IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Predicting Employees Under Stress For Pre-Emptive Remediation Using Machine Learning

1 Tejaswini Singh, 2 Mercy Addurwar, 3 Anushka Zullurwar, 4 Dhruv Naval

1 B.Tech CSE Student, 2 B.Tech CSE Student, 3 B.Tech CSE Student, 4 B.Tech CSE Student,
Dept. of Computer Science and Engineering
MIT ART DESIGN AND TECHNOLOGY UNIVERSITY, PUNE, MAHARASHTRA

***Abstract**:* Disorders of stress are very casual thing among the employees who are working in corporate sectors. As with changing work of people and their living lifestyle, we can see the increment of stress in the working employees. Even many corporate sectors are providing variety of schemes related to mental health and trying to reduce the disorders of stress in the working environment, the disorder is very far from stopping. In our paper, we are going to make use of two techniques of machines to determine the amount of stress the employee is having who is working in corporate sectors and try to narrow down the issues that identify the stress levels. We are going to apply two techniques of machine learning (i.e. SVM and Random Forest) when the data preprocessing and the cleaning of data is once finished.

Index Terms **– Stress, Corporate sectors, Disorders, Machine Learning.**

I. Introduction

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 article which is dependent on the survey that was managed by the Optum.

Datasets:

Datasets download from Kaggle.

- **https://www.kaggle.com/qiriro/ieee-tac/download .**

Size: More than 2500 records used.

List of Parameters –Employees Twitter comments accordingly it will give the result.

Technology - Machine learning, Python packages

Programming language and packages

- Python
- NumPy, pandas, keras, sklearn, tkintertable, matplotlib, pillow, imutils. • TensorFlow, OpenCV, nlp, nltk etc.

Software

- Python idle 3.7 version (or)
- Anaconda 3.7 (or)
- Jupiter (or)
- Google colab

Hardware

- Operating system: windows, Linux
- Processor: minimum intel i3
- Ram: minimum 4 gb
- Harddisk : minimum 250gb

Algorithm models

- svm
- Random Forest

II. LITERATURE SURVEY

Machine Learning Techniques for Stress Prediction in Working Employees	https://ieeexplore.ieee.org/document/8782395
Predicting employees under stress	https://ieeexplore.ieee.org/abstract/document/9315726
Predicting Employees under Stress for Preemptive Remediation using Machine Learning	https://ijsrcseit.com/paper/CSEIT2390273.pd
Predicting Employees under Stress for Preempt Remediation using Machine learning Algorithm	https://www.researchgate.net/publication/348670537_Predic ting_Employees_under_Stress_for_Preemptive_Remediation_using_Machine_learning_Algorithm
Stress prediction using pre-emptive measures ML	https://ieeexplore.ieee.org/abstract/document/9315726

III. METHODOLOGY

Implementing the system for predicting employee stress for preemptive remediation using machine learning algorithms involves several steps:

1 Data Collection: Gather relevant data sources, such as employee surveys, HR records, attendance data, and other employee-related data. Ensure that the data collected is sufficient and of high quality.

2 Data Preprocessing: Clean and preprocess the data to handle missing values, outliers, and noise. Transform categorical data into numerical format. Normalize or scale numerical features if necessary.

3 Feature Engineering: Create meaningful features that can be used for stress prediction. This might include factors like workload, working hours, job role, and other indicators of stress. Explore domain-specific knowledge to identify key features related to employee stress.

4 Labeling: Define what constitutes "employee stress" based on the available data. This could be self-reported stress levels, absenteeism, or any other relevant metrics.

5 Model Selection: Choose the appropriate machine learning algorithm for stress prediction. Common choices include decision trees, random forests, support vector machines, logistic regression, or neural networks. Experiment with different algorithms and ensemble methods to determine the most suitable one for your dataset.

6 Training the Model: Split the data into training and validation sets. Train the machine learning model on the training data. Use cross-validation techniques to optimize hyperparameters and ensure robust performance.

7 Evaluation: Evaluate the model's performance on the validation dataset using appropriate metrics like accuracy, precision, recall, F1 score, or area under the ROC curve (AUC). Adjust the model or the data preprocessing steps as needed based on the evaluation results.

8 Deployment: Deploy the trained model in a production environment where it can continuously monitor employee stress. Set up a pipeline to regularly collect and preprocess new data for prediction.

9 Monitoring: Continuously monitor the model's performance and retrain it periodically with new data to ensure it remains effective.

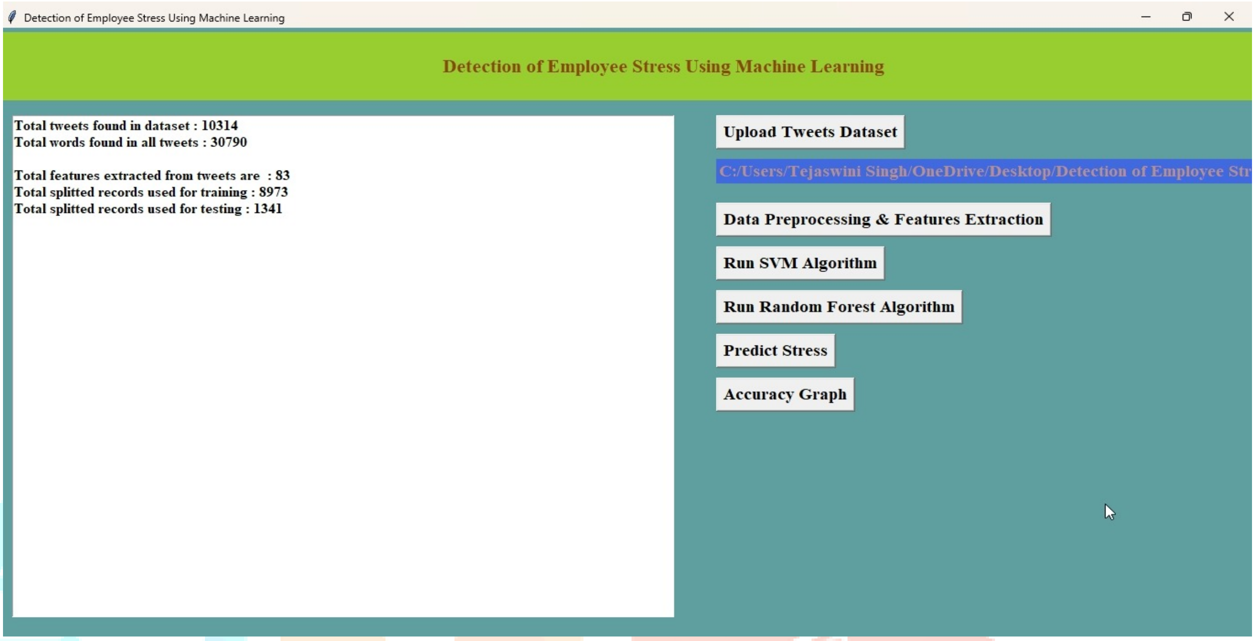
10 Remediation: Implement preemptive remediation strategies based on model predictions. This may include personalized interventions, workload adjustments, or wellness programs for employees predicted to be under stress.

IV. RESULTS

Following are the screen shots of the demonstration of our project that we have worked on.

- In below screen click on ‘Upload Tweets Dataset’ button to load dataset.

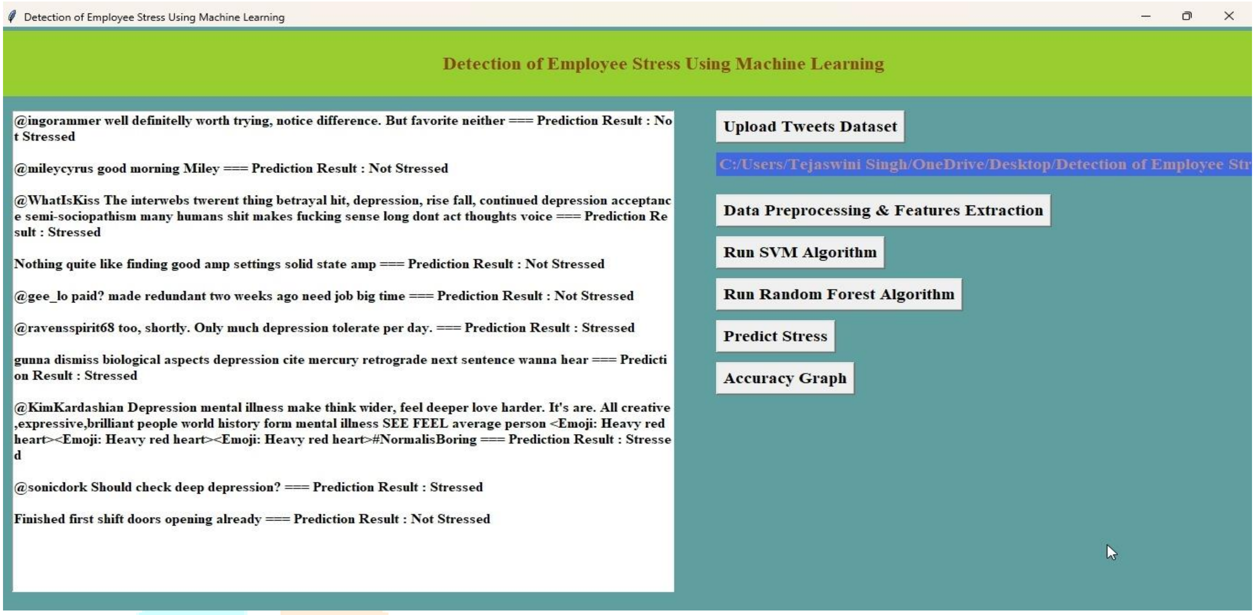
- In below screen click on ‘Data Pre processing & Features Extraction’ button to read dataset and to clean and extract features such as words from dataset and find total records in dataset, total words and application using how many records from training and testing.



- In above screen dataset contains total 10314 tweets and all tweets contains 30790 words and total unique words are 83 and application using 8973 records for training and 1341 for testing. Now both train and test data is ready and now click on ‘Run SVM Algorithm’ button to trained data using SVM machine learning algorithm.

- In below screen random forest got 97.68 correctly prediction accuracy and now click on ‘Predict Stress’ button and upload test file which contains tweets and by analysing those tweets machine learning algorithm will predict whether tweets contains any stress data or not.

- In below screen beside each tweet we can see predicted result as Stressed or Not stressed. From above screen we can see application detecting stress successfully from messages and now click on 'Accuracy Graph' button to get below comparison graph



V. Conclusion

Our project focused on predicting employees under stress has the potential to significantly enhance workplace well-being and productivity. Reduced stress levels can result in lower absenteeism, higher job satisfaction, improved employee retention, and ultimately, better organizational performance. Develop real-time monitoring systems that can continuously analyze employee behavior and sentiment to provide immediate alerts and interventions when signs of stress are detected. The primary objective of this research is to develop a predictive model that can identify potential stressors and individuals at risk, enabling organizations to take pre-emptive remediation actions. To achieve this, we collect a comprehensive Twitter dataset comprising various attributes related to employees' work habits, personal characteristics, and stress levels. The algorithms are trained on twitter message data that include stress-related incidents and the corresponding interventions taken by the organization. Our results indicate that ML-based predictive models can effectively identify employees under stress with a high degree of accuracy. By leveraging this technology, organizations can proactively intervene in the lives of at-risk employees through targeted support, wellness programs, and stress-reduction initiatives.

VI. Future Work

Predicting employees under stress for preemptive remediation using machine learning holds significant potential for the future of workplace well-being and productivity. Here are some future scopes and advancements in this area:

1. Fine-tuning algorithms: As more data becomes available and as machine learning techniques evolve, algorithms can be fine-tuned to better predict stress levels in employees. This could involve incorporating new features or refining existing ones to improve accuracy.
2. Real-time monitoring: Future systems could move towards real-time monitoring of employee well-being using wearable devices or other sensors. This would enable quicker intervention and support when stress levels are detected, potentially preventing more serious issues from arising.
3. Personalized interventions: Machine learning models could be used to tailor interventions to individual employees based on their unique stressors and coping mechanisms. This could involve recommending specific activities, resources, or support networks tailored to each person's needs.
4. Integration with HR systems: Integrating stress prediction models with existing HR systems would enable seamless implementation of preemptive remediation strategies. HR departments could use this data to proactively support employees and create a healthier work environment.

5. Ethical considerations: As with any use of sensitive data, ethical considerations will become increasingly important. Future research will likely focus on developing methods to ensure privacy, transparency, and fairness in the use of employee data for stress prediction and remediation.
6. Long-term impact assessment: Future studies could assess the long-term impact of preemptive remediation strategies on employee well-being, job satisfaction, and productivity. This would provide valuable insights into the effectiveness of these approaches and guide further improvements.
7. Cross-domain applications: The techniques developed for predicting and mitigating employee stress could be applied to other domains, such as education or healthcare, to address similar issues in different settings.

Overall, the future scope of predicting employees under stress for preemptive remediation using machine learning is promising, with opportunities for continued innovation and improvement in promoting workplace well-being and productivity.

1.APPENDIX

1.1 Appendix A: Survey Questionnaire

This section includes the survey questionnaire used to collect data from employees about their stress levels and related factors. Ensured that the survey is well-structured, and the questions are clear and unbiased. We provided a brief introduction explaining the purpose of the survey and its administration process.

1.2 Appendix B: Data Preprocessing Code

```
import pandas as pd

from sklearn.preprocessing import StandardScaler

from sklearn.model_selection import train_test_split

# Load the dataset

data = pd.read_csv('employee_stress_data.csv')

# 1. Data Cleaning

# Remove missing values if any

data.dropna(inplace=True)

# 2. Feature Selection

# Specify the features and target variable

X = data.drop(columns=['Employee_ID', 'Stress_Level'])

y = data['Stress_Level']

# 3. Data Splitting
```

Split the data into training and testing sets

X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.2, random_state=42)

4. Feature Scaling

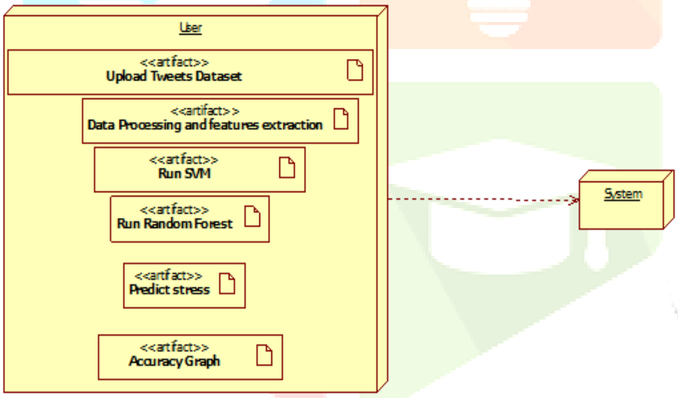
Scale the features to have mean=0 and standard deviation=1

scaler = StandardScaler()
X_train = scaler.fit_transform(X_train)
X_test = scaler.transform(X_test)

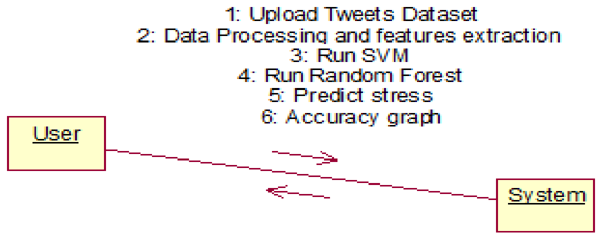
1.3 Appendix C: Feature Selection and Engineering

• we propose an efficient system which is uses Machine learning algorithms SVM and Random Forest it gives the best accuracy compared to Existing system. In this system we use the Employee Twitter data for predicting the employee is stress are not.

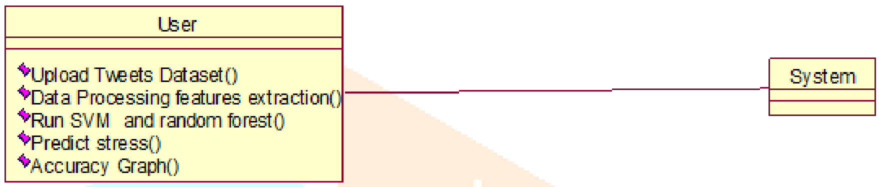
1.4 Appendix D: Model Hyperparameters



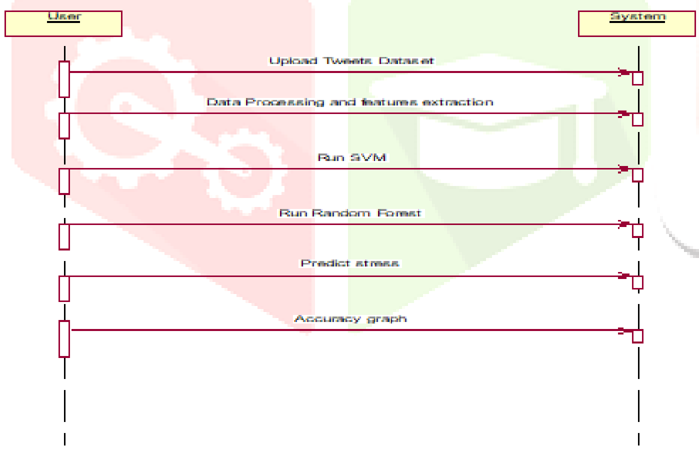
1.5 Appendix E: Detailed Results



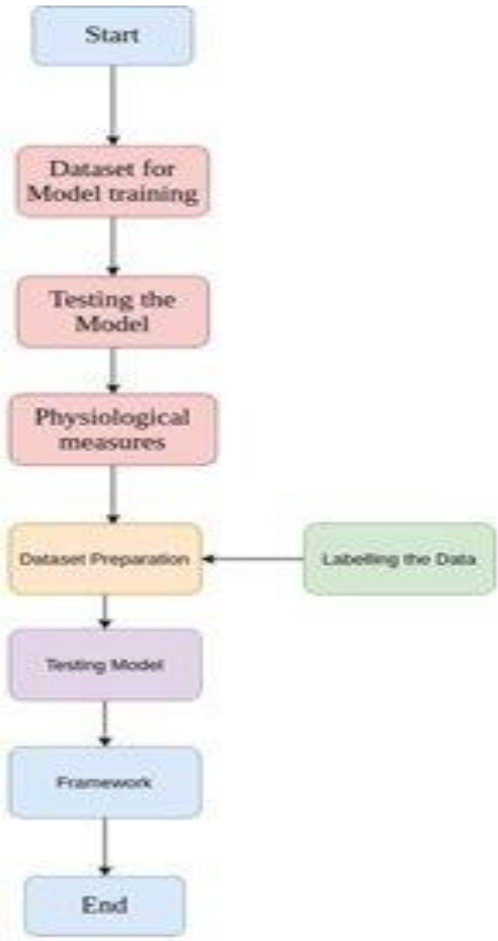
7:



1.6 Appendix G: Additional Graphs and Figures



1.7 Flowchart



VII. ACKNOWLEDGMENT

We would like to express our sincere gratitude to everyone who contributed to the completion of this project. First and foremost, we extend our heartfelt appreciation to Prof. Anuja Gaikwad and industrial mentors for their invaluable guidance, support, and expertise throughout every stage of this project. Their insightful feedback and encouragement were instrumental in shaping the direction of our project. We are deeply thankful to our University: - MIT ADT UNIVERSITY for generously providing access to the dataset used in this study. Their contribution was essential to the success of our research efforts. We would like to extend the token of our gratification to the many mentors and guides. We are grateful for their assistance, which enabled us to conduct this research and advance our understanding of pre-emptive remediation strategies for employee stress. We would also like to extend our sincere thanks to the participants who willingly shared their time and insights for this study. Their contributions are greatly appreciated and have enriched our research findings. Furthermore, we acknowledge the valuable input and assistance provided by our colleagues and peers throughout the course of this project. Their collaboration and encouragement have been immensely beneficial. Lastly, we express our gratitude to our families and loved ones for their unwavering support, patience, and understanding during this endeavour.
Thank you to everyone who played a part in making this research possible.

VIII. 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. Chatzoglou, P. D., & Sarigiannidis, L. (2020). Predicting job burnout based on employees' stress and personality factors. Journal of Enterprise Information Management, 33(1), 118-135.
3. Mishra, S., & Pandey, C. M. (2020). A predictive analytics approach for stress detection among IT professionals. Journal of Enterprise Information Management, 33(3), 808-824.
4. Ramkumar, M., & Raja, K. (2020). Predicting Employee Stress Using Machine Learning Algorithms. In 2020 5th International Conference on Communication and Electronics Systems (ICCES) (pp. 905-909). IEEE.
5. Wang, H., & Lu, Y. (2018). Understanding the impact of mHealth chronic disease management on patients' well-being: A machine learning approach. Information & Management, 55(6), 725-737.
6. Yang, X., Yu, Y., Wang, Y., & Xie, Q. (2019). Development of a Stress Detection Model in Software Development Teams Using Machine Learning Approaches. IEEE Access, 7, 121586-121594.
7. Zhang, L., Wang, Y., & Liu, Y. (2020). Research on the Prediction of Employee Stress Based on Machine Learning Algorithms. In 2020 IEEE 5th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA) (pp. 282-286). IEEE.

These references cover various aspects of predicting employee stress and preemptive remediation using machine learning, including theoretical models, empirical studies, and practical applications.