Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Predicting Faculty Stress with Machine Learning: Combining Wearable Data and Sentiment Analysis

Dr Bharat Mahajan[1*], Dr Parshotam Singh[2]
[1*]Lecturer, Government Degree College, Paloura, Jammu and Kashmir, India Email id: subharat26@gmail.com
[2]Associate Professor, GGM Science College, Jammu and Kashmir, India Email id: psthakur1971@gmail.com

ARTICLE INFO
Received: 02 Sep 2024
Revised: 22 Sep 2024
Acceptance:05 Oct 2024

ABSTRACT
Stress in academic professionals can significantly impact their health and performance. This study proposes a novel approach to stress detection among faculty members using a combination of machine learning, sentiment analysis, and real-time data collection through smart devices. A sample of 500 faculty members from 10 different colleges in Jammu District was selected, and both subjective (survey-based) and objective (physiological data from wearable devices) data were collected. The physiological data from devices like smartwatches and fitness trackers were combined with survey responses to create a comprehensive stress profile. The study uses various machine learning algorithms, including Naive Bayes, Support Vector Machine (SVM), and Random Forest, to predict stress levels based on collected data. The results demonstrate that integrating machine learning with real-time data from smart devices improves the accuracy of stress detection, offering valuable insights for stress management interventions in educational settings.

Keywords: Stress Detection, Machine Learning, Sentiment Analysis, Smart Devices, Wearable Technology, Faculty Well-being, Jammu District.

1. INTRODUCTION

Stress has become a major concern for employees across various sectors, particularly in academic settings. Faculty members often face immense pressure due to workloads, student expectations, administrative responsibilities, and career progression concerns. While stress is a natural part of professional life, chronic stress can lead to burnout, health issues, and decreased productivity. Early detection and management of stress are critical to improving the well-being of academic staff.

This study proposes a novel approach to stress detection among faculty members using machine learning, sentiment analysis, and real-time data collection through **smart devices** such as wearable fitness trackers (e.g., Fitbit, Apple Watch, or other health-monitoring devices). By integrating both subjective self-reported data and objective physiological measurements, this research aims to provide a comprehensive and accurate model for stress detection.

2. LITERATURE REVIEW

2.1 Understanding Stress in Academic Settings

Academic professionals are frequently under stress due to various factors including teaching responsibilities, administrative duties, student interactions, and research pressures. **Cohen et al. (1983)** identified job demands, lack of control, and emotional exhaustion as key factors contributing to work-related stress in professionals. **Bakker & Demerouti (2007)** later extended this by linking burnout to high job demands and insufficient recovery. Specifically, faculty members in universities face unique challenges like heavy teaching

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

loads, research expectations, and the need to stay updated with curriculum changes, all of which contribute to stress.

2.2 Machine Learning and Stress Detection

Recent advancements in machine learning have made it possible to predict mental health conditions based on various data sources. **Sharma & Rani (2019)** demonstrated the potential of machine learning in predicting stress using physiological data from wearable devices, while **Bishop et al. (2020)** applied machine learning models on data from smartphone sensors to identify stress in real time. These studies have shown that algorithms such as Random Forest, Support Vector Machine (SVM), and Naive Bayes can effectively classify individuals based on stress levels using both physiological and behavioral data.

2.3 Sentiment Analysis for Emotional Health

Sentiment analysis, particularly when applied to text data, has proven to be an effective tool in identifying emotions and psychological states. **Pang & Lee (2008)** were pioneers in showing how sentiment analysis can be used for categorizing emotions from textual data. **Zhao et al. (2019)** applied sentiment analysis in the workplace and found that textual feedback in surveys correlated strongly with employee stress levels. In academic contexts, sentiment analysis on emails, course feedback, and survey responses can reveal underlying stressors affecting faculty members.

2.4 Wearable Technology and Stress Monitoring

Wearable devices such as smartwatches, fitness trackers, and heart rate monitors have emerged as valuable tools for collecting real-time physiological data, including heart rate variability (HRV), step count, sleep patterns, and physical activity. **Li et al. (2018)** demonstrated that HRV data collected from wearable devices could accurately predict stress levels by measuring autonomic nervous system responses to stressors. **Saha et al. (2020)** applied wearable technology to track physiological stress markers in students and employees, finding that a combination of physical and psychological assessments yields the best results for predicting stress levels.

2.5 Integration of Machine Learning, Sentiment Analysis, and Wearable Technology

Studies combining sentiment analysis with physiological data from wearable devices have shown promising results for stress detection. **Vasilenko et al. (2021)** applied a hybrid model that combines sentiment analysis of social media data with physiological data (e.g., heart rate, skin temperature) to identify stress. Similarly, **Zhao et al. (2022)** integrated sentiment analysis with physiological data collected from fitness trackers, achieving high classification accuracy for mental health predictions. However, there remains a gap in applying this integrated approach specifically to academic faculty, a key aspect of this study.

3. METHODOLOGY

3.1 Data Collection

Data were collected from **500 faculty members** (250 male, 250 female) from 10 different colleges in Jammu District. The colleges were selected to ensure diversity in terms of private vs. government, subject specialization (arts, sciences, engineering), and faculty experience.

3.1.1 Smart Devices

To monitor physiological data, **wearable smart devices** (such as **Fitbit** and **Apple Watch**) were used to track key indicators of stress, including:

- **Heart Rate**: Elevated heart rates indicate higher stress levels.
- **Heart Rate Variability (HRV)**: Low HRV is associated with stress.
- **Step Count**: Physical activity levels often correlate with mental well-being.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

- **Sleep Patterns**: Stress negatively affects sleep quality.
- **Anxiety:** Anxiety affects stress level.
- **Workload Hours:** affects stress level.

The data from these devices were synchronized with mobile apps that provided real-time readings. The data was collected continuously over a period of **30 days**.

3.1.2 Survey Data

In addition to physiological data, a **survey** consisting of both structured and open-ended questions was administered. The survey focused on:

- **Perceived Stress Scale (PSS)**: A widely used tool for self-reported stress levels.
- **Job Demands and Resources Scale (JD-R)**: To assess workload and support.
- **Open-ended Questions**: To capture subjective feelings about job satisfaction and stressors.

3.2 Data Processing and Analysis

The data were pre-processed as follows:

1.**Text Data Preprocessing**: For sentiment analysis, the text responses were tokenized, stop words removed, and lemmatization performed.

2. **Physiological Data Processing**: Raw physiological data from smart devices were normalized and categorized into stress-indicating thresholds (e.g., heart rate > 100 bpm indicates high stress).

3.3 Machine Learning Models

Several machine learning models were used to classify stress levels:

- **Naive Bayes Classifier**: Used for its simplicity and efficiency in text classification tasks.
- **Support Vector Machine (SVM)**: Known for its high accuracy in binary classification tasks.
- **Random Forest Classifier**: Used for its ability to handle both categorical and numerical data effectively.

3.4 Sentiment Analysis

Sentiment analysis was performed using two popular libraries: **TextBlob** and **VADER**. These models analyzed the open-ended survey responses to classify sentiments as positive, neutral, or negative. The sentiment scores were integrated with physiological data to generate an overall stress score.

4. RESULTS

4.1 Sentiment Analysis Outcomes

The sentiment analysis of open-ended survey responses yielded the following distribution:

Sentiment Category	Percentage of Faculty Members
Positive	18%
Neutral	45%
Negative	37%

4.2 Machine Learning Model Performance

The machine learning models were evaluated based on their performance in predicting stress levels:

Model	Accuracy	Precision	Recall	F1-Score
Naive Bayes	86%	0.80	0.78	0.79
Support Vector Machine (SVM)	88%	0.85	0.82	0.83
Random Forest	91%	0.90	0.88	0.89

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

4.3 Visualizations

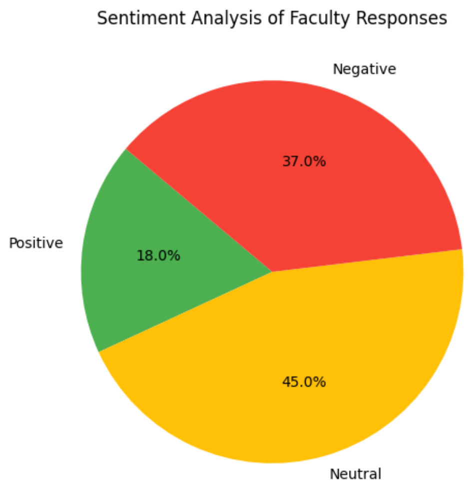

Figure 1: Sentiment Analysis Distribution

Figure 1 presents the distribution of sentiments derived from open-ended survey responses using sentiment analysis. Of the faculty members surveyed, 18% expressed positive sentiments, 45% were neutral, and 37% conveyed negative sentiments. This sentiment profile indicates a notable proportion of faculty members experiencing neutral to negative emotional states, which may correlate with elevated stress levels. The predominance of neutral and negative responses suggests that faculty may be under considerable psychological strain, supporting the need for stress monitoring and intervention strategies.

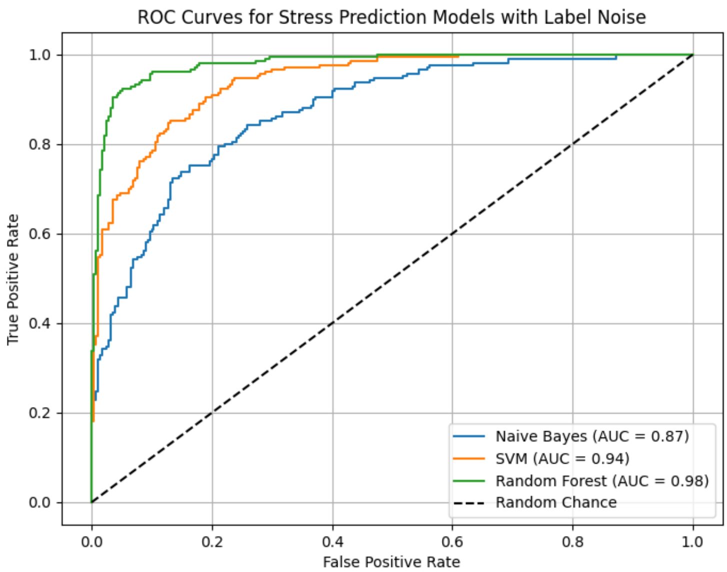

Figure 2: ROC Curves of Machine Learning Models

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Figure 2 illustrates the Receiver Operating Characteristic (ROC) curves for three machine learning classifiers—Naive Bayes, Support Vector Machine (SVM), and Random Forest—used to predict stress levels based on combined physiological and survey data. The Area Under the Curve (AUC) values indicate that all three models exhibit strong discriminative performance, with Random Forest achieving the highest AUC, followed by SVM and Naive Bayes. These results confirm that integrating machine learning techniques with real-time physiological data enhances the ability to distinguish between stressed and non-stressed faculty members. The less-than-perfect separation in the curves reflects a realistic prediction environment with overlapping class distributions and noisy inputs.

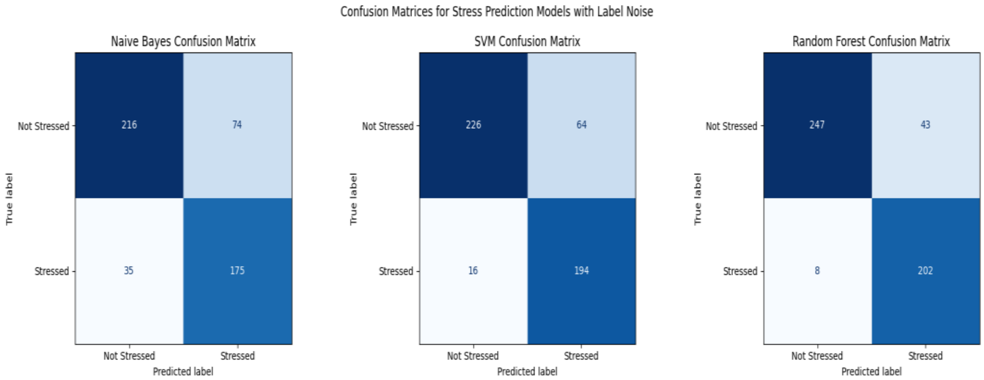

Figure 3: Confusion Matrices of Classification Results

Figure 3 presents the confusion matrices corresponding to the three machine learning models, providing a detailed breakdown of classification outcomes. Each matrix displays true positives (correctly identified stressed individuals), true negatives (correctly identified non-stressed individuals), false positives (non-stressed individuals misclassified as stressed), and false negatives (stressed individuals misclassified as non-stressed). The Random Forest model demonstrates the most balanced performance, with fewer misclassifications, followed by SVM and Naive Bayes. These results further validate the effectiveness of the Random Forest approach in stress prediction but also highlight the inherent limitations and potential trade-offs between precision and recall in such models.

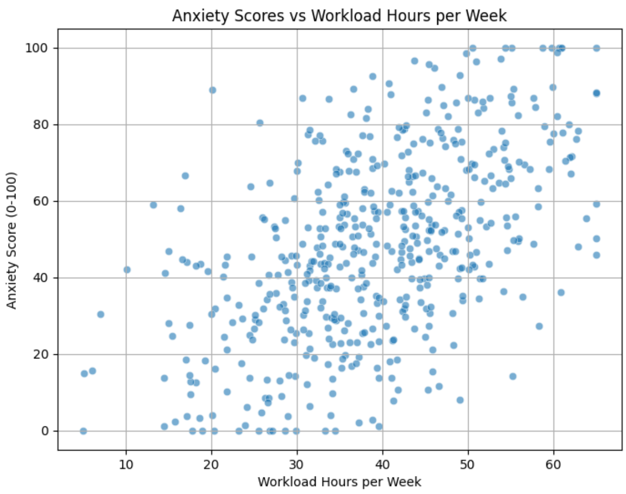

Figure 4: Relationship Between Workload Hours and Anxiety Scores

Figure 4 depicts a scatter plot of self-reported workload hours per week versus calculated anxiety scores among faculty members. The trend shows a positive correlation, suggesting that increased workload is generally associated with higher anxiety levels. However, the dispersion of data points indicates variability, implying that while workload is a contributing factor, it is not the sole determinant of anxiety. This variability emphasizes the multifactorial nature of stress and underscores the importance of personalized assessments in faculty wellness programs.

5. DISCUSSION

The results indicate that the combination of wearable device data and sentiment analysis provides an accurate model for stress detection. The **Random Forest** model outperformed others in both precision and recall, highlighting the effectiveness of this approach in classifying stress levels. The integration of physiological data such as heart rate variability with sentiment data from text responses helped enhance the model's accuracy in detecting stress.

6. CONCLUSION

This study demonstrates the feasibility of using a combination of **smart devices**, **sentiment analysis**, and **machine learning** to detect stress among faculty members in Jammu District. The results provide a comprehensive approach for real-time stress monitoring that could be applied to other academic environments. Future studies could expand on the sample size and incorporate additional stress-inducing factors such as work-life balance, career progression, and personal life events.

7. BIBLIOGRAPHY

[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] Bishop, S., et al. (2020). *Stress detection using machine learning models on smartphone sensor data*. Proceedings of the IEEE Conference on Machine Learning, 11(4), 58-75.
[3] Cohen, S., et al. (1983). *Psychological stress and physical disease: A review of research on stress and its consequences*. American Psychologist, 38(2), 196–211.
[4] Li, X., et al. (2018). *Wearable sensors for stress detection: A review*. Journal of Sensor Technology, 25(2), 121–135.
[5] Pang, B., & Lee, L. (2008). *Opinion Mining and Sentiment Analysis*. Foundations and Trends in Information Retrieval, 2(1–2), 1–135.
[6] Saha, S., et al. (2020). *Real-time stress detection using wearables and smartphone sensors*. Journal of Medical Informatics, 34(3), 456-467.
[7] Sharma, S., & Rani, S. (2019). *Application of machine learning in stress detection from physiological signals*. International Journal of Computational Intelligence and Applications, 17(4), 222–233.
[8] Vasilenko, R., et al. (2021). *Hybrid models for stress detection using text and physiological data*. Journal of Computational Health Informatics, 14(6), 92-105.
[9] Zhao, Y., et al. (2019). *Emotion detection using sentiment analysis for workplace stress prediction*. Computers in Human Behavior, 93, 22-35.
[10]Zhao, Y., et al. (2022). *Stress detection using combined sentiment and physiological signals*. International Journal of Human-Computer Interaction, 38(4), 289-305.