**TABLE OF CONTENTS**

**CONTENT PAGE NO**

ABSTRACT 2

TABLE OF CONTENTS

1.INTRODUCTION 3

1.1.Domain Description 4

1.2 About project 5

1.2.1 Problem Definition

1.2.2 Proposed Solution

1.3 Objective 5

2.RELATED WORK

2.1 Data Collection 6

2.2 Machine Learning Algorithms 7-8

2.3 Sentiment Analysis 9

2.4 Evaluation metrics

2.4.1 Precision

2.4.2 Recall

2.4.3 Confusion Matrix

2.4.4 Enhancing Accuracy

3. ARCHITECTURE DESIGN 10-11

3.1 Attributes

3.2Steps to prevent the stress detection

4.Flow of any algorithm

4.1 Algorithms 12-14

4.1.1 Decision Tree

4.1.2 Random Forest

4.1.3 K-Nearest Neighbour

*5*.Result16-19

6.Conclusion 20

**ABSTRACT**

We are doing this major project on **Human Stress Level Detection: Machine Learning Classification for Detecting Human Stress Levels.**

Stress is recognized as a critical factor in human life, impacting mental health globally according to the World Health Organization. Many individuals worldwide suffer from stress, a pervasive symptom affecting emotions, cognitive functions, and behavior.

This study leverages machine learning to predict stress levels from physiological data . Using libraries like NumPy, Pandas, Matplotlib, and Scikit-learn, it employs Decision Tree, Random-Forest, and K-Nearest Neighbor classifiers, aiming to uncover patterns for personalized stress management interventions.

Stress is recognized as a critical factor in human life, impacting mental health globally Effective stress management is essential to mitigate its adverse effects on health and well-being. This study aims to predict stress levels using machine learning techniques applied to a comprehensive dataset comprising seven physiological attributes.

This study leverages machine learning to predict stress levels from physiological data . Using libraries like NumPy, Pandas, Matplotlib, and Scikit-learn, it employs Decision Tree, Random-Forest, and K-Nearest Neighbor classifiers, aiming to uncover patterns for personalized stress management interventions.

# **LIST OF FIGURES**

Fig No Name of Figure

1 Image for Machine learning

2 Architecture Design of the project

3 Flow of any algorithm

4 Flow of a Decision Tree algorithm

5 Flow of a Random Forest algorithm

6 Dataset list of the stress detection

7 Accuracy and the confusion matrix of Decision Tree

8 Accuracy and confusion matrix of Random- Forest

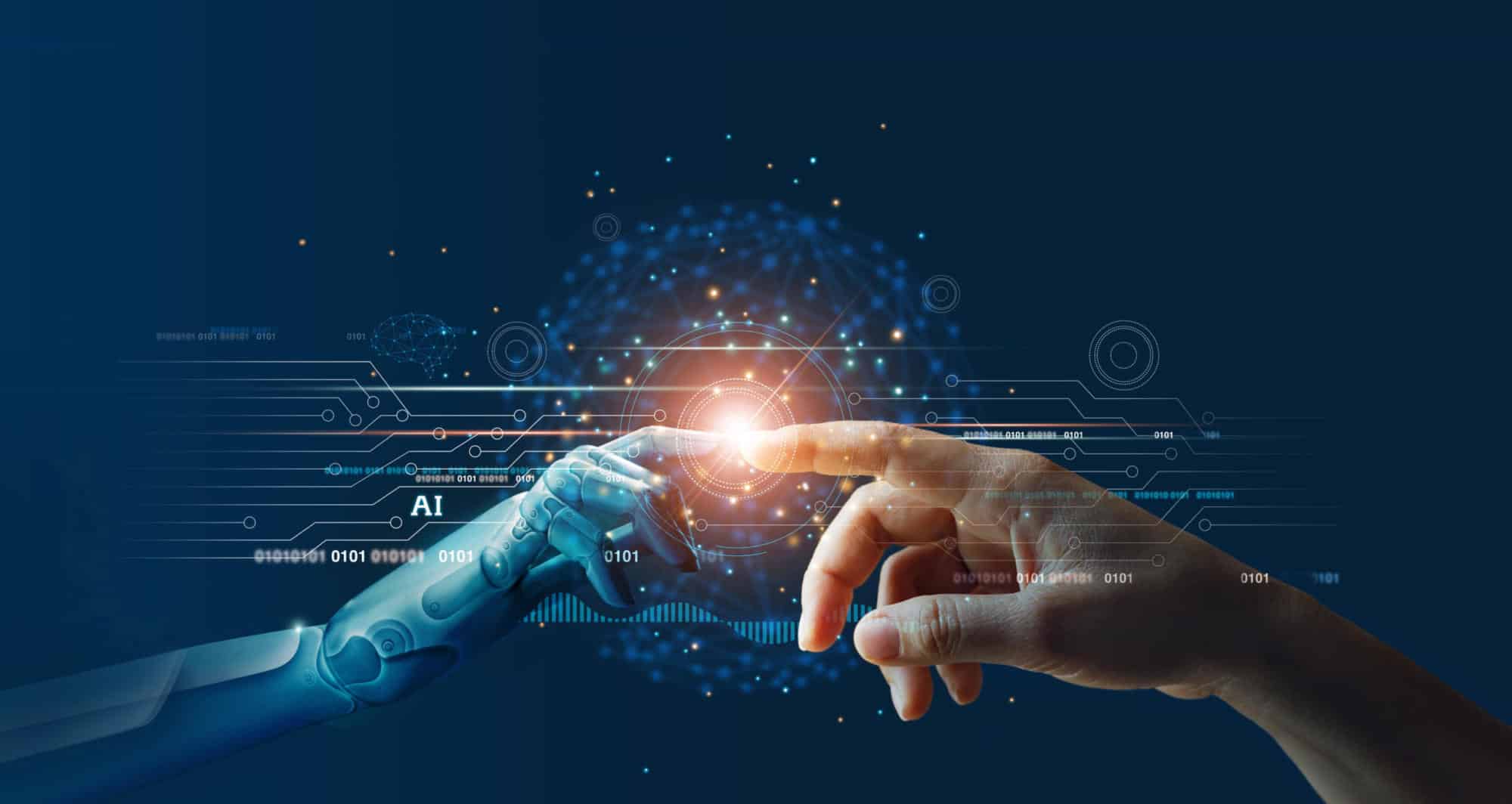
9 Accuracy and confusion matrix of KNN

**1.INTRODUCTION**

**1.1 DOMAIN DESCRIPTION**

**WHAT IS MACHINE LEARNING ?**

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.



**Fig1:Image for Machine learning**

Machine Learning (ML) is coming into its own, with a growing recognition that ML

can play a key role in a wide range of critical applications, such as data mining,Natural language processing, image recognition, and expert systems. ML provides potential solutions in all these domains and more, and is set to be a pillar of our future civilization.

“A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.” -- Tom Mitchell, Carnegie Mellon University

 **Some machine learning methods:**

Machine learning algorithms are often categorized as supervised or unsupervised.

**Supervised machine learning:**

Supervised machine learning algorithms can apply what has been learned in the past

to new data using labelled examples to predict future events.

**Unsupervised machine learning:**

Unsupervised machine learning algorithms are used when the information used to

train is neither classified nor labelled.

**1.2 ABOUT PROJECT**

**1.2.1 PROJECT DEFINITION:**

Stress is one of the predominant symptoms in human life, playing a significant role in affecting people's health and well-being. It can lead to numerous diseases such as cardiovascular problems, respiratory issues, cancer, and other ailments. The increasing global population has contributed to rising stress levels among individuals, making stress a common issue in today’s society.

Stress can cause life-threatening problems worldwide. When individuals suffer from stress, their normal behavior changes, leading to disruptions in their daily routines. Continuous monitoring of their behavior using machine learning algorithms allows us to gather real-time data, accurately analyzing their condition. In India, a staggering 85% of people experience stress, often leading to depression.

A stress detection system can differentiate between normal and stressed behavior. Physiological devices continuously monitor stressed individuals, providing valuable parametric data for doctors to predict their stress levels. These devices store data in a database for further analysis, enabling doctors to determine whether a person is stressed or not based on the collected values.

**1.2.2 PROPOSED SOLUTION**

The proposed system uses physiological devices to monitor daily values, storing this information in a database. Doctors can then use this weekly or monthly data to diagnose diseases and determine the stress level in individuals. This is particularly important for people over 60 years, who often suffer from multiple diseases. The system helps in predicting stress and other ailments, facilitating early treatment based on the recorded data.

Our project employs machine learning classification algorithms to predict whether a person is exhibiting normal or stressed behavior. Stress is defined as a state where an individual feels overstrained, uncomfortable with tasks, sits silently, and is preoccupied with thoughts. This altered behavior affects relationships and overall happiness. Consulting a doctor is essential for individuals experiencing stress or depression. Doctors can assess the stress level and recommend further analysis and treatment from neurologists and psychiatrists.

Early detection of stress can lead to effective treatment, but chronic stress can result in severe health issues like heart disease, cancer, and asthma. Stressed individuals often isolate themselves and are constantly absorbed in their thoughts. The proposed system aims to provide a comprehensive solution for stress detection and management using advanced machine learning techniques.

The remainder of this document is structured as follows: Section II discusses related work and various schemes proposed for stress detection. Section III details the experimental approach for stress detection using machine learning. Section IV presents the experimental results and compares different stress detection models. Finally, Section V concludes the paper.

Impact of Stress on Health:

Stress affects individuals in multifaceted ways, leading to both physical and mental health issues. Prolonged stress can disrupt the immune system, making individuals more susceptible to illnesses. It can exacerbate conditions like hypertension and diabetes, and also contribute to mental health disorders such as anxiety and depression. Understanding the far-reaching impact of stress is crucial for developing effective detection and intervention strategies.

**1.3 Objective**

Machine learning algorithms are instrumental in analyzing complex data patterns to predict stress levels. By leveraging large datasets, these algorithms can identify subtle changes in behavior and physiological responses that indicate stress. Decision trees, k-nearest neighbors (k-NN), and random forests are among the algorithms used to analyze data and determine stress levels. These algorithms can provide accurate predictions, which are essential for timely intervention and treatment.

**Physiological Monitoring Devices**

Physiological devices, such as wearable sensors, play a critical role in stress detection. These devices monitor various parameters like heart rate, skin conductance, and cortisol levels, which are indicative of stress. Continuous monitoring provides real-time data that can be analyzed using machine learning models to assess the stress level of an individual. This data-driven approach enables proactive healthcare, allowing for early intervention and better management of stress-related health issues.

**Data Analysis and Storage**

The data collected by physiological devices is stored in a database, creating a comprehensive record of an individual’s health over time. This historical data is invaluable for doctors, enabling them to track trends, identify triggers, and provide personalized treatment plans. The integration of data storage and machine learning analysis creates a robust system for ongoing health monitoring and stress management.

**Benefits for Healthcare Professionals**

For healthcare professionals, the proposed system offers several benefits. It provides a continuous stream of accurate data, enabling better diagnosis and treatment. Doctors can use this information to tailor interventions, monitor treatment efficacy, and adjust care plans as needed. This data-driven approach enhances the quality of care and improves patient outcomes.

**Future Directions**

The proposed system represents a significant advancement in stress detection and management. Future work could focus on integrating additional machine learning algorithms, exploring new physiological markers, and enhancing the user interface for better accessibility. Expanding the system to include predictive analytics for other health conditions could further improve its utility and impact.

**Further Sections**

Section II: Related Work

Discuss various existing schemes and methodologies proposed for stress detection, including their strengths and limitations.

Section III: Experimental Approach

Detail the experimental setup, including the data collection process, machine learning algorithms used, and evaluation metrics.

Section IV: Results

Present the results of the experiments, comparing the performance of different machine learning models in stress detection.

Section V: Conclusion

Summarize the key findings, discuss the implications of the research, and suggest future directions for further improvement.

By expanding and detailing each section, the document provides a comprehensive overview of the stress detection system, its methodology, and its potential impact on healthcare.

**2.Related Work**

The concept of a stress detection system leveraging machine learning algorithms has gained significant traction in recent years, offering a promising avenue for healthcare professionals to predict and manage stress levels. Researchers have been actively working on developing methodologies to accurately detect stress using various datasets and machine learning techniques.

**2. 2 Data Collection**

One of the primary sources of data for stress detection comes from social media platforms. Websites like Kaggle provide extensive datasets that are invaluable for research purposes. Social media platforms such as Facebook, Twitter, and Instagram are particularly useful because individuals often express their emotions, including stress and depression, through posts, quotes, images, and discussions. When people experience stress or depression, they tend to share their feelings on these platforms, providing a rich source of data for analysis.

**2.2 Machine Learning Algorithms**

Researchers have employed various machine learning algorithms to analyze these datasets and predict stress levels. Prominent algorithms used in this context include Decision Trees, k-Nearest Neighbors (k-NN), Random Forest.

**Decision Tree:** A Decision Tree is a flowchart-like structure used for classification and regression tasks. Each internal node represents a feature, each branch represents a decision rule, and each leaf node represents the outcome. Decision Trees are easy to interpret and visualize, making them useful for understanding the decision-making process of the model. However, they can be prone to overfitting, especially with small datasets.

**k-Nearest Neighbors (k-NN):** k-NN is a simple, non-parametric algorithm used for classification and regression. It classifies a data point based on the majority class of its k-nearest neighbors. k-NN is easy to implement and understand, but it can be computationally expensive, especially with large datasets. In the context of stress detection, k-NN has shown to provide accurate results by considering the similarity between social media posts.

**Random Forest:** Random Forest is an ensemble learning method that operates by constructing multiple decision trees during training and outputting the mode of the classes for classification tasks. It improves accuracy and controls overfitting by averaging multiple decision trees. Random Forests are robust and provide high accuracy, making them suitable for stress detection.

**2.3 Sentiment Analysis**

Sentiment analysis plays a crucial role in classifying stress and depression from social media posts. This technique involves analyzing the emotional tone of the text to determine whether it reflects positive, negative, or neutral sentiments. In the context of stress detection, sentiment analysis helps identify posts that convey stress, anxiety, or depressive thoughts. Researchers use various sentiment analysis techniques to preprocess the data, extract relevant features, and feed them into machine learning models for classification.

**2.4 Evaluation Metrics**

To assess the performance of the stress detection models, researchers use several evaluation metrics, including precision, recall, and the confusion matrix.

**2.4.1 Precision:** Precision measures the accuracy of the positive predictions made by the model. In stress detection, it indicates the proportion of posts correctly identified as stress-related out of all posts predicted as stress-related.

**2.4.2 Recall:** Recall measures the model's ability to identify all relevant instances. In this context, it indicates the proportion of actual stress-related posts that were correctly identified by the model.

**2.4.3 Confusion Matrix:** The confusion matrix provides a detailed breakdown of the model's performance, showing the true positives, true negatives, false positives, and false negatives. It helps researchers understand the model's strengths and weaknesses and identify areas for improvement.

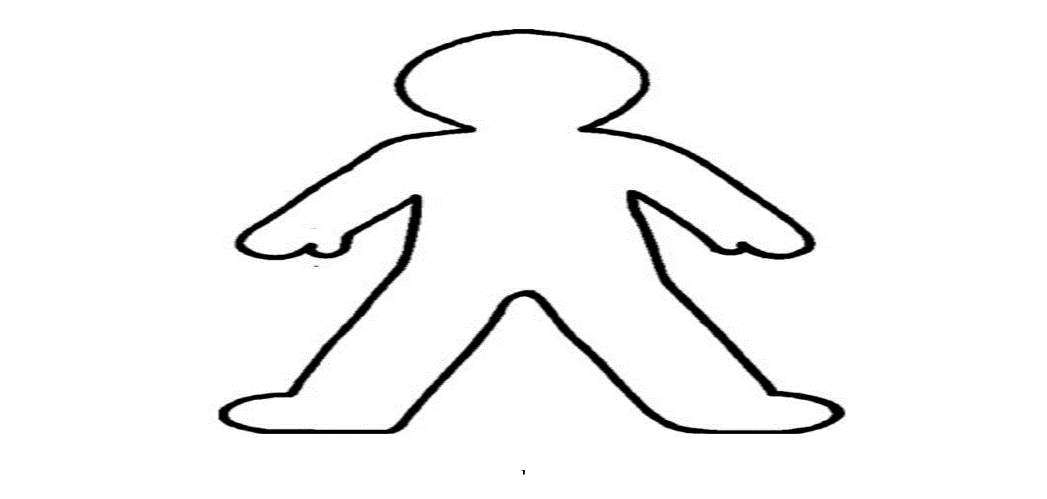
**2.4.4 Enhancing Accuracy**

To improve the accuracy of stress detection, researchers often combine multiple techniques and algorithms. By using an ensemble of methods, they can achieve better results and more robust predictions. Techniques such as feature engineering, hyperparameter tuning, and cross-validation are employed to fine-tune the models and enhance their performance The integration of machine learning algorithms with social media data offers a powerful approach to stress detection. By analyzing the emotional content of social media posts, researchers can develop models that accurately predict stress levels. The use of evaluation metrics such as precision, recall, and the confusion matrix ensures that these models are reliable and effective. As research in this area continues to evolve, it holds great potential for improving mental health monitoring and providing timely interventions for individuals experiencing stress and depression.

In summary, the related work in stress detection demonstrates the potential of machine learning algorithms and sentiment analysis in identifying stress from social media data. By leveraging these techniques, researchers can develop accurate models that aid healthcare professionals in predicting and managing stress levels, ultimately contributing to better mental health outcomes.

**3. Architecture Design**

Stress detection is one of the important topics in machine learning. Due to the stress the people will not happy and there will not doing their work correctly. Based the values we can easily predict the values Stress detection is an increasingly vital area of research in machine learning, particularly given the pervasive nature of stress in modern life. Understanding and identifying stress can significantly enhance people's well-being, productivity, and overall quality of life. Machine learning offers robust tools to analyze and predict stress levels based on various data inputs, such as physiological signals, social media activity, and behavioral patterns.



Model using different machine learning algorithms

0-(low/normal)

1-(medium low)

2-(medium)

3-(medium high)

4-(high)

**Stress Levels**

Snoring range

Respiratory rate range

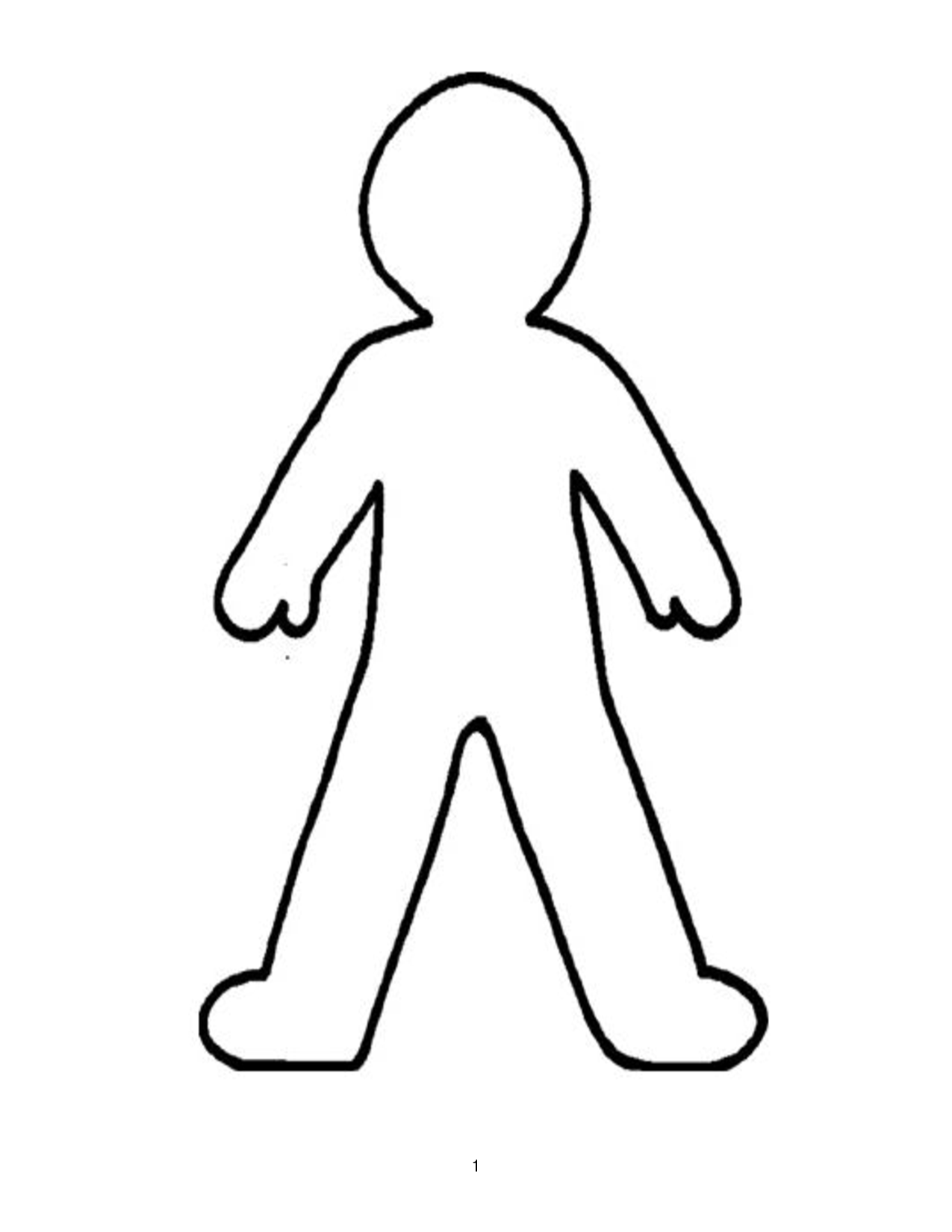
Heart,limb rate range

Oxygen in blood range

Eye movement rate

Body temperature

**Insert Details**



*1.Snoring range*

**Fig 2: Architecture Design of this project**

**3.1 ATTRIBUTES**

*1.Snoring rate*

Snoring affects about 40% of adult men and 24% of adult women, often caused by obstructed airflow due to factors like obesity, nasal congestion, or alcohol consumption.

2.*Respiratory rate range*

A normal respiratory rate for a healthy adult at rest is 12 to 20 breaths per minute; deviations can indicate stress or respiratory/metabolic issues, with increased rates often linked to anxiety or panic.

3.*Heart rate range*

The normal resting heart rate for adults is 60 to 100 beats per minute; during stress, heart rate often increases as the body responds to stressors, signaling heightened arousal and the "fight or flight" response.

4.*Limb rate range*

Limb movement rates, such as fidgeting or restlessness, often stress as the body attempts to release nervous energy, indicating heightened anxiety or tension.

*5.Oxygen in blood range*

Blood oxygen levels typically range from 95-100%; stress can cause rapid breathing, potentially leading to hyperventilation and altered oxygen and carbon dioxide balance, affecting overall oxygenation.

*6.Eye movement rate*

Eye movement rate often increases during stress, with rapid or erratic movements indicating heightened alertness or anxiety. These changes can be part of the body's fight-or-flight response, reflecting increased cognitive and sensory processing.

*7.Body temperature*

Stress can cause body temperature to fluctuate, often resulting in sweating or chills as the body's autonomic nervous system responds. These changes are part of the fight-or-flight response, preparing the body to handle perceived threats.

**3.2 Steps to prevent the stress detection**

• Step 1: Data collection means collecting the data in my project I have use the statistical datasets where I have considered the seven attribute based on these attributes we will collect the data.

• Step 2: Data pre-processing where we will be cleaning the data and also see the any missing data is available in the values are to be checked.

• Step 3: Feature extraction is done by the values where in the values we have encoded the values means the original data is encoded into unreadable format.

• Step 4: We have extract the features by some of the algorithms are decision tree,Random-forest and KNN that value predicts the person is in stress or not stress , K-Nearest Neighbour is used for classification technique and predicts the nearest neighbour in the data.

• Step 5: The dataset is divided into training dataset and testing dataset, always the training dataset is more than the testing datasets.

• Step 6: Based on the trained value we will test the result as 0 as not stressed and 1 is stressed. When we consider the trained value we have to compare with the threshold value based on these values we can predict and detect the person is in stress or not stress.

**4.FLOW OF ANY ALGORITH**

Steps in the Flow of Any Machine Learning Algorithm

1. \*\*Problem Definition\*\*:

- Clearly define the problem you are trying to solve. Understand the business context, objectives, and constraints. Determine the type of problem (e.g., classification, regression, clustering).

2. \*\*Data Collection\*\*:

- Gather data from various sources relevant to the problem. This could involve scraping data from websites, extracting information from databases, or using pre-existing datasets.

3. \*\*Data Cleaning\*\*:

- Clean the collected data by handling missing values, removing duplicates, and correcting inconsistencies. This step ensures the quality and reliability of the data.

4. \*\*Data Exploration and Visualization\*\*:

- Perform exploratory data analysis (EDA) to understand the patterns and relationships in the data. Use visualizations like histograms, scatter plots, and box plots to gain insights and identify trends or outliers.

5. \*\*Feature Selection and Engineering\*\*:

- Select relevant features that contribute to the prediction. Create new features by combining or transforming existing ones to improve the model’s performance. This step might include normalization, encoding categorical variables, and scaling.

6. \*\*Splitting the Dataset\*\*:

- Split the dataset into training and testing sets. The training set is used to build the model, while the testing set evaluates its performance. Typically, an 80-20 or 70-30 split is used.

7. \*\*Choosing the Algorithm\*\*:

- Select an appropriate machine learning algorithm based on the problem type and data characteristics. Common algorithms include linear regression, decision trees, k-nearest neighbors (KNN), random forest, support vector machine (SVM), and neural networks.

8. \*\*Training the Model\*\*:

- Train the selected algorithm using the training data. This involves feeding the data into the algorithm and adjusting its parameters to minimize errors and improve accuracy.

9. \*\*Model Evaluation\*\*:

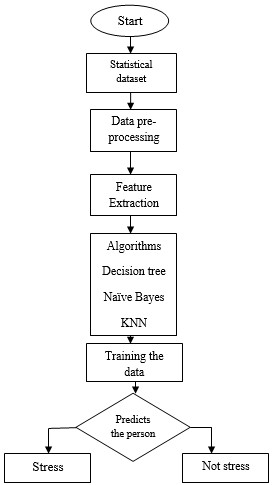
- Evaluate the model’s performance using the testing data. Common evaluation metrics include accuracy, precision, recall, F1 score, mean squared error (MSE), and area under the curve (AUC). Use cross-validation techniques to ensure the model’s robustness.

10. \*\*Hyperparameter Tuning\*\*:

- Optimize the model’s performance by tuning hyperparameters. Use techniques like grid search or random search to find the best combination of hyperparameters that yield the highest accuracy or lowest error.

11. \*\*Model Deployment\*\*:

- Once satisfied with the model’s performance, deploy it in a production environment. This involves integrating the model into an application or system where it can make real-time predictions.

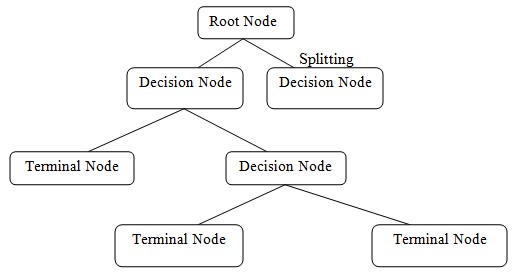


**Fig 3: Flow of any algorithm**

# ***4.1 Algorithms***

# ***1) Decision Tree***

Random Forest is a widely-used machine learning algorithm that falls under the category of supervised learning. It is versatile, capable of handling both classification and regression tasks with a high degree of accuracy. The foundation of Random Forest lies in the concept of ensemble learning, which aims to improve the performance and robustness of models by combining multiple weak learners to form a strong learner.



**Fig 4:Flow of a Decision Tree algorithm**

A decision tree as further constituents is:

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 splits into sub node and that node is divided into sub node are called as leaf or child node.

**Decision Tree Advantages**

1. **Easy to Understand and Interpret**:
   * Decision trees are straightforward to understand and interpret, with a clear visual representation that makes the decision-making process transparent.
2. **Handles Both Numerical and Categorical Data**:
   * They can work with both numerical and categorical data without requiring much preprocessing.
3. **Feature Importance**:
   * Decision trees highlight which features are most important, providing valuable insights into the data.

**Decision Tree Disadvantages**

1. **Prone to Overfitting**:
   * They can easily overfit the training data, especially if the tree is allowed to grow too complex.
2. **High Variance**:
   * Small changes in the training data can lead to very different trees, resulting in high variance.

2*.****Random-Forest Algorithm***

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of **ensemble learning*,****which is a process of*combining multiple classifiers to solve a complex problem and to improve the performance of the model. supervised learning techniques. It is versatile and can be applied to both classification and regression problems, making it a valuable tool for a broad range of applications. The foundation of Random Forest lies in the concept of ensemble learning, a powerful method that enhances the performance of models by aggregating the predictions of multiple classifiers.

Ensemble learning leverages the strength of various individual models to produce a collective output that is often more accurate and robust than any single model on its own. In the case of Random Forest, this is achieved by creating a multitude of decision trees during the training phase. Each tree in the forest is constructed using a random subset of the training data and a random subset of features. This randomness introduces diversity among the trees, which helps in capturing different patterns in the data and reducing the risk of overfitting.

Once the forest of decision trees is built, the Random Forest algorithm combines the predictions of these individual trees to make a final decision. For classification tasks, it employs a majority voting mechanism, where each tree votes for a class, and the class with the most votes is chosen as the final prediction. For regression tasks, the algorithm takes the average of the predictions from all the trees. This aggregation process ensures that the final model is less sensitive to the noise and biases present in the individual trees, resulting in improved generalization and performance.

Random Forest Advantages

1. \*\*Reduced Overfitting\*\*:

- Random forests reduce the risk of overfitting compared to individual decision trees by averaging multiple trees, which improves generalization.

2. \*\*High Accuracy\*\*:

- They generally offer high accuracy and robust performance on a wide range of tasks, benefiting from the ensemble approach of combining multiple decision trees.

3. \*\*Handles Missing Data\*\*:

- Random forests can handle missing data well by maintaining the overall performance, even when some data points are missing.

Random Forest Disadvantages

1. \*\*Complexity and Interpretability\*\*:

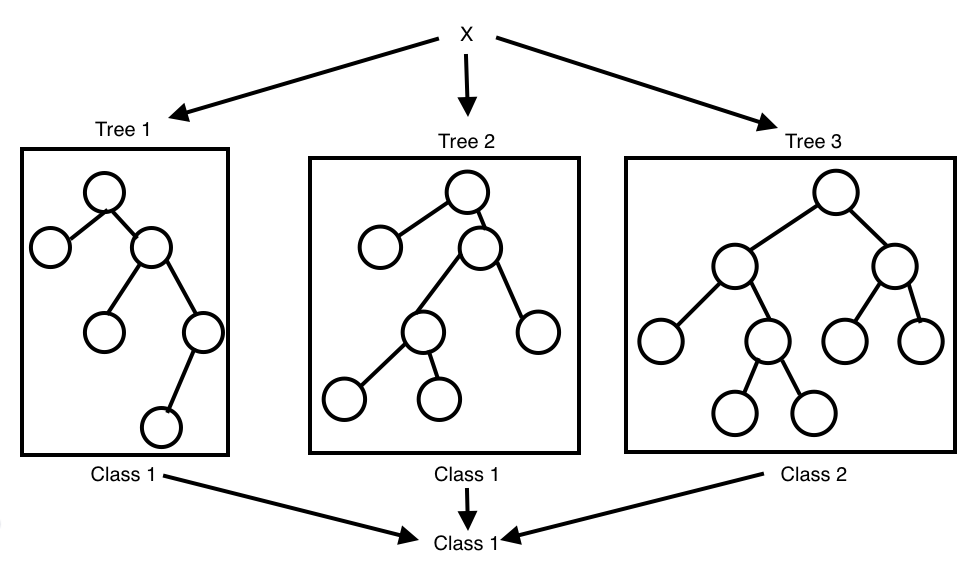
- The model can be complex and less interpretable compared to a single decision tree due to the ensemble of many trees.

2. \*\*Computationally Intensive\*\*:

- Training and predicting with random forests can be computationally intensive and require more memory, especially with large datasets.

3. \*\*Less Effective for Extrapolation\*\*:

- Random forests are less effective at extrapolating beyond the range of the training data since they are based on averages of many trees, which can smooth out extreme values.



**Fig 5:Flow of a Random Forest algorithm**

**Step-1:** Select random K data points from the training set.

**Step-2:** Build the decision trees associated with the selected data points

**Step-3:** Choose the number N for decision trees that you want to build.

**Step-4:** Repeat Step 1 & 2.

**Step-5:** For new data points, find the predictions of each decision tree, and assign the new data points to the category that wins the majority votes.

# ***3) K-Nearest Neighbour***

K-Nearest Neighbor is one of the important classification algorithms. K-Nearest Neighbor stores all the feasible cases and classifies new cases based on a similar measure of the data. KNearest Neighbor measures the nearest neighbour by the surrounding. K-Nearest Neighbour is divided into classification and regression. The neighbours are taken from a group of objects.

*Working of K-Nearest Neighbour 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 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.

Advantages of k-Nearest Neighbors (k-NN) Algorithm

1. \*\*Simplicity and Intuition\*\*:

- k-NN is simple to understand and easy to implement. It relies on basic concepts of distance and nearest neighbors, making it intuitive for many applications.

2. \*\*No Training Phase\*\*:

- k-NN is a lazy learning algorithm, meaning it does not require a training phase. All computations are deferred until classification, which can be advantageous when training time is a constraint.

3. \*\*Versatility\*\*:

- k-NN can be used for both classification and regression tasks, making it a versatile algorithm for various types of problems.

## 4. Result

The results of the stress detection where the data set is collected from the website where collect the statistical data and also consider the attributes for the dataset. Based on the statistical data train the dataset and compare with the threshold values and also use the classification algorithm from the machine learning techniques. Based on the threshold values the doctors will predict the person is in stress or not stress.

Stress detection is an important area of study where datasets are often collected from various sources, including websites that gather statistical data. These datasets typically contain attributes relevant to stress levels, such as heart rate, sleep patterns, physical activity, and other physiological or behavioral indicators. By compiling and organizing this data, researchers and practitioners can create a comprehensive dataset suitable for analysis.

Once the dataset is collected, the next step involves training the dataset using machine learning techniques. This training process involves using the collected data to build a model that can accurately predict stress levels. Various machine learning algorithms can be employed for this purpose, such as decision trees, support vector machines, or neural networks. These algorithms analyze the relationships and patterns within the data to learn how different attributes contribute to stress levels.

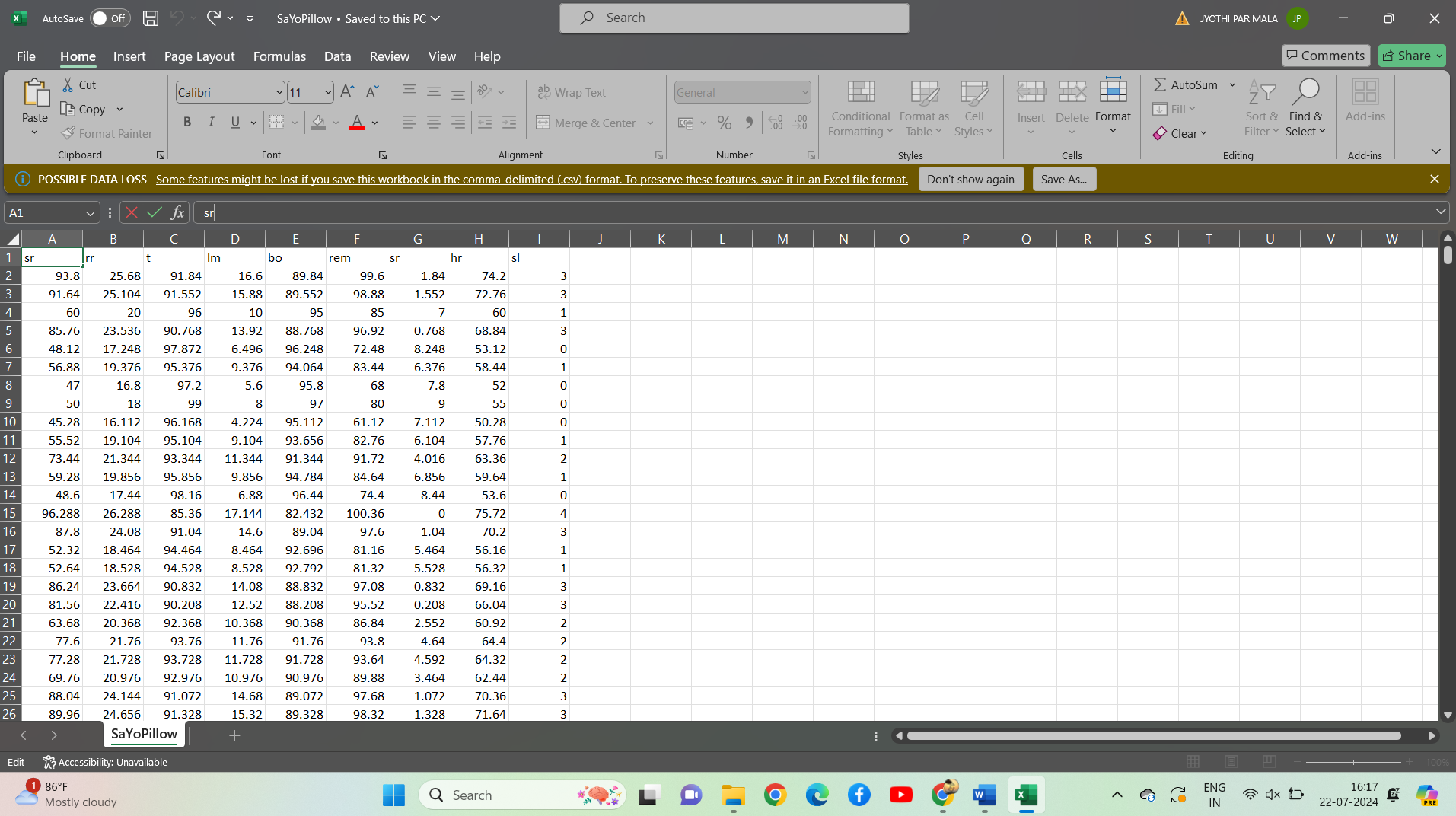
After training the model, it is essential to compare the results with predefined threshold values. These threshold values act as benchmarks to determine whether a person's stress level is above or below a critical point. By comparing the model's predictions with these thresholds, it becomes possible to classify individuals as either stressed or not stressed. This classification process involves analyzing the output of the machine learning model and categorizing it based on the threshold criteria.

The classification results can then be utilized by healthcare professionals, such as doctors, to make informed decisions about a person's stress level. If the model indicates that a person's stress level exceeds the threshold, doctors can intervene and provide appropriate care or recommendations. Conversely, if the stress level is below the threshold, it may indicate that the person is not experiencing significant stress. By leveraging machine learning techniques and threshold-based classification, doctors can make more accurate and timely predictions about stress, ultimately improving patient outcomes and well-being.

Accuracy measured from the datasets and accuracy measured for normalized dataset are 100% accuracy got in Decision tree, In Naïve- Bayes algorithm and K-Nearest Neighbor the results shows the classification accuracy, classification error, sensitivity, specificity, False positive rate and Precision of the dataset.

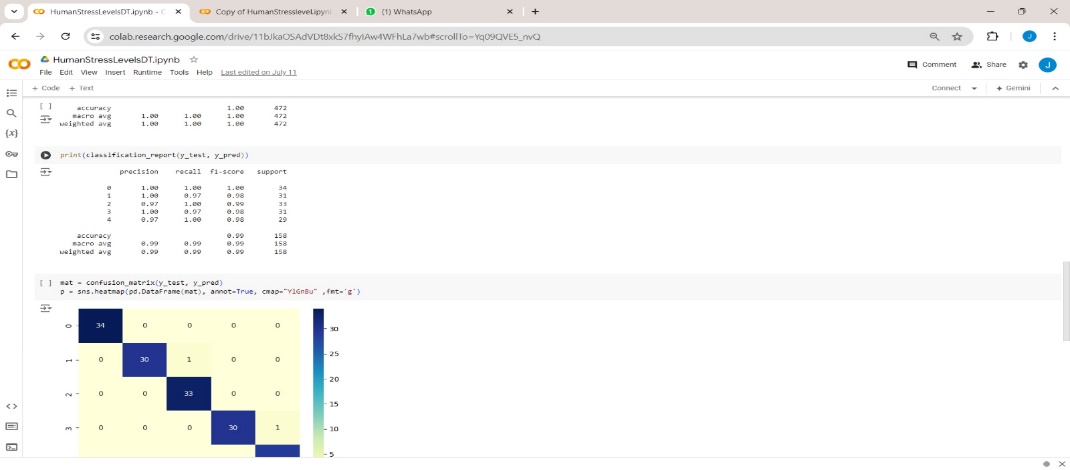
The predicted values for stress detection taken with the help of classification machine learning algorithms and get the values from the confusion matrix.

* Classification accuracy – 1.0
* Classification Error – 0.0
* Sensitivity – 1.0
* Specificity – 1.0
* Precision – 1.0
* False Positive Rate – 0.0
* True Target values – [ 0 1 0 0 0 1 0 0 1 0 0 0 1 1]
* Predicted Target value – [ 0 1 0 0 0 1 0 0 1 0 0 1 1]
* Accuracy measured for normalized dataset – 100%



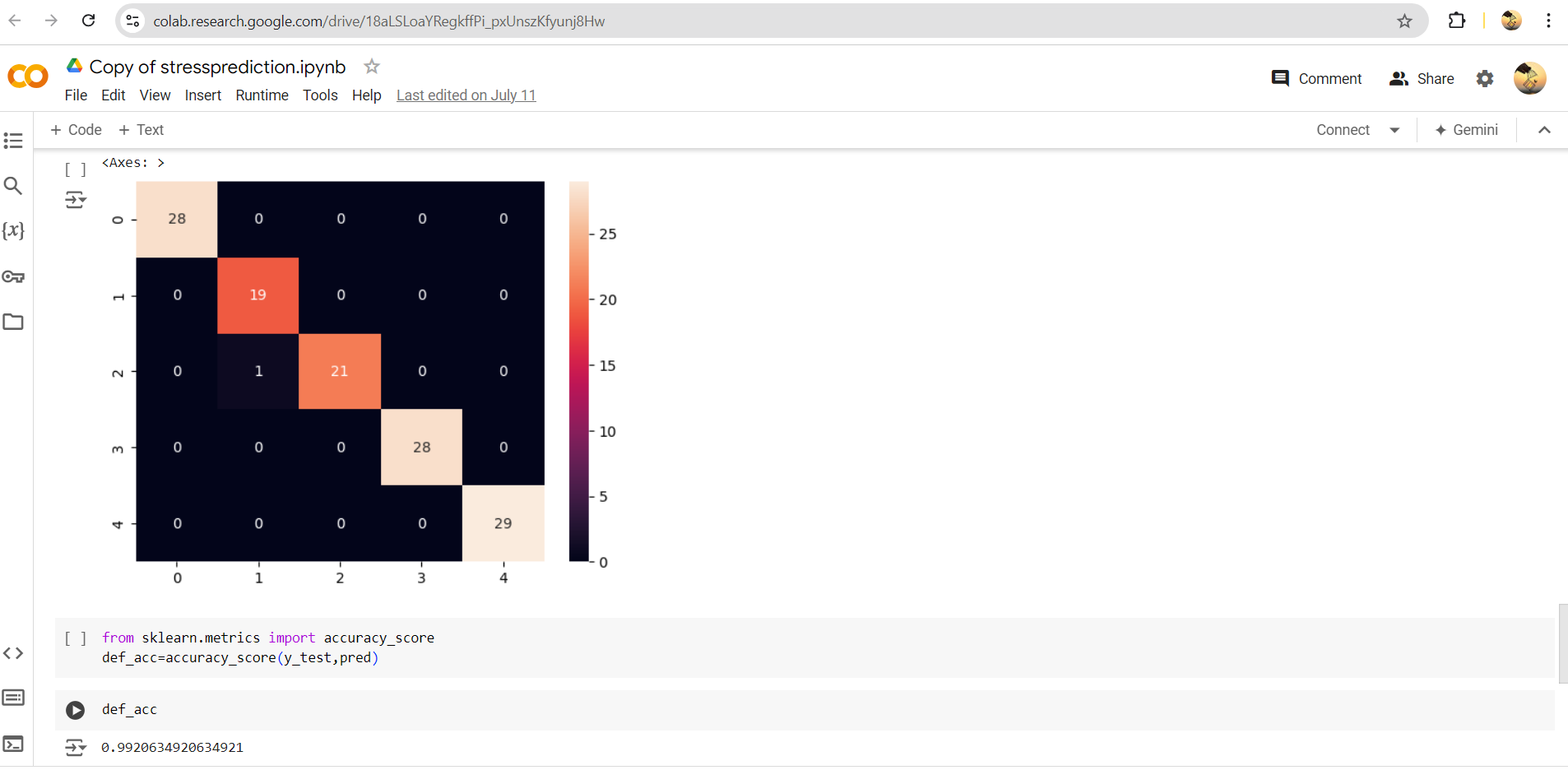
**Fig 6:** **Dataset list of the stress detection**

The above figure shows that the dataset having the values where we have considered the six attributes based on this attributes we are consider the values and compare with the threshold values.



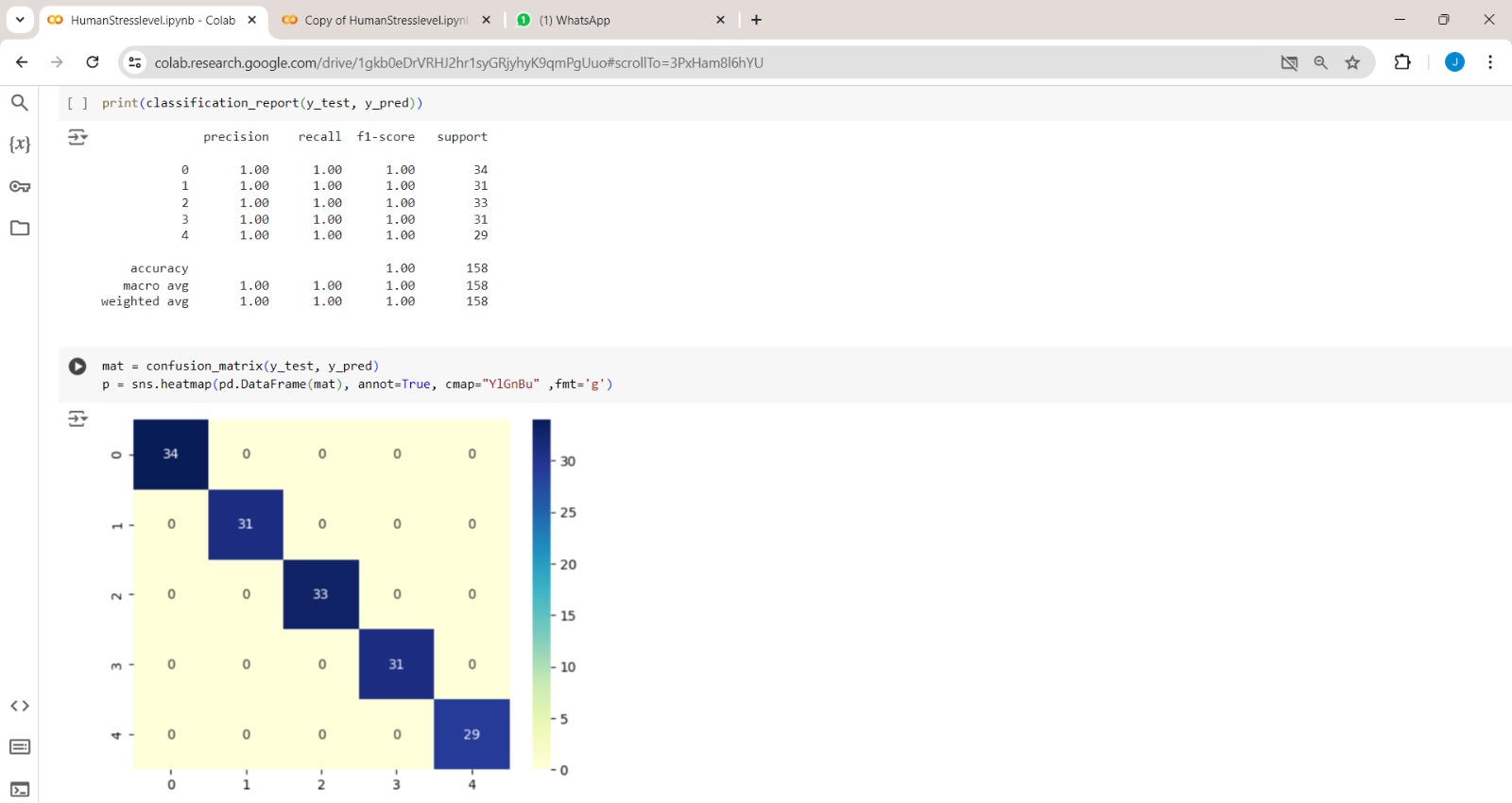
**Fig 7:Accuracy and the confusion matrix showned by the Decision Tree**

In above figure shows that the accuracy measured by the decision tree algorithm. When we use more algorithms we can accuracy. Based on the accuracy we can easily detect the stress.



**Fig 8: Accuracy and confusion matrix shown by the Random- Forest**

In above figure shows that the accuracy measured by the random forest algorithm. When we use more algorithms we can accuracy. Based on the accuracy we can easily detect the stress.



**Fig 9 :Accuracy and confusion matrix showned by the KNN**

In figure 8 shows the predicted values are measured by Naïve Bayes and K-Nearest Neighbour. In predicted values it shows the target and predicted values, classification accuracy and error, sensitivity, specificity, false positive rate and precision.

## 5. Conclusion

The proposed system may help in predicting the stressed people, thereby helpful for the society in solving the serious existing problem of stress by knowing the rate of the stressed level and taking necessary steps and preventive measures to further decrease the stressed level. The proposed systems have taken statistical data and have taken some attributes based on the values have used some of the algorithms are Decision Tree, K-Nearest Neighbor, Random forest to get accuracy because when use more algorithms can get best results will train that data and the data is also divided into training data and testing data and compare with the threshold values and that value declares whether the person is in stress or not stress.Among these algorithms K-Nearest Neighbor algorithm has given 100 percentage accuracy after training and testing of the dataset.