A PROLIFIC SYSTEM TO DETECT STRESS USING MACHINE LEARNING

Prof. Miruna Joe Amali[1], Roshni Sri KA[2],Priyadharshini T[2], Rithika R[2]

[1]*Professor and Head of the Dept.ofComputerScienceandEngineering,KLNCollegeofEngineering,Pottapalayam, Sivagangai, Tamil Nadu, India*
[2]*B. E Student, Dept. of Computer Science and Engineering, K L N College of Engineering, Pottapalayam ,Sivagangai, Tamil Nadu, India*

---***---

Abstract - *The proposed model presents a comprehensive system for detecting and assessing stress in employees based on text messages, using machine learning techniques. The system utilizes the Naive Bayes algorithm to classify text inputs as stress-related or non-stress-related, and it calculates a stress percentage using both a machine learning model and a keyword-based approach. The analysis is further refined by asking employees to input their weekly working hours and daily sleep duration. These additional factors, coupled with the text analysis, provide a holistic picture of the employee's stress levels. By incorporating machine learning and personal input data, the system ensures a highly accurate and context-sensitive assessment of stress.*

The system evaluates stress at three different levels: low, moderate, and high. If stress is detected in the text, but the working hours remain within the company's set average and the sleep duration is 6 hours or more, the system classifies the stress level as low. In such cases, employees are regarded as managing stress effectively, and the system offers general suggestions to maintain balance based on their working hours. When stress is identified in the text and either the working hours exceed the average or sleep is insufficient (less than 6 hours), the system flags a moderate stress level. In this case, it provides personalized suggestions, helping the employee manage their workload or improve sleep patterns. Finally, if the text indicates stress and both factors—excessive working hours and insufficient sleep— are present, the stress level is labeled high, requiring immediate attention and more focused intervention. If no stress is detected in the text, the system offers suggestions based on working hours alone to promote well-being and maintain a healthy work-life balance. This stress detection system aims to enhance workplace well-being by providing tailored, real-time recommendations that help employees manage stress effectively.

***KEYWORDS*: Machine Learning, Naive Bayes Algorithm, Stress Detection, Text Analysis, Employee Well-being**

1.INTRODUCTION

In today's fast-paced work environment, employee well-being has become a critical factor for both individual productivity and organizational success. Stress, if left unaddressed, can significantly impact an employee's performance, mental health, and overall job satisfaction. To tackle this issue, a comprehensive stress detection and assessment model has been developed, leveraging machine learning to analyze employee text inputs alongside their working hours and sleep patterns. By integrating these factors, the model offers a highly accurate, personalized approach to identifying and managing stress. The core of this system utilizes the Naïve Bayes algorithm to classify text messages as either stress-related or non-stress-related. This classification is further refined by calculating a stress percentage through a combination of machine learning and a keyword-based analysis. However, the system goes beyond simple text analysis by incorporating additional key factors—employees' weekly working hours and daily sleep duration. By considering these objective metrics alongside the subjective text input, the model provides a holistic picture of each employee's stress level. The system evaluates stress across three levels: low, moderate, and high. When stress is detected through text but is mitigated by balanced working hours and sufficient sleep, the system flags a low stress level, indicating that the employee is managing well. If either working hours exceed the company's set average or sleep falls below 6 hours, moderate stress is identified, and personalized recommendations are provided. In cases where both excessive working hours and poor sleep patterns are present, the stress level is classified as high, triggering immediate attention and intervention. By combining machine learning with real-time data input, this model ensures a context-sensitive and dynamic approach to stress management, aimed at enhancing employee well-being and promoting a healthier work-life balance.

2.LITERATURE SURVEY

[1] Analyzing Perceived Psychological and Social Stress of University Students: A Machine Learning Approach - Ishrak Jahan Ratul, Mirza Muntasir Nishat, Fahim Faisal, Sadia Sultana, Ashik Ahmed, Md Abdullah Al Mamun. The concept of the study revolves around using machine learning techniques to better understand the factors contributing to stress among university students. Traditional methods of stress assessment often rely on self-reported surveys,

which may not fully capture the complexity of stress experiences. This study leverages machine learning algorithms to analyze diverse data points, such as academic performance, social relationships, and demographic factors, to identify key predictors of psychological and social stress. By using both supervised and unsupervised learning, the study uncovers patterns and trends that offer a more nuanced understanding of the stressors affecting university students.

[2]Early Mental Stress Detection Using Q-Learning Embedded Starling Murmuration Optimiser-Based Deep Learning Model - S. K. R. Moosavi, M. H. Zafar, F. Sanfilippo, M.N.AkhterandS.F.Hadi. This system integratesQ-learning with a Starling murmuration-inspired optimization technique within a deep learning model for early mental stress detection. Q-learning helps the model adapt in real-time, while the murmuration technique optimizes performance. Deep learning identifies complex stress patterns from physiological data. This approach enables proactive, real-time stress monitoring and can be integrated into wearable technology for continuous mental health support.

[3]Ensemble Hybrid Learning Methods for Automated Depression Detection -M. N. Ali,M. AmerandM. Elsisi. This system examines hybrid and ensemble methods for automated depression detection using text data from sources like social media. Hybrid models combine classifiers to better capture depressive patterns, while ensemble methods enhance accuracy. Key challenges include symptom variability, privacy concerns, and potential biases in training data. Performance is evaluated using metrics such as accuracy, precision, and recall to assess depression detection effectiveness.

[4]Calibration of Transformer-Based Models for Identifying Stress and Depression in Social Media -L.Ilias, S.Mouzakitis and D. Askounis. This study leverages advanced natural language processing (NLP) techniques, particularly transformer models, to detect mental health conditions like stress and depression through social media data. By fine-tuning pre-trained models on accurate data set, the authors enhance the model's ability to identify subtle linguistic cues indicative of emotional distress.

The research demonstrates that transformer-based models outperform traditional machine learning methods in predictive accuracy, offering valuable insights for early mental health intervention. The paper also addresses ethical concerns, such as privacy and responsible AI use in mental health diagnostics.

3. EXISTING SYSTEM

The existing system for identifying stress in text messages primarily utilizes the Support Vector Machine (SVM) algorithm as its core classification method, known for its effectiveness in high-dimensional spaces. It analyzes various forms of written communication, such as social media posts and chat messages, to classify them as indicative of stress or not. The system begins with data collection, followed by an extensive preprocessing phase that includes tokenization, normalization, and stop word removal. After preprocessing, feature extraction techniques like Term Frequency-Inverse Document Frequency (TF-IDF) are employed to convert text into a numerical format suitable for SVM input. The SVM model is then trained on labeled datasets to identify the optimal hyperplane separating stress and non- stress messages. While the approach demonstrates a systematic methodology, it faces significant challenges, including high computational costs, dependency on effective feature extraction, and a lack of transparency in decision-making.

The complexity of the SVM algorithm can result in difficulties in interpreting classifications, particularly critical in mental health contexts. Additionally, the system risks over fitting due to high dimensionality and has limitations in handling non-linear relationships. Overall, while effective in detecting stress, the system highlights the need for ongoing research to address its limitations and improve performance.

4.PROPOSED SYSTEM

The proposed system is a comprehensive tool designed to monitor and assess employee stress using a combination of text analysis and personal data inputs, primarily employing the Naïve Bayes algorithm for text classification. It analyzes written communication, such as emails and messages, to detect stress- related indicators while incorporating a keyword-based detection approach for enhanced accuracy.

Alongside text analysis, employees provide personal metrics like weekly working hours and daily sleep duration, allowing for a nuanced evaluation of stress contributors. The system categorizes stress into three levels—low, moderate, and high—based on text analysis and personal data, facilitating targeted interventions. Employees receive tailored recommendations for managing stress, which include general well-being tips and specific strategies for work load and sleep improvement. This proactive approach helps identify early signs of stress, enabling timely interventions to prevent burnout. Moreover, the system promotes a healthy work-life balance by offering suggestions even when no stress is detected, fostering long-term well-being.

By integrating psychological and lifestyle factors, the system provides a holistic understanding of employee stress, ultimately enhancing workplace well-being and productivity.

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 10 | Oct 2024 www.irjet.net p-ISSN: 2395-0072

5.SYSTEMOVERVIEW

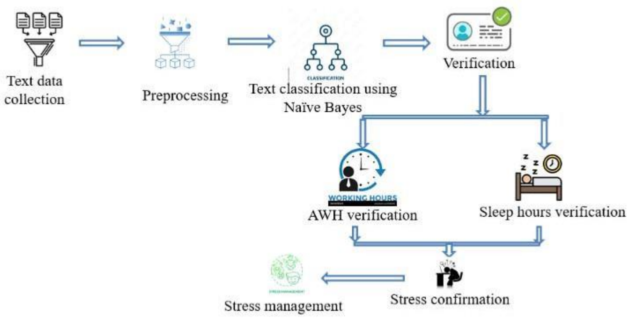

Fig 1-Architecture of the system

A system designed for stress detection and management through text analysis and lifestyle monitoring. It begins with text data collection, where user inputs, such as text messages, provide insights into mental states. The preprocessing stage cleans and formats the data for analysis, removing irrelevant information and handling missing values.

The system employs a Naive Bayes classifier to analyze the text and detect stress levels based on linguistic patterns. Verification follows, checking lifestyle factors like average working hours and sleep hours to identify potential stress triggers. If stress is confirmed, the system recommends management techniques such as time management and relaxation exercises. This integrated approach combines emotional analysis with lifestyle monitoring, offering personalized feedback for effective stress management.

5.1 Text data collection

Text Data Collection and Processing involves gathering text messages from various sources. Sources include SMS, social media platforms, and messaging apps like Whats App and Facebook. API integration and automated data collection tools facilitate efficient data gathering. Collected data must ensure accuracy, completeness, and sufficient volume. Data quality assurance checks are performed to validate data integrity. Machine learning algorithms, specifically Natural Language Processing (NLP), extract relevant features. Text classification and language modeling tasks benefit from this processed data. Scikit-learn and NLTK libraries provide essential tools for NLP tasks. Effective text analysis enables applications like sentiment analysis and topic modeling. By leveraging large-scale text data businesses can gain valuable insights into customer opinions and behaviors

5.2Text classification using Naive Bayes

Text preprocessing phase aimed at enhancing model accuracy by removing noise from the input data. This involves converting all text to lowercase, tokenizing words into manageable units, and eliminating stop words that do not add significant meaning to the analysis. The cleaned and processed text is then used to define two class labels: "Stress" and "No Stress," allowing the system to classify text messages accordingly. The core of the classification process integrates a keyword detection approach with a Naive Bayes training model, which effectively categorizes the messages into either "Stress" or "No Stress" based on the presence of stress-related terms and patterns. Additionally, the system calculates a stress percentage by applying a stress probability function, providing a more nuanced understanding of the detected stress levels. To ensure the effectiveness of stress detection, the performance of the Naive Bayes model is evaluated using various metrics, including accuracy, precision, and recall. This evaluation helps identify the model's strengths and weaknesses in detecting stress. Furthermore, the system assigns points based on the stress level: five points are given if stress is detected, while two points are assigned if no stress is found. This scoring mechanism serves as an intuitive way to gauge overall employee stress levels, intuitive way to gauge overall employee stress levels, contributing to a more comprehensive assessment of mental well-being in the workplace.

5.3 Verification with AWH and sleephours

The module assesses employee stress by comparing weekly working hours to a set average threshold. If hours exceed this threshold, two stress points are assigned, while sleep entries under six hours also receive two stress points. The total stress level points are calculated by adding these points to those derived from text analysis. Employees are notified when their working hours or sleep falls short, accompanied by reminders and suggestions for improvement. Additionally, repeated instances of low sleep are tracked for further review, allowing for targeted interventions. Based on the total points accumulated, the module recommends personalized stress-relieving techniques, promoting proactive management of stress and encouraging healthier lifestyle choices.

5.4 Stress Analysis

Stress analysis is based on a comprehensive analysis of several key factors, including text messages classified as "Stress" or "No stress," along with the employee's working hours and sleep duration. The module aggregates points from these factors, providing a numerical representation of stress levels. If the total points amount to 2, it indicates effective stress management, suggesting that the employee is maintaining a healthy balance. A score of 5 signifies low stress, which may warrant attention to ensure it does not escalate.With7points,theresultiscategorizedasmoderate stress, indicating a need for proactive measures.

A score of 9 reflects high stress, signaling an urgent requirement for intervention. Based on the total points accumulated, tailored suggestions are provided to help employees manage their stress effectively, enhancing their overall well-being and productivity.

5.5 Stress Management

Stress management is facilitated by offering personalized suggestions tailored to the individual employee's needs, based on the points collected from their stress analysis. By evaluating the number of working hours and other contributing factors, the system can provide specific recommendations to help alleviate stress. For instance, employees may be encouraged to listen to calming music, which can enhance relaxation and improve mood.

Additionally, suggestion may include engaging in meditation and yoga, both of which are effective practices for reducing anxiety and promoting mental clarity. Going on a vacation or taking short breaks can also be recommended to help employees recharge and disconnect from work-related pressures. These personalized strategies aim to empower employees to take proactive steps in managing their stress levels, ultimately fostering a healthier work environment. By implementing such tailored interventions, the organization demonstrates its commitment to employee well-being and productivity.

Fig 2-Stress Detection

6.RESULT

This analysis implements a stress detection and management system using a data set of text messages labeled for stress levels. Initially, it preprocesses the text by cleaning it and removing noise, applying stemming, and filtering out stop words. The code uses a Naïve Bayes classifier trained on TF-IDF features derived from the cleaned text to classify user input as "Stress" or "No Stress." After predicting stress levels, the code calculates a stress level score based on user input regarding working hours and sleep duration, with points assigned for excessive work hours or inadequate sleep. The score helps categorize stress into four levels: great management, low, moderate, or high stress. Personalized suggestions for Stress management techniques are generated based on the number of working hours. The final output provides the user with their stress classification, stress percentage, and tailored techniques to manage their stress effectively, promoting overall well-being.

7.FUTUREENHANCHEMENTS

The system can be transformed into a website that helps more corporate employees manage their stress and well-being. By integrating health metrics like heart rate, physical activity, and sleep quality, alongside work-related factors such as workload and deadlines, the platform would offer a more accurate and comprehensive understanding of stress triggers. Expanding the model to detect a broader range of emotions, including frustration, anxiety, and burnout, would provide deeper insights into employee well-being.

This approach allows for more personalized interventions and solutions, promoting better emotional and mental health management across the workforce, ultimately improving productivity and job satisfaction.

8. CONCLUSION

The system effectively detects employee stress by integrating Naive Bayes text analysis with real-world data, such as working hours and sleep patterns. This hybrid approach enhances accuracy in stress identification through machine learning and empirical metrics. Upon confirming stress, data-driven interventions like workload reduction are implemented to alleviate pressure and improve employee well-being. By continuously analyzing sentiment and behavioral data, the system provides a comprehensive view of mental health, helping employees avoid burnout and maintain focus. This innovative AI-driven model fosters a healthier work environment, benefiting both employees and organizations.

REFERENCES

[1] M. Milyavskaya, M. Saffran, N. Hope, and R. Koestner, “Fear of missing out: Prevalence, dynamics, and consequences of experiencing FOMO,”Motivat. Emotion, vol. 42, no. 5, pp. 725–737,Oct.2018.

[2] L. Tomczyk and E. Selmanagic-Lizde, “Fear of missing out (FOMO)Among us in Bosnia and Herzegovina—Scale and selected mechanisms,” Children Youth ServicesRev., vol.88,pp. 541–549, May 2018.

[3] S. Mahmud, S. Hossain, A. Muyeed, M. M. Islam, and M. Mohsin,“The global prevalence of depression, anxiety, stress, and, insomnia and its changes among health professionals during COVID-19 pandemic: A rapid systematic review and meta-analysis,” Heliyon,vol. 7, no. 7, Jul. 2021, Art. no. e07393.

[4] M.H.Akhand,“Factors contributing to mental health problems in Kazakhstan: Literature review,” B.S. thesis, Dept. Appl.Sci., Turku Univ.,Turku, Finland, 2019.

[5] P.W.Corrigan and A.C.Watson, “Understanding the impact of stigma on people with mental illness,”Proc.Comput.Sci.,vol. 1, pp. 16–20, 2022.

[6] H.J.Lim,L.Moxham,C.Patterson,D.Perlman,V.Lopez,andY. S. Goh, ‘‘Students’ mental health clinical placements, clinical confidence and stigma surrounding mental illness: A correlational study,’’ Nurse Educ. Today, vol. 84, Jan. 2020, Art. no. 104219.

[7] K. Masood, B. Ahmed, J. Choi, and R. Gutierrez-Osuna, ‘‘Consistency and validity of self-reporting scores in stress measurement surveys,’’ in Proc. Annu. Int.Conf. IEEEEng. Med. Biol. Soc., Aug. 2012, pp. 4895–4898

[8] T. W. Wong, Y. Gao, and W. W. S. Tam, ‘‘Anxiety among university students during the SARS epidemic in Hong Kong,’’ Stress Health, vol. 23, no. 1, pp. 31–35, Feb. 2007.

[9] K. Shamsuddin, F. Fadzil, W. S. W. Ismail, S. A. Shah, K. Omar,N. A. Muhammad, A.Jaffar, A.Ismail, andR.Mahadevan,”Correlates of depression,anxiety and stress among Malaysian University students”,Asian J.Psychiatry,vol.6,no.4,pp-318–323,Aug.2013.

[10] M.M.Al-Shahrani, B.S.Alasmri, R.M.Al-Shahrani, N.M.Al-Moalwi,A.A.AlQahtani,andA.F.Siddiqui,‘‘The prevalence and associated factors of academic stress among medical students of King Khalid University: An analyticalcross-sectionalstudy,’’Healthcare,vol.11,no. 14, p. 2029, Jul. 2023.

[11] T.McCloud, S.Kamenov, C.Callender, G.Lewis,and G. Lewis, ‘‘The association between higher education attendance and common mental health problems among young people inengland: Evidence from two population-based cohorts,’’ Lancet Public Health, vol. 8, no. 10, pp. e811–e819, Oct. 2023.

[12] J. Friedrich, A. Bareis, M. Bross, Z. Bürger, Á. Cortés Rodríguez, N.Effenberger, M. Kleinhansl, F. Kremer, and C. Schröder, ‘‘How is your thesis going?’ Ph.D. students’ perspectives on mental health and stress in academia,’’ PLoS ONE, vol. 18, Jul. 2023, Art. no. e0288103.

[13] H. Almerekhi, H. Kwak, and B. J. Jansen, ‘‘Investigating toxicity changes of cross-community redditors from 2 billion posts and comments,’’ PeerJComput. Sci., vol. 8, p. e1059, Aug. 2022, doi: 10.7717/peerj-cs.1059.

[14] M. M. Tadesse, H. Lin, B. Xu, and L. Yang, ‘‘Detection of depression-related posts in Reddit social media forum,’’ IEEE Access, vol. 7,pp. 44883–44893, 2019.

[15] R. Shounak, S. Roy, V. Kumar, and V. Tiwari, ‘‘Reddit comment toxicity score prediction through BERT via transformer based architecture,’’ in Proc. IEEE 13th Annu. Inf. Technol., Electron. MobileCommun. Conf.(IEMCON), Oct.2022, pp.0353–0358.

[16] M. Chauhan, S. V. Vora, and D. Dabhi, ‘‘Effective stress detection using physiological parameters,’’ in Proc. International Conference on Innovations in Information, Embedded and Communication Systems, vol. 2017, pp-014-018.