International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 08 | Aug - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Stress Detection using NLP

Supriya Mallad [2],Deepthi K [1]

[1]*Student, 4th Semester MCA, Department of MCA, BIET, Davangere*

[2] *Assistant Professor, Department of MCA, BIET, Davangere*

Abstract

This paper presents a novel web application that leverages advanced machine learning and natural language processing techniques to detect stress from textual data, such as social media posts or survey responses. The system employs robust text preprocessing methods—including stemming, stopword removal, and cleaning—and utilizes a Naive Bayes classifier to accurately categorize input text as "Stress" or "No Stress." Developed using the Flask framework, the application offers an interactive platform where users can assess their stress levels, access curated stress-reduction tips, and communicate with administrators who manage user queries and resources. This integrated solution not only enables early detection and proactive management of stress but also increases awareness of its impact on mental health, providing practical support for individuals, mental health professionals, and organizations to enhance well-being and productivity.

Keywords: Stress Detection, Machine Learning, Natural Language Processing, Text Classification, Naive Bayes,

I. INTRODUCTION

Stress is an inevitable aspect of modern life, affecting individuals across all age groups and professions. In today's fast-paced world, people are exposed to various stressors, including academic pressures, workplace demands, social expectations, and personal challenges. While a certain amount of stress can be motivating and enhance performance, chronic or excessive stress is widely recognized as a major contributor to a range of physical and mental health problems. Prolonged exposure to stress has been linked to conditions such as anxiety, depression, cardiovascular diseases, weakened immune function, and a general decline in quality of life. The World Health Organization (WHO) and numerous health agencies have highlighted stress as a growing public health concern, emphasizing the urgent need for effective strategies to detect, manage, and mitigate its impact.

The ability to recognize stress early and take preventive action is crucial for maintaining both mental and physical well-being. Traditional approaches to stress detection often rely on self-reporting, clinical interviews, or physiological measurements such as heart rate variability and cortisol levels. While these methods can be effective, they are often invasive, time-consuming, and not scalable for large populations. In recent years, the proliferation of digital communication—especially through social media, forums, and messaging platforms—has provided a unique opportunity to analyze textual data for signs of psychological distress. People frequently express their emotions, concerns, and experiences in written

form, making text analysis a promising avenue for unobtrusive and scalable stress detection.

Advancements in artificial intelligence (AI), particularly in the fields of machine learning (ML) and natural language processing (NLP), have revolutionized the way we interpret and classify textual information. Machine learning algorithms can be trained to recognize complex patterns in large datasets, while NLP techniques enable the extraction of meaningful features from raw text. Text classification, a core task in NLP, has been successfully applied to sentiment analysis, spam detection, and topic categorization. However, the application of these techniques to stress detection remains a relatively new and evolving area of research. Existing studies have demonstrated the feasibility of using machine learning models to identify stress-related language patterns, but challenges remain in achieving high accuracy, interpretability, and real-world applicability.

One of the key challenges in textual stress detection is the variability and subjectivity of language. People express stress in diverse ways, using different vocabularies, tones, and linguistic structures. Cultural, social, and personal factors further influence how stress is communicated in text. Moreover, the subtlety of stress indicators—such as sarcasm, metaphor, or indirect references—can complicate the classification process. Therefore, robust preprocessing steps, including text cleaning, stemming, and stopword removal, are essential to ensure that machine learning models focus on the most relevant features. Among various classification algorithms, the Naive Bayes classifier has shown promise due to its simplicity, efficiency, and effectiveness in handling text data, especially in scenarios with limited computational resources.

In addition to the technical challenges, there is a pressing need to translate research advancements into practical tools that can benefit individuals, mental health professionals, and organizations. Web-based applications offer an accessible and user-friendly platform for delivering such solutions. By integrating machine learning-powered stress detection with interactive features—such as personalized feedback, stress-reduction tips, and direct communication with administrators—web applications can empower users to monitor their mental health proactively and seek timely support. For administrators and mental health practitioners, these platforms provide valuable insights into user concerns and facilitate the dissemination of evidence-based resources.

This paper introduces a state-of-the-art web application designed to detect stress from textual input using advanced machine learning and NLP techniques. The system accepts user-generated text, applies comprehensive preprocessing, and utilizes a Naive Bayes classifier to categorize the input as “Stress” or “No Stress.” Built on the Flask framework, the application features a dual-interface: administrators can manage users, curate stress-reduction tips, and respond to queries, while users can assess their stress levels, access resources, and communicate with administrators. The goal of this project is not only to provide accurate and timely stress detection but also to raise awareness about the importance of mental health and offer practical support for stress management.

By bridging the gap between cutting-edge AI research and real-world application, this project aims to contribute to the growing field of digital mental health. The proposed system demonstrates how machine learning and NLP can be harnessed to address one of the most pressing health challenges of our time. Through continuous refinement and user engagement, such tools have the potential to transform the way we understand, detect, and manage stress, ultimately enhancing well-being and productivity in society.

II. RELATED WORK

1. Text-Based Stress Detection and Classification using Machine Learning, Authors: Harika Juluri,

This study proposes a system that analyzes text messages to detect and classify stress, leveraging natural language processing techniques. The methodology includes cleaning and preprocessing a text dataset, extracting relevant linguistic features, and training classification models—such as Random Forest and XGBoost. Evaluated using accuracy, precision, recall, and F1 score metrics, the model aims to effectively identify stress levels and provide actionable recommendations to users. The goal is to empower individuals with timely insights into their mental state, thereby supporting better stress management strategies.

2. The Stress Detection using Naive Bayes and Decision Tree Classifiers, Authors: N Krishna Chowdary, Shaik Rasool Basha

This research presents an automated approach for stress detection by applying Naive Bayes and Decision Tree classifiers to multiline text data, enhancing earlier methods limited to single-line inputs. The study addresses the growing challenge of identifying stress, especially in individuals unaware of their condition. It emphasizes the importance of early detection in preventing severe health outcomes. The proposed model processes user-generated content, categorizes it as stressed or unstressed, and demonstrates high accuracy in classification, showing promise for real-world applications in mental health monitoring.

3. Stress Detection Using Natural Language Processing and Machine Learning, Authors: Jerripotula Priyanka.

This paper presents a system for detecting psychological stress from social media text, primarily Twitter posts, using a suite of NLP and machine learning algorithms—including k-Nearest Neighbors, Bernoulli Naïve Bayes, Random Forest, Decision Trees, and Support Vector Machines. Implemented on a large real-world dataset, the approach involves data collection, preprocessing, model training, and classification to distinguish between stressed and non-stressed users. The authors also explore how an individual's social connections can provide valuable context, establishing that stress signals among friends often correlate. The proposed model aims to enable timely stress recognition and support mental well-being through proactive care.

4. Stress Detection Using Classification Algorithm, Authors: J. S. Kanchana.

This paper explores an automated stress detection model that leverages linguistic features from social media text—specifically tweets—to classify users

into stressed or non-stressed categories. Using Naive Bayes as the classification algorithm, the study underscores the advantages of web-mined textual content over traditional detection methods, which are often time-consuming and expensive. The proposed method demonstrates that mining informal textual data can provide efficient, scalable insights into an individual's psychological state, supporting proactive stress management initiatives.

5. Stress Detection Using Machine Learning Techniques, Authors: R. Swarna Malika.

This paper presents a machine learning approach to detect stress using a labeled dataset from Kaggle. It evaluates the performance of three classification algorithms—Logistic Regression, Decision Tree, and Random Forest—for identifying stress levels among individuals. Among them, the Random Forest classifier achieved the highest accuracy, making it a promising tool for early stress detection. By leveraging such predictive models, the study underscores the potential to intervene before stress leads to serious mental or physical health issues.

6. Stress Detection Using Natural Language Processing and Machine Learning Over Social Interactions, Authors: Tanya Nijhawan.

This paper explores the use of social media analytics and machine learning for stress detection based on users' social interactions and posts. By applying sentiment analysis, emotion classification using BERT, and topic modeling with Latent Dirichlet Allocation (LDA), the authors aim to identify psychological stress in online content. The study shows that emotional states like sadness, fear, and anger can be effectively classified using advanced NLP models. The proposed system combines BERT-based deep learning with traditional ML techniques to deliver high-accuracy predictions of emotional and mental health indicators from social media behavior.

7. Stress detection using natural language processing and machine learning over social interactions, Authors-Tanya Nijhawan, Girija Attigeri & T. Ananthakrishna,

Cyberspace, especially social media, serves as a platform where individuals share daily experiences, making it a rich source for sentiment and emotion analysis. This research extends traditional sentiment analysis to detect stress in individuals based on their social media posts and comments. The approach uses large-scale tweet datasets, applying machine learning algorithms and the deep learning model BERT for sentiment classification. Additionally, Latent Dirichlet Allocation (LDA), an unsupervised machine learning method, is used to identify topics within the textual data, helping to link specific topics to user emotions. By combining these models, the system can detect online emotions, which can then be analyzed for signs of stress or depression. The models, evaluated with various metrics, show high detection rates, making this research valuable for monitoring and supporting mental health through social media analysis.

8. Machine Learning Driven Mental Stress Detection on Reddit Posts Using Natural Language Processing, Authors:Shaunak Inamdar, Rishikesh Chapekar, Shilpa Gite, Biswajeet Pradhan,

This paper explores how social media activity, particularly on Reddit, can reveal users' mental health conditions. By applying machine learning models and various NLP techniques—such as ELMo, BERT, and Bag of Words—the authors classify posts as stressful or non-stressful. Their approach, using only preprocessed text, achieved strong results with an F1 score of 0.76, Precision of 0.71, and Recall of 0.74. These findings highlight the potential of such methods for early detection of mental stress on social platforms, offering valuable tools for addressing mental health issues in real-world applications.

9. Stress detection using deep neural networks, Authors-Russell Li & Zhandong Liu,

Previous studies have shown that physiological signals from body-worn sensors can predict stress, but traditional machine learning methods relying on hand-crafted features have produced mixed results. To overcome this, the authors developed two deep neural networks—a 1D convolutional neural network for chest-worn sensors and a multilayer perceptron for wrist-worn sensors—that automatically extract features from raw data. These networks were used for both binary stress detection and three-class emotion classification, achieving improved performance on publicly available datasets.

10. Deep Learning Approach for Detecting Work-Related Stress Using Multimodal Signals, Authors: Wonju Seo, Namho Kim, Cheolsoo Park, Sung-Min Park, Stress is a natural response to various life challenges, and recognizing its symptoms early is crucial for maintaining health. This research explores deep learning methods, particularly Convolutional Neural Networks (CNNs), to detect stress by analyzing image and video data for physical signs like facial expressions and posture. Using CNNs can significantly improve the accuracy of stress detection compared to traditional methods.

III. METHODOLOGY

The proposed system for textual stress detection is designed as a web-based application that integrates advanced machine learning and natural language processing techniques. The methodology comprises several key stages: data collection, text preprocessing, feature extraction, model training and evaluation, system implementation, and user interaction. Each stage is described in detail below.

1. Data Collection

The foundation of the stress detection system is a labeled dataset containing text samples annotated as either "Stress" or "No Stress." These samples can be sourced from social media posts, online forums, survey responses, or other platforms where users express their thoughts and emotions. The dataset is curated to ensure diversity in language, context, and user demographics, providing a robust basis for model training and evaluation. Data privacy and ethical considerations are strictly adhered to, with all personally identifiable information anonymized or removed.

2. Text Preprocessing

Raw text data is inherently noisy and unstructured, necessitating comprehensive preprocessing to

enhance model performance. The preprocessing pipeline includes the following steps:

- Text Cleaning: Removal of punctuation, special characters, numbers, and extraneous whitespace.
- Lowercasing: Conversion of all text to lowercase to ensure uniformity.
- Stopword Removal: Elimination of common words (e.g., "the," "is," "and") that do not contribute significant meaning to the classification task.
- Stemming/Lemmatization: Reduction of words to their root forms (e.g., "running" to "run") to consolidate similar terms and reduce dimensionality.
- Tokenization: Splitting sentences into individual words or tokens for further analysis.

3. Feature Extraction

Following preprocessing, the cleaned text is transformed into a numerical representation suitable for machine learning algorithms. The system employs the Bag-of-Words (BoW) model or Term Frequency-Inverse Document Frequency (TF-IDF) vectorization to convert text into feature vectors. These methods capture word frequency and importance within the dataset, enabling the classifier to discern patterns indicative of stress.

4. Model Training and Evaluation

The core of the system is a Naive Bayes classifier, chosen for its simplicity, efficiency, and proven effectiveness in text classification tasks. The labeled dataset is divided into training and testing subsets, typically using an 80-20 split. The Naive Bayes model is trained on the feature vectors extracted from the training set, learning to associate specific word patterns with the "Stress" or "No Stress" labels. Model performance is evaluated using metrics such as accuracy, precision, recall, and F1-score on the testing set, ensuring robust and reliable classification.

5. System Implementation

The entire workflow is implemented as a web application using the Flask framework, providing a user-friendly and interactive platform. The backend handles text input, preprocessing, feature extraction, and classification, while the frontend offers intuitive interfaces for both administrators and users. Administrators can manage user accounts, add or update stress-reduction tips, and respond to user queries. Users can register, log in, input text for stress prediction, view curated tips, and communicate with administrators.

6. User Interaction and Feedback

The web application is designed to facilitate seamless interaction between users and administrators. Users receive immediate feedback on their stress prediction, along with personalized tips and resources for stress management. The system also allows users to submit queries or concerns, which administrators can address through a dedicated dashboard. This interactive approach not only enhances user engagement but also supports ongoing improvement of the system based on real-world feedback.

7. Security and Privacy

To ensure user trust and compliance with data protection regulations, the application incorporates robust security measures, including secure authentication, encrypted data storage, and access

controls. User data is handled with strict confidentiality, and all communications between users and administrators are protected.

This methodology provides a comprehensive framework for developing, deploying, and maintaining a machine learning-based web application for textual stress detection. By combining advanced NLP techniques, effective classification algorithms, and an interactive user interface, the proposed system offers a practical and scalable solution for stress awareness and management.

IV TECHNOLOGY USED

1. **Natural Language Processing (NLP)**:

The foundation of the system is built on Natural Language Processing (NLP), which is essential for handling and analyzing user-generated text data. NLP techniques such as text cleaning, tokenization, stopword removal, and stemming or lemmatization are implemented to preprocess and standardize the input, ensuring that the text is in a suitable form for further analysis.

2. **Feature extraction**

For feature extraction, the system employs methods like Bag-of-Words (BoW) and Term Frequency-Inverse Document Frequency (TF-IDF) vectorization. These techniques convert textual data into numerical vectors, capturing the frequency and importance of words, which are then used as input features for the machine learning model.

3. **Naive Bayes algorithm**:

At the core of the classification process is the Naive Bayes algorithm, a probabilistic machine learning model known for its effectiveness and efficiency in text classification tasks. This model is trained on labeled datasets to distinguish between “Stress” and “No Stress” texts based on learned word patterns and features.

4. **Web application**:

The web application is developed using the Flask framework, which provides a lightweight and flexible environment for building interactive web interfaces. Flask enables the integration of backend machine learning logic with a user-friendly frontend, allowing both users and administrators to interact seamlessly with the system.

5. **User interface**:

For the user interface, standard web technologies such as HTML, CSS, and JavaScript are used to create intuitive forms for text input, display prediction results, and facilitate communication between users and administrators. The interface is designed to be accessible and easy to use for individuals with varying levels of technical expertise.

6. **Security and data privacy**

Security and data privacy are prioritized through secure authentication mechanisms, encrypted data storage, and strict access controls. These measures ensure that user information and stress prediction data are handled confidentially and in compliance with data protection standards.

7. **Administrative dashboard**

Finally, the system incorporates an administrative dashboard that allows administrators to manage user accounts, curate stress-reduction tips, and respond to user queries efficiently. This dashboard streamlines the management of resources and enhances the overall support provided to users.

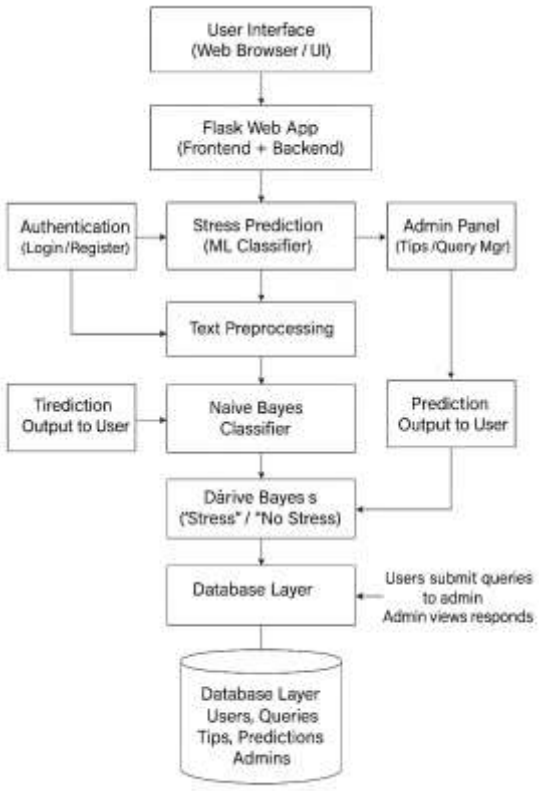
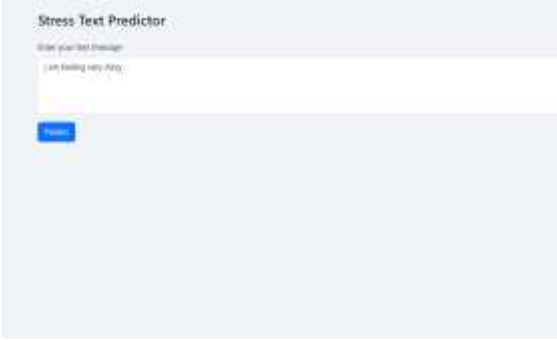


Fig 1. Proposed Methodology

V Result



User can upload image for prediction in this page.

Prediction result is displaying in this page.

VI. CONCLUSION

The proposed web-based stress detection system effectively harnesses the power of natural language processing and machine learning to identify stress indicators from textual data. By employing comprehensive text preprocessing techniques and leveraging the Naive Bayes classifier, the system achieves reliable classification of user input as "Stress" or "No Stress." The integration of this intelligent backend with a user-friendly Flask-based web application provides an accessible platform for individuals to monitor their mental health proactively. Additionally, the inclusion of an administrative dashboard facilitates efficient user management and dissemination of stress-reduction resources, enhancing the overall support framework. Emphasizing security and privacy, the system ensures safe handling of sensitive user data. This holistic approach not only advances the capabilities of digital mental health tools but also promotes early stress detection and management, ultimately contributing to improved well-being and productivity. Future enhancements may include exploring more sophisticated machine learning models and expanding the dataset to increase accuracy and adaptability across diverse user populations.

REFERENCES

1. Text-Based Stress Detection and Classification using Machine Learning, Authors: Harika Juluri, Deepender Singh, T R Yograj Singh,G Sai Kiran Yadav, Journal: International Journal of Creative Research Thoughts (IJCRT), Volume 12, Issue 7 (July 2024);,

ISSN: 2320-2882, Publisher: IJPUBLICATION (IJCRT)

2. Stress Detection using Naive Bayes and Decision Tree Classifiers, Authors: N Krishna Chowdary, Shaik Rasool Basha, Taraswi Jampana, Vishnu Sathvik Reddy, Meghanadh Reddy, Dr. M Kavitha, Affiliation: Department of Computer Science & Engineering, Koneru Lakshmaiah Education Foundation (KLEF), Andhra Pradesh, India, DOI: 10.48047/IJFANS/11/Sp.Iss5/062, ISSN: Not explicitly mentioned in the abstract, but (IJFANS) commonly uses ISSN 2319-1775, Publisher: IJFANS (International Journal of Food and Nutritional Sciences)

3. Stress Detection Using Natural Language Processing and Machine Learning, Authors: Jerripotula Priyanka, Prof. K. Venkata Rao, Journal: (IJNRD), Vol. 8, Issue 9, September 2023; ISSN: 2456-4184, Paper ID: IJNRD2309181, Publisher: IJNRD (IJ Publication).

4. Stress Detection Using Classification Algorithm, Authors: J. S. Kanchana, H. Thaqneem Fathima, R. Surya, R. Sandhiya, Affiliation: Department of Information Technology, K.L.N. College of Engineering, Madurai, India, Journal & Year: Published in (IJERT), Volume 7 Issue 04, April 2018; ISSN: 2278-0181, Publisher: IJERT – International Journal of Engineering Research & Technology

5. Stress Detection Using Machine Learning Techniques, Authors: R. Swarna Malika, Ravi (Assistant Professor), Affiliation: Department of Computer Science, K.R. College of Arts and Science, India, Journal: (JETIR), Vol. 10 Issue 3, March 2023; ISSN: 2349-5162

6. Stress Detection Using Natural Language Processing and Machine Learning Over Social Interactions, Authors: Tanya Nijhawan, Girija Attigeri, T. Ananthakrishna, Journal: Journal of Big Data, Volume 9, Article number: 33 (2022), DOI: https://doi.org/10.1186/s40537-022-00560-z Publisher: SpringerOpen.

7. Stress detection using natural language processing and machine learning over social interactions, Authors-Tanya Nijhawan, Girija Attigeri & T. Ananthakrishna, Publisher- Journal of Big Data, 20 March 2022, Article number: 33 (2022)

8. Machine Learning Driven Mental Stress Detection on Reddit Posts Using Natural Language Processing, Authors:Shaunak Inamdar, Rishikesh Chapekar, Shilpa Gite, Biswajeet Pradhan, DOI:10.1007/s44230-023-00020-8, March 2023

9. Stress detection using deep neural networks, Authors-Russell Li & Zhandong Liu, Publisher-BMC Medical Informatics and Decision Making volume 20, Article number: 285 (2020),

10. Deep Learning Approach for Detecting Work-Related Stress Using Multimodal Signals, Authors: Wonju Seo, Namho Kim, Cheolsoo Park, Sung-Min Park, Publisher: IEEE Sensors Journal, Publication Date: June 2022 (Volume 22, Issue 12), DOI: 10.1109/JSEN.2022.3170915