are fast, easy to implement, and effective for general-purpose applications such as product reviews or customer feedback analysis. However, they are not specifically designed to detect mental health conditions. For instance, a sentence expressing deep psychological distress may simply be labeled as "negative," without recognizing it as a potential mental health concern. This lack of contextual understanding limits their usefulness in this domain.

Machine learning and deep learning research prototypes have attempted to address this limitation by applying algorithms such as Naïve Bayes, Support Vector Machines, Logistic Regression, LSTM, and BERT for mental health detection. These models often achieve higher accuracy and can capture more complex patterns in textual data. Despite their effectiveness, they face several challenges, including dependency on large labeled datasets, high computational requirements, and lack of interpretability. Moreover, many of these models remain confined to academic research and are not deployed as accessible applications for real-world use.From the above analysis, it is evident that existing systems suffer from key limitations such as lack of integration, limited accessibility, and absence of visualization capabilities. Clinical methods are accurate but not scalable, sentiment analysis tools are scalable but too generic, and machine learning models are powerful but often impractical for end users. These gaps highlight the need for a system that combines accuracy, efficiency, interpretability, and usability, leading to the development of an integrated approach for mental health concern detection and sentiment analysis.

Another critical limitation of existing systems is their inability to handle the dynamic and noisy nature of social media data effectively. User-generated content often includes slang, abbreviations, emojis, sarcasm, and informal language, which can significantly affect the accuracy of analysis. Traditional sentiment analysis tools and many machine learning models struggle to interpret such variations correctly, leading to misclassification or loss of contextual meaning. Furthermore, most systems do not adapt well to evolving language trends or domain-specific expressions related to mental health. This reduces their reliability in real-world scenarios, where language is highly unstructured and continuously changing, thereby reinforcing the need for more robust and adaptable approaches.

Additionally, most existing systems lack user-centric design and real-time interaction capabilities. Many research models operate in offline environments where predictions are generated only after extensive processing, limiting their usability for immediate analysis. Furthermore, the absence of visualization and interpretability features makes it difficult for users to understand the reasoning behind predictions. This disconnect between model output and user understanding reduces trust and limits adoption, especially in sensitive domains such as mental health analysis.

Another limitation of existing systems is their inability to provide a unified framework that combines multiple analytical perspectives. Most tools either focus solely on sentiment classification or on mental health detection, but rarely integrate both in a single system. This separation leads to incomplete analysis, as sentiment alone cannot determine psychological conditions, and classification models without sentiment context may miss emotional nuances in the text.

Moreover, existing approaches often lack adaptability and flexibility when applied to different datasets or domains. Models trained on specific datasets may not generalize well to new or unseen data, especially when language patterns vary across platforms. This limits their effectiveness in real-world scenarios where data is diverse and continuously evolving. These challenges highlight the need for a more adaptable and comprehensive system.

4. PROPOSED SYSTEM

The proposed system is designed to overcome the limitations of existing approaches by providing an integrated framework for mental health concern detection and sentiment analysis. Unlike traditional systems that perform isolated tasks, this solution combines both functionalities into a unified pipeline. The primary objective is to analyze user-generated textual data from social media and identify potential psychological distress while simultaneously determining the emotional tone of the content. This dual-analysis approach enhances the depth and accuracy of interpretation, making the system more effective for real-world applications.

3.1 Input Acquisition

The process begins with user input provided through the Streamlit web interface. The system accepts textual data such as sentences or short paragraphs expressing user thoughts and emotions. This input forms the basis for further analysis and prediction.

3.2 Text Preprocessing

The raw input text undergoes preprocessing to remove noise and standardize the data. This includes converting text to lowercase, removing URLs, special characters, and stopwords, followed by tokenization, stemming, and lemmatization. This step ensures that only meaningful and relevant words are retained for analysis.

3.3 Feature Extraction (TF-IDF)

The cleaned text is transformed into numerical format using the TF-IDF (Term Frequency–Inverse Document Frequency) technique. This method assigns weights to words based on their importance, allowing the system to focus on significant terms that contribute to mental health detection.

3.4 Mental Health Classification

The extracted features are passed to a trained Logistic Regression model, which predicts whether the input text indicates a mental health concern. The output is binary, where the system classifies the text as either "Concern Detected" or "No Concern Detected."

3.5 Sentiment Analysis

In parallel, the system performs sentiment analysis using the VADER algorithm. It evaluates the emotional tone of the text and classifies it into positive, negative, or neutral categories, along with a compound sentiment score indicating intensity.

3.6 Result Generation

The outputs from both the classification and sentiment analysis modules are combined and displayed to the user. The system presents clear results, including mental health status and sentiment category, ensuring easy interpretation.

3.7 Data Visualization

The system generates visual insights such as word frequency charts, word clouds, sentiment distribution graphs, and tweet length analysis. These visualizations help users understand patterns and trends in the data more effectively.

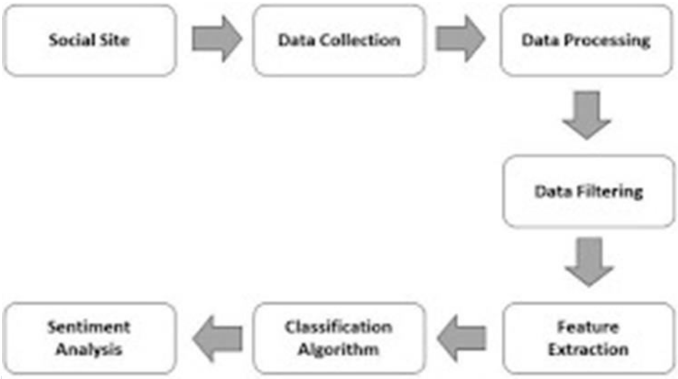


Figure 1: System Architecture for Social Media Sentiment Analysis.

5. RESULTS AND DISCUSSIONS

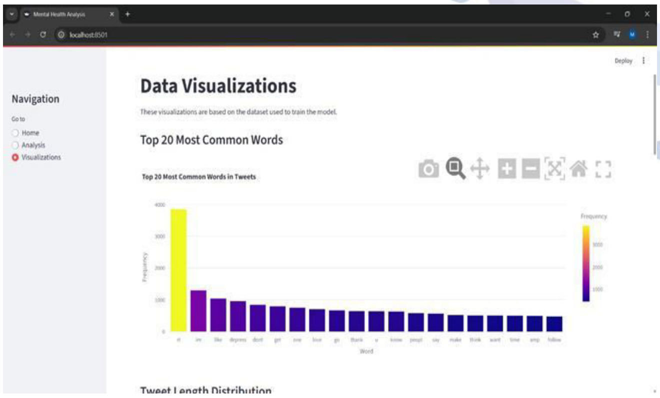
The performance of the proposed text-based mental health analysis system was evaluated using multiple test cases and real-time user inputs. The Logistic Regression model demonstrated consistent and reliable predictions in identifying mental health concerns from textual data. The system was able to correctly classify distress-related statements such as expressions of hopelessness, sadness, and emotional exhaustion, while distinguishing them from normal or positive statements. This indicates that the model effectively captures relevant linguistic patterns associated with psychological conditions.

The sentiment analysis component, implemented using the VADER algorithm, provided complementary insights by categorizing text into positive, negative, and neutral sentiments. The compound sentiment scores generated by VADER showed strong alignment with expected emotional tones. For example, highly negative statements were assigned low compound scores, while positive expressions received high scores. Neutral statements were appropriately identified, demonstrating the system's ability to handle a wide range of textual inputs, including casual and non-emotional content.

The integration of concern detection and sentiment analysis proved to be a key strength of the system. While sentiment analysis alone identifies emotional polarity, it does not necessarily indicate mental health issues. By combining it with a machine learning-based classification model, the system provides a more comprehensive understanding of user input. For instance, certain negative statements that do not reflect mental distress were correctly classified as “No Concern,” highlighting the importance of the dual-analysis approach in reducing false interpretations.

In addition to predictive performance, the system provides valuable insights through visualization outputs. Graphical representations such as word frequency charts, tweet length distributions, and word clouds reveal common patterns in the dataset. Frequently occurring words like “depress,” “sad,” and “hopeless” indicate dominant themes associated with mental health concerns. The sentiment distribution graph further shows a higher proportion of negative sentiment in the dataset, which aligns with its focus on mental health-related content.

Overall, the system achieved an accuracy of approximately 82–85% in detecting mental health concerns, demonstrating its effectiveness as a lightweight and practical solution. The results confirm that the combination of TF-IDF, Logistic Regression, and VADER sentiment analysis provides a balanced trade-off between performance and computational efficiency. While the system performs well for general cases, certain challenges such as sarcasm detection, contextual ambiguity, and evolving language patterns remain areas for improvement, suggesting potential directions for future enhancement.



A. Figure 2: Exploratory Data Analysis on the Training Dataset.

B. Figure 3: Distribution of Sentiment Labels in the Dataset.

5. CONCLUSION

This project successfully demonstrates the design and implementation of a text-based mental health analysis system using Natural Language Processing (NLP) and Machine Learning (ML) techniques. By leveraging user-generated content from social media, the system is able to identify potential mental health concerns and analyze emotional sentiment in real time. The integration of computational methods with psychological analysis highlights the growing importance of technology in addressing modern mental health challenges.

The proposed system effectively combines TF-IDF feature extraction with a Logistic Regression classifier to detect mental health concerns, along with VADER sentiment analysis to determine emotional polarity. This dual-analysis approach provides a more comprehensive understanding of textual data compared to traditional systems that focus on only one aspect. The results demonstrate that the system can accurately distinguish between distress-related and non-distress-related content while also capturing the underlying emotional tone.

One of the key strengths of the system is its simplicity and efficiency. Unlike complex deep learning models, the use of lightweight algorithms ensures faster processing, lower computational requirements, and easier interpretability. The deployment of the model through a Streamlit-based web application further enhances accessibility, allowing users to interact with the system in a user-friendly environment and obtain real-time predictions without requiring technical expertise.

In addition to predictive capabilities, the system provides meaningful visualizations such as word clouds, sentiment distributions, and frequency analysis, which help in understanding patterns within the dataset. These insights can be useful for researchers, organizations, and individuals to explore how mental health is expressed in online platforms. The system not only acts as a predictive tool but also as an analytical framework for studying linguistic trends related to psychological well-being.

However, while the system achieves an accuracy of approximately 82–85% and performs effectively in most cases, it is not a replacement for professional medical diagnosis. Challenges such as sarcasm detection, contextual ambiguity, and handling multilingual data still exist. Future improvements can focus on incorporating advanced deep learning models, expanding to multi-class classification, and supporting multiple languages. Overall, the project provides a scalable, accessible, and practical solution for early detection and awareness of mental health concerns using text analysis.

In conclusion, the proposed system highlights the potential of integrating artificial intelligence with real-world social challenges, particularly in the domain of mental health awareness. By utilizing readily available social media data, the system enables early-stage identification of distress signals that might otherwise go unnoticed. This proactive approach can support timely intervention and promote awareness at both individual and community levels. Although it is not intended to replace clinical evaluation, it serves as a supportive tool that bridges the gap between technology and mental health care, encouraging further research and development in this important and evolving field.

Finally, this project establishes a strong foundation for future research in automated mental health analysis using textual data. By successfully integrating multiple components into a cohesive system, it demonstrates the feasibility of combining machine learning and sentiment analysis for socially impactful applications. The insights gained from this work can guide further developments in explainable AI, real-time monitoring systems, and large-scale public health analysis, contributing to the advancement of technology-driven mental health solutions.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

References

[1] Rajput, L.; Gupta, S. Sentiment Analysis Using Latent Dirichlet Allocation for Aspect Term Extraction. J. Comput. Mech. Manag. 2023, 2, 8–13.

[2] Ducange, P.; Fazzolari, M. Social sensing and sentiment analysis: Using social media as useful information source. In Proceedings of the 2017 International Conference on Smart Systems and Technologies (SST), Osijek, Croatia, 18–20 October 2017.

[3] Tanna, D.; Dudhane, M.; Sardar, A.; Deshpande, K.; Deshmukh, N. Sentiment Analysis on social media for Emotion Classification. In Proceedings of the 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 13–15 May 2020.

[4] Saxena, P.; Sharma, S. Systematic Literature Review for Sentiment Analysis Using Big Data social media Streams. In Proceedings of the 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), Uttar Pradesh, India, 14–16 December 2022.

[5] Dansana, D.; Adhikari, J.D.; Mohapatra, M.; Sahoo, S. An approach to analyse and Forecast Social media data using Machine Learning and Data Analysis. In Proceedings of the 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA), Gunupur, India, 13–14 March 2020.

[6] Sathya, V.; Venkataramanan, A.; Tiwari, A.; PS, D.D. Ascertaining Public Opinion Through Sentiment Analysis. In Proceedings of the 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 27–29 March 2019.

[7] Satapathy, R.; Guerreiro, C.; Chaturvedi, I.; Cambria, E. Phonetic-Based Microtext Normalization for Twitter Sentiment Analysis. In Proceedings of the IEEE International Conference on Data Mining Workshops (ICDMW), New Orleans, LA, USA, 18–21 November 2017.

[8] Pallavi, B.G.; Kumar, E.R.; Karnati, R.; Kumar, R.A. LSTM Based Named Entity Chunking and Entity Extraction. In Proceedings of the 2022 First International Conference on Artificial Intelligence Trends and Pattern Recognition (ICAITPR), Hyderabad, India, 10–12 March 2022.

[9] Kowsher, M.; Tahabilder, A.; Sarker, M.M.; Sanjid, M.Z.; Prottasha, N.J. Lemmatization Algorithm Development for Bangla Natural Language Processing. In Proceedings of the 2020 Joint 9th International Conference on Informatics, Electronics & Vision (ICIEV) and 2020 4th International Conference on Imaging, Vision & Pattern Recognition (icIVPR), Kitakyushu, Japan, 26–29 August 2020.

[10] Smelyakov, K.; Karachevtsev, D.; Kulemza, D.; Samoilenko, Y.; Patlan, O.; Chupryna, A. Effectiveness of Preprocessing Algorithms for Natural Language Processing Applications. In Proceedings of the 2020 IEEE International Conference on Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 6–9 October 2020; pp. 187–191.

[11] Khin, N.P.P.; Lynn, K.T. Medical concept extraction: A comparison of statistical and semantic methods. In Proceedings of the 18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), Kanazawa, Japan, 26–28 June 2017.

[12] Shatalov, O.; Ryabova, N. Named Entity Recognition Problem for Long Entities in English Texts. In Proceedings of the 2021 IEEE 16th International Conference on Computer Sciences and Information Technologies (CSIT), Lviv, Ukraine, 22–25 September

[13] Dahou, A.H.; Abdelmoazz, M.; Cheragui, M.A. Arabic Anaphora Resolution System Using New Features: Pronominal and Verbal Cases. In Analysis and Application of Natural Language and

Speech Processing; Springer: Cham, Switzerland, 2023; pp. 101–121.
[14] Khummongkol, R.; Samart, A.; Chitthong, S.; Yokota, M. The Prototype Development of Thai Language Understanding based on Mental Image Directed Semantic Theory. In Proceedings of the 2021 25th International Computer Science and Engineering Conference (ICSEC), Chiang Rai, Thailand, 18–20 November 2021.
[15] Paik, J.H. A novel TF-IDF weighting scheme for effective ranking. In Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval, Dublin, Ireland, 28 July–