



Unveiling Personnel Tension Disorder Using Machine Learning Algorithms

V.PAVANI ¹, P.ROHIT ², T.JYOTHI KUMARI ³, N.SIVALEELA⁴

¹Associate Professor, Department of Electronics & Communication Engineering, Chaitanya College of Engineering, JNTU-GV University, Vishakhapatnam, India

²Assistant Professor, Department of Electronics & Communication Engineering, Chaitanya College of Engineering, JNTU-GV University, Vishakhapatnam, India

³Assistant Professor, Department of Electronics & Communication Engineering, Chaitanya College of Engineering, JNTU-GV University, Vishakhapatnam, India

⁴Assistant Professor, Department of Electronics & Communication Engineering, Chaitanya College of Engineering, JNTU-GV University, Vishakhapatnam, India

Abstract

Stress disorders are a rather common occurrence among business sector workers. As people's job and lifestyles change, we can observe an increase in stress among working employees. Even while many business sectors offer a range of mental health programs and attempt to lessen stress-related problems in the workplace, the disorder is still very much alive and well. In this paper, we will apply two machine learning techniques to ascertain the level of stress experienced by corporate sector employees and attempt to identify the problems that contribute to the stress levels. Following the completion of data preprocessing and cleaning, we will use two machine learning approaches (SVM and Random Forest). Our trained model's accuracy was examined and evaluated in detail. Using these two machine learning techniques, it is discovered that the primary factors that lead to stress problems are an employee's sex, family background, and ease of access to health benefits at work. Because of these outcomes, corporate sectors can now reduce stress and create a very welcoming workplace for their personnel. In order to determine if an employee is in a relaxed or stressed state, the author uses social media datasets, such as tweets, where employees can express their opinions. However, manually assessing these opinions may require a significant amount of human labor. Therefore, we are utilizing machine learning techniques, and the results of our experiment indicate that these algorithms have an accuracy of over 80% in detecting stress.

Keywords: Employee Stress, Machine Learning, Support Vector Machine, Random Forest Algorithm

I. INTRODUCTION

Employees in corporate sectors frequently suffer from stress-related mental health disorders. Previous analyses have raised some concerns about the same thing. According to research conducted by the Association of Industry, Assocham, over 42% of professional workers in India's corporate and private sectors experience stress or common anxiety disorders as a result of their late-night work schedules and set hours. According to an article in the Economic Times from 2018, this percentage of singles is increasing and is based on a poll that Optum oversaw. A study is being conducted that takes into account the responses of about eight

lakh working employees from over seventy large organizations, each of which employs over 4,500 working professionals.

A stress-free workplace should be prioritized for increased productivity and a comfortable working environment for employees. We can help employees develop a stress disorder for mental health well-being in a number of ways, such as by providing career guidance, counseling, stress management sessions, and raising awareness of health issues. By identifying working employees who will require such assistance, we can increase the rates of such measures to become successful. We attempt to achieve this by utilizing our machine learning methods to develop a model that forecasts the rate at which stress is achieved. In addition to giving HR managers a greater understanding of their working professionals, this strategy will assist them in adopting the necessary safety measures to lessen the likelihood that their employees will experience stress.

II LITERATURE SURVEY

Numerous studies utilizing various machine learning algorithms have been conducted in relation to employee stress problems.

Mohana Priya D., et.al [1] has mentioned that our design's main objective is to lessen IT professionals' stress by fusing image processing methods with graphical technology literacy. The prior stress discovery techniques that attempted to avoid live discovery and specific consoling have been continuously improved upon by our system.

Ramya Srikanteswara., et.al [2] has suggested employing machine learning and survey data to identify stress in IT workers. Classification algorithms are essential for classifying data into pre-established groups in machine learning.

Laavanya Mohan., et.al [3] has talked about the Perceived Stress Scale (PSS) method of predicting stress. The PSS technique was chosen because it is a simple questionnaire with proven psychometric qualities. The questions are ranked using a weighted average method, and PSS scores are calculated by inverting responses, which further increases accuracy.

N. BalaKrishna., et.al [4] has concentrated on using Convolution Neural Networks (CNN), more especially the effective MobileNet framework, to interpret facial expressions associated with stress. The foundation of this research study is an extensive collection of facial photographs that reflect a range of stress intensities.

Kazi.A., et.al [5] has used a machine learning approach to discuss the stress levels that staff and students report. This study explores the complex elements impacting stress in these different cohorts by utilizing machine learning approaches.

III MACHINE LEARNING ALGORITHMS

The area of machine learning (ML), which sits at the intersection of statistics and computer science and at the heart of data science and artificial intelligence (AI), is currently growing at a rapid rate since it is still relatively new. Recent advances in machine learning have been fueled by the creation of new theories for learning algorithms as well as the continuous increase in the availability of large amounts of data (commonly referred to as "big data") and inexpensive computing. In numerous fields, including healthcare, biomedicine, manufacturing, education, financial modeling, data governance, police, and marketing, the use of machine learning (ML)-based techniques is spreading throughout science, technology, and business, resulting in more evidence-based decision-making. Even while interest in these areas has grown over the last 10 years, we are just now starting to explore the possibilities of these machine learning techniques for researching systems that get better with time. Machine learning is the process of creating and implementing algorithms that, instead of being programmed to assign particular outputs in response to particular environmental inputs, use statistical techniques to assess the data and its attributes and determine the action. Two machine learning algorithms, such as the Support Vector Machine and Random Forest algorithms, are presented in this study.

SUPPORT VECTOR MACHINE ALGORITHM

A classic classification-based machine learning technique is the Support Vector Machine (SVM). In order to achieve the principle of structural risk minimization, the classifier only needs a limited number of support vectors, which is derived from the idea of addressing the dual form of large-dimensional problems. Nonlinear and local minimum problems were resolved by statistical learning theory. The system call can transform the sequence into a call sequence with a specific length of vectors in a high-dimensional space by using the short system frequency. Consequently, support vector machines can be used to carry out anomaly detection. SVM, a fairly traditional two-classification model, operates by identifying an appropriate hyperplane to divide the gathered data samples. The idea behind segmentation is to optimize the interval, including the hard and soft intervals, and then turn it into a unique quadratic programming problem. Use the linear separable support vector machine by maximizing the hard interval if the training sample is linearly time-sharing; use the linear support vector machine by maximizing the soft interval and choosing the appropriate kernel function if the training sample is approximately linearly time-sharing; and use a nonlinear support vector machine to enable the selection of the appropriate kernel function and the maximization of the soft interval if the training sample is linearly non-time-sharing.

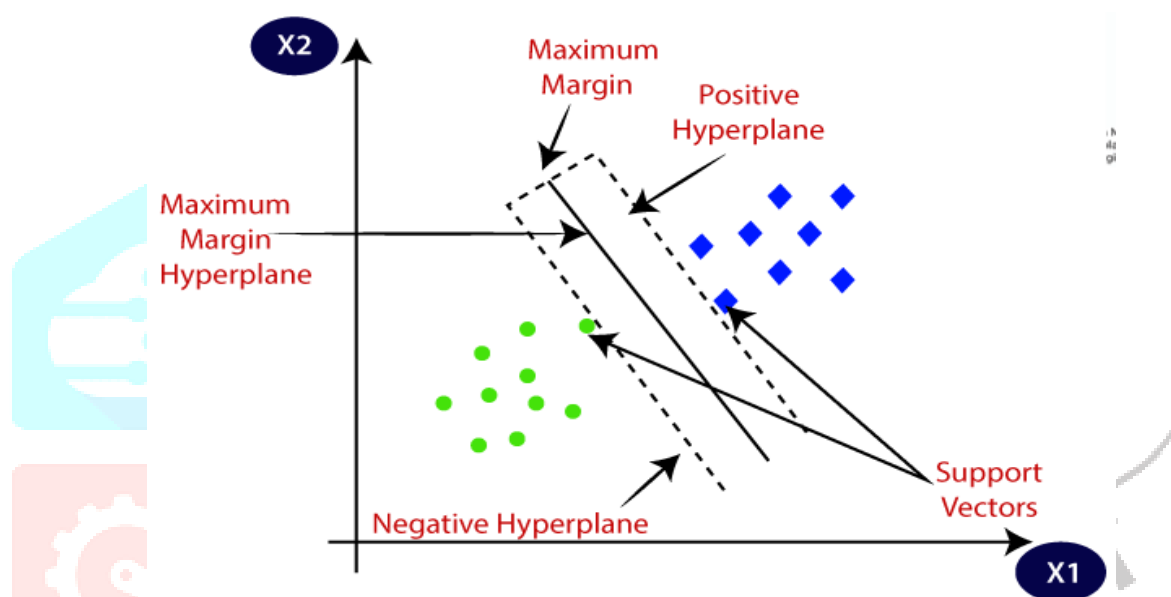


Figure 1: Structure of Support Vector Machine (SVM) Algorithm

RANDOM FOREST ALGORITHM

One ensemble learning technique that falls under the homogeneous base learner category in terms of the kind of constructive classifiers is Random Forest. One of the primary obstacles in combining the basic classifiers to create a stronger classifier in ensemble learning is weighting the base learners. Since all base learners are decision trees, as the name suggests, their structure is less complicated than that of comparable techniques. There are two benefits to the random forest structure. A computational perspective is used in the first category, whereas a statistical perspective is used in the second. Both regression and classification problems can be handled from the random forest's computational point of view. Because the classifier's train and prediction operations are carried out quickly, the random forest is regarded as one of the quick classic classifiers. The random forest's direct application to high-dimensional problems is another benefit. From a statistical perspective, random forests fall into the second type. These include feature prioritizing, assigning distinct weight coefficients to various classes, visualization, and unsupervised learning capabilities.

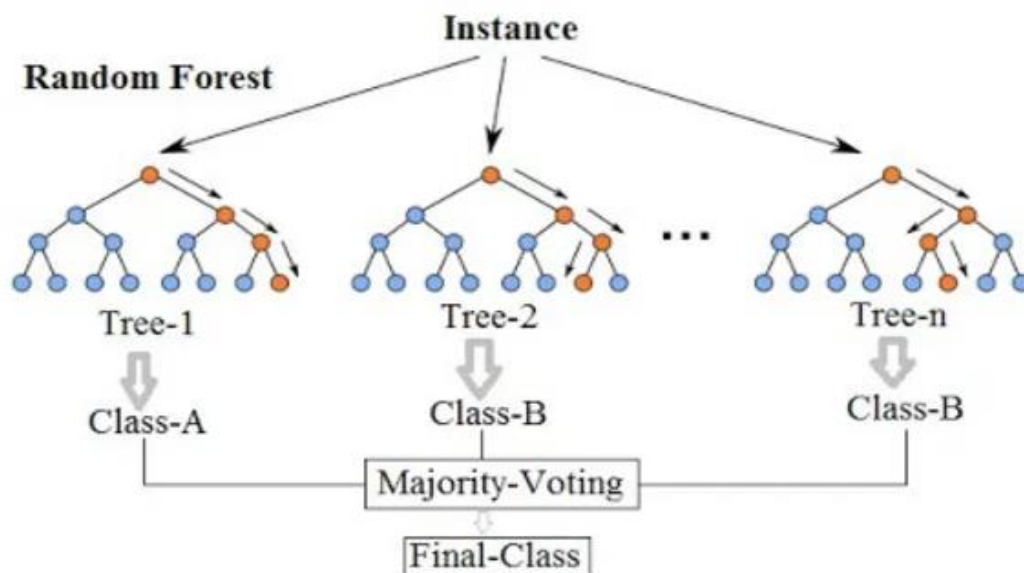


Figure 2: Random Forest Algorithm Simplified Structure

IV PROPOSED METHOD

Using Machine Learning Classifiers like Random Forest and Support Vector Machine algorithms, we identify employee stress problems in this suggested approach. In order to improve prediction accuracy, machine learning techniques maximize features found in datasets used to train classification algorithms. The systematic implementation of the Employee Stress Disorders Using Machine Learning is shown in Figure 3.

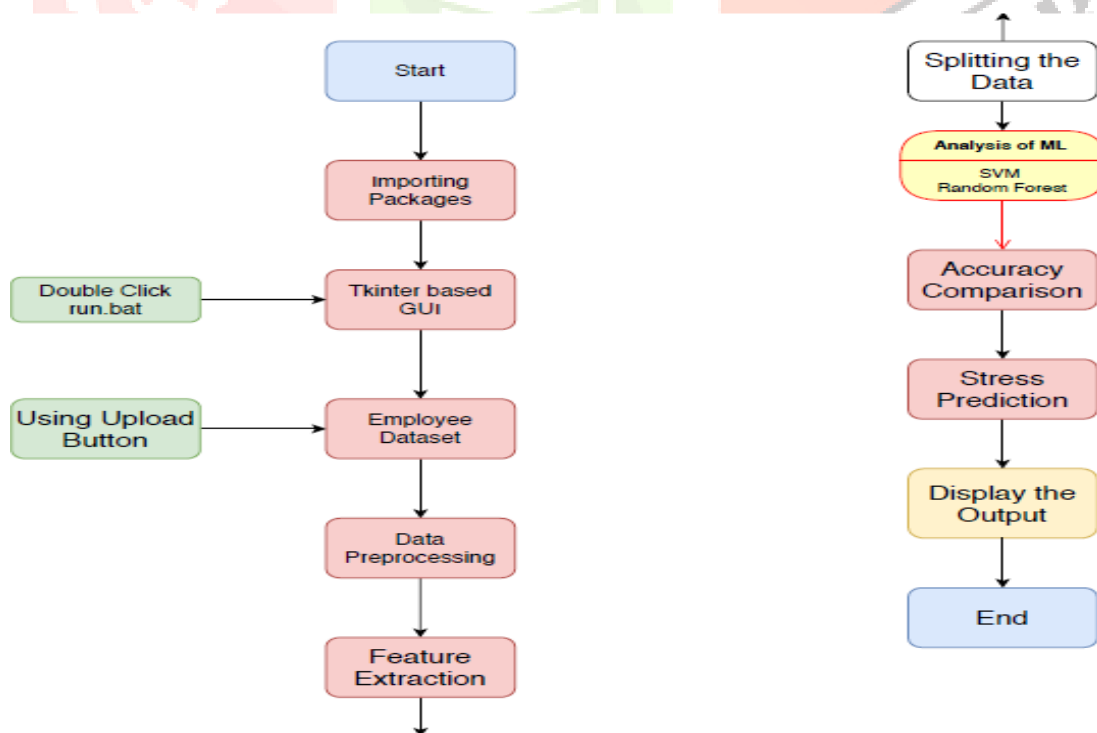


Figure 3: Proposed Approach of Employee Stress Disorders Using Machine Learning

V RESULTS & DISCUSSION

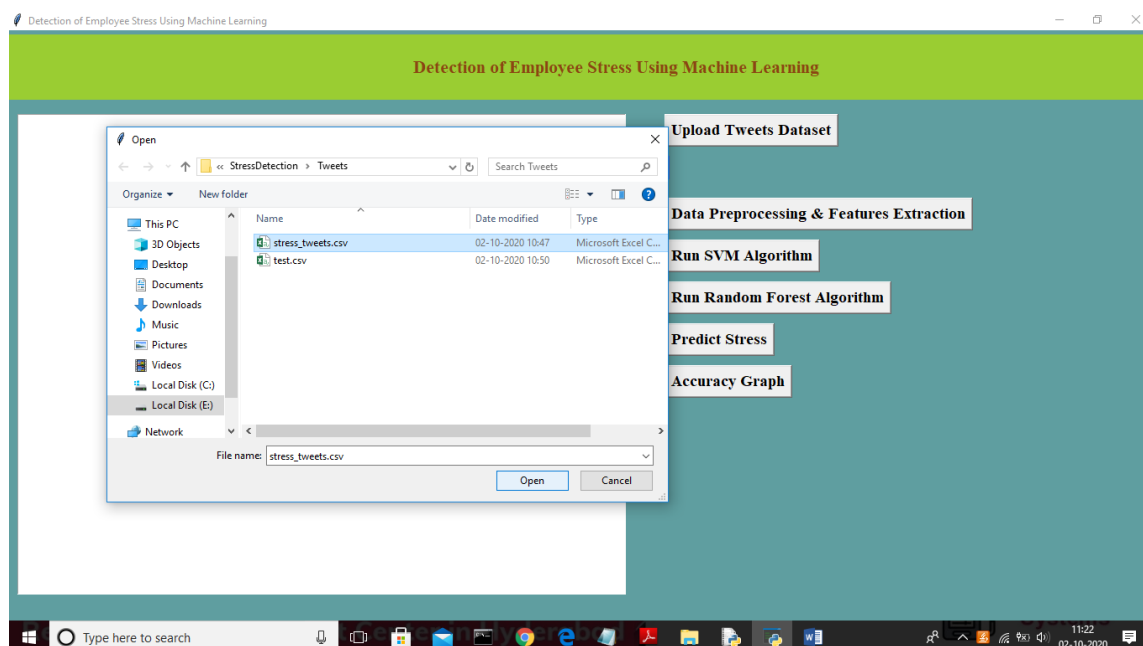


Figure 4: Upload Dataset File

To determine whether or not all of the values in the dataset are connected, machine learning algorithms will be employed; if an attribute is determined to be irrelevant, it will be eliminated. Figure 4 demonstrates how to choose the "stress_tweets.csv" dataset, click the "Open" button to load the dataset, and see the screen below.

The "Data Preprocessing & Features Extraction" button in Figure 5 demonstrates how to read a dataset, clean it up, and extract features like words. It also illustrates how to identify the total number of records in the dataset, the total number of words, and the application by calculating the number of records needed for training and testing.

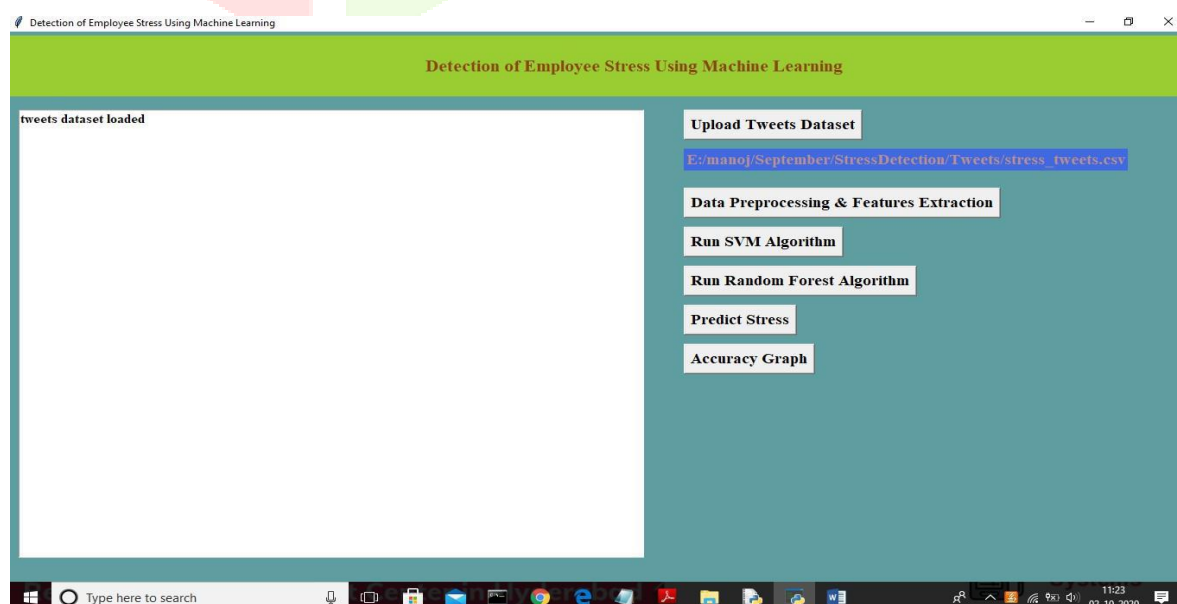


Figure 5: Data Preprocessing & Features Extraction'

Figure 6 displays the dataset, which consists of 10314 tweets in total. Each tweet has 30790 words, with 83 unique words overall. The application uses 1341 records for testing and 8973 records for training. Now that the test and train data are prepared, click the "Run Random Forest Algorithm" and "Run SVM Algorithm" buttons to begin training the data using machine learning techniques.

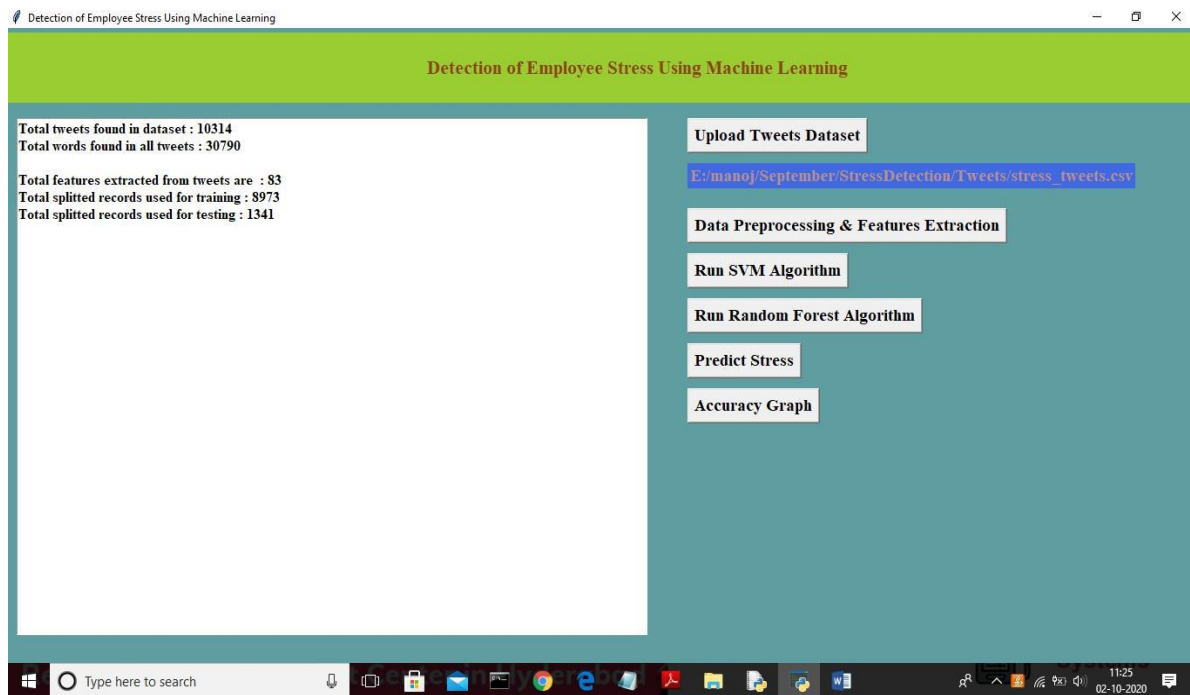


Figure 6: Run Machine Learning Algorithms

Accuracy of Support Vector Machine Algorithm is displayed in Figure 7. 89.70% of the test data was projected to be accurate by the Support Vector Machine Algorithm.

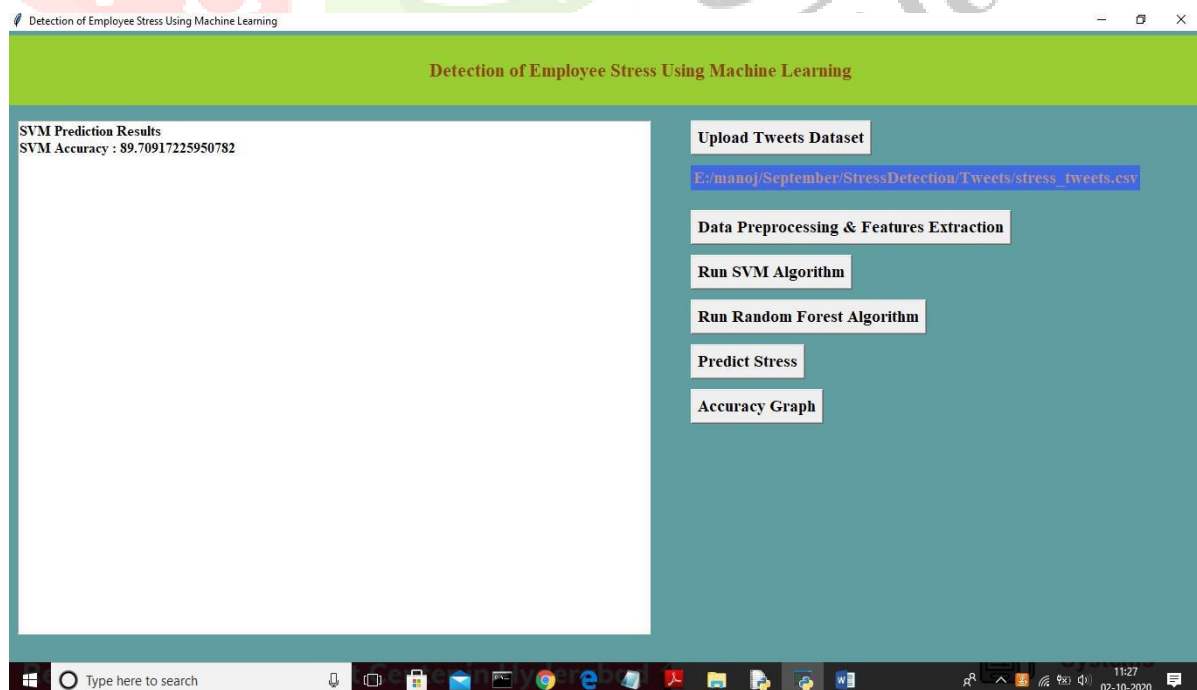


Figure 7: Support Vector Machine Algorithm Accuracy

The Random Forest Algorithm Accuracy is displayed in Figure 8. The Random Forest Algorithm's prediction accuracy was 97.68.

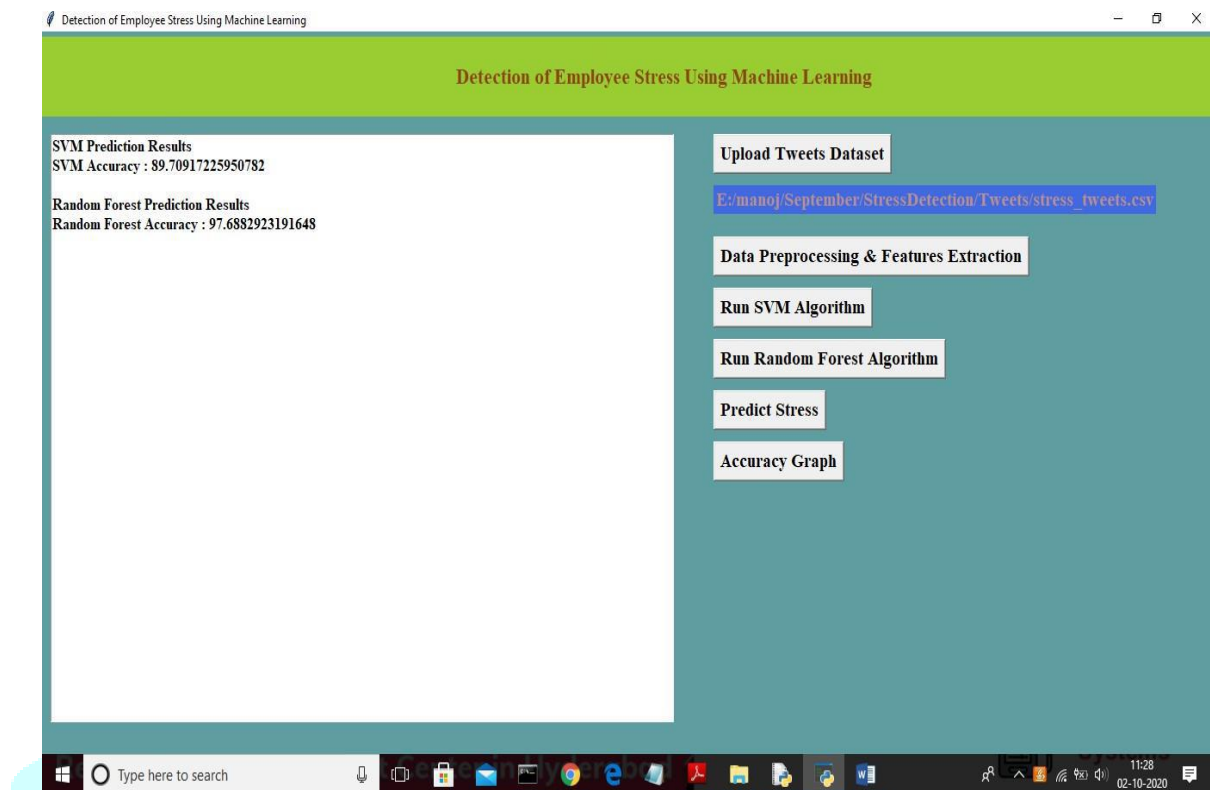
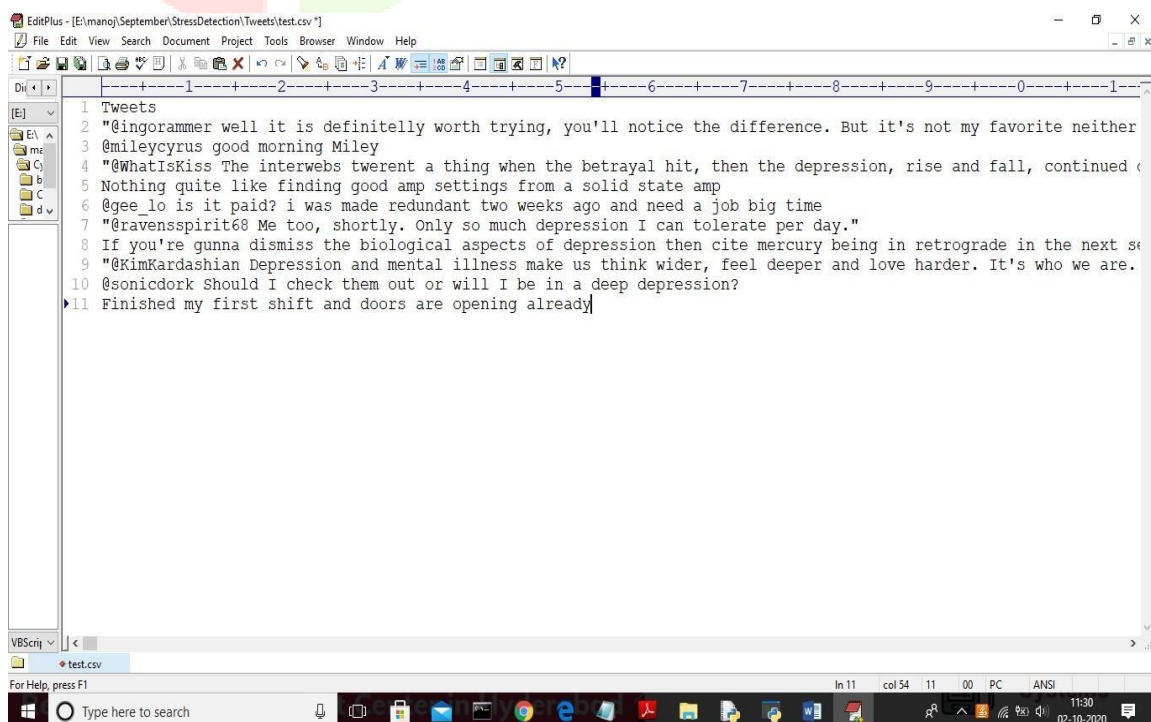


Figure 8: Random Forest Algorithm Accuracy

Figure 9 demonstrates how to click the "Predict Stress" button and upload a test file with tweets. A machine learning algorithm will then analyze the tweets to determine whether or not they contain stress data. Once the "test.csv" file has been uploaded, click the "Open" button to forecast stress. The projected outcome is displayed next to each tweet in the screen above as either worried or not.



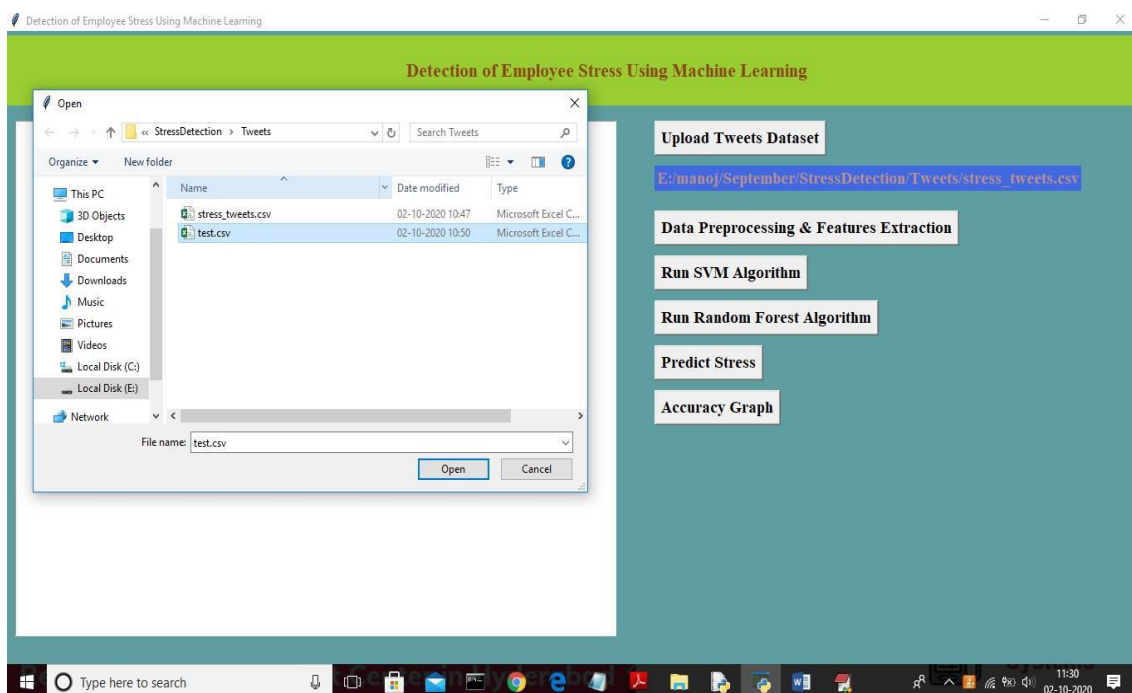


Figure 9: Upload Test.CSV File

The accuracy comparison of the suggested algorithms is displayed in Figure 10. The SVM algorithm is shown by the x-axis above, while the Random Forest algorithm's accuracy is represented by the y-axis. Based on the graph above, we can conclude that Random Forest is superior to SVM.

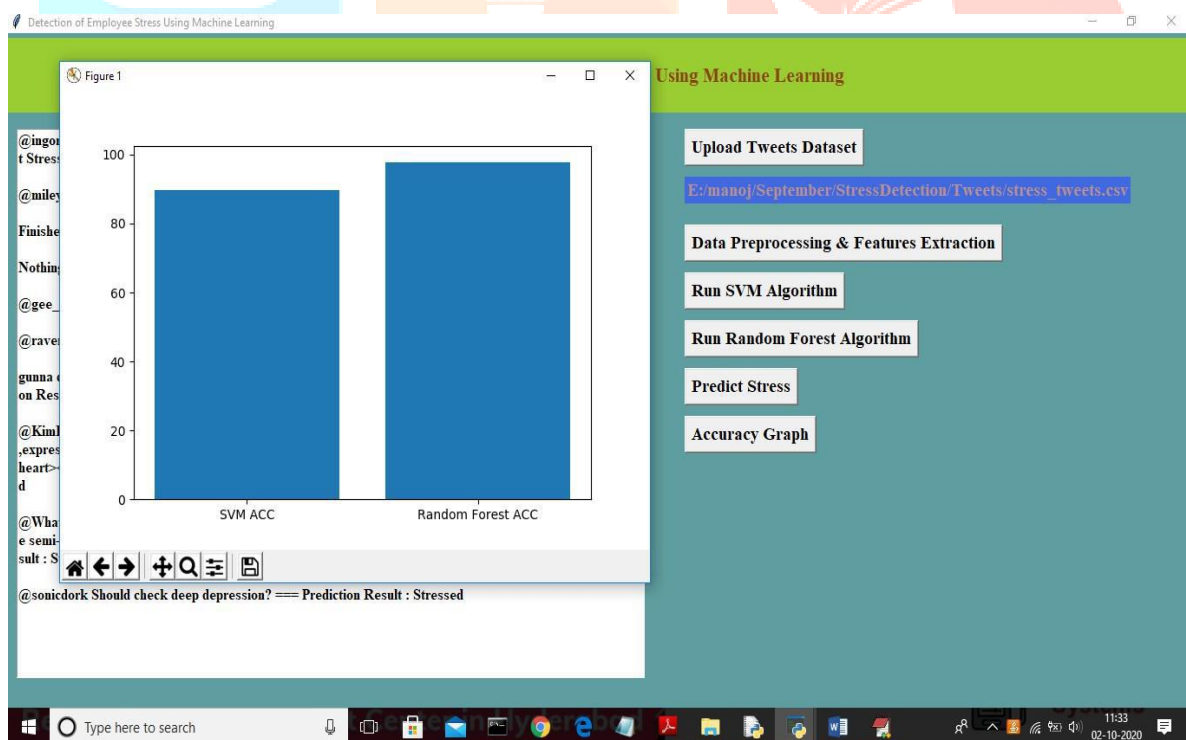


Figure 10: Accuracy Comparison of Proposed Algorithms

VI CONCLUSION

When evaluating whether an employee can acquire conceptual health-related issues, gender, the family history of the illness, and whether a single employer offers the conceptual health benefits to their employees were more important than the other factors. Our research revealed that, despite their non-technical job roles, employees in tech organizations are more likely to experience stress. Businesses could successfully leverage these beliefs to create HR strategies that are more appealing to working personnel. With a 75% accuracy rate, the use of two machine learning approaches (SVM and Random Forest) to predict stress and conceptual health issues yields worthy results that might be further investigated, so fulfilling the purpose of this work.

REFERENCES

1. M. P. D, H. H. S, B. R. B, H. M and J. K. P, "Detection of Stress by Machine Learning in IT Industry," *2023 7th International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, 2023, pp. 53-56, doi: 10.1109/ICCMC56507.2023.10083884.
2. R. Srikanteswara, P. Rithicka, Y. V. Shilpha Kala, S. Rangaraj and V. Devaiah, "Machine Learning-Based Stress Detection in IT Employees: A Data-Driven Approach for Workplace Well-being," *2024 Second International Conference on Data Science and Information System (ICDSIS)*, Hassan, India, 2024, pp. 1-7, doi: 10.1109/ICDSIS61070.2024.10594118
3. L. Mohan and G. Panuganti, "Perceived Stress Prediction among Employees using Machine Learning techniques," *2022 International Conference on Communication, Computing and Internet of Things (IC3IoT)*, Chennai, India, 2022, pp. 1-6, doi: 10.1109/IC3IoT53935.2022.9768026.
4. N. BalaKrishna, M. B. M. Krishnan, B. Veerabhadru, A. Rishika, B. Sruthika and B. S. Kumar, "Facial Image Analysis for Stress Detection in IT Professionals using Machine Learning Techniques," *2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, Salem, India, 2024, pp. 571-577, doi: 10.1109/ICAAIC60222.2024.10575715.
5. Kazi, A., Jagtap, J., Jayaswal, R., Kolhar, S., Choudhury, T., Kotecha, K. (2024). Machine Learning for Multimodal Stress Detection – A Case-Study. In: Lapina, M., Raza, Z., Tchernykh, A., Sajid, M., Zolotarev, V., Babenko, M. (eds) AISMA-2024: International Workshop on Advanced Information Security Management and Applications. AISMA 2024. Lecture Notes in Networks and Systems, vol 863. Springer, Cham. https://doi.org/10.1007/978-3-031-72171-7_12
6. S. R, S. Sharma, P. B. N. Kiran, S. K. Gupta, V. Sindhu and R. V. Srinivas, "Novel Method for Detection of Stress in Employees using Hybrid Deep learning Models," *2023 8th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, 2023, pp. 984-989, doi: 10.1109/ICCES57224.2023.10192609.
7. U. S. Reddy and A. V. Thota, "Workplace Stress Prediction Using Machine Learning Techniques", *2018 IEEE International Conference on Computational Intelligence and Informatics Innovation (ICCIC)*.
8. Y. Deng, Z. Wu, C.H. Chu, Q. Zhang and D.F Hsu, "Combinatorial fusion sensor feature selection and fusion for stress identification", *D.F. World Journal of Advanced Robotic Systems*, vol. 10, no. 8, pp. 306, 2013.
9. A. Haldorai, A. Ramu and M. Suriya, "Organization Internet of Things (IoT s): Supervised Unsupervised and Reinforcement Learning", *EAI/Springer Innovations in Communication and Computing*, pp. 27-53, 2020.

10. H. Anandakumar and R. Arulmurugan, "Artificial Intelligence and Machine Learning for Enterprise Management", *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)*, pp. 1265-1269, 2019.
11. M. Shruti, Harshini J. Philip and I. V. S. L. Haritha, "Stress Level Detection of IT Professionals Using Machine Learning", *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pp. 1-5, 2022.
12. R. K. Kushwaha, M. S. A. Ansari, J. V. N. Ramesh and P. Rohit, "MIMO Optimized Algorithm to Develop the Energy Efficiency Underwater," *2023 International Conference on New Frontiers in Communication, Automation, Management and Security (ICCAMS)*, Bangalore, India, 2023, pp. 1-5, doi: 10.1109 / ICCAMS 60113.2023 .10525874.
13. Samanta, Pravin & Logesh, Dr & MOHAPATRA, Dr & Sivakumar, Dr & Krishna, K & Tamizhselvi, Dr & AMUTHA.N, & KUMAR, RANJEET & Vemulapati, Pavani & Thota, Jyothi Kumari & Penki, Rohit. (2024). INNOVATIVE APPROACHES TO ENHANCE BREAST CANCER DETECTION USING MACHINE LEARNING AND IMAGE PROCESSING.
14. Singh, Charanjeet & Prasad, Dr & Swathi, Dr & M, Thanga & Karthiyayini, S. & Malar, Dr & Penki, Rohit & Chappa, Hemanthakumar & Russia, Dr & Murugan, Dr. (2023). MACHINE LEARNING & ARTIFICIAL INTELLIGENCE APPROACH FOR TRAINING SIGNAL IN 6G WIRELESS SENSOR NETWORKS.
15. Nisha Raichur, Nidhi Lanikai and Priyanka Mural, *Detection of Pressure Employing Image Analysis and Techniques for Machine Learning*, vol. 9, no. 3S, July 2017.
16. Penki, Rohit & Annamalai, Kaliappan & S, Muthuraj & .C, Thilagavathi & F, Ravindaran & Namdeo, Kritika & T, Satheesh & E, Dinesh & T, Kannan & Fernandes, Anthony & Sahu, Swati. (2023). Title of Invention: "MACHINE LEARNING BASED PEST DETECTION AND ALERT SYSTEM FOR FARMERS USING INTERNET OF THINGS TO IMPROVE AGRICULTURE YIELD".
17. Thota, Jyothi Kumari & Penki, Rohit & Vemulapati, Pavani. (2024). EARLY DETECTION OF MELONAMA IN SKIN CANCER USING PROBABILISTIC NEURAL NETWORKS. 9. 592-600.
18. V. Laijawala, A. Aachaliya, H. Jatta and V. Pinjarkar, "Classification Algorithms based Mental Health Prediction using Data Mining", *2020 5th International Conference on Communication and Electronics Systems (ICCES)*, pp. 1174-1178, 2020.
19. Vemulapati, Pavani & Thota, Jyothi Kumari & Penki, Rohit. (2024). GAN ALGORITHM USING IMAGE SEGMENTATION OF BRAIN CANCER OF MRI IMAGES USING GLCM FEATURES. 9. 2403139.