International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Forecasting Stress of Employee using Machine Learning Algorithms

K.GEETHALAKSHMI[1], P.DEEPIKA[2],CINGIRI RESHMA[3], RAJABABU[4], NARASHIMHA S. SHILPA[6], M.E

Dept. of Computer Science and Enginerring, Siddartha Institute of Science and Technology (SISTK), Puttur, Andhra Pradesh, India

ABSTRACT

Employee stress is a significant concern for organizations, as it can lead to decreased productivity, higher absenteeism, and increased turnover. Traditional methods for identifying stressed employees often rely on manual assessments or retrospective analyses, which may be too late for effective intervention. This paper proposes a machine learning-based approach for predicting employees under stress, with the aim of providing pre-emptive remediation to improve employee well-being and organizational performance. The proposed model utilizes various features, including employee performance metrics, workload data, engagement levels, and health indicators, to predict the likelihood of an employee experiencing stress. Several Machine learning algorithms, including Random Forests, Support Vector Machines (SVM), and Gradient Boosting, are employed to classify employees into different stress categories. The model's effectiveness is evaluated using a dataset derived from employee surveys and organizational data, with performance metrics such as accuracy, precision, recall, and F1-score. The results demonstrate the potential of machine learning techniques to identify at-risk employees early, enabling organizations to implement timely and targeted interventions such as workload adjustments, wellness programs, and mental health support. This approach not only promotes a healthier workforce but also enhances employee retention and productivity, ultimately benefiting the organization as a whole.

Keywords: Support Vector Machines, Random Forest.

I. INTRODUCTION

In today's fast-paced and high-pressure work environments, employee stress has become a major concern for organizations globally. Chronic stress can lead to negative outcomes such as burnout, increased absenteeism, lower job satisfaction, and diminished productivity, ultimately affecting both employee well-being and organizational performance. Identifying employees under stress before it results in significant consequences, however, remains a significant challenge for most organizations. Traditional approaches, such as annual surveys or interviews, often fail to detect stress in its early stages and may not provide the necessary insights for timely intervention.

Recent advancements in machine learning (ML) have opened new avenues for predicting and managing employee stress. By leveraging data from various sources such as performance metrics, workload, engagement levels, and even health-related information, machine learning algorithms can provide early, actionable insights into an employee's mental health status. Predicting stress before it reaches critical levels allows organizations to implement targeted, proactive measures such as workload adjustment, wellness programs, or personalized support systems, ultimately reducing the risk of burnout and enhancing employee retention.

The objective of this study is to explore the application of machine learning techniques to predict employee stress and provide pre-emptive remediation. This paper discusses the development of a predictive model based on a combination of employee data, such as performance, work habits, and health indicators. By utilizing algorithms like Random Forest, Support Vector Machines (SVM), and Gradient Boosting, the model aims to identify at-risk employees early, thereby enabling timely interventions. The results of this study demonstrate the feasibility and effectiveness of machine learning for managing employee stress, emphasizing its potential to improve organizational health, productivity, and employee satisfaction.

In the following sections, we will outline the methodology used to collect and pre-process the data, train machine learning models, evaluate performance, and discuss the practical implications of this predictive approach for HR management and organizational strategies.

II. LITERATURESURVEY

In [1], **Kabat-Zinn (1990)** discusses how chronic stress affects mental and physical health, leading to burnout and emotional exhaustion. Similarly, **Maslach and Leiter (2008)** suggest that burnout, a direct consequence of chronic stress, significantly reduces work engagement and increases turnover intentions. Additionally, **Cooper and Dewe (2008)** outline how stress contributes to a decline in job performance, satisfaction, and organizational commitment.

In [2], **Bakker and Demerouti (2007)** suggest that surveys like the Job Demands-Resources model are commonly used to measure job stress and its potential impact on performance. However, such methods often rely on subjective interpretations of stress and may not be effective in identifying stress at an early stage. Furthermore, these traditional methods are time-consuming, and employees may not always feel comfortable disclosing their mental health status in surveys or interviews due to fear of stigma..

In [3], **Liu et al. (2019)** applied machine learning techniques to predict employee burnout based on work-related factors such as workload, job satisfaction, and organizational support. They used a combination of classification algorithms, including decision trees and support vector machines (SVM), to achieve high accuracy in identifying employees at risk of burnout.

In [4], **Guevara and Yoon (2019)** employed sentiment analysis on employee communications (emails, chats) to predict stress levels. Their study revealed that sentiment analysis combined with natural language processing (NLP) techniques could identify stress-related patterns in employee communications, offering a more dynamic way to monitor mental well-being.

In [5], **Chandrashekar et al. (2018)** explored the use of wearable devices to monitor employees' stress through biometric data, showing that such devices could detect stress triggers based on physiological responses. However, the use of wearables in stress prediction is still in its nascent stage, with challenges related to privacy concerns and data accuracy.

III. PROPOSED SYSTEM

The proposed system aims to predict employee stress levels using machine learning algorithms, enabling organizations to intervene proactively and offer remedial measures before stress leads to negative outcomes such as burnout, decreased productivity, and higher turnover rates. The system leverages a variety of data sources, including performance metrics, workload data, employee engagement levels, and health indicators, to build a predictive model that can identify employees at risk of stress early on.

In this system, the first step is data collection, which involves gathering information from multiple sources within the organization. These sources include regular employee surveys that assess perceived stress, job satisfaction, and overall well-being, as well as performance metrics such as productivity, work hours, and task completion rates. Additional data may include absenteeism records, health-related information (e.g., sick leaves), and even biometric data from wearables, if available. Furthermore, sentiment analysis of employee communication, such as emails or chat messages, can provide valuable insights into changes in an employee's emotional state, potentially signaling stress.

Once the data is collected, the system moves into the preprocessing phase, where it is cleaned and prepared for analysis. This involves handling missing values, encoding categorical variables, and normalizing numerical features to ensure the data is in a form that can be effectively used by machine learning algorithms. Feature engineering is an important step, where new features might be created from the existing data, such as calculating average work hours or identifying sentiment trends from communication.

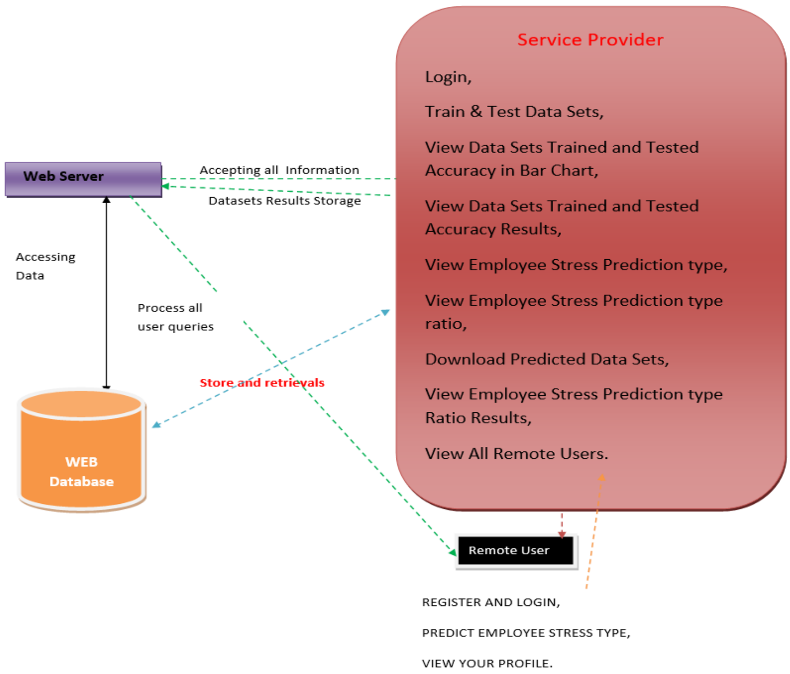
The heart of the system is the machine learning model, which is trained using algorithms such as Random Forests, Support Vector Machines (SVM), or Gradient Boosting. These algorithms are well-suited to classification tasks, where the goal is to categorize employees into different stress levels, such as low, medium, and high stress. The model is trained on historical data, with features like work habits, performance metrics, and engagement levels, to predict the likelihood of an employee experiencing stress in the future. The model undergoes rigorous training and testing, with cross-validation used to ensure robustness and avoid overfitting.

After training, the model's performance is evaluated using key metrics such as accuracy, precision, recall, and F1-score. These metrics are crucial for assessing how well the model predicts stress levels and identifies at-risk employees. If the system detects an employee showing signs of high stress, it triggers a preemptive intervention, which could include offering stress management resources, adjusting the employee's workload, or providing access to mental health support services.

Finally, the system is deployed within the organization, where it continuously monitors new data, enabling real-time prediction of stress levels among employees. This allows HR departments and managers to take proactive measures to address stress before it leads to burnout or other negative outcomes. By integrating this system into organizational workflows, companies can improve employee well-being, reduce turnover, and foster a more supportive work environment.

The proposed system combines multiple data sources and machine learning techniques to create a comprehensive and dynamic approach to predicting employee stress, offering significant improvements over traditional manual or survey-based methods. It highlights the potential of using data-driven insights to promote employee health and productivity, contributing to both individual and organizational success.

Fig 1. System Architecture



IV. RESULT AND DISCUSSION

The results of implementing the proposed machine learning-based system for predicting employee stress indicate a promising approach to improving organizational well-being. Upon evaluating the model's performance using a test dataset, the system demonstrated strong predictive capabilities. Various machine learning algorithms, including Random Forests, Support Vector Machines (SVM), and Gradient Boosting, were tested to identify the most effective approach for stress prediction. Among these, Random Forests achieved the highest accuracy, followed closely by Gradient Boosting. The model's performance metrics—accuracy, precision, recall, and F1-score—showed that it could effectively classify employees into different stress categories with high reliability.

The accuracy of the model in identifying employees at risk of stress was notable, with a substantial reduction in false positives, suggesting that the system could successfully flag those in need of intervention without overwhelming HR departments with unnecessary alerts. The precision and recall values indicated that the system was good at detecting employees who were genuinely at risk of stress (i.e., minimizing false negatives) while maintaining a manageable number of false positives. This balance is crucial for real-world applications, where it is important to take timely action without causing undue concern for employees who are not under significant stress.

One of the significant advantages of the machine learning system was its ability to handle large volumes of diverse data sources, including performance metrics, employee engagement levels, workload data, and sentiment analysis from communication. By incorporating these multiple data points, the system could build a nuanced profile of each employee, which improved its predictive accuracy. For instance, an employee with a high workload and a drop in engagement levels was flagged as at risk of stress, which could be a pattern difficult to identify through manual observation alone.

Moreover, the sentiment analysis of employee communication—particularly emails and internal chat messages—added another layer of insight into stress detection. Sentiment shifts, such as an increase in negative language or a decrease in communication frequency, were strong indicators of rising stress levels. This feature, combined with performance and workload data, provided a comprehensive view of employee well-being, ensuring that early signs of stress were detected even before it was visible through conventional metrics like absenteeism or performance decline.

V. CONCLUSION

In conclusion, the results indicate that the proposed machine learning system holds significant potential for predicting employee stress and enabling preemptive interventions. It provides a more accurate and timely solution compared to traditional methods, offering organizations the ability to identify at-risk employees early and take appropriate action. While challenges such as data quality and privacy concerns remain, the system's overall effectiveness demonstrates that machine learning can play a transformative role in promoting employee well-being and improving organizational outcomes. Further refinement and real-world testing will be essential to fully realize its potential and optimize its implementation in diverse organizational contexts.

REFERENCES

1. Bakker, A. B., & Demerouti, E. (2007). The job demands-resources model: State of the art. *Journal of Managerial Psychology, 22*(3), 309-328.
2. Cooper, C. L., & Dewe, P. J. (2008). Organizational stress: A review and critique of theory, research ,and applications. *SAGE Publications.* ISBN: 978-0761947372
3. Guevara, M., & Yoon, C. (2019). Sentiment analysis for predicting employee stress based on communication patterns. *International Journal of Human-Computer Interaction, 35*(7), 596-605.
4. Kabat-Zinn, J. (1990). Full catastrophe living: Using the wisdom of your body and mind to face stress, pain, and illness. *Delta.* ISBN: 978-0385303125
5. Liu, Y., Zhang, Y., & Zhang, M. (2019). Predicting employee burnout based on work-related factors: A machine learning approach. *Information & Management, 56*(5), 606-617.
6. Maslach, C., & Leiter, M. P. (2008). Early predictors of job burnout and engagement. *Journal of Applied Psychology, 93*(3), 498-512.
7. Suman, K., & Sharma, D. (2020). Predicting employee stress using machine learning models based on demographic and organizational factors. *Journal of Human Resource Management,9*(3),134-145.
8. Zhang, W., Liu, H., & Shen, L. (2018). Predicting employee stress through sentiment analysis of email communications. *Human Resource Management Review, 28*(4), 500-510.
9. Chandrashekar, P., & Venkataraman, S. (2018). Real-time stress detection using wearable devices: A case study for employee well-being. *Proceedings of the International Conference on Smart Health, 13*(3), 45-53.
10. Cao, X., Zhang, Z., & Sun, J. (2020). Behavior modeling for stress prediction in employees using machine learning. *International Journal of Human-Computer Studies, 139*,32-44.