Original Article

Stress Prediction in Working Employees using Artificial Intelligence of Things

SUHAS KS[1], **PHANEENDRA HD**[2]

[1] Student, MTech IT, The National Institute of Engineering, Mysore.

[2] Professor, Department of Computer Science and Engineering, The National Institute of Engineering, Mysore.

Email: suhasksgbd90@gmail.com [1], hdphanee@nie.ac.in [2]

DOI: 10.47750/pnr.2022.13.S01.237

Abstract

Stress issues are a common issue among today's working IT professionals. As people's lifestyles and workplace cultures change, employees are more prone to encounter stress. In this project, we will use IoT and machine learning approach like supervised learning to examine stress in working employees. After proper data cleaning and preprocessing, we used a variety of Machine Learning approaches like KNN, Decision Tree and Naïve Bayes algorithm to train our model. The accuracy of the above-mentioned models was determined and compared. Among the models used, KNN Algorithm had the best accuracy. Significant factors that affect stress were found using KNN, Decision Tree, and Naive Bayes algorithms. With these findings, organizations can set their sights on reducing stress and providing a much more comfortable working environment for their employees.

Keywords: Employees, Machine Learning, KNN, Decision Tree, Naïve Bayes, Stress.

INTRODUCTION

Working people frequently experience stress-related mental health illnesses such as depression, pressure, stress, interpersonal sensitivity, fear, anxiousness, etc. As people's lifestyles and workplace cultures change, stress among employees is more likely to occur. Although many corporate sectors and industries offer mental health-related programs and seek to improve the workplace culture, the problem is still out of control. In order to comprehend and recommend therapies with a deviated mental behavior, the evaluation of mental health is absolutely essential. In India's private sector, over 42% of working professionals experience depression or general anxiety disorder as a result of job-related stress brought on by long hours and stressful deadlines, according to research by the industry association Assocham [4]. Prioritizing the maintenance of a stress-free workplace will raise productivity and improve employee wellbeing. Working professionals can benefit from counselling, career advice, stress management classes, and health awareness initiatives as ways to manage stress and maintain their mental health. Early detection of employees who will require such assistance will increase the likelihood of such actions being successful. By applying machine learning techniques, we can develop a model to forecast the risk of stress in working individuals by considering some of his or her professional and personal aspects as parameters and are collected through meticulously obtained questionnaires, by this we can simplify this process. This method will assist HR managers in better understanding their employees and in taking preventative action to minimize the possibility that a worker would perform below expectations. We are using machine learning techniques in this system to assess employee stress patterns and identify the factors that have the biggest impact on stress levels. IoT Sensors are used to extract some parameters required for stress prediction.

Sensors such as PIR Sensor, Hear-Beat sensor, and Temperature Sensor. After thorough data cleaning and pre-processing, our model is trained using a variety of machine-learning approaches.

Furthermore, the paper is organized as described below.

Section 1 contains the Introduction and it covers the brief overview of employee stress and Machine Learning Techniques used.

Section 2 contains the literature survey related to the proposed work. It illustrates the existing system and it also illustrates

enhancement of the existing system. Overall abstract of each paper is explained in a brief.

Section 3 describes the proposed system.

Section 4 describes the various Machine Learning techniques used for building the model.

Section 5 covers methodology.

Section 6 covers results and the performance analysis of the different models used for stress prediction.

Section 7 Covers conclusion and future enhancements.

RELATED WORK

In [1], G. Azar et al. aimed to predict the psychological disorder. They have used various machine learning approaches and intelligent genetic algorithm to build a semi-automated system. They have compared the person's mental health with the DSM-IV-TR. The further aim is to make the system fully automated. Through this experiment, they have proved that genetic algorithms can be applied for many real-time applications.

In [2], The goal of Fang Li was to forecast student stress. In light of this, the author integrated resources for mental health education into the cloud using data mining and a cloud platform, allowing them to share each other's high-quality resources. The author also covered three elements that affects the mental health of college students i.e., students, society and education. The author proposed several ways to support the management of college students' psychological health through the examination of the management system for the psychological health of students.

In [3], A.R. Subhani et al. used various machine learning frameworks to analyze and predict the levels of stress. The analysis of stress included the use of EEG signals. They proposed to implement Logistic Regression, Support Vector Machine, and Naive Bayes, as well as EEG feature extraction. The experiment's findings have provided the best accuracy for stress prediction. In the experiment, accuracy for level 2 stress was 94.6 percent, and accuracy for multi-level stress was 83.4 percent.

In [4], Aditya Vivek Thota et al. aimed at predicting stress of IT employees. Working professionals in the tech industry who participated in the OSMI mental health survey provided the data. The best accuracy was achieved by boosting (75%), while the lowest accuracy was attained by bagging (69.43%). Among the other models Logistic Regression acquired 73%, KNN attained an accuracy of 73%, Decision Tree attained 70%, Random Forest attained 73%. The cross-validated AUC value for the random forest classifier was higher, indicating a more stable model.

In [5], Sandhya P et al. used various machine learning approaches to predict stress in IT Employees. The dataset was taken in the form of a questionnaire where employees were asked to fill in the details. The best accuracy was achieved by boosting (81.7%), while the lowest accuracy was attained by bagging (77.7%). Among the other models Logistic Regression acquired 79.9%, KNN attained an accuracy of 80.4%, Decision Tree attained 80.6%, and Random Forest attained 81.2%.

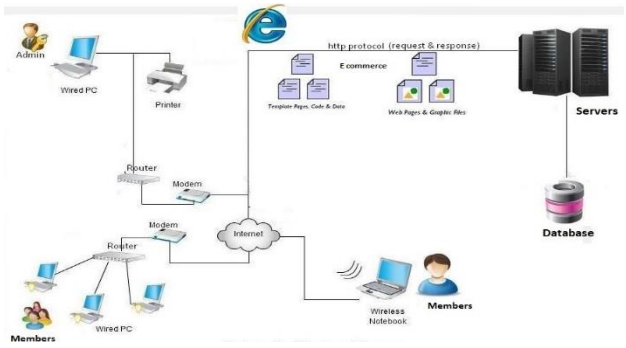
In [6], the goal of Monisha S. et al. is to forecast student stress. In order to assess reliable data and organize the components most likely to cause stress based on probabilistic characteristics, the authors employed the Naive Bayes technique. Authors have examined stress patterns using a variety of machine learning techniques.

In [7], Vidit Laijawala et al. aim at mental health using data mining techniques. Authors have collected data from online-available datasets. To forecast the mental health of individuals, various machine learning techniques are applied. In that, decision tree attained the highest accuracy of 82.2%, while random forests attained an accuracy of 79.3% and Naïve Bayes of 78.7%. The data has been analyzed using WEKA tool.

PROPOSED SYSTEM

The algorithm determines factors that significantly affect stress levels. Employees' stress levels were determined based on their pulse, temperature, family history, and the availability of health insurance at work. The system's major goal is to identify risk variables that have an impact on employee mental health as shown in Fig1.

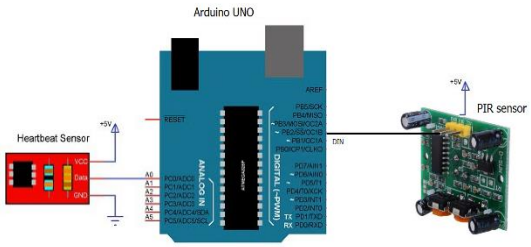
Fig1. Proposed System



Gender, Family History, Colleague History, Illness, Working Hours, and other parameters are used by the system. Using IoT sensors, the system also harvests real-time information such as pulse, temperature, and staff mobility as shown Fig2.

Some of the reference ranges considered for prediction are heart rate <=100 is Normal and >100 and without any symptoms is considered as abnormal which is related to stress.[8]

Fig2. Circuit Diagram



Sensors used: PIR Sensor, Heartbeat Sensor and LM-35 sensor.

MACHINE LEARNING TECHNIQUES USED

Artificial intelligence, known as machine learning, enables computers and computing systems to autonomously learn from the past and advance over time without explicit human programming. The construction of computer programs with the capability to access data and learn on their own is the cornerstone of machine learning. When properly given to an intelligent system and educated effectively, this is very helpful in the healthcare industry, where there is a great amount of data. The final prediction model will be superior, free of human errors, and minimize the diagnostic process time.

KNN Classifier: The K-Nearest Neighbor (KNN) classifier is one of the supervised learning techniques that can be applied to labelled data. In this instance, it was used to ascertain whether the employee was under stress. KNN classifies the dependent variable based on how much an example from the known data that is comparable to the dependent variable resembles the independent variables [4].

Decision Tree - A decision tree can be used to model several options, if-else statements, or decisions in a way that resembles a tree. Decision trees are used in this instance to determine the 15 attributes that contribute most frequently [4].

Naïve Bayes- One of the simplest and most effective classification techniques for developing machine learning models that can predict outcomes is the Nave Bayes Classifier. In this case, Naïve Bayes is used to compare the values and classify the attribute values to one of the predefined set of classes [6].

METHODOLOGY

For the benefit of IT firms, the system can be implemented as a real-time application. We utilize Visual Studio and SQL Server for application development since they are more real-time and application-friendly.

1. **Data Collection:** We gather data on stress at this stage of the employee stress prediction procedure. Data has been gathered from a variety of sources and includes variables like gender, age, financial_problems, family, working_ hours, learning_method, health_problems, partiality_fix,colleague_issue,pressure,regular_ interaction, etc.
2. **Data Preparation**: Only pertinent data was extracted after analysis of the stress data. Data that is needed for processing is extracted and segmented in accordance with the specifications. The essential data extraction is carried out since the entire data is not required for processing and processing would take too long if we entered all the data.
3. **Data Splitting:** At this point, the data will be split into a training dataset and a testing dataset in a ratio of 90:10.
4. **Model Education:** The machine learning algorithm is trained by providing datasets to it at this phase. Regular training can increase the machine learning model's prediction rate greatly. Various machine learning algorithms are used to train the model. In this model we have used.
5. **Stress Prediction:** After the model is trained, the system predicts the stress of the working employee based on the parameters and value driven by IoT sensors.

RESULTS AND PERFORMANCE ANALYSIS

To forecast the stress of a working employee, all of the above-mentioned models were trained. The results are tabulated in the table below.

The classification performance of all trained models is shown in Figure 3, with KNN having the highest accuracy (87.2%) and Naive Bayes having the lowest accuracy (56.8%).

Fig 3. Performance analysis of each method

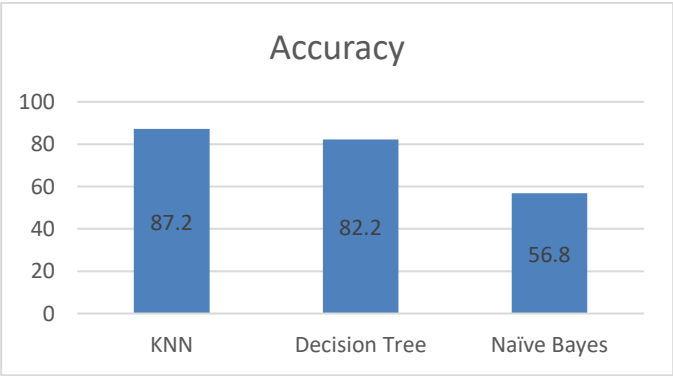


Figure 4 shows the time taken for execution of each model being the KNN has the least execution time.

Fig4. Execution Time

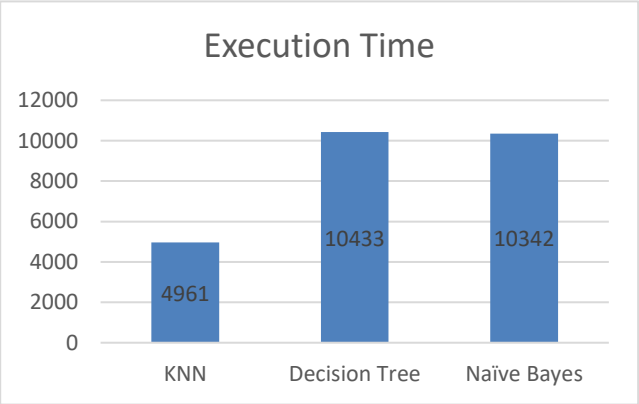


Table 1 contains the evaluation and tabulation of the performance of several trained models. KNN outperformed the other models in terms of accuracy and execution time after the model was trained using a variety of machine learning approaches.

Table 1. Performance of different models trained

Algorithm	Accuracy	Execution Time
KNN	87.2	4961
Decision Tree	82.2	10433
Naïve Bayes	56.8	10342

CONCLUSION AND FUTURE WORK

Family medical history, idle time during workdays, in terms of relationships with coworkers, more important than other considerations is whether a firm offers its employees mental health benefits. Using real-time parameters adds a step in predicting stress in working employees. Because of the strict deadlines and extended working hours, those who work for IT organizations are somewhat more likely to experience stress [4]. Using some IoT sensors, capturing some parameters like pulse and temperature will play a major role in predicting stress.

From the above Algorithms used KNN Algorithm acquired good results of 87.2% whereas the Decision tree algorithm acquired a result of 82.2% and Naïve Bayes acquired the least accuracy of 56.8%.

The goal of this article was achieved by using machine learning techniques to predict stress and mental health conditions. These techniques produce notable findings and can be further investigated.

In the future, one can use a Convoluted Neural network and various deep learning techniques to analyze the accuracy of the models. One can consider a dataset released by some healthcare providers or the questionnaires released from various institutes

One can use different machine learning techniques to build a model where the model is capable of giving a solution if the employee is undergoing some stress attacks.

REFERENCES

1. Azar, G., Gloster, C., El-Bathy, N., Yu, S., Neela, R. H., & Alothman, I. (2015, May). Intelligent data mining and machine learning for mental health diagnosis using genetic algorithm. In 2015 IEEE International Conference on Electro/Information Technology (EIT) (pp. 201-206). IEEE.
2. Li, F. Research on the College Students' Psychological Health Management based on Data Mining and Cloud Platform.
3. Subhani, A. R., Mumtaz, W., Saad, M. N. B. M., Kamel, N., & Malik, A. S. (2017). Machine learning framework for the detection of mental stress at multiple levels. IEEE Access, 5, 13545-13556.
4. Reddy, U. S., Thota, A. V., & Dharun, A. (2018, December). Machine learning techniques for stress prediction in working employees. In 2018 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC) (pp. 1-4). IEEE
5. Sandhya, P., & Kantesaria, M. (2019). Prediction of mental disorder for employees in IT industry. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8(6S), 374-376.
6. Monisha, S., Meera, R., & Swaminath, V. (2020, January). Predictive analysis of student stress level using Naïve Bayesian classification algorithm. In 2020 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-7). IEEE.
7. Laijawala, V., Aachaliya, A., Jatta, H., & Pinjarkar, V. (2020, June). Classification algorithms based mental health prediction using data mining. In 2020 5th International Conference on Communication and Electronics Systems (ICCES) (pp. 1174-1178). IEEE.
8. https://www.health.harvard.edu/heart-health/what-your-heart-rate-is-telling-you