Employee Stress Management System Using Machine Learning

Dr. P. Vijaya Bharati
Associate Professor
Department of Computer Science and Engineering
Vignan's Institute of Engineering for Women
Visakhapatnam, India

Chaitanya Lakshmi Chava
Department of Computer Science and Engineering
Vignan's Institute of Engineering for Women
Visakhapatnam, India

Gandi Anusha
Department of Computer Science and Engineering
Vignan's Institute of Engineering for Women
Visakhapatnam, India

A Vandana Satya
Department of Computer Science and Engineering
Vignan's Institute of Engineering for Women
Visakhapatnam, India

Alajangi Dharani
Department of Computer Science and Engineering
Vignan's Institute of Engineering for Women
Visakhapatnam, India

Abstract- The aspiration of the "Employee Stress Management System" is to identify the employees under stress within companies of various work environments and embedded remote work adaptations to help raise the balance in their work, life, and health. Since the COVID-19 epidemic, most companies transformed their working styles into unusual spaces such as working from home leaving employees with ambiguity and stress in managing their working goals. So, it is need of the hour for the company executives to undertake this prediction as assistance to conduct appropriate remediation to help employees balance their work and manage performance outcomes. Thus, this project is pivotal in the work-life which uses machine learning algorithms to analyze the database and to perform prediction analysis in determining stressed employees. The main framework of the project relies on python with provided Graphical User Interface including visual graphs and heatmap for scrutiny by the company's management along with prediction results.

Keywords – Data Science, Random Forest Classifier, Stress Management, Machine Learning.

I. INTRODUCTION

Stress has a constant and continuous pattern in the workspace of every company as a result of boundaries drawn for the representation of employees' work including deadlines, on-time deployments, etc. As this not only affects the employee's professional life but also interrupts his personal life which in turn deviates from their performance. In order to help employees sustain the inevitably changing work environments within a company, the management of the company is expected to take pre-emptive measures to support and provide employees with stress-free work life. This approach needs a real-time analysis of data that can be used by the executives to take actions to balance the workforce by conducting stress remediation activities as needed.

II. PROPOSED ALGORITHM

A. Working

To perform future prediction of Employee Stress Management System, a machine learning classification algorithm named RFC (Random Forest Classifier) model is used. This model is applied to the employee dataset holding several unique features and records obtained from the GitHub website. Python's Tkinter is used to provide users with a Graphical user interface that also provides Visual Graphs and Feature's Heatmap displayed on features of employee data with the help of python's Seaborn.
Random Forest is a classification model that on fitting a dataset performs bagging where a set of decision trees (DT) are created randomly from the training data and then it scrutinizes the votes of each tree to consider the majority as the final prediction.

B. RFC's future predictions: To predict the stressed employees within a company, the system relies on RFC's predictions.

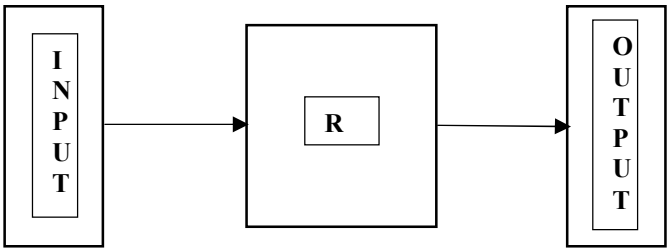

Figure 1: Applying RFC to the dataset for the future prediction of stress in employees

The working model of the Employee Stress Management System involves steps beginning from the data pre-processing to the resulting interface that allows users to perform predictions on new test datasets by fitting into the model.

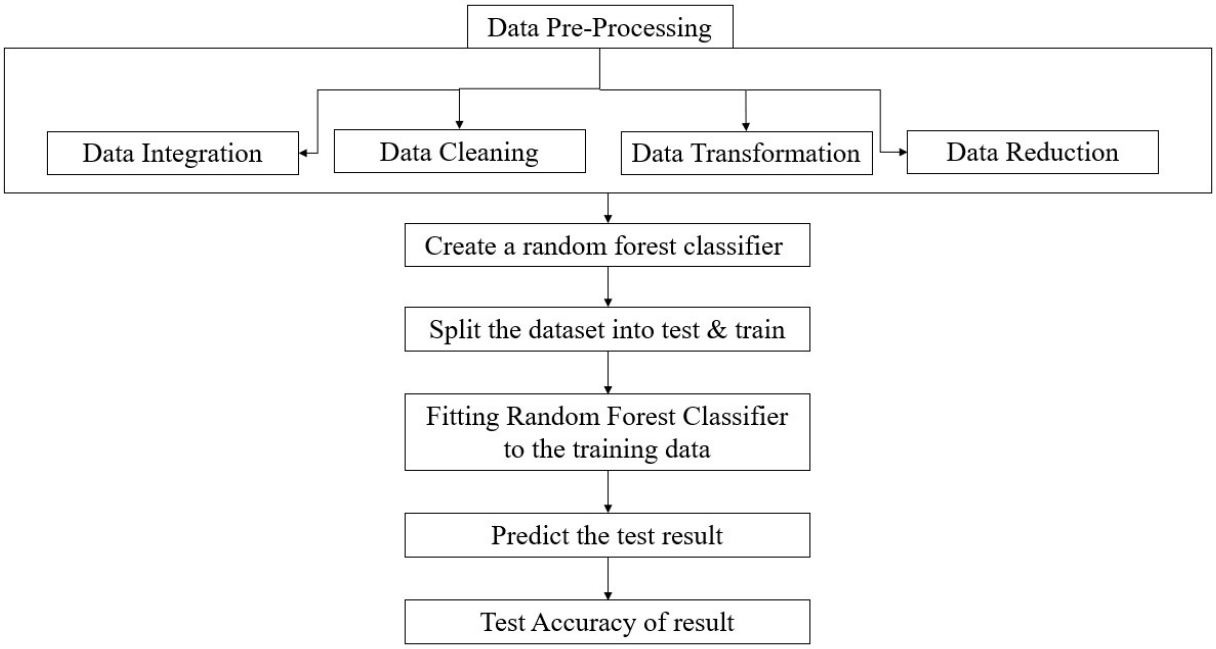

Figure 2: Working model of Employee Stress Management System

C. Dataset

The dataset used in this project is obtained from the GitHub with a size of 16000 records and 30 unique employee features of a company with varied attributes to support the model.

III. EXPERIMENT AND RESULT

The Stress Management System is provided with a GUI Interface which allows the operations such as the addition of new employee details, and verification if he is a stress victim with the prediction analysis of the entire dataset. The feature selection is also used in this system to help users identify the vital features that affect an employee in becoming a stress victim so as to provide support in balancing professional and personal lives. The users can also view heatmap and graphs vividly on unique employee features. The Spyder Anaconda is used to experiment with the project and is run on a PC for the experiment which is equipped with a Personal laptop embedded with an Intel Core i5 8th Gen processor along with 1TB HDD, 512GB SSD, and 8GB RAM.

This project predicts the employees who are becoming the stress victims to help take pre-emptive measures of stress remediation.

Figure 3. Employee Stress Analysis Main Page

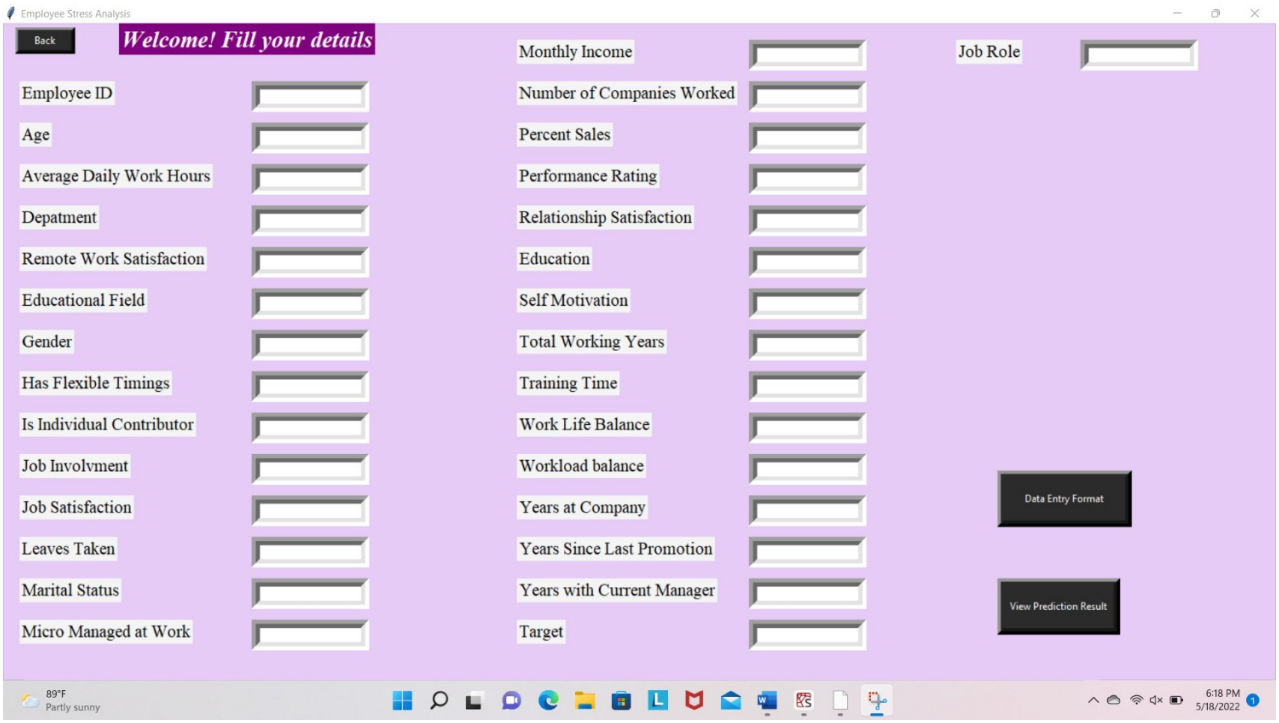

Figure 4. Page to Add and Predict new employee data as a single record

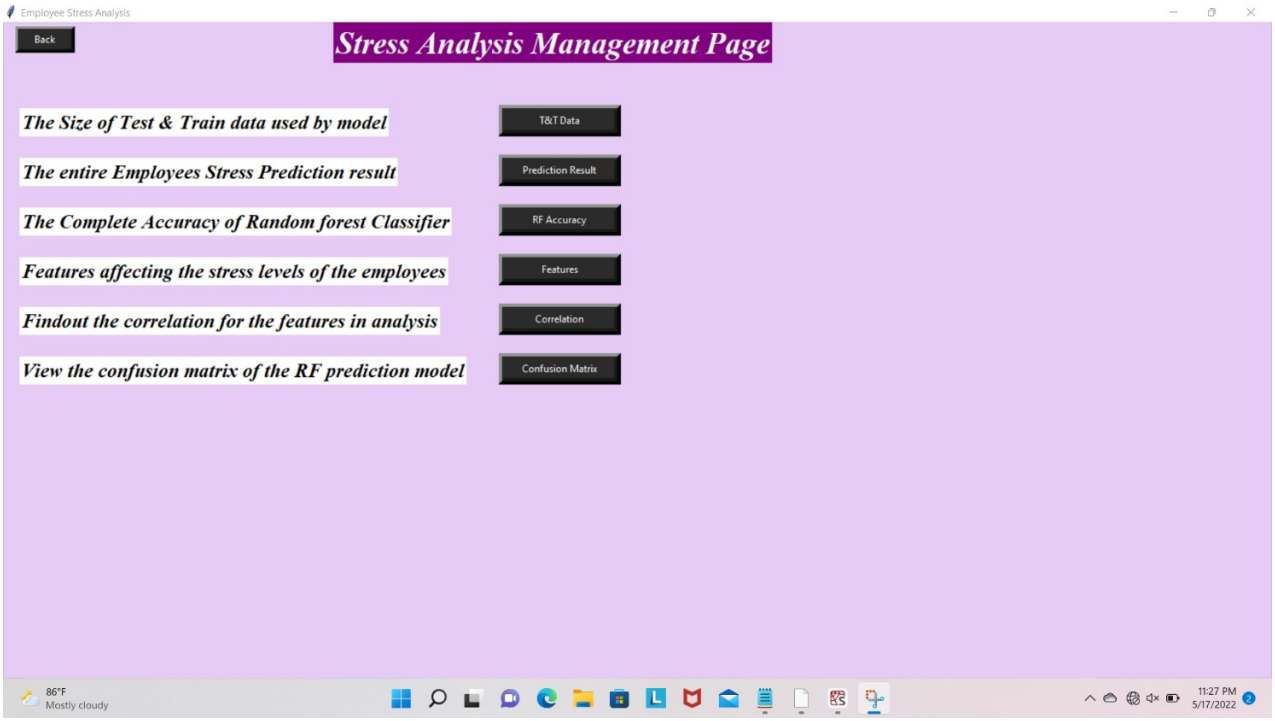

Figure 5. Stress Analysis Management Page to verify several RFC functions such as Prediction

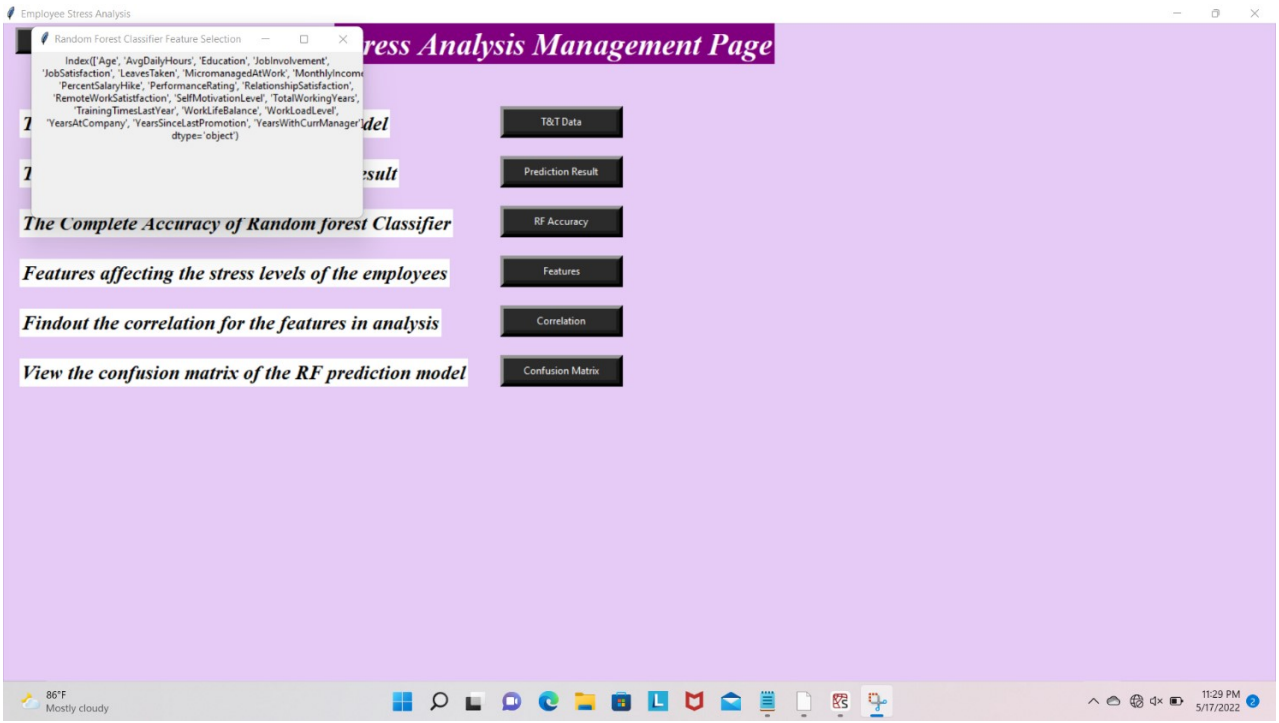

Figure 6. Display of Features affecting the stress levels of an employee

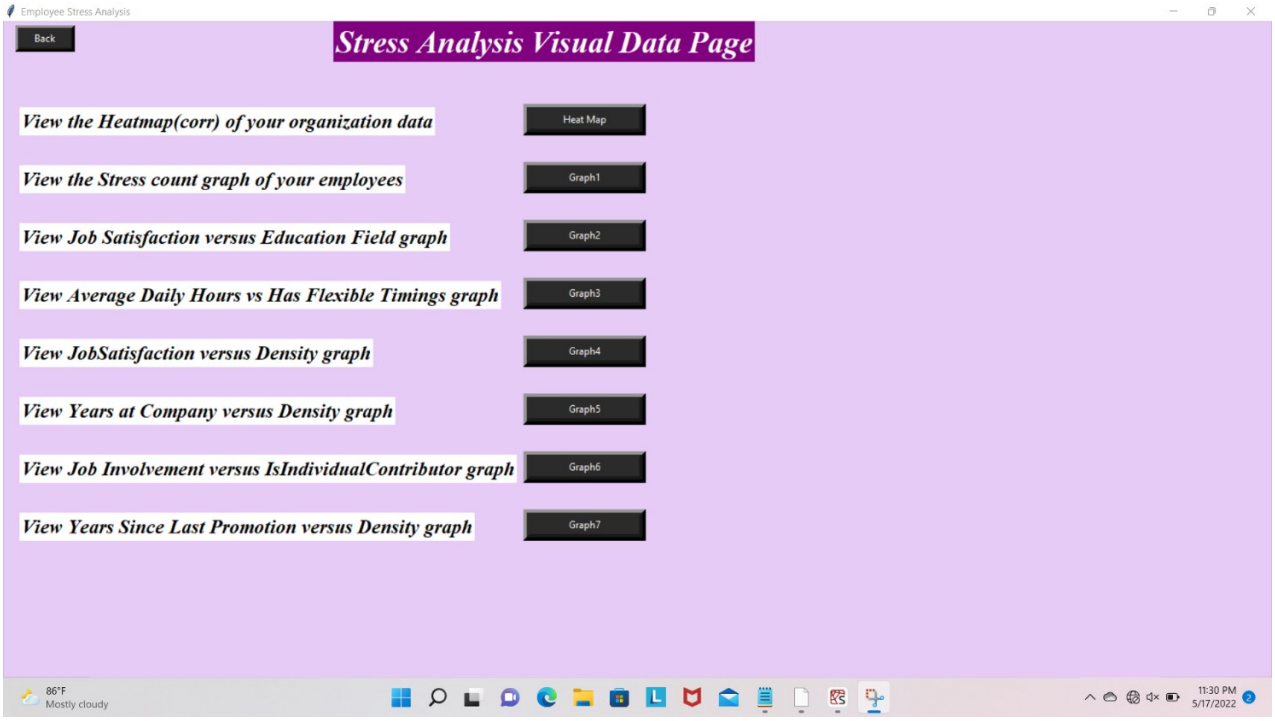

Figure 7. Page representing the choice of visual data (Graphs and Heatmap)

Figure 8. Heatmap representing correlation of features against features

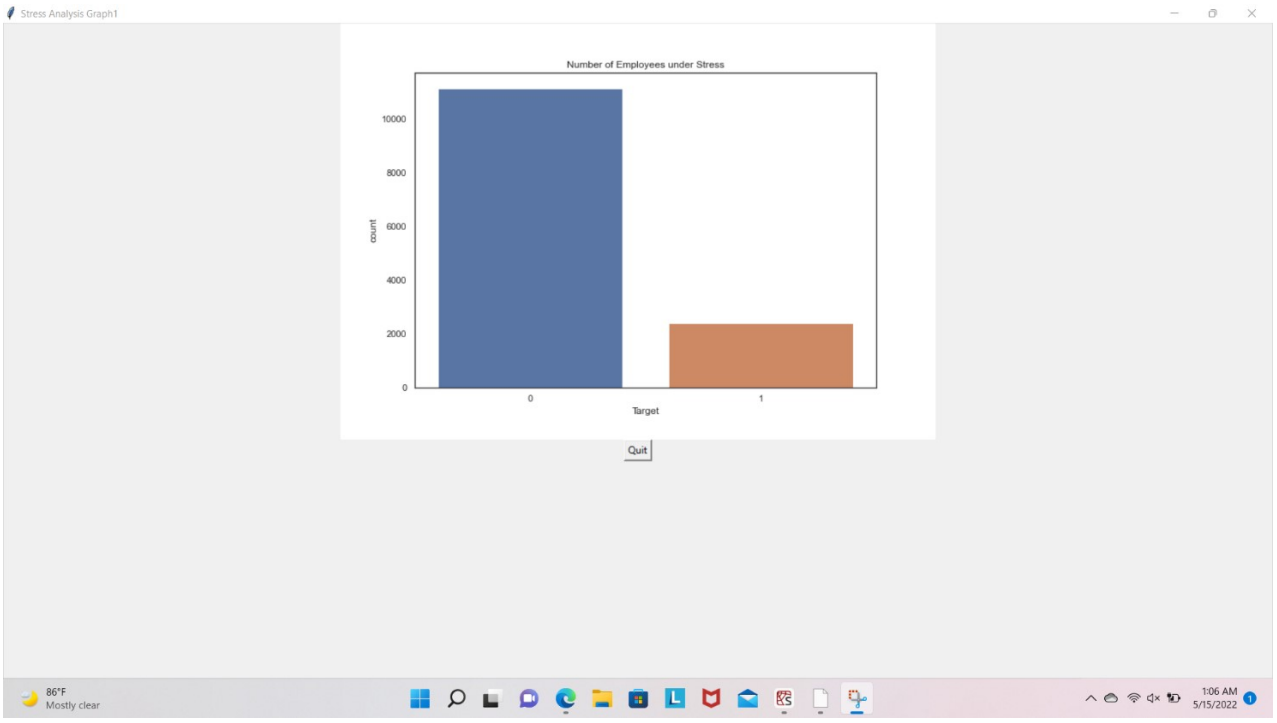

Figure 9. X vs Y Graph representing Count-plot of Number of Employees (Count) on Y-axis and Stress (Target) on X-axis

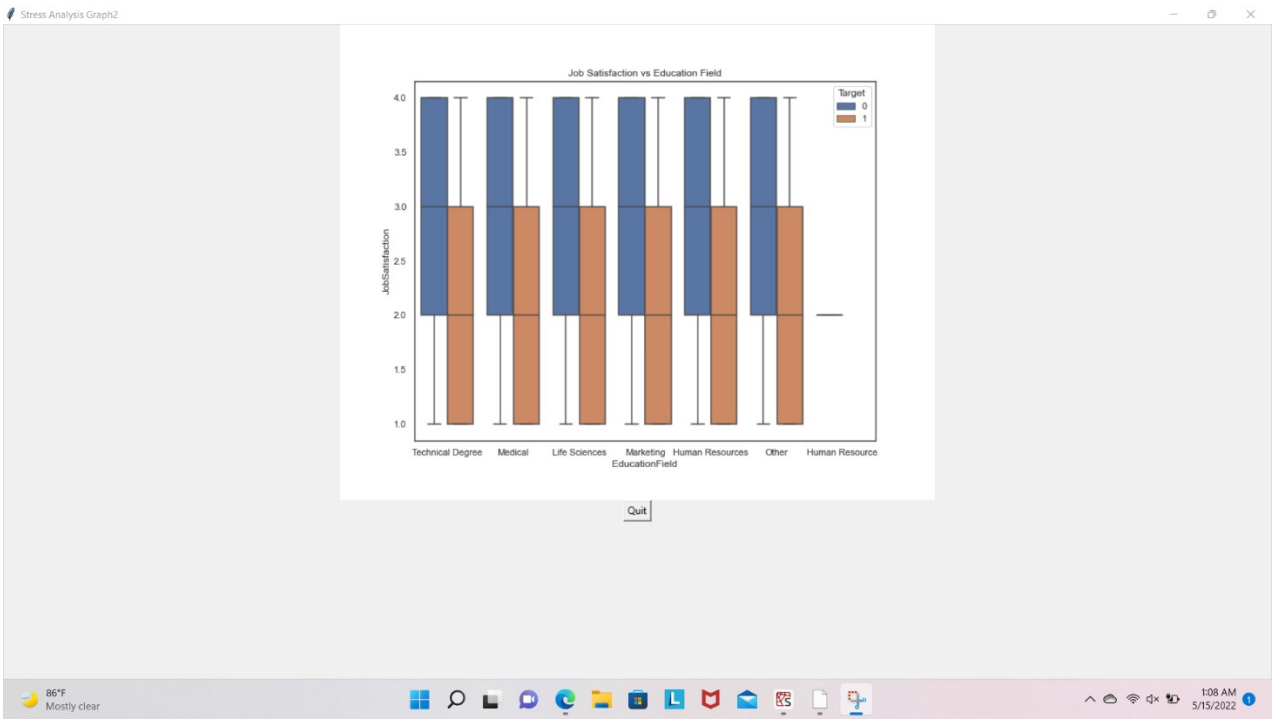

Figure 10. X vs Y Graph representing Education Field on X-axis and Job Involvement on Y-axis

IV. CONCLUSION

Bringing balance to work, life, and health of employees has become challenging in an ever-changing working environment, so the prediction of stress and visualization of data has become the main reason behind the development of this project. As this prediction analysis can help a company to take preemptive measures to reduce stress among employees either by altering the features detected to be of higher effect or by conducting necessary stress remediation programs within the workspace. As "Pre-emptive analysis is better than loss", we can predict the stress effects before any loss occurrences that might cause the company to lose the track of its goals and targets.

V. REFERENCES

[1] R. Bhattacharyya and S. Basu, "India Inc looks to deal with rising stress in employees", *Retrieved from 'The Economic Times'*, 2018.

[2] A. R. Subhani, W. Mumtaz, M. N. B. M. Saad, N. Kamel and A. S. Malik, "Machine Learning Framework for the Detection of Mental Stress at Multiple Levels", *IEEE Access*, vol. 5, pp. 13545-13556, 2017.

[3] Dr. G V Garje, Apoorva Inamdar, Harsha Mahajan, "Stress Detection And Sentiment Prediction: A Survey" International Journal of Engineering Applied Sciences and Technology, 2016.

[4] Reshmi Gopalakrishna Pillai, Mike Thelwall, Constantin Orasan, "Detection of Stress and Relaxation Magnitudes for Tweets", International World Wide Web Conference Committee ACM, 2018.

[5] Huijie Lin, Jia Jia, JiezhonQiu, " Detecting stress based on social interactions in social networks", IEEE Transactions on Knowledge and Data Engineering, 2017

[6] Zainudin Z Shamsuddin S M and Hasan S 2019 The International Conference on Advanced Machine Learning Technologies and Applications (AMLTA2019) 921 43-5

[7] L. Breiman, "Random forests", Mach. Learn., vol. 45, no. 1, pp. 5-32, 2001.

[8] Gregorutti, B. (2016). Rfgroove: Importance measure and selection for groups of variables with random forests. https://CRAN.R-project.org/package=RFgroove, R package version 1.1, (2016).

[9] Gregorutti, B., Michel, B., and Saint-Pierre, P. (2016). Correlation and variable importance in random forests. Statistics and Computing, pages 1–20.

[10] Gregorutti, B., Michel, B., and Saint-Pierre, P. (2015). Grouped variable importance with random forests and application to multiple functional data analysis. Computational Statistics and Data Analysis, 90:15–35.

[11] Deng, Y., Wu, Z., Chu, C., and Yang, T. (2012). Evaluating feature selection for stress identification. In Information Reuse and Integration (IRI), 2012 IEEE 13th International Conference on, pages 584–591.

[12] Breiman, L., Friedman, J., Stone, C.-J., and Olshen, R.-A. (1984). Classification and Regression Trees. The Wadsworth and Brooks-Cole statistics-probability series. Taylor & Francis.

[13] Kai Bi, Xiaodan WANG, Xu Yao et al., "Adaptive selective Integration based on Bagging and Confounding Matrix [J]", Acta electronica Sinica, vol. 04, pp. 711-716, 2014.

[14] F. A. Faria, J. A. dos Santos, S. Sarkar et al., "Classifier Selection Based on the Correlation of Diversity Measures: When Fewer Is More[C]", 2013 XXVI Conference on Graphics Patterns and Images Arequipa, pp. 16-23, 2013.

[15]]S. Sriramprakash, V. D. Prasanna, and O. V. R. Murthy, “Stress Detection in Working People,” Procedia Comput. Sci., vol. 115, pp. 359–366, 2017, doi: 10.1016/j.procs.2017.09.090.

[16] S. Gedam and S. Paul, ”A Review on Mental Stress Detection Using Wearable Sensors and Machine Learning Techniques,” in IEEE Access, vol. 9, pp. 84045-84066, 2021, doi: 10.1109/ACCESS.2021.3085502.

[17] S. M. Rajbhoj and S. F. Lakudzode, ”Human stress monitoring system using electrocardiogram,” 2016 International Conference on Inventive Computation Technologies (ICICT), 2016, pp. 1-5, doi: 10.1109/INVENTIVE.2016.7823216.

[18] A. Liaw and M. Wiener, "Classification and regression by randomForest", R news, vol. 2, no. 3, pp. 18-22, 2002.

[19] C. Zhang and Y. Ma, *Ensemble Machine Learning: Methods and Applications*, Springer, New York, NY, USA, 2012.

[20] B. Pang, L. Lee and S. Vaithyanathan, "Thumbs up? stress classification using machine learning techniques" in Language Processing (EMNLP), Philadelphia, pp. 79-86, July 2002.