Application of Machine Learning in Analyzing the Psychological Well Being amongst the Employees in the Private Sector

Vinod Kumar
Research Scholar,
Amity Business School,
Amity University, Noida
vinod.2360@gmail.com

Vijit Chaturvedi
Professor,
Amity Business School,
Amity University,Noida
vchaturvedi@amity.edu
https://orcid.org/0000-0001-6757-1340

Babu Lal
Ex Director HR,
BHEL, Noida
lalsumang@gmail.com

Sohaib Alam
Assistant Professor,
Department of English,
College of Sciences and
Humanities in Al-Kharj,
Prince Sattam bin Abdulaziz University,
Kingdom of Saudi Arabia
s.alam@psau.edu.sa
https://orcid.org/0000-0002-9972-9357
Corresponding Author

Abstract

As a result of a wide range of influencing factors, the amount of stress experienced by Indian workers has grown, which is a major source of worry for companies. This research used various machine learning techniques to acquire its true data, and it was conducted with the participation of Indian working professionals. The goal was to better understand the elements that influence the psychological health of professionals. This research looks at things from two different angles. The first step is to recognise the significant causal factors that bring about stress in the lives of working professionals, as well as the second step is to anticipate the degrees of stress that they experienced. Experiments have been done using a variety of supervised machine learning methods. The KNN Algorithm has the best accuracy in comparison to Decision Tree Algorithm and the naïve Bayes Algorithm. This study predicted stress and psychological issues using machine learning.

Keywords: Psychological Well-being, Machine Learning (ML), Internet of Thing (IoT), Artificial Intelligence

Introduction

Problems of psychological wellbeing that are brought on by stress are unfortunately rather widespread among the working class. In the past, a number of researches have shown worry on the same issue. A survey that was conducted by an industry group called Assocham found that more than 42 percent of working individuals in the Indian private industry struggle from depression or general chronic anxiety as a direct result of the long work hours and stringent deadlines they are required to meet. According to a report that appeared in the Economic Times in 2018, which was derived from a study that was carried out by Optum, it was found that half of the working Indian professionals experience stress, this proportion of people is growing (Morel et al., 2020). The replies of as many as 8 lakh workers from around 70 significant organisations, each of which had a headcount of 4,500 or more people, were taken into consideration for the study. For the sake of increased employee productivity and general well-being, it is imperative that efforts be made

to keep the workplace as stress-free as possible.

It is possible to take a number of actions to assist working professionals in coping with stress in order to improve their mental health. These steps include providing counselling aid, career advice, stress management workshops, and health awareness initiatives. Increasing the likelihood of these types of initiatives being effective may be accomplished by early identification of workers who will need such assistance. Researchers hope to make this procedure easier by using machine learning techniques to construct a model that can predict the risk of excessive stress and if an individual needs treatment by taking some of the individual's personal and professional variables as parameters collected through the use of carefully drafted surveys (Reddy et al., 2018). In doing so, researchers will be able to determine whether or not an individual needs treatment for their stress. A strategy of this kind will not only assist HR managers in gaining a better understanding of their staff members, but it will also assist in the implementation of preventative steps that will reduce the likelihood of an employee quitting their job or underperforming. They are also able to undertake early prediction in order to determine whether or not a person needs care for their psychological health.

Figure 1 Machine learning for analysing employees psychological well being

The mental health of employees in relation to their jobs has emerged as a critical concern for companies all over the world. A Labour Force Survey that was endorsed by the government in the United Kingdom discovered that the overall number of instances of depression, stress, or anxiety connected to work in 2018 and 2019 was 602,000, which resulted in a loss of 12.8 million days of work (Glaz et al., 2021). According to the findings of a recent research that was sponsored by the WHO, depression and anxiety disorders are responsible for one trillion dollars' worth of lost productivity per year throughout the world.

The real kicker is that once they reach their full potential, artificial intelligence (AI) as well as machine learning (ML) have the prospects to be of great assistance in the workplace. The primary function of artificial intelligence in its present iteration is that of a support mechanism. When seen from the perspective of the future, however, it is possible that this will have a substantial influence on the mental health of the workforce.

Managing and Prioritising Workload

One of the factors that contributes to the mental health of workers in the workplace is the problem of working long hours and overtime. According to the findings of a recent survey conducted by Citrix in nine different nations, it is typical practise for 86 percent of office employees to put in extra hours outside of their chosen working hours (Srividya et al., 2018). A little less than half of these workers, or 47%, are putting in these extra hours each and every day or the most of the time.

Artificial intelligence is able to assist humans better prioritise their tasks in the midst of an epidemic of working extra as well as a mushrooming workload. After the smart machine has had the opportunity to learn about a person's employment function, it will be able to provide suggestions about the sequence in which they should undertake activities and projects, with a primary emphasis on the person's health and happiness. For instance, after completing a difficult or time-consuming piece of work, the framework may decide to suggest a less taxing and more enjoyable activity in order to assist the individual in unwinding and resetting their focus, or in order to guarantee that they will be able to leave the workplace on time.

Automation of Everyday Tasks

A lack of employee involvement is a further major cause of job-related mental health disorders. This is often the consequence of a job that is monotonous and demoralising, as well as work overload and increasing mental stress. A

recent study conducted by Gallup discovered that 85 percent of people working in organisations around the globe are either disengaged or not engaged (Cho et al., 2019). This means that employees are, at best, not supplying their innovative strategies or hard work, and, at worst, are actively resisting the organisation that they work in. The annual cost to enterprises is estimated to be $7 trillion due to these issues.

Notwithstanding, once the smart workspace has been customized to an individual's role and requirements, there is the possibility for the computer to take on elements of a person's everyday tasks, removing interruptions and mundane, unsatisfying work (Shafiee and Mutalib, 2020). This would allow the person to focus more on the tasks that bring them the most satisfaction.

It is crucial that the system learns these distinctions and only automates what is useful to the person since there may be minor variances or more visible variations between one employee and another. Not all tasks will be mechanized, and if a person appreciates certain lighter duties that are less stressful on their body and mind, they will have the option to remain doing such things.

Simplified Technology Experience

Because the IT experience of today is centred on apps and devices, there are restrictions placed on the ways in which an individual's digital workplace may be customised. This can lead to irritation and unneeded stress in certain cases. Nevertheless, as the level of complexity of AI technology increases, it will become more contextualised. This means that it will place more of its attention on the activity or result that the user is attempting to accomplish rather than on the program or tool that they are using.

In addition to this, the general experience of using technology will be made less complicated. The user will only be provided with a subset of microapps and applications that are relevant to the work at hand, and this will be done for them in a single interface that is personalised for them (Shinde and Rajeswari, 2018). All of this will aid toward lowering stress levels among employees that are tied to technology and preventing them from becoming so overwhelmed.

Monitoring Stress and Mental Health

Research is now being conducted to determine how artificial intelligence technology may be utilised to identify and diagnose mental health disorders. This has the potential to be useful in the workplace as well. In the not too distant future, AI systems will be able to monitor an employee's enunciation of words and tone of voice for signs of stress, identify linguistic cues that may predict depression, analyse spikes in heartrate, enhanced typing pace, or a slowdown in economic productivity, all of which may serve as warning signs for an employer. The use of biometric data or smart watches, such as facial expression analysis, for instance, might potentially be a factor in the timely diagnosis of mental health problems. With this knowledge, it would be possible to give activities that are proportionate to the amounts of stress experienced by a person.

Machine Learning Techniques Used

Machine learning is a subfield of AI that gives computers and other computing systems the capacity to learn on their own and develop based on their prior experiences without being specifically taught by a person. This ability is known as "unsupervised learning." The concept of machine learning revolves on the creation of computer systems that are able to retrieve data and educate themselves independently. This is especially useful in the medical field because there is a massive amount of data available; if this data is correctly inputted into an intelligent system and trained in the appropriate manner, the resulting prediction model will be unrivalled, free from human mistakes, and lessen the amount of time needed for diagnostics (Mohr et al., 2017). As a consequence of this, the answers included in the OSMI 2017 dataset were employed to train the following machine learning models, all of which have been validated via application to healthcare-related categorization issues.

Logistic Regression

The logistic regression approach is a kind of predictive analysis, much like other types of regression methods. It is put to use in situations in which the value of one binary variable is determined by the values of one or more independent factors. In this case, researchers consider the

14 pertinent characteristics to be independent variables, and then consider the likelihood of an employee experiencing stress and the consequent need for therapy to be a dependent variable that is to be forecasted by the trained model.

KNN Classifier

The K-Nearest Neighbor (KNN) classifier is an example of an algorithm for supervised learning that may be used to data that has been labelled. It was used in this setting for the purpose of determining whether or not a person need therapy (Kumar and Chong, 2018). KNN assigns a classification to the dependent variable based on the degree to which its independent variables resemble an existing instance that is analogous to one that is previously known.

Decision Trees

The K-Nearest Neighbor (KNN) classifier is an example of an algorithm for supervised learning that may be used to data that has been labelled. It was used in this setting for the purpose of determining whether or not a person need therapy. KNN assigns a classification to the dependent variable depending on the extent to which its independent variables resemble an existing instance that is analogous to one that is previously known.

Random Forest Classifier

Random Forests are a collection of decision trees that collaborate with one another to provide better results than those produced by a single decision tree could ever achieve on its own. Random Forest is a machine learning method that is adaptable (Tate et al., 2020), simple, and easy to use. It consistently generates decent results, even when it is not hyper tuned.

Boosting

In addition to this, scholars added ensembled approaches, which improve the performance of already existing models. Boosting is a very useful and widespread ensemble classifier that has a high level of accuracy. The primary goal of boosting is to mitigate the effects of bias introduced by the model.

Bagging

Bootstrap Aggregating is another powerful approach of ensemble construction (or bagging). It entails training a model by using the same technique but doing so on a variety of distinct subsets of the dataset's data (Garcia-Ceja et al., 2018). This aids in enhancing the model's accuracy and stability while also contributing to a reduction in the model's overall variance.

Research Methodology

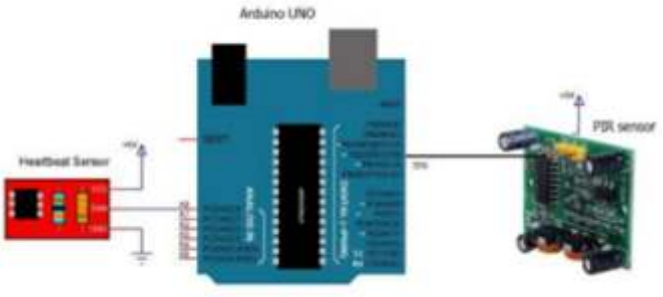
Proposed System

The algorithm identifies elements that have a substantial impact on a person's degree of stress. The employees' heart rates, temperatures, medical histories, as well as the accessibility of medical insurance at work were taken into consideration in order to estimate their stress levels. As can be seen in Figure 1, the primary purpose of the system is to determine the risk factors that have an effect on the mental health of the workforce.

Figure 2 Proposed System

The method takes into account a variety of characteristics, such as family history, gender, illness, colleague history, working hours, and much more. As seen in Figure 2, the system also collects real-time data using Internet of Things sensors to monitor things like pulse, temperature, and employee movement (Sandhya and Kantesaria, 2019). Some of the reference ranges that are taken into consideration for prediction include the following: a heart rate of =100 is regarded to be normal, but a heart rate of >100 in the absence of any symptoms is thought to be abnormal and is associated with stress.

Figure 3 Circuit Diagram



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

Machine Learning Techniques Used

Machine learning is a subfield of artificial intelligence that gives computers and other computing systems the ability to independently learn from their own history and progress over time without being explicitly programmed by humans. The foundation of machine learning is the process of building computer programmes that have the capacity to acquire information and learn on their own without any human intervention. In the healthcare business, where there is a significant deal of data, this may be of tremendous assistance when it is appropriately provided to a smart system and trained in an efficient manner (Su et al., 2020). The ultimate prediction model will be better, devoid of human mistakes, and optimised to reduce the amount of time spent on the diagnostic procedure.

Methodology

Real-time applications are a useful tool for businesses in the information technology industry, and this system may be configured to support them. For application development, they make use of Visual Studio and SQL Server due to the fact that these tools are more real-time and user-friendly respectively.

KNN Classifier: One of the methods of supervised learning that may be used with data that has been labelled is called the K-Nearest Neighbor (KNN) classifier. In this particular situation, it was used to determine whether or not the worker was experiencing stress. KNN assigns a classification to the dependent variable on the basis of the degree to which an instance from the existing data that is analogous to the dependent variable has similarities with the independent variables (Priya et al., 2020). A decision tree is a graphical representation of several choices, if-then statements, or other kinds of decisions that resembles a tree. It may be used to describe these kinds of things. In this particular scenario, decision trees are used in order to ascertain which 15 qualities contribute most often to the outcome (Joshi et al., 2018). Inexperienced Bayes- The Nave Bayes Classifier is both one of the simplest and one of the most successful classification approaches that may be used in the process of constructing machine learning systems that can predict outcomes. In this particular scenario, the Naive Bayes algorithm is used in order to compare the results and categorise the attribute values into one of the predetermined set of categories.

Data Collection

At this point in the technique for predicting employees' levels of stress, researchers collected data on stress. The data has been collected from a wide range of sources, and it takes into account factors such as age, gender, difficulties with money or family, working time, learning technique, health issues, colleague issue, partiality fix, pressure, frequent contact, and so forth.

Data Preparation

After doing a study on the stress data, researchers pulled just the relevant information. The specifications are used to determine which data must be collected and segmented in order for the processing to be successful. The processing only requires a subset of the data, so only the important data are extracted. Otherwise, the processing would take much longer than necessary since they have to input all of the data.

Data Splitting

At this stage, the data will be divided into a training dataset as well as a testing dataset in the proportion of ninety to ten percent.

Model Education

At this point, the machine learning algorithm is being taught by the datasets that are being given to it. Training the machine-learning model on a regular basis may significantly improve its accuracy of prediction. In order to train the model,

many different machine learning methods are used. In this particular model, researchers have employed.

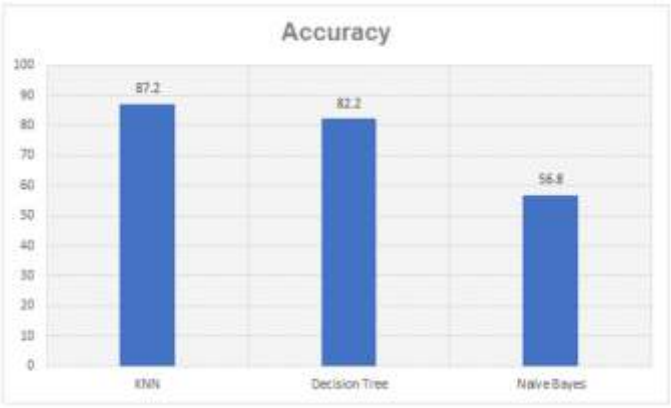
Prediction of Psychological Wellbeing

Following the training of the model, the system is able to make predictions about the levels of stress experienced by working employees by using the parameters and values provided by IoT sensors.

Result and Discussion

All of the aforementioned models were trained with the goal of predicting the stress that a working person would feel. The tabulated results may be seen in the following table. Figure 3 displays the classification performance of all trained models, revealing that KNN has the best accuracy (87.2%), while Naive Bayes has the lowest accuracy (56.8%).

Figure 4: Performance examination of each technique



Figu 4 depicts the amount of time needed to carry out the procedures for each model, with the KNN requiring the shortest amount of time.

Figure 5: Execution Time

The results of evaluating and tabulating the performance of a number of trained models are shown in Table 1. After being trained using a range of different machine learning strategies, the KNN model demonstrated superior results than those of the other systems in terms of precision as well as the amount of time required to complete the process.

Table 1. Performance of different models trained

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

Conclusion

Whether or not a company provides its employees with benefits for mental health is more relevant than other factors, including the health information of the employee's family, the amount of free time they have during the workday, and their connections with their colleagues. When trying to anticipate stress levels among working personnel, using real-time metrics adds another step. Those who work in the information technology industry are considerably more prone to feel stress than the general population due to the stringent deadlines and prolonged working hours. The collection of certain characteristics using certain Internet of Things sensors, such as pulse and temperature, will play a significant part in the process of anticipating stress. The KNN Algorithm obtained the best results, with an accuracy of 87.2%, out of the algorithms described above. In comparison, the Decision Tree Algorithm obtained an accuracy of 82.2%, while the Naive Bayes Algorithm obtained the least accuracy, with 56.8%. The objective of this work was accomplished by making use of methods from machine learning in order to make predictions about stress and psychological concerns. These methodologies result in noteworthy findings, which may then be examined further. In the not too distant future, one will be able to evaluate the models' precision by using a Convoluted Neural network (CNN) in conjunction with a variety of deep learning strategies. One may take into consideration a dataset that was made public by a few healthcare professionals or the surveys that were made public by a number of institutions. Building a model that is capable of providing a solution in the event that an

employee is experiencing certain stress attacks may be accomplished via the use of a variety of machine learning approaches.

Acknowledgements

Funding

"This study is supported via funding from Prince Sattam bin Abdulaziz University project number (PSAU/2023/R/1444)"

Authors' contributions

All authors contributed toward data analysis, drafting and revising the paper and agreed to be responsible for all the aspects of this work.

Declaration of Conflicts of Interests

Authors declare that they have no conflict of interest.

Data Availability Statement

The database generated and /or analysed during the current study are not publicly available due to privacy, but are available from the corresponding author on reasonable request.

Declarations

Author(s) declare that all works are original and this manuscript has not been published in any other journal.

References:

- Didier Morel, Kalvin C. Yu, Ann Liu-Ferrara, Ambiorix J. Caceres-Suriel, Stephan G. Kurtz and Ying P. Tabak,(July 2020) "Predicting hospital readmission in patients with mental or substance use disorders: A machine learning approach", International Journal of Medical Informatics, vol. 139, pp. 1-11. https://doi.org/10.1016/j.ijmedinf.2020.104136
- A Srinivasulu Reddy, Aditya Vivek Thota and A Dharun,(2018) "Machine Learning Techniques for Stress Prediction in Working Employees", IEEE International Conference on Computational Intelligence and Computing Research, pp. 1-4.
- A Le Glaz, Y Haralambous, D Kim-Dufor, P Lenca, R Billot, TC Ryan, (2021), "Machine Learning and Natural Language Processing in Mental Health: Systematic Review", J Med Internet Res, vol. 23, no. 5, pp. 1-20. https://doi.org/10.2196/15708
- M. Srividya, S. Mohanavalli and N. Bhalaji, "Behavioral Modeling for Mental Health using Machine Learning Algorithms", Journal of Medical Systems, vol. 42, no. 201842, pp. 88-12. https://doi.org/10.1007/s10916-018-0934-5
- Gyeongcheol Cho, Jinyeong Yim, Younyoung Choi, Jungmin Ko and Seoung-Hwan Lee,(, 2019) Review of Machine Learning Algorithms for Diagnosing Mental Illness, vol. 16, no. 4, pp. 262-269. https://doi.org/10.30773/pi.2018.12.21.2
- Nor Safika Mohd Shafiee and Sofianita Mutalib,(2020) "Prediction of Mental Health Problems among Higher Education Students Using Machine Learning", International Journal of Education and Management Engineering (IJEME), pp. 1-9. https://doi.org/10.5815/ijeme.2020.06.01
- Santosh Shinde and P.R. Rajeswari,(2018) "Intelligent health risk prediction systems using machine learning: A review", International Journal of Engineering & Technology, vol. 7, no. 3, pp. 1019-1023. https://doi.org/10.14419/ijet.v7i3.12654
- S Kumar and I Chong,(2018) "Correlation Analysis to Identify the Effective Data in Machine Learning: Prediction of Depressive Disorder and Emotion State", International Journal of Environmental Research and Public Health, vol. 15, no. 12, pp. 2907-24. https://doi.org/10.3390/ijerph15122907
- David C. Mohr, Mi Zhang and Stephen M. Schueller,(2017) "Personal Sensing: Understanding Mental Health Using Ubiquitous Sensors and Machine Learning", Annual Review of Clinical Psychology. https://doi.org/10.1146/annurev-clinpsy-032816-044949
- AE Tate, RC McCabe, H Larsson, S Lundström, P Lichtenstein and R Kuja-Halkola,(2020) "Predicting mental health problems in adolescence using machine learning techniques", PLoS ONE, vol. 15, no. 4, pp. e0230389. https://doi.org/10.1371/ journal.pone.0230389

- Enrique Garcia-Ceja, Michael Riegler, Tine Nordgreen, Petter Jakobsen, Ketil J. Oedegaard and Jim Tørresen,(2018) "Mental health monitoring with multimodal sensing and machine learning: A survey", Pervasive and Mobile Computing, pp. 1-26. https://doi.org/10.1016/j.pmcj.2018.09.003
- Chang Su, Zhenxing Xu, Jyotishman Pathak and Fei Wang,(2020) "Deep learning in mental health outcome research: a scoping review", Translational Psychiatry, vol. 10, no. 116, pp. 1-26. https://doi.org/10.1038/s41398-020-0780-3
- Deepali J. Joshi, Mohit Makhija, Yash Nabar, Ninad Nehete and Manasi S. Patwardhan,(2018) "Mental health analysis using deep learning for feature extraction", ACM India Joint International Conference on Data Science and Management of Data, pp. 356-359. https://doi.org/10.1145/3152494.3167990
- Anu Priya, Shruti Garga and Neha Prerna Tigga,(2020) "Predicting Anxiety Depression and Stress in Modern Life using Machine Learning Algorithms", Procedia Computer Science, vol. 167, pp. 1258-1267. https://doi.org/10.1016/j.procs.2020.03.442
- P Sandhya and Mahek Kantesaria,(2019) "Prediction of Mental Disorder for Employees in IT Industry", International Journal of Innovative Technology and Exploring Engineering (IJITEE).