National Research Journal of Information Technology & Information Science
Volume No: 11, Issue No: 1, Year: 2024 (January-June)
PP: 27-33
ISSN No: 2350-1278 (Print)
Peer Reviewed & Refereed Journal
National Research Journal

MACHINE LEARNING: A WAY FORWARD TO STRESS DETECTION

Shikha Verma

Assistant Professor, Department of Computer Science, Khalsa College for Women, Civil Lines, Ludhiana

Inderpreet Kaur

Assistant Professor, Department of Computer Science, Khalsa College for Women, Civil Lines, Ludhiana

ABSTRACT

Stress has become a significant aspect of everyone's life in today's competitive world. It has a wide range of direct and indirect effects and has become a major issue for a lot of individuals these days. The purpose of this research is to determine the causes of stress, anxiety, and depression that impact individuals from professional backgrounds. Our study seeks to characterize and identify how the development of technology has affected people's lives and stress levels. Modern ailments that impact a person physically, emotionally, cognitively, and socially may now be attributed mostly to stress. This article aims to educate people about early stress identification and treatment to prevent life-threatening situations using machine learning techniques. This paper suggests various supervised machine learning algorithms in detecting stress efficiently and effectively among a large number of people .Thus this paper will be useful in providing information about the use of machine learning techniques in the analysis of stress detection.

KEYWORDS: Artificial Intelligence, Machine Learning, Stress Detection

1. INTRODUCTION

Stress is a natural human response that prompts us to address challenges and threats in our lives. People from a variety of socioeconomic backgrounds have been impacted by stress in the ever-changing context of a nation like India. Everybody experiences stress and there are various reasons such as struggle for survival, maintaining a lifestyle, peer pressure etc. Stress is the body's reaction to the external stimuli or internal factors. It is a tense state of mind brought on by a challenging circumstance. Numerous things, including pressure from the workplace, money problems, relationship troubles, health concerns, or significant life events, can cause stress. The goal of machine learning-driven stress detection insights is to predict people's levels of stress. Predictive models use a range of data sources, including behavioural data, physiological measures, and environmental variables, to find trends and risk factors related to stress.

In the case of mock-up and questionnaire rounds, the antiquated methods of stress detection were dependent upon the individual's response. In the case of medical tests, the methods were dependent on the time at which the samples were obtained, making them prone to human error. These techniques were opaque and imprecise. As a result, in order to successfully address the issue of stress. These days, there are tools, techniques, and machine learning algorithms that can accurately, consistently, and precisely identify stress at its root cause and manage it at broader scope.

Machine learning, a branch of artificial intelligence (AI), allows computers to learn and grow on their own without requiring to be explicitly programmed. The focus of machine learning is on developing computer programs with the ability to explore data and acquire knowledge on

ISSN No: 2350-1278 (Print)
Peer Reviewed & Refereed Journal
National Research Journal

their own. In order to look for patterns in the data and utilize the examples to guide future decisions, the learning process begins with observations or data, such as examples, first-hand experience, or teaching. The basic idea is that computers should be able to learn without human assistance and adapt their behaviour accordingly.

2. LITERATURE REVIEW

The topic of stress detection has received much attention in health and wellness research. Automatic stress detection can become a useful tool to inspire people to change their behaviour and lifestyle to achieve a healthy stress balance. In recent years, machine learning models developed using physiological responses to stress and emotional stimuli have been used to automate stress prediction and detection. Researchers have proposed several ways to identify stress and analyse the conditions that lead to this problem. Many people use machine learning algorithms to study and analyse stress. Traditionally, stress has been identified through questionnaires, medical tests etc. These methods were very specific, targeted and people-centred, which in one way or another affected the aim and purpose of the measures. Those methods lacked transparency and accuracy. Therefore, modern technologies and methods were needed to effectively solve the stress problem. Today, there are machine learning techniques or algorithms that detect stress at a basic level and deal with it on a broader level with accuracy, consistency and precision.

ML-based techniques hold a lot of potential for mental diagnosis since they can use very objective methods to harness observable affect-related behaviours. Bayesian networks, decision trees, and SVM, among other methodologies, have been used by scientists. System and trained appropriately, the proposed system will be unparalleled and free of human errors, resulting in a reduction in diagnosis time (Narzary et al., 2022).

Stress detection and analysis in working environment using SVM classifier mainly aims to help the organization in identifying the stress caused by employees at different levels, so that the administrator can monitor their crew members, and recast them to improve the job satisfaction which enthuses them to be more productive and organization can have fruitful accomplishments (Swathi et al., 2022).

With the help of proposed method it is possible to identify people that have anxiety and depressive illnesses by utilizing prediction models to identify user language on social media, which has the potential to supplement conventional screening. Predictive models based on machine learning technique may provide the possibility to diagnose symptoms sooner, perhaps before psycho-social effects become serious (Kumari et al., 2022).

Some authors report the impact of different Machine Learning techniques for stress detection. The Performance analysis of three algorithms, namely Random forest, Logistic Regression and Decision Tree is performed and accuracy is obtained. When it comes to accuracy, it is observed that the Random forest algorithm outperforms more than the Logistic Regression and Decision Tree algorithms. Hence, by using the Random forest, the prediction of stress detection is performed and the result is obtained (Malika et al., 2023).

In recent years, one of the primary methods for identifying stress has been machine learning. Pre-processing, cleaning, and data collecting are steps in the machine learning process of detecting stress. In order to extract relevant data or generate new features that can record stress-related patterns, feature engineering approaches are applied. In order to gather physiological or behavioural markers of stress, this may entail time series analysis, frequency domain analysis, or statistical metrics extraction. To enhance performance, significant traits

ISSN No: 2350-1278 (Print)
Peer Reviewed & Refereed Journal
National Research Journal

are produced or extracted. They can assist in determining and classifying the person's mental state as stressed.

While there are several machine learning methods available for stress prediction, all strategies have the same fundamental application. When each method is employed independently, the results may differ in form or ability to be distinguished.

3. METHODOLOGY

While there are several machine learning algorithms available for stress prediction, the basic working of all the methods is the same. The final result's relevance remains constant throughout all scenarios, even if the structure and method of differentiation across stress levels may vary depending on the specific algorithm used. As a result, the actions listed below are used to identify stress while utilizing many machine learning algorithms. The following steps are executed for detecting stress while using multiple machine learning algorithms:

Step 1:

Input stress dataset:

The initial phase involves gathering stress data from all accessible sources, including surveys, fast test rounds, and performance-related data. The particular needs of the user are taken into account while creating attributes and class IDs for the database. Information about stress levels may be found in the "Stress" section. Quantitative measures such as age, blood pressure, heart rate, or stress level assessed on a scale may be represented by numerical variables in the dataset. Additionally, it may include categorical variables that reflect qualitative traits like gender, professional classifications, or stress levels that are divided into low, middle, and high categories.

Step 2:

Data pre-processing:

The data is checked for any unnecessary fields or attributes which do not form an essential part of stress detection process in the second stage. The process of transforming unstructured text data into a more organized and clean format that is appropriate for analytical purposes is known as text pre-processing. Specifically, it involves many stages to eliminate background noise, standardize text, and extract significant elements. Additionally, null, duplicate, and missing values are checked in the data. The tuples containing any type of the mentioned values will be removed from the dataset so as to avoid any illogical mistakes while processing the dataset for prediction.

Step 3:

Train-Test Split:

In the third step, the data is divided into training and test data. Generally, 30% of the data is used as test data and the remaining 70% is divided as training data. Using training data, the model can detect stress levels and set thresholds for different scenarios. Applying the trained model to test data allows detection of over fitting and under fitting problems. The machine is trained using the training data according to the data. The accuracy of the predictions is verified by test data.

ISSN No: 2350-1278 (Print)
Peer Reviewed & Refereed Journal
National Research Journal

Step 4:

Applying the stress detection model:

Depending on the type of user being evaluated, the desired algorithm is implemented in the fourth stage to determine the stress level based on the real-time attributes or features entered. To categorize stress levels, researchers use labelled data to train machine learning models such as logistic regression, SVM, decision trees or random forests.

Step 5:

Performance Evaluation:

The accuracy, precision, F-score, and other metrics are used to assess the model's performance in the last stage, which confirms the algorithm's correctness. Real-time stress monitoring is made possible by integrating the learned model into practical applications. In order to increase accuracy, user input, updates, and on-going monitoring are essential.

Stress detection/prediction methods in stress management using machine learning:

Stress has become a major cause of many modern diseases that directly or indirectly affect a person physically, emotionally, mentally and socially. Therefore, it is very important for us to develop or design systems, create methods and apply the right techniques in the right way to detect stress early to avoid this devastating threat before it gets out of our control.

4. CLASSIFICATION ALGORITHMS

There are various algorithms used to detect stress.. These are Decision tree, Random forest, Logistic Regression, Support Vector machine, KNN and Naïve Bayes Classifier:

4.1 DECISION TREE

Decision tree, also known as tree classifier, is the most widely used and efficient classification and prediction method. The tree structure of a decision tree is similar to a flowchart. Each leaf node is a class identifier. Each branch is a test result. Each internal node is a test for an attribute. Decision trees are one of the best and most useful guided methods. Decisions trees can be used for regression and classification tasks. Decisions trees are a tree based classifier. The internal nodes represent the characteristics of the dataset. The branches represent the decision procedures. The leaves represent the optimal outcomes. The root node, sub node, splitting node, and terminal node, sometimes known as the child node, make up a decision tree. Decision tree is represented as tree like structure. Each of the nodes are further divided into additional nodes. A decision tree has the nodes:

1. Root node: Root node represents the main node and it divides into so many sub nodes.
2. Splitting: The root node is split into one or more nodes.
3. Decision node: When sub node is further divided into sub node is called as decision Node.
4. Leaf node or Terminal node: Nodes do not split into terminal node
5. Branch node or sub node: The sub node of a complete tree is called as sub node.
6. Parent and Child Node: Parent node is a root node, root node is further split into sub node and that node is divided into sub node are called as leaf or child node

4.2 LOGISTIC REGRESSION ALGORITHM

One of the most popular machine learning techniques is logistic regression. Logistic

ISSN No: 2350-1278 (Print)
Peer Reviewed & Refereed Journal
National Research Journal

regression fits within supervised learning group. Logistic regression predicts the output of the dependent categorical variable by using a predefined set of independent variables. The output must therefore be a discrete value or a categorical value. Instead of providing an exact value from 0 to 1, it provides probabilistic values from 0 to 1. These probabilistic values can be True, False, Yes, or No, 0 to 1 etc. Logistic regression is a classifying procedure used to classify observations into discrete sets of classes. An example of a classifying issue is whether an email is spam or not.

4.3 RANDOM FOREST

Random Forest machine learning algorithm can be used for both regression and classification based machine learning problems. It is one of the best machine learning models for predictive statistics. The algorithm's strength lies in its ability to handle complex datasets and mitigate over fitting, making it a valuable tool for various predictive tasks in machine learning. In case of regression, it can handle the data set containing continuous variables and handles categorical variables in case of classification. It performs better for classification and regression tasks.

4.4 SUPPORT VECTOR MACHINE

A class of linear methods known as support vector machines (SVM) is useful for a variety of tasks, including regression, novelty detection, density estimation, and classification. SVM creates a prediction model through the use of classification algorithms. SVMs are a fantastic tool for using the process of processing data according to the appropriate classification approach to categorize stress into distinct levels according to severity, such as low, medium and high.

4.5 NAÏVE BAYES CLASSIFIER

NaïveBayes is one of the supervised learning techniques. Naïve Bayes is classified as a probability classifier and it is based on the bayes theorem. Naive Bayes Classification Naive Bayes algorithm is the algorithm that learns the probability of an object with certain features belonging to a particular group or class. Bayes theorem provides a method of calculating the posterior probability,

P(c|x), from P(c), P(x), and P(x|c).

Naive Bayes classifier assumes that the effect of the value of a predictor (x) on a given class (c) is independent of the values of other predictors. This assumption is called class conditional independence. P(x) is the Predictor prior probability

$P(c|x) = P(x|c)P(c)\ P(x)$

Where P(c|x) is the Posterior probability

P(x|c) is the likelihood

P(c) is the class prior probability

4.6 K-NEAREST NEIGHBORS:

An approach for regression analysis and classification is called K-Nearest Neighbors (KNN). It's a method for learning under supervision helps to assess if someone needs therapy. Based on how similar the dependent variable is to a comparable instance from previously gathered data, KNN classifies the data. KNN Classification is the statistical model with a binary dependent variable. In classification analysis, KNN is used to estimate the parameters of a KNN model. Mathematically, a binary KNN model consists of a dependent variable with two

ISSN No: 2350-1278 (Print)
Peer Reviewed & Refereed Journal
National Research Journal

possible values, denoted by the labels "0" and "1" on an indicator variable. Among the crucial methods for classification is K-Nearest Neighbor. K-Nearest Neighbor keeps track of every possible instance and uses a comparable metric to categorize new situations.

Working of K-Nearest Neighbor Algorithm:

K-Nearest Neighbours algorithm uses feature similarity to predict the values of new datasets which is having new data point will be assigned a value based on the points in the training set.

· Step 1: For implementing the algorithm, we need to collect the dataset during the first step of K-Nearest Neighbour and load the training as well as test data.

· Step 2: Choose the value of K i.e. the nearest data points and k can be any integer.

· Step 3: Calculate the distance between test data and each row of training data with the help of method is to calculate the distance.

· Step 4: Based on the distance values, sort them in ascending order

.· Step 5: Choose the k rows for from the sorted array.

· Step 6: Assign a class to the test point is based on the class of each row.

· Step 7: End

5. PERFORMANCE METRICS

Variety of performance indicators can be used to assess effectiveness of machine learning algorithms. Accuracy, precision, recall, and many more measures are among the numerous performance metrics. Choosing the appropriate Performance metrics will be based upon our chosen methodology and forecasts. Here, the performance criteria that we have chosen are f1-score, accuracy, precision, and recall. The percentage of accurate predictions that a strategy achieves is what is known as its accuracy.

The equation for accuracy is shown below:

Accuracy =True Positives + True Negatives

Total Population

The F1 score of a particular approach is defined as the harmonic mean of the precision and recall. The equation for the F1 score is shown below:

F1 Score = 2 . Precision · Recall

Precision + Recall

The equations for precision and recall are shown below:

Precision = True Positives

True Positives + False Positives

Recall = True Positives

True Positives + False Negatives

6. CONCLUSION

Applying machine learning techniques to predict stress levels offers personalized insights into mental well-being. By analysing multiple factors, such as numerical measurements (blood pressure, heart rate) and categorical characteristics (e.g. gender, occupation), machine

ISSN No: 2350-1278 (Print)
Peer Reviewed & Refereed Journal
National Research Journal

learning models can learn patterns and make predictions about individual stress levels. Machine learning can accurately detect and monitor stress levels, thereby helping to develop proactive strategies and interventions to manage and improve mental health of a person. The automation of stress detection using actual real time data from health care organizations and agencies, which can be built, using big data, will be useful. It can be fed as streaming data and by using the data, investigation of the patients in real time can be prepared. We can detect the stress by using the sensor. Hence, the study holds relevance in the context of ever increasing stress levels.

REFERENCES

Razavi, M., Ziyadidegan, S., & Sasangohar, F. (2022). Machine learning techniques for prediction of stress-related mental disorders: A scoping review. Journal of Mental Health Research. Retrieved from: https://journals.sagepub.com/doi/pdf/10.1177/1071181322661298

Mittal, S., Mahendra, S., Sanap, V., & Churi, P. (2022, August 24). How can machine learning be used in stress management: A systematic literature review of applications in workplaces and education. Journal of Applied Psychology. Retrieved from: https://www.sciencedirect.com/science/article/pii/S2667096822000532

Malika, S. R., & Ravi. (2023, March). Stress detection using machine learning techniques. Journal of Applied Machine Learning. Retrieved from: https://www.jetir.org/papers/JETIRFV06012.pdf

Hanchate, R., Narute, H., Shavage, S., & Tiwar, K. (2023, May). Stress detection using machine learning. Journal of Machine Learning Research.Retrieved from: https://www.irjet.net/archives/V10/i5/IRJET-V10I5105.pdf

Kumari, K. (2023, July 12). Machine learning unlocks insights for stress detection. Journal of Artificial Intelligence Research. Retrieved from: https://www.analyticsvidhya.com/blog/2023/06/machine-learning-unlocks-insights-for-stress-detection/

Gupta, T. (2023, December 27). Rise of mindfulness apps: AI's contribution to stress management. Journal of Technology and Wellness. Retrieved from: https://news.abplive.com/technology/rise-of-mindfulness-apps-ai-s-contribution-to-stress-management-1652680