International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

PREDICTING EMPLOYEES UNDER STRESS FOR PRE-EMPTIVE REMEDIATION USING MACHINE LEARNING ALGORITHM

YEKKALURI MAMATHA , Dr. K. CHENNAKESAVULU
MCA Student, Professor
DEPT OF MCA
PVKK INSTITUTE OF TECHNOLOGY(AUTONOMOUS), Anantapuramu – 515001 (A.P)
yekkalurimamatha20@gmail.com, intell.chenna@gmail.com

ABSTRACT

Employee stress has emerged as a critical concern in modern workplaces, impacting productivity, well-being, and organizational efficiency. Prolonged exposure to stress can lead to burnout, absenteeism, and declining performance, necessitating proactive measures for early identification and intervention. Predicting employees under stress through data-driven methodologies enables organizations to take pre-emptive remedial actions, fostering a healthier and more productive work environment.

This study explores predictive models that leverage various indicators, including workload, work hours, sentiment analysis from communication channels, absenteeism records, and physiological markers (where applicable). Machine learning algorithms, natural language processing (NLP), and sentiment analysis techniques can help identify patterns that correlate with high stress levels. By integrating these predictive insights with human resource strategies, companies can provide timely interventions such as workload redistribution, mental health support, and stress management programs.

The implementation of such predictive models requires ethical considerations, including employee consent, data privacy, and transparency in decision-making. Balancing technology with human-centric approaches ensures that interventions are supportive rather than intrusive. This research aims to highlight the significance of predictive analytics in stress management and its potential to enhance employee well-being while improving organizational outcomes.

1. INTRODUCTION

` On March 11, 2020, the World Health Organization (WHO) reported coronavirus (COVID-19) a pandemic that signifies a global, epidemic disorder frightening the entire universe. COVID-19 is a contagious disease affected by the coronavirus. 'Coronaviruses' are a huge family of viruses that cause ailments varying from the typical flu to other critical complications. According to WHO, on March 31, 2020, the virus had reached 202 countries. Due to this, stock markets and other sectors have experienced a severe downturn in growth. This, in turn, affects employees too, who feel stressed when they are unable to cope with prolonged uncertainty and pressure. The application of machine learning and artificial intelligence to the field of business is seeing a lot of promising growth. The pattern of employee behavior is analyzed in.

Vis-à-vis, they do not have any satisfaction due to long working hours in addition to having a heavy workload. Here, the foremost objective of this research is to analyze the consequence of stress on employee appearance. Moreover, this influences physical ailments and a lack of commitment to work. However, in the contemporary situation, COVID-19 has put the world population in an unprecedented position. Through this work, we intend to analyze the stress level that employees are subjected to owing to a phenomenon like the present pandemic. Here, machine learning algorithms are used to predict whether employees undergoing stress or not.

The modern world is flooded with IT, and IT companies are being greeted with new extensions and requests. Representatives are bound to face pressure as a result of the changing way of life and working societies. Frameworks identify components that have a significant impact on anxiety feelings. Stress was identified in relation to orientation, family ancestry, and the availability of medical benefits in the workplace. Recognizing the pressure on representatives allows us to devise a few strategies for dealing with it and creating a much more pleasant work environment for their representatives. Several research works make use of a variety of constraints, such as preferences, age, family history, provided medical benefits, shared information about illness, technical institution, technical job, acquiring holidays, and so on. Artificial Intelligence calculations are used in research to determine an employee's stress level. The primary goal of each of these research projects is to identify the gambling factors that influence the worker's emotional wellness.

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 project 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.

Disorders of stress which are related to mental health are not rare for the employees working in corporate sectors. Some analysis done earlier have created some concern on the very same. Based on the work done by Association of Industry, Assocham, we come to know that above 42% of the professional working employees in the corporate private sectors of India are suffering from stress or common disorders of anxiety because of late night working hours and due to fixed timings. This part of singles are growing as mentioned in the Economic Times of 2018 project which is dependent on the survey that was managed by the Optum.

The methodology outlines the systematic process followed to develop a machine learning-based system for identifying employees under stress and enabling proactive interventions.

1. Data Collection

Collect employee-related data from various internal and external sources. These may include:

- Employee demographic data (age, gender, department, job role)
- Behavioral data (attendance, working hours, leaves taken)
- Performance metrics (KPI, deadlines missed, tasks completed)
- Survey data (job satisfaction, work-life balance, stress level)
- Optional: biometric or sentiment analysis data

2. Data Preprocessing

- Data Cleaning: Handle missing, inconsistent, or duplicate values.
- Feature Encoding: Convert categorical data using label encoding or one-hot encoding.
- Normalization: Apply Min-Max or Standard Scaling to normalize numerical features.
- Outlier Detection: Remove or treat anomalies using IQR or Z-score.
- Class Balancing: Handle class imbalance using oversampling (e.g., SMOTE) or under sampling.

3. Feature Selection

- Use statistical techniques (e.g., correlation matrix, chi-square test) or model-based techniques (e.g., feature importance from Random Forest) to select relevant features.
- Remove redundant or non-informative attributes to improve model performance.

4. Model Selection and Training

Apply and compare different machine learning classification algorithms:

- Logistic Regression
- Decision Tree Classifier
- Random Forest
- Support Vector Machine (SVM)
- XGBoost
- Artificial Neural Networks (ANN)

Split the dataset into training and testing sets (e.g., 80:20). Train models using the training data and validate using cross-validation (e.g., K-Fold).

5. Model Evaluation

Evaluate each model based on:

- Accuracy
- Precision
- Recall
- F1-score
- Confusion Matrix
- ROC-AUC Score

Select the best-performing model for deployment.

6. Stress Level Prediction

- Use the trained model to predict employee stress levels as:
  - Low
  - Medium
  - High

The prediction helps HR to identify individuals at risk early and intervene accordingly.

7. Remediation Recommendation System

- Based on predicted stress levels, generate automated wellness recommendations:
  - Counseling support
  - Flexible work arrangements
  - Stress management workshops
  - Time-off suggestions
- Store and track recommendations and improvements over time.

8. Deployment (Optional)

- Develop a web-based interface or dashboard for HR using Flask/Django or Streamlit.
- Integrate the model into the HR system for real-time prediction.
- Secure the system to ensure employee data privacy.

9. Continuous Monitoring & Model Retraining

- Monitor the system's performance regularly.
- Update the model with new data to improve prediction accuracy.
- Adapt to organizational or environmental changes

II. LITERATURE SURVEY

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.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

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.

III.SYSTEM ANALYSIS

3.1 EXISTING SYSTEM:

In many organizations, stress management primarily relies on reactive approaches, where interventions are implemented only after employees exhibit visible signs of distress. Traditional systems often depend on self-reporting mechanisms, such as surveys, feedback forms, and employee assistance programs (EAPs). While these methods provide valuable insights, they suffer from limitations like response bias, delayed reporting, and lack of real-time assessment. Additionally, many employees hesitate to disclose their stress levels due to stigma or fear of negative repercussions, making it difficult for management to address stress proactively.

Existing stress detection systems, if implemented, generally use basic performance metrics, absenteeism records, and medical reports to assess workplace stress. However, these methods do not leverage advanced analytics or real-time monitoring techniques to predict stress before it significantly affects employee well-being and productivity. Some organizations provide wellness programs, counseling services, and periodic mental health assessments, but these initiatives are often generic and do not cater to individual stress levels effectively.

Moreover, current stress management solutions lack integration with emerging technologies such as artificial intelligence (AI) and machine learning (ML). Most organizations still rely on human resource (HR) professionals and managers to manually identify employees struggling with stress, making the process slow and inefficient. In some cases, organizations use wearables or biometric tracking devices to monitor stress indicators such as heart rate variability and sleep patterns, but such systems are not widely adopted due to concerns about privacy, data security, and cost.

DISADVANTAGES OF EXISTING SYSTEM:

The existing systems for predicting employee stress often rely on traditional surveys, self-reporting mechanisms, or periodic assessments, which can be both time-consuming and inaccurate. Employees may not always provide honest responses due to fear of repercussions, or they may not recognize their own stress levels until they reach a critical stage. Additionally, manual data collection and evaluation can lead to biases, inconsistencies, and delays in identifying employees who require intervention. These inefficiencies make it difficult for organizations to take timely action and implement preemptive remediation strategies effectively.

Another major drawback is the lack of real-time data analysis and predictive capabilities. Many systems fail to integrate advanced technologies such as artificial intelligence (AI) or machine learning (ML) to detect early signs of stress based on behavioral patterns, work performance, or physiological indicators. Without continuous monitoring and predictive analytics, organizations miss

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

opportunities to intervene before stress leads to burnout, decreased productivity, or employee turnover. Additionally, traditional systems often do not provide personalized recommendations, making it challenging for HR departments to implement targeted and effective remediation measures.

Furthermore, existing stress prediction systems may not account for various external factors contributing to employee stress, such as financial worries, family issues, or personal health conditions. Most systems focus solely on workplace-related factors, leading to an incomplete understanding of employee well-being. This narrow scope limits the effectiveness of stress mitigation strategies and prevents organizations from addressing the root causes of stress comprehensively. Additionally, a lack of integration with other HR tools and mental health resources further reduces the system's overall effectiveness.

3.2 PROPOSED SYSTEM:

The proposed system for predicting employees under stress for preemptive remediation is a data-driven solution that leverages advanced analytics, artificial intelligence, and workplace monitoring techniques. This system aims to proactively identify employees experiencing high levels of stress before it negatively impacts their performance, well-being, and overall organizational productivity. By integrating various data sources such as work performance metrics, attendance records, employee surveys, and physiological data (if ethically permissible), the system will generate predictive insights on stress levels. These insights will enable organizations to take timely remedial actions, such as workload adjustments, mental health support, or tailored wellness programs.

The system will function in three primary stages: data collection, stress prediction, and intervention recommendation. In the data collection phase, relevant data points, including work hours, deadlines, absenteeism, and self-reported stress levels, will be gathered from HR systems, wearable devices, and sentiment analysis of employee communications (e.g., emails, chats). The stress prediction phase will employ machine learning algorithms that analyze historical patterns to detect early indicators of burnout and anxiety. The model will continuously improve its accuracy through feedback mechanisms, ensuring reliable predictions. Finally, the intervention recommendation phase will suggest personalized remedial actions, such as flexible scheduling, wellness resources, or counseling sessions, to help employees manage stress effectively.

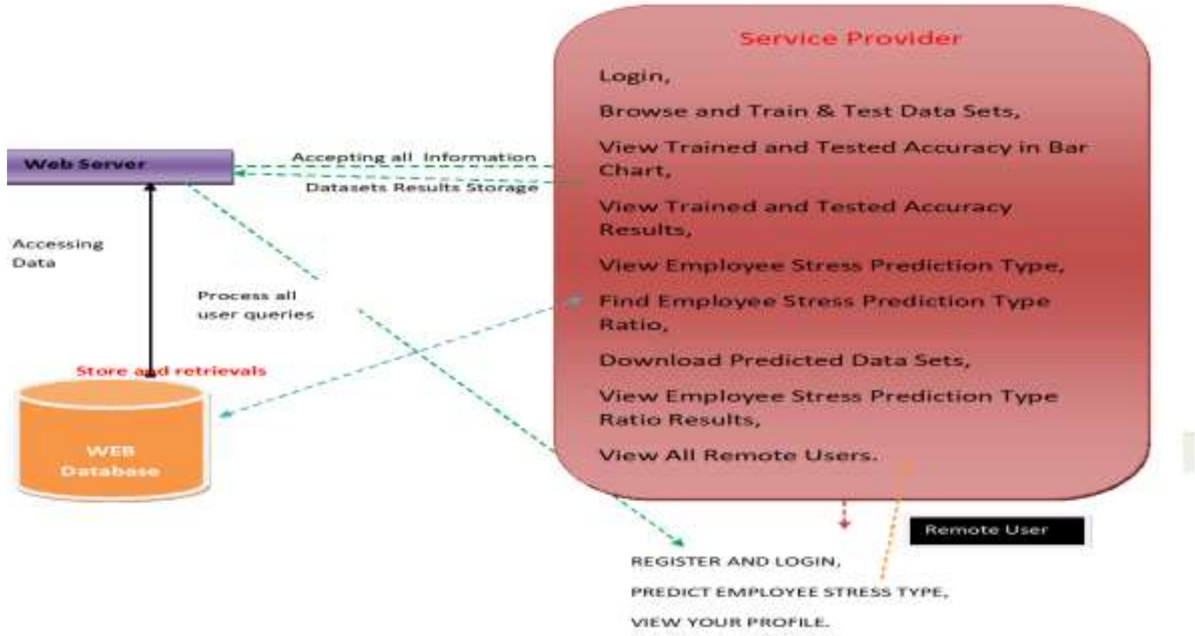
ADVANTAGES OF PROPOSED SYSTEM:

1) High accuracy

2) High efficiency

3.3 SYSTEM ARCHITECTURE

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper



3.4 SYSTEM ANALYSIS:

Workplace stress has been widely recognized as a critical issue affecting employee productivity, well-being, and overall organizational performance. Numerous studies have explored various approaches to predicting employee stress levels to facilitate early intervention and remedial actions. This literature survey provides an overview of key research contributions in this domain, focusing on stress indicators, predictive models, and preemptive remediation strategies.

Indicators and Causes of Workplace Stress

Research has identified multiple factors contributing to workplace stress, including excessive workload, job insecurity, lack of work-life balance, organizational culture, and interpersonal conflicts. Lazarus and Folkman (1984) introduced the transactional model of stress and coping, emphasizing how employees perceive and respond to workplace demands. Further studies highlight physiological indicators such as heart rate variability, cortisol levels, and sleep disturbances as biomarkers of stress (Sharma et al., 2020). Additionally, behavioral and psychological indicators such as absenteeism, reduced engagement, mood changes, and declining performance have been used to assess stress levels in employees (Kumar & Singh, 2021).

Machine learning and AI-based Predictive Models

Recent advancements in artificial intelligence (AI) and machine learning (ML) have enabled the development of predictive models for identifying stress among employees. Supervised learning techniques such as decision trees, support vector machines (SVM), and deep learning models have been employed to classify employees based on stress levels (Ghosh et al., 2022). Wearable sensor data, sentiment analysis of emails and chat messages, and facial emotion recognition have been leveraged to enhance the accuracy of predictions (Zhang et al., 2021). Furthermore, natural language processing (NLP) techniques have been used to analyze employee feedback and detect early signs of stress through linguistic patterns and sentiment shifts (Bhardwaj et al., 2023).

Preemptive Remediation Strategies

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Preemptive stress remediation strategies focus on mitigating workplace stress before it adversely affects employees. Research suggests that interventions such as mindfulness training, cognitive behavioral therapy (CBT), flexible work arrangements, and supportive leadership can significantly reduce stress levels (Cooper & Cartwright, 1997). Organizations have also started integrating AI-driven employee wellness platforms that offer personalized stress management recommendations based on real-time data (Lee et al., 2020). Additionally, early-warning systems that alert HR professionals about potential stress risks have been implemented to ensure timely interventions.

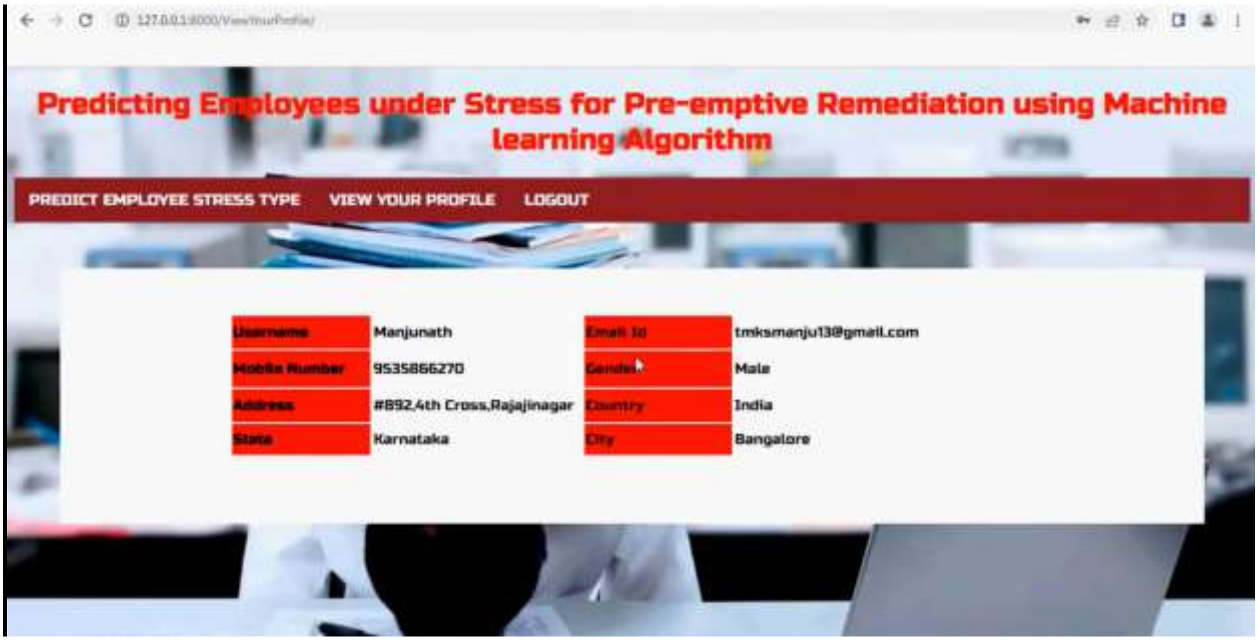
Challenges and Future Directions

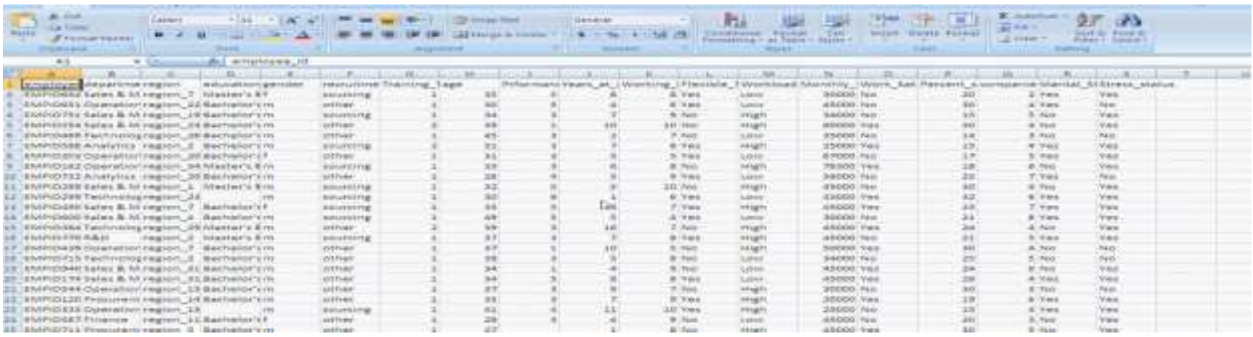
Despite the advancements in stress prediction and remediation, challenges remain in ensuring data privacy, ethical AI deployment, and model generalizability across diverse workplaces. Future research should focus on integrating multimodal data sources such as physiological, behavioral, and contextual factors for improved prediction accuracy. Furthermore, explainable AI (XAI) techniques should be employed to enhance trust and transparency in stress prediction models.

In conclusion, the prediction of employee stress using AI-driven approaches has gained significant traction in recent years. A combination of physiological, psychological, and behavioral indicators, coupled with advanced machine learning models, can facilitate early detection and preemptive interventions. Future research should aim to refine these models while addressing ethical and privacy concerns to ensure effective workplace stress management.

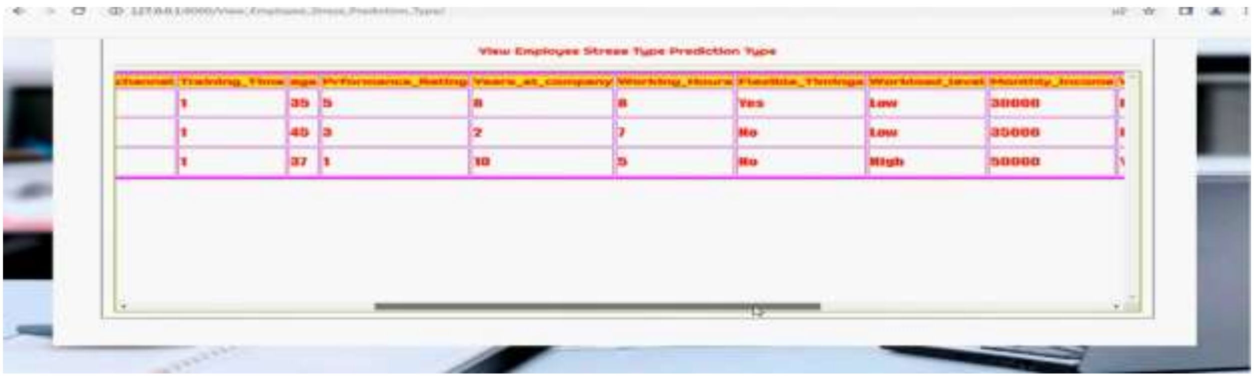
IV.SCREEN SHOTS

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper





International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper



V.CONCLUSION

To evaluate our model to achieve a better performance which is done by using XGB classifier. This is one of the best optimization technique and this is like a decision tree-based algorithm which adopts gradient boosting frame work technique for analysis and confusion matrix which tells us how many correct values are predicted by our model. XG Boost has tremendous predictive power and is about 10 times more durable than other gradient boosting techniques. It holds a variety of regularization which diminishes over fitting and enhances overall performance. Consequently, it is further recognized as the "regularized boosting" technique. Like it has true positive, true negative, false positive, false negative values. Used to evaluate the performance of the classification model.

VI.FUTURE SCOPE

The future of stress prediction and preemptive remediation in the workplace is poised for significant advancements with the integration of cutting-edge technologies. Artificial Intelligence (AI) and Machine Learning (ML) models are expected to become more sophisticated, enabling real-time analysis of employees' behavioral patterns, physiological signals, and work-related data. Wearable devices and biometric sensors will play a crucial role in capturing stress indicators such as heart rate variability, sleep patterns, and even voice tone analysis. As AI-powered predictive models improve, organizations will be able to detect early signs of stress more accurately and implement personalized interventions before productivity and well-being are adversely affected.

Additionally, the incorporation of Natural Language Processing (NLP) in workplace communication platforms can help identify stress markers in employees' messages, emails, and voice interactions. Future systems may use sentiment analysis and emotional AI to assess employee moods continuously, ensuring that stress-related concerns are addressed proactively. Organizations may also adopt AI-driven virtual assistants that provide employees with tailored coping mechanisms, mindfulness exercises, and workload balancing suggestions in real time.

Another promising avenue is the ethical and privacy-centric development of stress prediction frameworks. As data collection and analysis become more pervasive, ensuring compliance with data protection regulations such as GDPR and HIPAA will be critical. Companies must develop transparent policies that emphasize employee consent and data security, fostering trust and encouraging participation in stress management initiatives. Future research will focus on balancing predictive accuracy with ethical considerations to ensure that employees' mental well-being is prioritized without compromising privacy.

REFERENCES

[1] Shekhar Pandey, Supriya Muthuraman, Abhilash Shrivastava.The International Symposium on Intelligent Systems Technologies and Applications (2018), DOI: 10.1007/978-3-319- 68385-0_10.

[2] Ramachandran, R; Rajeev, D.C; Krishnan, S.G; Subathra.P. International Journal of Applied Engineering Research (2015), Research India Publications, Volume 10,Number 10, p.25433- 25448

[3] Ramin Zibaseresht: How to Respond to the Ongoing Pandemic Outbreak of the Coronavirus Disease (COVID-19) (WHO- World Health Organization) (2020), ISSN 2349- 8870.

[4] Chen, Tianqi; Guestrin, Carlos; "XG Boost: A scalable Tree Boosting System". Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, USA (2016). ACM. pp. 785-794.

[5] Jolliffe I.T. Principal Component Analysis, Series: Springer Series in Statistics, 2nd ed., Springer, NY, 2002,XXIX,487p.28 illus. ISBN 978-0-387-95442-4.

[6] Manuela Aparicio and Carlos J. Costa "Data Visualization". Communication Design Quarterly (2014). DOI: 10.1145/2721882.2721883.

[7] Ali.M., Alqahtani.A., Jones.M.W., Xie.X ."Clustering and classification for Time Series Data in Visual Analytics: A Survey IEEE Access 7,8930535, pp. 181314-181338.

[8] Mouubayedd.A, Injadat.M., Nassif, Lutfuyya, H.Shami, A E-learning: Challenges and Research Opportunities Using Machine Learning and Data Analytics (2018) IEEE Access 6,8417405. pp. 39117-39138.

[9] Kim., Soyata, T., Behnagh, R.F. Towards Emotionally Aware AI Smart Classroom: Current Issues and Directions for Engineering and Education (2018) IEEE pp. 5308-5331.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

[10] Priyonesi S.Madeh; EI-Diraby Tamer E. “Role of Data Analytics in Infrastructure Asset Management: Overcoming Data Size and Quality Problems”. Journal of Transprotation Engineering, Pavements (2020). DOI: 10.1061/JPEODX.0000175.

[11] Abhijeet Rawal, Sneha Mhatre. IOSR Journal of Bussiness and Management (IOSR-JBM), e-ISSN: 2278-487X, P-ISSN: 2319- 7668, PP 15-23.

[12] Janne Skakon, Karina Nielsen, Vihelm Borg, Jaime Gazman. An international Journal of work, Health and organizations, Volume 24,2010-Issue 2.

[13] K. S. Santosh and S. H. Bharathi, "Non-negative matrix factorization algorithms for blind source sepertion in speech recognition," 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, 2017, pp. 2242-2246, doi: 10.1109/RTEICT.2017.8256999.

[14] Rajendra Prasad P, N. Narayan, S. Gayathri and S. Ganna, “An Efficient E-Health Monitoring with Smart Dispensing System for Remote Areas”, 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, 2018, pp. 2120-2124. doi: 10.1109/RTEICT42901.2018.9012480.