Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

Human stress detection

Predictive Analytics with case studies
(CSE3085)

SLOT:- F1

Team Members

20MIA1031 - SANJAY.M
20MIA1039- SRIGANTH R
20MIA1106- MADHU KIRAN P
20MIA1117- PILLARAM MANOJ
20MIA1162 - GUNA SHANKAR.S

TABLE OF CONTENTS

Abstract:

In today's fast-paced world, many people ignore the myriad benefits that sleep brings to the human body. The SaYoPillow (Smart-Yoga Pillow) is a proposed solution aiming at enhancing people's knowledge of the relationship between stress and sleep, as well as bringing the notion of "Smart-Sleeping" to reality via an edge device. The suggested gadget incorporates an edge processor with a model that evaluates physiological changes that occur during sleep as well as sleeping habits. The device can forecast the user's stress levels for the following day based on the data collected. To ensure the confidentiality of the studied stress data and average physiological changes, the gadget conducts a secure data transfer to the IoT cloud for storage.

A secure data transfer from the cloud to third-party applications is also suggested by the device. Users can regulate the visibility and accessibility of data through a user interface. The SaYoPillow is a distinctive and ground-breaking solution since it not only takes into account sleeping patterns for stress relief but also contains security safeguards to safeguard user data. The SaYoPillow.csv file allows users to view the correlation between a number of parameters, including the snoring range, respiration rate, body temperature, limb movement rate, blood oxygen levels, eye movement, number of hours of sleep, heart rate, and stress levels, all of which are classified as 0 (low/normal), 1 (medium low), 2 (medium), 3 (medium high), and 4 (high/extreme) (high). The tool can be extremely accurate in forecasting stress levels. Up to 96% of the time, the device is accurate in estimating stress levels.

INDEX TERMS: Stress dection, stress levels,Machine leaning algorithms, deep learning algorithms, user interface

Introduction:

Stress is a widespread issue that affects people everywhere. Several things, including difficulty at job, money troubles, marital problems, and health issues, can contribute to it. Anxiety, depression, and cardiovascular illnesses are just a few of the physical and mental health issues that stress can cause. In order to properly manage stress, it is crucial to understand what causes it in people.

Stress is a very fascinating, intensely emotional state that is defined as "vague reaction of the body to any request upon it." Because to the negative effects of long-haul pressure, which can range from headaches and sleep disturbances to an increased risk of cardiovascular diseases, this is the case. Position requests are a major source of pressure on people. Pressure can be caused by situations like ongoing vulnerability to danger, tight deadlines, involved tasks, or simply dull chores.

In order to begin pressure decrease mediations, non-intrusive pressure detecting tools that continuously monitor sensations of anxiety with little impact on professionals' daily activities could be used. These applications couldn't provide more effective and affordable mediations in challenging work environments, much alone more beneficial circumstances where workers would be more willing to handle their responsibilities.

The physical response to pressure is stress. The sensation is squeezed. Pressure can be brought on by a variety of events or conditions in life. Each everyone has their own unique technique of dealing with pressure. Their ability to adjust to it depends on their ancestry, early life events, character, social situation, and financial situation. But, too much pressure can lead to unfavourable outcomes that make people feel depressed. Stress may be a mental health issue that shortens one in four people's lives, according to the World Health Organization (WHO).

In addition to causing mental health problems, human pressure also leads to socioeconomic problems, a lack of trust in the workplace, strained professional relationships, melancholy, and in some extreme situations, the long-term compulsion to commit suicide. They could end up in a very long period of constant anxiety and terror as a result. Over time, it may negatively impact both physical and mental health. People are hence uneasy despite their thriving. The strain may be profound, physical, or even mental in origin.

Literature review:

[1]Reshma Radheshamjee Baheti and Supriya Kinariwala in their paper "Detection and Analysis of Stress using Machine Learning Techniques," a machine learning-based stress detection system is proposed. In order to detect stress, the system gathers physiological information such as heart rate, blood pressure, and skin conductance level and uses machine learning algorithms. The effectiveness of various machine learning algorithms, including Naive Bayes, K-Nearest Neighbors, Support Vector Machines, and Random Forest for stress detection, is compared in the study. A thorough review of stress levels across various age groups and genders is also presented in the study.

Overall, the work shows promise for utilising machine learning algorithms to detect stress in people and emphasises the potential of leveraging physiological cues to do so. Insightful information on how gender and age affect stress levels is also provided by the study. To validate the results on a larger and more diverse population, additional research is required. It is crucial to keep in mind that the study was conducted on a relatively small sample size.

[2]Sayali Shelke in her paper"Stress Detection Using Machine Learning" they conveyed that the use of machine learning algorithms to identify people's degrees of stress is explored in this research. The writers talk about how crucial it is to recognise stress levels precisely and promptly in order to avoid a variety of physical and mental health problems. They concentrate on measuring stress levels using physiological signs including heart rate, blood pressure, and skin conductance.

The performance of various machine learning methods, such as Support Vector Machines (SVM), Random Forest, and Deep Neural Networks (DNN), is assessed by the authors using a dataset of physiological signals gathered from people who were under various degrees of stress. They get to the conclusion that SVM and DNN are more effective at detecting stress levels than Random Forest.

[3]The authors, Russell Li and Zhandong Liu, concentrated on using artificial intelligence (AI) models called deep neural networks to identify stress. Early detection can help prevent or lessen the consequences of stress, which is a common disorder that can have detrimental implications on both physical and mental health.Most likely, the researchers discussed their research on the application of deep neural networks for stress detection. For applications like stress detection that call for the analysis of varied and multi-dimensional data, deep neural networks are a form of machine learning model that is well-suited. Deep neural networks can analyse enormous volumes of data and uncover complicated patterns.

The research methods, including the kind of deep neural network architecture utilised, the data used for training and validation, and the performance measures used to assess the efficacy and accuracy of the model, may have been covered by the authors.The results of this study could have an impact on the field of stress detection and management since they could help create more precise and effective techniques for identifying stress in a variety of situations, including the workplace and hospital settings. The efficacy of their strategy and its potential for real-world applications would probably require more investigation and validation.

[4]The paper "Mental Stress Detection in University Students Using Machine Learning Algorithms" by Ravinder Ahuja and Alisha Banga probably describes their investigation into the use of machine learning algorithms for identifying mental stress in university students. They might have discussed their technique, accuracy results, and findings in the publication.The efficacy of the suggested strategy for identifying mental stress in university students may have been indicated by the precision of the machine learning algorithms utilised in the research. Different performance indicators, such as precision, recall, F1-score, or area under the receiver operating characteristic (ROC) curve, could be used to assess the accuracy.

The research methodology, including specifics on the data used for training and validation, the features or variables used in the machine learning algorithms, and the particular machine learning methods used, may have also been covered in the publication. The justification for the authors' algorithm selection and their methods for addressing any flaws or biases in their methodology may have been covered in their discussion.

[5]The research conducted by Z. Zainudin et al. on stress detection using machine learning and deep learning approaches is likely to be covered in the publication titled "Stress Detection using Machine Learning and Deep Learning" published in 2021. It's possible that the paper discussed the researchers' approach, conclusions, and research algorithms.There may have been a description of the methodology employed in the study, together with information on the data utilised for training and validation, the data preprocessing techniques used, and the experimental design. The methods employed by the authors to gather the data and any restrictions or considerations with regard to the data used for their study may have been included.

The deep learning and machine learning techniques that were employed in the study may have also been covered in the publication. The precise algorithms used, such as decision trees, support vector machines, neural networks, convolutional neural networks (CNNs), recurrent neural networks (RNNs), or other pertinent methods, could be described in this. It's possible that the authors provided justification for the algorithms they selected and the methods via which they were used in their research.

[6]The paper probably focuses on multimodal physiological data-based machine learning and deep learning methods for stress detection. A model that can accurately identify stress levels based on physiological information from several modalities, such as heart rate, skin conductance, cortisol levels, and others, may have been developed by the authors.The research's approach may have included gathering and preprocessing the multimodal physiological data from study

subjects. The methods used to preprocess and normalise the data, as well as any considerations or restrictions on the data utilised for their study, may have all been discussed by the authors.

The paper might have additionally covered the deep learning and machine learning techniques they employed in their study. The precise algorithms used, such as support vector machines, random forests, deep neural networks, recurrent neural networks (RNNs), convolutional neural networks (CNNs), or other pertinent methods, could be described in this. The authors might have discussed how they used the multimodal physiological data to train and assess their models.

[7]The paper "Stress Detection Using Machine Learning and Deep Learning" is most likely a research paper focusing on identifying stress levels using machine learning and deep learning techniques. Using these methods, the authors may have proposed a novel approach to accurately detect stress.The methodology employed in the research may have been stated in the publication, which could have included information on the data used for training and evaluation. This may include gathering physiological or other stress-related data from study participants, as well as preparation activities such as data normalisation or feature extraction.

Furthermore, the publication may have discussed the outcomes of the experiments or evaluations carried out during their research. This could contain performance indicators such as accuracy, precision, recall, F1-score, or other relevant metrics used to analyse the effectiveness of their models. The authors may have compared their findings to those of other researchers, or they may have acknowledged the limitations and potential consequences of their findings.

[8]Cem Ersoy, Bert Arnrich, and Yekta "Stress detection in daily life scenarios using smartphones and wearable sensors: A survey" The paper is most likely a survey research focusing on stress detection in everyday life scenarios using smartphones and wearable sensors. The authors may have undertaken a survey of existing literature and research papers on this topic, providing a complete overview of the field's current state.It might have also gone into the various types of sensors used for stress detection, such as accelerometers, heart rate monitors, electrodermal activity sensors, and so on. The authors may have presented an overview of these sensors' capabilities and limitations in recording physiological or behavioural stress signals.

In addition, the paper may have explored the many machine learning or deep learning techniques utilised in stress detection using smartphones and wearable sensors. A explanation of the specific algorithms used, such as classification algorithms, feature extraction approaches, and data fusion methods, could be included. The authors may have explored the obstacles and opportunities connected with applying machine learning or deep learning to detect stress in real-world circumstances.

[9]V. R. Archana and B. M. Devaraju's paper "Stress Detection Using Machine Learning Algorithms" most likely describes a research study on stress detection using machine learning techniques. To properly measure stress levels, the authors may have presented a fresh approach or used existing machine learning algorithms.The methodology employed in the research may have been stated in the publication, which could have included information on the data used for training and evaluation. This may include gathering physiological or other stress-related data from study participants, as well as preparation activities such as data normalisation or feature extraction.

The authors may have talked about the machine learning techniques they employed in their study. A description of the specific methods used, such as support vector machines, decision trees, random forests, k-nearest neighbours, logistic regression, or other relevant algorithms, should be included. The authors may have mentioned how these algorithms were used in their study to detect stress levels based on data collected.

Furthermore, the publication may have discussed the outcomes of the experiments or evaluations carried out during their research. This could include evaluating the performance of their suggested strategy or comparing other machine learning algorithms in terms of accuracy, precision, recall, F1-score, or other relevant metrics used to assess the efficacy of their models. The authors may have discussed the limits and possible consequences.

[10]E. Padma, Talapaneni Praveen, and Shaik Karimulla's "Stress Detection Using Machine Learning and Image Processing" is most likely a research work focusing on stress detection using machine learning and image processing techniques. The authors may have proposed a novel method for accurately detecting stress levels that combines machine learning and image processing. The methodology employed in the research may have been stated in the publication, which could have included information on the data used for training and evaluation. Collecting photos or visual data connected to stress, such as facial expressions, physiological signs, or other pertinent visual cues, may be part of this process. The authors may have described the image preprocessing stages, such as image enhancement, feature extraction, or image segmentation.

The researchers usage of image processing methods and machine learning algorithms may have been covered by the authors. This could describe the particular algorithms used, such as decision trees, convolutional neural networks (CNNs), support vector machines (SVMs), or other pertinent algorithms. It's possible that the authors described how these algorithms were used in their study to identify stress levels from photographs or other visual information. Additionally, the report might have covered the outcomes of the tests or assessments made during the investigation. This could involve evaluating the performance of their suggested strategy or contrasting various machine learning methods in terms of accuracy, precision, recall, F1-score, or other pertinent metrics used to gauge the efficacy of their models.

Dataset Description:

Snoring range: This feature measures the user's snoring range while sleeping. It is most usually measured with a microphone or other comparable equipment.

Respiration rate: This option counts the number of breaths the user takes per minute while sleeping. A respiratory belt or a pulse oximeter can be used to measure it.

Body temperature: This setting measures the user's body temperature when sleeping. It can be measured with a thermistor or a thermal camera, for example.

Limb movement rate: This parameter measures the frequency and extent of the user's limb movements while sleeping. It can be measured with an instrument like an accelerometer.

Blood oxygen levels: This parameter checks the user's blood oxygen saturation when sleeping. A pulse oximeter is often used to measure it.

Eye movement: This characteristic monitors the frequency and amount of the user's eye movements while sleeping. It can be measured with an instrument called an electrooculogram (EOG).

Number of hours slept: This parameter calculates the total amount of time the user slept throughout the measurement period. It might be derived from other measurements or self-reported by the user.

Heart rate: This setting measures the user's heart rate when sleeping. It can be measured with an instrument like an electrocardiogram (ECG).

Stress levels: This parameter assesses the user's perceived stress levels when sleeping and is graded from 0 (low/normal) to 4 (extreme) (high). The user most likely self-reported it.

The dataset appears to be looking into the connections between these numerous characteristics and the user's perceived stress levels while sleeping. It can be used to discover patterns or relationships between various parameters and stress levels, as well as to inform interventions or treatments for increasing sleep quality and stress reduction.

Methodology:

Data collection: Wearable devices, EEG, EOG, and EMG sensors are used to gather physiological signals such as heart rate, respiration, movement, and brain activity.

Data Preprocessing : The data are preprocessed to reduce noise and artefacts and to segment the data into distinct time periods, such as epochs.

Feature extraction: Using a variety of signal processing techniques, pertinent features, such as frequency content, amplitude, and variability, are retrieved from the preprocessed data.

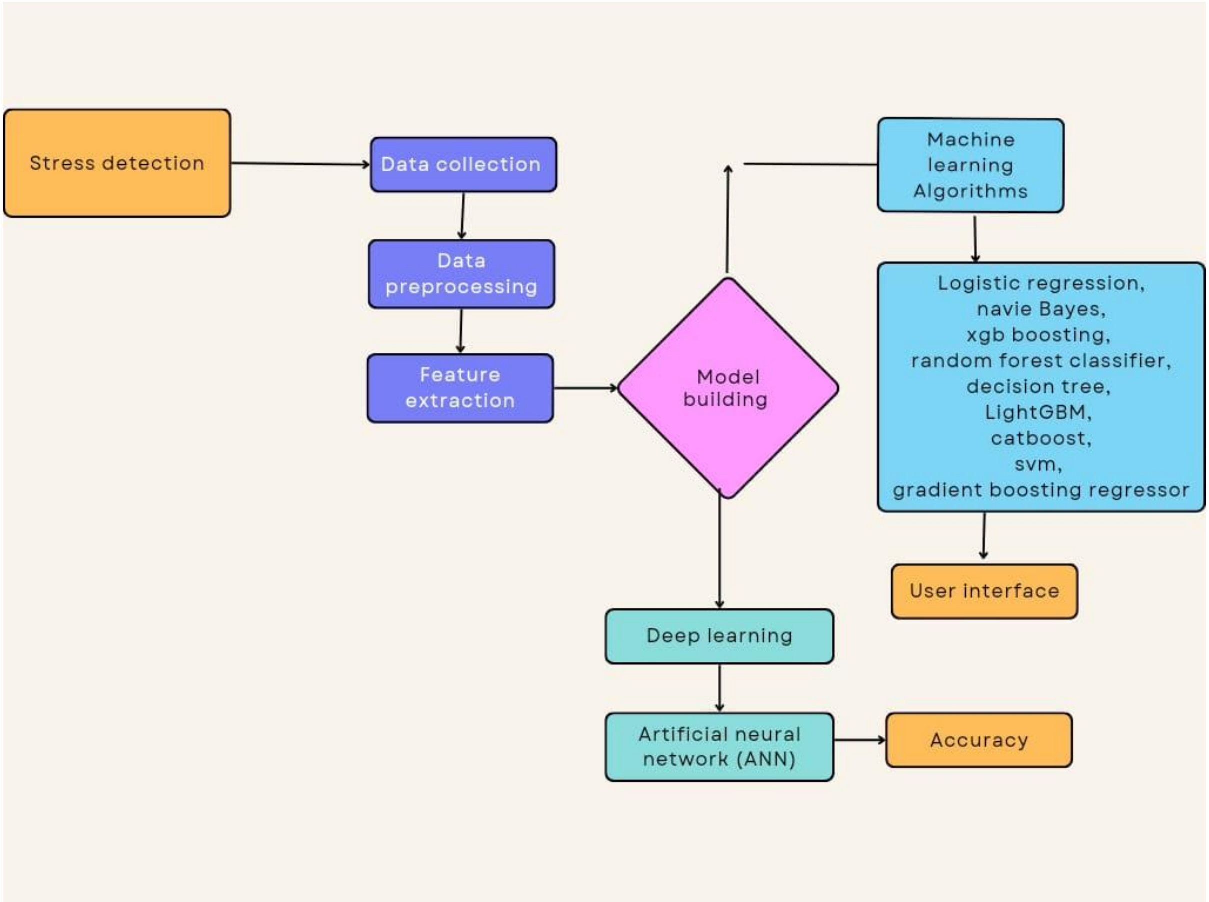
Development of a machine learning model: Using the extracted features and a labelled dataset that contains examples of both stressed and non-stressed states, a machine learning model is created.

Model training and validation: To assess the machine learning model's accuracy and generalisation performance, a subset of the labelled dataset is used for training and the remaining subset for validation.

Stress detection: By applying the trained model to fresh, unexplored data and predicting the stress state based on the extracted features, it is possible to identify stress in real-world data.

Stress management: Based on the findings of the stress detection process, stress management solutions, such as meditation, cognitive behavioural therapy, or medication, can be developed.

ARCHITECTURE



Results and Discussions:

Logistic regression:

The goal of logistic regression is to forecast stress levels using physiological cues including heart rate, respiration rate, blood oxygen level, body temperature, and eye movements. A binary logistic regression is used to predict whether a person is under or over stress. The model is trained using physiological signal data and stress level ratings, and its performance is measured using measures such as precision, recall, and F1 score. The logistic function is used to convert the linear relationship between the independent and dependent variables into a sigmoidal curve that predicts the likelihood of the result.

The accuracy for logistic regression algorithm is 100%.

XGB classifier:

The XGB classifier is a binary classification implementation of the XGBoost algorithm. To reduce overfitting and increase generalisation performance, it employs an ensemble of decision trees and regularisation algorithms. The XGB classifier is trained on a dataset that combines physiological signal data and stress level evaluations and is used to predict whether a person is likely to be in the high-stress category. The model's performance is measured using measures like as accuracy, precision, recall, and F1 score.

The accuracy for XGB classifier algorithm is 98%.

Random forest classifier:

Random Forest Classifier is a binary classification implementation of the Random Forest technique. To reduce overfitting and increase generalisation performance, it employs an ensemble of decision trees and random sampling. The Random Forest Classifier predicts the chance of a person being in the high-stress group based on a dataset that contains physiological signal data and stress level evaluations. Metrics including as accuracy, precision, recall, and F1 score are used to assess the model's performance. The process of finding the best collection of hyperparameters for the model is known as hyperparameter tuning.

The accuracy for Random forest classifier regression algorithm is 98%.

Navie bayes:

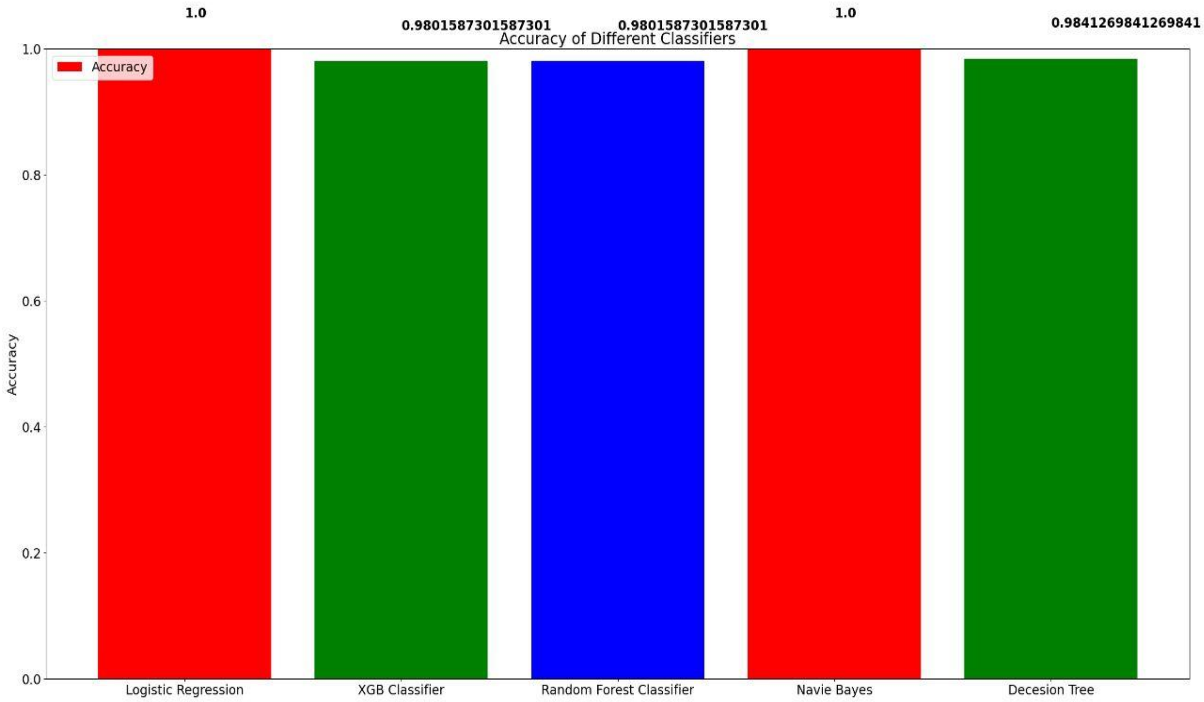
The Naive Bayes classifier is a binary classification implementation of the Naive Bayes algorithm. Given the class variable, it assumes that the features are independent of one another and computes the posterior probability of each class given the observed features. The Naive Bayes classifier predicts the chance of an individual being in the high-stress group based on a dataset that contains physiological signal data and stress level evaluations. Metrics including as accuracy, precision, recall, and F1 score are used to assess the model's performance.

The accuracy for Navie bayes algorithm is 100%.

Decision tree:

Decision Tree algorithm is a machine learning strategy for categorization analysis. It builds a decision tree based on the values of input variables and predicts the value of a target variable using this model. The algorithm is trained on a dataset combining physiological signal data and stress level assessments and is used to predict whether or not a person would fall into the high-stress group. Metrics including as accuracy, precision, recall, and F1 score are used to assess the model's performance.

The accuracy for Decision tree regression algorithm is 98.4%.



KNN:

The k-nearest neighbours algorithm is a non-parametric, supervised learning classifier that uses distance to create classifications or predictions about an individual data point's grouping. While it can be used for either regression or classification problems, it is most commonly used as a classification algorithm, based on the assumption that similar points can be found nearby.
The accuracy for KNN algorithm is 100%

Hyperparameters:

Searching for hyperparameters in KNN and DT models. The K 5, 10, and 15 hyperparameter settings produced the same accuracy values in KNN testing. K is therefore decided to be 5 as the value.

The accuracy value from the DT hyperparameter test with max_depth values of 20, 40, and 60 is the same. Therefore, a value of 40 is chosen for max_depth.

LightGBM:

LightGBM is a framework for gradient boosting that use tree-based learning techniques.
It is intended to be widely dispersed and efficient, with the following benefits:

- Faster training speed and higher efficiency.
- Lower memory usage.
- Better accuracy.
- Support of parallel, distributed, and GPU learning.
- Capable of handling large-scale data.

Gradient-based One-Side Sampling (GOSS), a cutting-edge method used by LightGBM, is used to filter out the data instances and get the split value,while XGBoost computes the optimal split using a pre-sorted method and a Histogram-based approach.
The accuracy for LightGBM algorithm is 100%.

CatBoost:

CatBoost (Categorical Boosting) is a machine learning technique that is specifically developed to operate with categorical information and solve some of the limitations of classic gradient boosting methods. CatBoost can handle both numerical and categorical data and recognises which features are categorical automatically. It employs a gradient boosting variation called gradient-based one-side sampling (GOSS), which lowers the amount of data that must be processed during training. As a result, it is faster and uses less memory than classic gradient boosting approaches.

CatBoost can handle category features automatically, without the requirement for one-hot encoding, by converting them into numerical representations.

CatBoost is able to determine the value of a feature based on how well it performs when combined with other features. CatBoost may be used with GPUs, greatly reducing the amount of time needed to train huge datasets.
The accuracy for CatBoost algorithm is 100%.

Gradient Boosting Regressor:

The Gradient Boosting Regressor (GBR), a machine learning technique, is a member of the ensemble boosting method family. In an iterative process, it turns a number of poor learners into a single strong learner. The algorithm creates a new model during each iteration in an effort to fix the mistakes caused by the prior models.It can handle both linear and nonlinear correlations between the data and the target variable, GBR is an effective technique for regression tasks. Additionally, it can deal with data anomalies and missing values. Gradient descent, which minimises the loss function by shifting the model's parameters in the direction of the negative gradient, is the foundation of GBR.

The number of trees in the ensemble, the learning rate, and the maximum depth of the trees serve as the three primary hyperparameters of GBR. The model's complexity is governed by the number of trees, and the gradient descent's step size is governed by the learning rate. The extent of interaction between the features is governed by the trees' maximum depth. The GBR algorithm, which is effective, can do regression jobs with a high degree of accuracy.

The accuracy for Gradient Boosting Regressor algorithm is 99.65%.

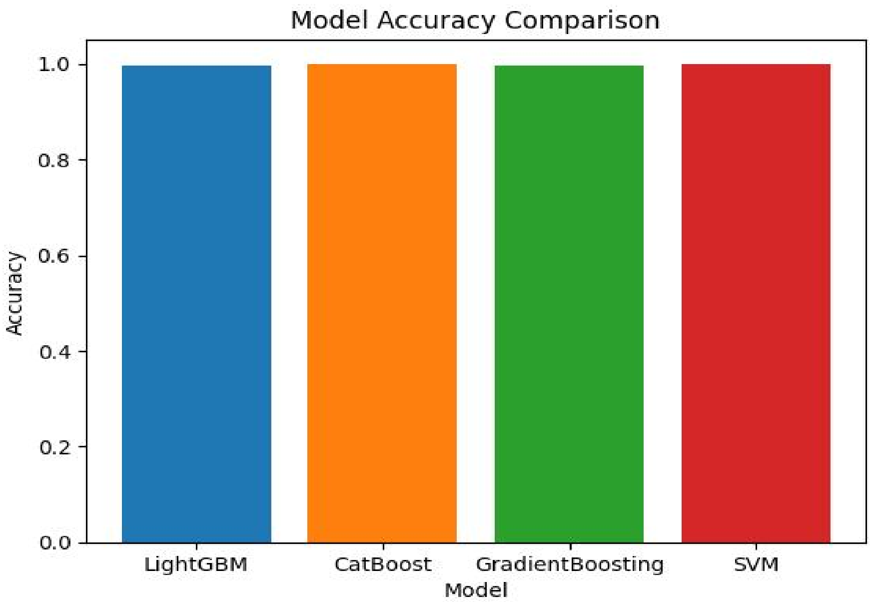
SVM:

The Support Vector Machine (SVM) is a well-known machine learning technique for classification and regression analysis. It is a supervised learning algorithm that analyses data and recognises patterns before categorising or grouping fresh data points. SVM works by locating a hyperplane in a high-dimensional space that divides the various classes. The hyperplane is chosen so that the margin between the closest points of the different classes is maximised.

The distance between the hyperplane and the nearest data points in each class is represented by this margin. Support vectors are the spots nearest to the hyperplane. SVM also has the ability to handle high-dimensional data with a minimal number of samples.

As it concentrates on the data points that are closest to the decision border and disregards the remaining data points, SVM performs well on datasets with many features but fewer samples.

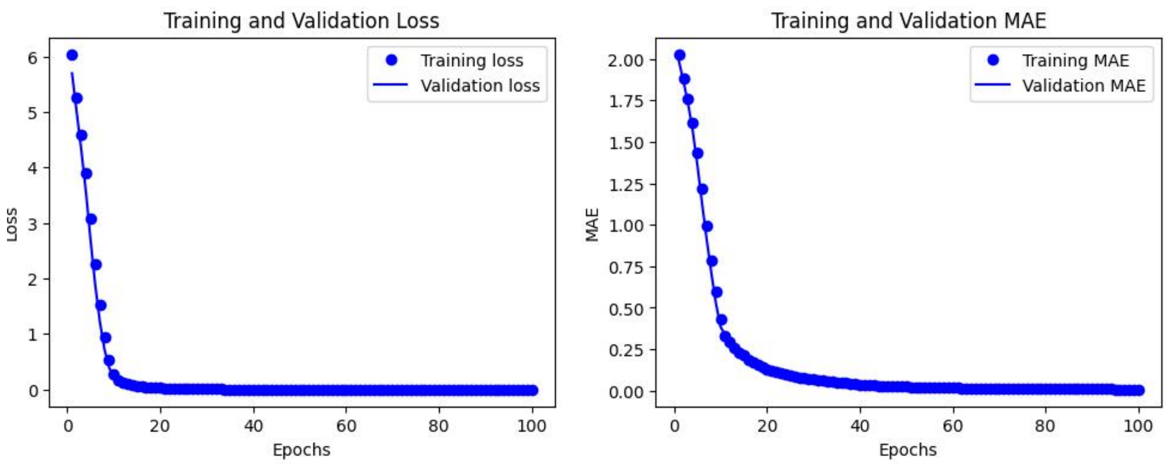
SVM has several uses in a variety of industries, including banking, bioinformatics, image classification, and text classification. It is an effective machine learning tool due to its excellent accuracy and capacity for handling complex data.
The accuracy for SVM algorithm is 100%.



ANN:

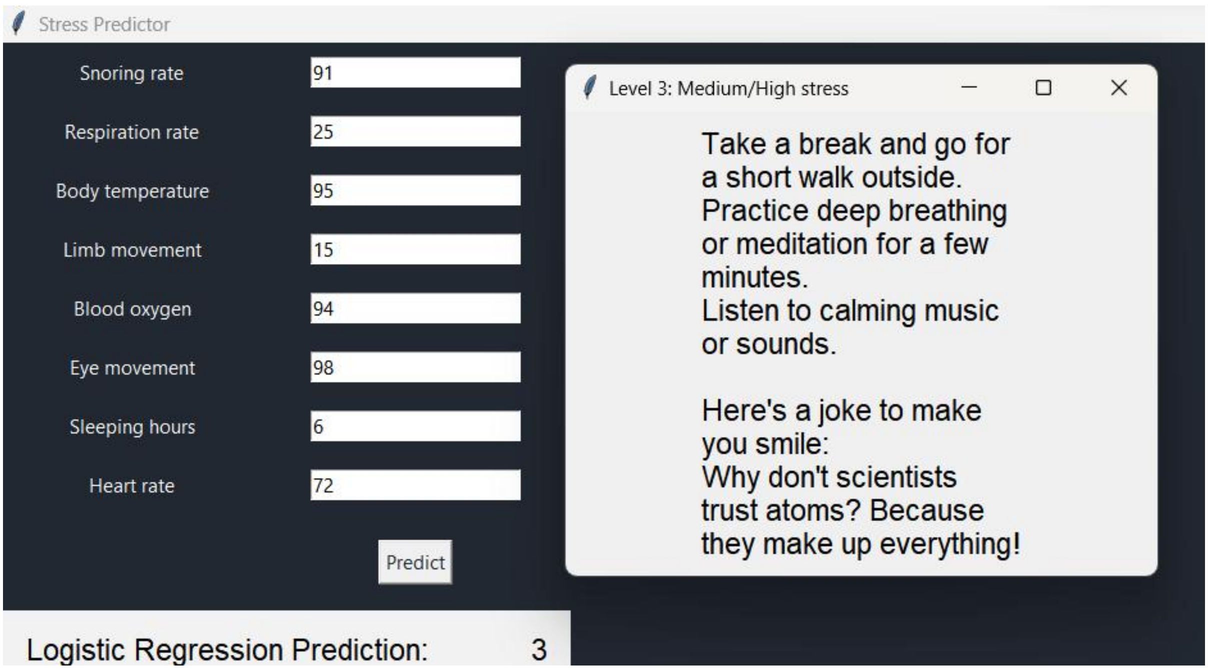
Artificial Neural Networks, or ANNs, are a subset of machine learning models that take their cues from the design and operation of biological neural networks seen in the human brain.An input layer, one or more hidden layers, and an output layer are the three layers that make up an artificial neural network (ANN). Each neuron takes in information from the neurons in the layer below and then transforms it to create an output that is sent to the neurons in the layer above. ANNs have been used successfully in a variety of applications, including picture and audio recognition, natural language processing, and predictive modelling in fields such as finance, healthcare, and marketing.
The accuracy for ANN algorithm is 73%.

```
Epoch 1/100
16/16 [==============================] - 2s 24ms/step - loss: 6.0395 - mae: 2.0290 - val_loss: 5.6941 - val_mae: 1.9903
Epoch 2/100
16/16 [==============================] - 0s 4ms/step - loss: 5.2620 - mae: 1.8862 - val_loss: 4.9967 - val_mae: 1.8658
Epoch 3/100
16/16 [==============================] - 0s 4ms/step - loss: 4.5870 - mae: 1.7605 - val_loss: 4.3111 - val_mae: 1.7312
Epoch 4/100
16/16 [==============================] - 0s 3ms/step - loss: 3.9009 - mae: 1.6156 - val_loss: 3.5546 - val_mae: 1.5663
Epoch 5/100
16/16 [==============================] - 0s 3ms/step - loss: 3.0841 - mae: 1.4350 - val_loss: 2.6844 - val_mae: 1.3609
Epoch 6/100
16/16 [==============================] - 0s 4ms/step - loss: 2.2577 - mae: 1.2210 - val_loss: 1.8804 - val_mae: 1.1309
Epoch 7/100
16/16 [==============================] - 0s 3ms/step - loss: 1.5176 - mae: 0.9969 - val_loss: 1.2170 - val_mae: 0.9134
Epoch 8/100
16/16 [==============================] - 0s 3ms/step - loss: 0.9354 - mae: 0.7857 - val_loss: 0.7182 - val_mae: 0.7117
Epoch 9/100
16/16 [==============================] - 0s 3ms/step - loss: 0.5317 - mae: 0.5970 - val_loss: 0.3926 - val_mae: 0.5259
Epoch 10/100
16/16 [==============================] - 0s 3ms/step - loss: 0.2848 - mae: 0.4315 - val_loss: 0.2252 - val_mae: 0.3845
```



The epoch is 100 for ANN and the above plot is for training, validation loss and training ,validation MAE for ANN algorithm.

User Interface:



Comparision Table :

Comapring performance of different machine learning algorithms on two datasets, namely "sleeping" and "non-sleeping".

Table-1-For Research Paper

S.no	Dataset	Algorithms	Best Accuracy(paper)
1,4	sleeping	Svm, knn,DT, NavieBayes, PCA+linear SVM	SVM, Navie bayes-100%
3	Non sleeping	Random forest,Navie Bayes, SVM,KNN	Svm-85.71%

Table-2-sleeping

S.no	algorithms	accuracy
1	svm	100%
2	knn	87.5%
3.	DT	98%
4.	Navie Bayes	100%
3	PCA+linear SVM	75.0%

Table-3-non sleeping

S.no	Algorithms	Accuracy
1.	Random Forest	83.33%
2.	Navie Bayers	71.42%
3.	SVM	85.71%
4.	KNN	55.55%

Table-4 our models with accuracy-sleeping data.

S.no	Algorithms	Accuracy
1.	Logistic regression	100%
2.	XGB classifier	98%
3.	Random forest classifier,	98%
4.	Navie bayes	100%
5.	Decision tree	98.4%
6.	LightGBM	100%
7.	CatBoost(Categorical Boosting)	100%
8.	GradientBoosting Regressor	99.65%
9.	SVM	100%
10.	ANN	73%
11.	KNN	100%

Table-1 :

Table-1 showing the performance of different machine learning algorithms on two datasets, namely "sleeping" and "non-sleeping".

For the "sleeping" dataset, the algorithms used were SVM, KNN, DT, NavieBayes, and PCA+linear SVM. The best accuracy reported in the paper was achieved by SVM and Navie Bayes, both with an accuracy of 100%.

For the "non-sleeping" dataset, the algorithms used were Random forest, Navie Bayes, SVM, and KNN. The best accuracy reported in the paper was achieved by SVM with an accuracy of 85.71%.

Table-2:

SVM (Support Vector Machines) and Naive Bayes have the highest accuracy of 100%, indicating that they may be the most suitable algorithms for the task or dataset.

Decision Tree (DT) has a relatively high accuracy of 98%.

KNN (K-Nearest Neighbors) has an accuracy of 87.5%, which is lower than the top three algorithms.

PCA+Linear SVM has the lowest accuracy of 75%, indicating that it may not be as suitable for the task or dataset as the other algorithms.

Table-3:

SVM (Support Vector Machines) has the highest accuracy of 85.71%, indicating that it may be the most suitable algorithm for the task or dataset.

Random Forest also has a relatively high accuracy of 83.33%.

Naive Bayes has a lower accuracy of 71.42% compared to SVM and Random Forest.

KNN (K-Nearest Neighbors) has the lowest accuracy of 55.55%, indicating that it may not be as suitable for the task or dataset as the other algorithms.

Table-4:

Logistic Regression: It is a statistical approach for analysing a dataset and predicting the likelihood of a binary result (i.e., success/failure). Logistic Regression received a flawless accuracy score of 100%, indicating that it properly identified all occurrences.

XGB Classifier: XGB is a gradient boosting method with decision trees as its foundation model. The XGB Classifier received a 98% accuracy score, indicating that it successfully identified 98% of occurrences.

Random Forest Classifier: Random Forest is an ensemble method that makes a final prediction by combining numerous decision trees. The Random Forest Classifier had an accuracy score of 98%, indicating that it correctly identified 98% of the occurrences.

Naive Bayes: It is a probabilistic algorithm that estimates the likelihood of each class given the input attributes by applying the Bayes theorem. The Naive Bayes algorithm accurately classified every occurrence, earning a perfect accuracy score of 100%.

A decision tree: It is a model of decisions and potential outcomes that resembles a tree and is used to categorise input data. The Decision Tree method identified 98.4% of instances correctly, yielding an accuracy score of 98.4%.

LightGBM: Using tree-based learning techniques, LightGBM is a gradient boosting framework. The LightGBM algorithm accurately classified every occurrence, earning a perfect accuracy score of 100%.

CatBoost: CatBoost is an algorithm for gradient boosting that is specifically made for categorical data. The CatBoost system accurately classified every occurrence, earning a perfect accuracy score of 100%.

Gradient Boosting Regressor: For regression issues, a boosting approach is utilised called Gradient Boosting Regressor. A 99.65% accuracy rating suggests that 99.65% of cases were properly identified by the Gradient Boosting Regressor.

Support Vector Machine (SVM): It is a machine learning algorithm that is used for regression analysis and classification. The SVM system accurately classified every instance, earning a perfect accuracy score of 100%.

Artificial Neural Network (ANN): An algorithm for machine learning, ANN is based on how the human brain works and is structured. The ANN algorithm categorised 73% of occurrences correctly, earning it an accuracy score of 73%.

Conclusion:

The prediction of Human stress is extremely important to know while investing. From this study it can be concluded that algorithms such as Logistic Regression, XGB classifier, Random forest classifier, Navie bayes, Decision tree, LightGBM, CatBoost, Gradient Boosting Regressor, KNN,SVM and ANN can be employed to predict the human stress. We used totally Ten algorithms for accuracy comparison- logistic regression-100%, XGB classifier-98%, random forest classifier-98%, navie bayes-100% ,decision tree-98.41% , LightGBM-100%, CatBoost-100%, Gradient Boosting Regressor-99.65%, SVM-100% and ANN-73%.There is 100% accuracy from the Logistic Regression, Navie Bayes, LightGBM, Catboost and SVM. Another models that we could depend on is the Gradient Boosting Regressor and Decision Tree. It also works relatively fine.

Compare to deep learning algorithm ANN , Machine learning algorithms like SVM, logistic regression, Catboost, lightGBM KNN,and Navie Bayes got highest accuracy(100%).For ANN algorithms we have took less batch layers so we got less accuracy, if we take more batch layers to ANN we will get highest accuracy like Ml algorithms.Advantage of ml over deep learning models: when we compare to deep learning, ml algorithms will get good performance (Accuracy)for smaller datasets, where as deep learning will perform better for large datasets compare to smaller datasets

We are getting the inputs from the user(manually) or using any instruments to measure and based on the inputs we are predicting stress level of the person and each level will increase by 25% and there will be 4 levels. Our research shows that machine learning systems can effectively anticipate stress levels. We are concluding that parameters limb movement, snorting rate, respiratory rate, body temperature, spo2 level, eye movement, sleeping hours, heart rate also able to predict the human stress level as we get similar results to the non-sleeping dataset.

We can adopt effective stress management techniques by identifying the variables that create stress in humans. Our study's findings may be valuable to scholars and healthcare practitioners interested in investigating stress-related disorders in humans.

Future Work:

Future work shall include more classification models and the effectiveness of various stress detection techniques, such as physiological signals, speech analysis, or facial expression identification could be compared. This would enable you to recognise the advantages and disadvantages of each technique and choose the best strategy for specific situations.

References:

https://www.ijeat.org/wp-content/uploads/papers/v9i1/F8573088619.pdf
https://www.irejournals.com/formatedpaper/1702649.pdf
https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-020-01299-4
https://www.sciencedirect.com/science/article/pii/S1877050919306581
https://iopscience.iop.org/article/10.1088/1742-6596/1997/1/012019/pdf
https://www.researchgate.net/publication/344052913_Stress_Detection_with_Machine_Learning_and_Deep_Learning_using_Multimodal_Physiological_Data
https://www.researchgate.net/publication/354183013_Stress_Detection_using_Machine_Learning_and_Deep_Learning
https://www.sciencedirect.com/science/article/pii/S1532046419300577
https://journals.resaim.com/ijresm/article/download/171/154/315
https://jespublication.com/upload/2022-V13I3053.pdf

GITHUB LINK

https://github.com/Sanjay3369/Predictive.git

Colab link
https://colab.research.google.com/drive/1LaV_Epaxw6bSKBY6WmcA50Vggnv5K_tB?usp=sharing

YOUYTUBE video link of project execution video

https://youtu.be/DTGxVO9ogHk