IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

"Stress Detection In Humans Using Sleeping Habits Based On Machine Learning Algorithms"

Chaitra P

Department of Computer Science and Engineering

Atria institute of technology, Bangalore

Abstract

Stress, an increasingly prevalent aspect of modern life, can significantly impact an individual's physical and mental well-being. Hence, understanding and monitoring stress levels play a crucial role in promoting overall health and quality of life. This paper "Human Stress Detection Based on Sleeping Habits Using Machine Learning Algorithms" presents a novel and effective approach to detect human stress levels by analyzing their sleeping habits. These parameters include the user's snoring range, respiration rate, body temperature, limb movement rate, blood oxygen levels, eye movement, the number of hours of sleep, heart rate, and stress levels categorized into five classes: 0 (low/normal), 1 (medium low), 2 (medium), 3 (medium high), and 4 (high). The inclusion of these diverse parameters ensures a comprehensive analysis of sleep patterns and their correlation with stress levels. The model is trained using the dataset, and its performance is evaluated on a separate test dataset to ensure generalization and unbiased assessment. The results of the experiments reveal a Training score of 100% and an impressive Test score of 97%, demonstrating the effectiveness and robustness of the proposed methodology. The achieved high accuracy showcases the model's capability to learn intricate patterns from the dataset and make accurate stress predictions based on the user's sleeping habits.

Keywords—Machine learning, training score, Sleeping patterns, human stress

Introduction

Stress is either an emotional or physical pressure. Any incident or study that gives you anxiety, rage, or frustration can trigger it. The body's reaction to a demand or challenge is stress. Similar to when it helps you avoid danger or fulfill a deadline; short-term stress

can be beneficial. Everyone experiences stress; it is a natural human reaction. In actuality, the mortal body is built to observe stress and respond to it. When experience changes or challenges (stressors), body adapt to novel circumstances. Stress can help us stay awake, motivated, and prepared to avert danger. For instance, if you have a big test coming up, your body may work harder and stay awake longer as a result of a stress response. But when stresses persist without relief or rest periods, stress becomes an issue.

The stress is detected by using the machine learning algorithm. Nowadays, stress is a major problem for many youngsters. The time period that was formerly thought to be the most carefree is currently under a lot of stress. The stress increased a wide variety of problems such as depression, suicide, heart attack, and stroke. The body's natural response to change is a series of physical, emotional, and intellectual responses. Three classification algorithms, ensemble models are applied. The stress dataset was downloaded from Kaggle. By comparing three algorithms, Majority Voting gives the best accuracy.

The major contributions of the project are summarized as follows:

1. In this paper we can detect the stress by using the machine learning algorithm. Nowadays, stress is a major problem for many youngsters.
2. Three classification algorithms, XGBoost, AdaBoost and Majority Voting algorithm are applied.
3. The stress dataset was downloaded from Kaggle.

METHODOLOGY

There are three algorithms used in this model. They are Xgboost, Adaboost and Majority Voting. Compare the algorithm and find the accuracy value.

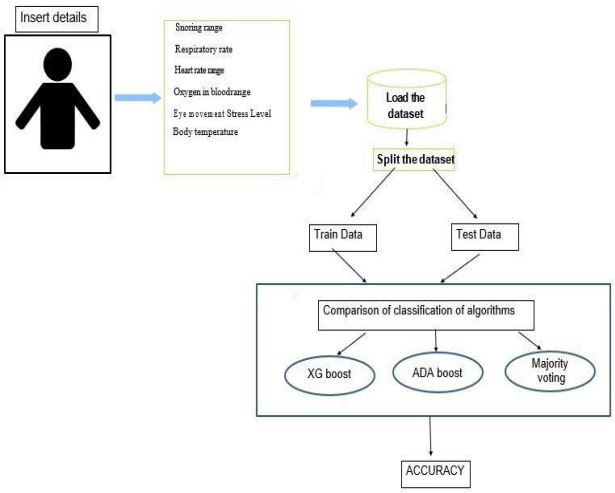


Fig : Work flow of the System

In the proposed approach we are passing sleeping habit dataset and user details(snoring range, repository rate, body temperature, oxygen level in blood, limb movement rate, eye movement rate)as input system will preprocess and apply algorithms and predict the stress level and show the performance comparison of algorithms. There are three algorithms used in this model. They are Xgboost, Adaboost and Majority Voting. Compare the algorithm and find the accuracy value. In this project Majority Voting algorithm gives the best accuracy.

1) **XGBOOST:** In XGBoost is an optimized distributed gradient boosting library designed for efficient and scalable training of machine learning models. It is an ensemble learning method that combines the predictions of multiple weak models to produce a stronger prediction. XGBoost stands for “Extreme Gradient Boosting” and

it has become one of the most popular and widely used machine learning algorithms due to its ability to handle large datasets and its ability to achieve state-of-the-art performance in many machine learning tasks such as classification and regression.

Mathematically, we can write our model in the form

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \epsilon \mathcal{F}$$

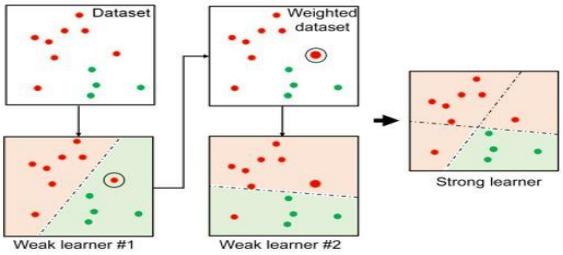
where, K is the number of trees, f is the functional space of F, F is the set of possible CARTs.The objective function for the above model is given by:

$$obj(\theta) = \sum_i^n l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$

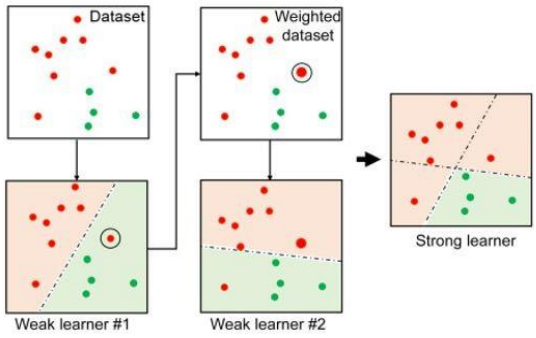
One of the key features of XGBoost is its efficient handling of missing values, which allows it to handle real-world data with missing values without requiring significant pre-processing. Additionally, XGBoost has built-in support for parallel processing, making it possible to train models on large datasets in a reasonable amount of time.

2) **AdaBoost:** AdaBoost is adaptive in the sense that subsequent weak learners are tweaked in favor of those instances misclassified by previous classifiers. In some problems it can be less susceptible to the over fitting problem than other learning algorithms. The individual learners can be weak, but as long as the performance of each one is slightly better than random guessing, the final model can be proven to converge to a strong learner.

Although AdaBoost is typically used to combine weak base learners (such as decision stumps), it has been shown that it can also effectively combine strong base learners (such as deep decision trees), producing an even more accurate model. Every learning algorithm tends to suit some problem types better than others, and typically has many different parameters and configurations to adjust before it achieves optimal performance on a dataset. AdaBoost (with decision trees as the weak learners) is often referred to as the best out-of-the-box classifier.



3) **Majority Voting:** A Voting Classifier is a machine learning model that trains on an ensemble of numerous models and predicts an output (class) based on their highest probability of chosen class as the output. Suppose you have trained a few classifiers, each one achieving about 80% accuracy. You may have a Logistic Regression classifier, an SVM classifier, a Random Forest classifier, a K-Nearest Neighbors classifier, and perhaps a few more, A very simple way to create an even better classifier is to aggregate the predictions of each classifier and predict the class that gets the most votes. This majority-vote classifier is called a hard voting classifier.



Different modules in problem solving

1. Data collection
2. Data preprocess
3. Building ml model
4. Prediction
5. Performance evaluation

1. Data Collection

In this module we are collecting sayopillow dataset from kaggle.com.people merely sleep these days, disregarding the health benefits sleep offers the body. The smart-yoga pillow (sayopillow) is a suggested edge gadget that aims to fully materialize the concept of "smart-sleeping" and assist in understanding the relationship between stress and sleep. It is suggested to use an edge processor to analyze both sleeping habits and the physiological changes that take place when we sleep the influence of stress on the next day is hypothesized based on these alterations.

2. Data Preprocess

This stage aims to enhance the quality of the utilized dataset, as it include several missing values and outliers. Medical datasets commonly suffer from such issues due to various causes, including device failure, network loss, irregular time recording, etc. Unfortunately, Several ML models are sensitive to outliers; most cannot handle missing values. Data preprocessing includefilling in missing data and data encoding.

3. Filling missing values

Many statistical approaches exist to deal with missing data, but itmainly depends on how much data are missing and the importance of the feature missing . When the fraction of the missing data is between 5% and 10%, traditional statistical approaches, such as mean, max, and mode, work exceptionallywell. When the fraction of missing values is 20–50%, sophisticated approaches, such as hot-deck and expectation maximization , are appropriate. To ensure data reliability in our used data, we choose to remove featureswith more than 30% missing values. Features with missing values that are less than 30% are imputed using feature means.

4. Data Encoding

Categorical and numeric features are combined in the utilized dataset. Numeric features perform better with ML and DL than categorical ones, unfortunately. Therefore, we encoded all categorical features using the label encoder module of the Scikit-learn library.

5. Sampling data

SMOTEENN is used for re-sampling data. The SMOTE-ENN method combines the SMOTE and ENN techniques. SMOTE is an oversampling method, and ENN is an edited closest neighbor under sampling method (ENN). In the ENN approach, the observation and its KNN are removed from the dataset if the majority class of the KNN and the observation's class are different. Due to this, information about the minority class in themajority class is lost. By doing this, the bias towards the majority class is lessened, which enhances the performance of machine learning models.

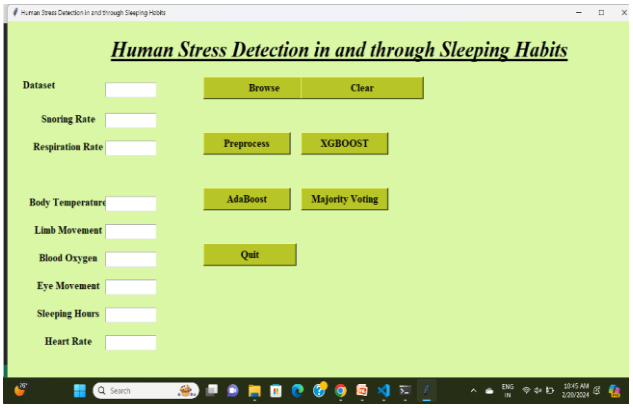
Building and Training modules

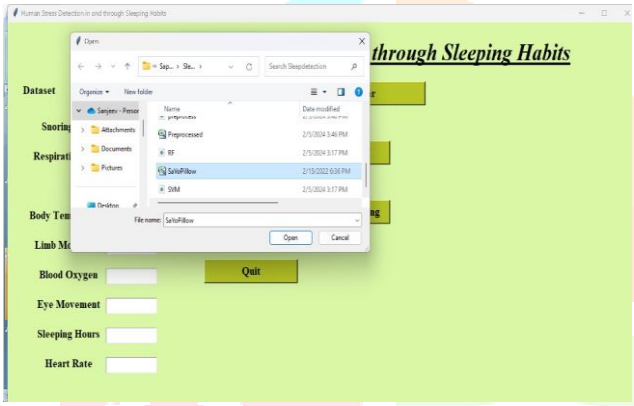
In this module we are splitting the preprocessed data into20% testing and 80% training data.

Then we will build various machine learning algorithms likeXgboost, Adaboost

and Majority Voting. Train the models with data and save trained model.

Output





Conclusion

Personalized machine learning models of sleep-wake states outperform their generalized. Counterparts in terms of estimating sleeping patterns are indistinguishable. Results of the experiments reveal a Training score of 100% and an impressive Test score of 97%, demonstrating the effectiveness and robustness of the proposed methodology. The achieved high accuracy showcases the model's capability to learn intricate patterns from the dataset and make accurate stress predictions based on the user's sleeping habits.

REFERENCES

[1] Suja SreeithPanicker,PrakasamGayathriA survey of machine learning techniques in physiology based mental stress detection systems,Volume 39, Issue 2, April–June 2019.

[1] Aishwarya Bannore; Tejashree Gore; Apoorva Raut; Kiran Talele (10 November 2021)Mental stress detection using machine learning algorithm, International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME).

[2] RavinderAhuja,AlishaBanga (2019) Mental Stress Detection in University Students using Machine Learning Algorithms Volume 152.

[3] V. C. Magaa and M. Muoz-Organero, "Reducing Stress on Habitual Journeys," in Proc. 5thIEEE Int. Conf. Consum. Electron. – Berlin (ICCE-Berlin), 2015, pp. 153–157.

[4] R. W. McCarley, "Neurobiology of REM and NREM Sleep," Sleep Med., vol. 8, no. 4, pp. 302 – 330, 2007.

[5] L. Ciabattoni, F. Ferracuti, S. Longhi, L. Pepa, L. Romeo, and F. Verdini, "Real-Time Mental Stress Detection based on Smartwatch," in Proc. IEEE Int. Conf. Consum. Electron. (ICCE), 2017, pp. 110–111.

[6] SergioMuñoz, Carlos A.Iglesias "A text classification approach to detect psychological stress combining a lexicon-based feature framework with distributional representations" Volume 59,Issue 5, September 2022.

[7] Sami elzeiny, marwa qaraqemachine learning approaches to automatic stress detection 2018 ieee/acs 15th international

conference on computer systems and applications (AICCSA).

[8] GiorgosGiannakakis,Dimitris GrigoriadisKaterinaGiannakaki,Olympia Simantiraki,Alexandro "Review on Psychological Stress Detection Using Biosignals" IEEE Transactions on Affective Computing (Volume :13, ISSUE:1,01 JAN-MARCH 2022).

[9] Jacqueline Wijsman , Bernard Grundlehner,Hao Liu, Hermie Hermens, Julien Penders "Towards mental stress detection using wearable physiological sensors" on IEEE Engineering in Medicine and Biology Society.

[10] Jing Zhai , Armando Barreto "Stress Detection in Computer Users Based on Digital Signal Processing of Noninvasive Physiological Variables" 2006 International Conference of the IEEE Engineering in Medicine and Biology Society.

[11] Shalom Greene,Himanshu Thapliyal,Allison Caban-HoltA Survey of Affective Computing for Stress Detection: Evaluating technologies in stress detection for better health IEEE ConsumerElectronics Magazine (Volume: 5, Issue: 4, October 2016).

[12] Alberto de Santos Sierra,Carmen Sanchez Avila,Javier Guerra Casanova, Gonzalo Bailador delPozo

[13] [11] Peng Cen1 , Kexin Zhang1 and Desheng Zheng1, *" Sentiment Analysis Using DeepLearning", Approach Journal on Artificial Intelligence, vol.2, no.1, pp.17-27, 2020.

[14] Dirash A R1, Dr. S K Manju Bargavi2, "LSTM Based Sentiment Analysis", International Journal of Trend in Scientific Research and Development (IJTSRD) Volume 5 Issue 4.

[15] Ditiman Hazarika 1 , Gopal Konwar 1 , Shuvam Deb 1 , Dr.Dibya Jyoti Bora 2, "Sentiment Analysis on Twitter by Using TextBlob for Natural Language Processing", Proceedings of the International Conference on Research, Vol. 24observation's