iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.61865

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

Stress Detection in IT Professionals Using Machine Learning

Induja S[1], Pavan Kumar G R[2], Rita mary[3], Rama Dhananjaya V[4]
[1, 2, 3]*Students,* [4]*Associate Prof. & Head, Dept. CS & E, ICEAS, Bangalore-92*

Abstract: ***In today's fast-paced technology landscape, stress management. The work environment in the IT field is often characterized by long hrs, and high expectations, which can leading to elevated stress levels. Unchecked stress not only impacts the health and well-being of professionals but also affects in job satisfaction. This study aims to predicting the stress levels of IT professionals using machine learning techniques, thereby aiding in proactive stress management. We utilize a range of features indicative of work stress, including Heart Rate, Skin Conductivity, Hours Worked, Number of Emails Sent, and Meetings Attended. These features provide a physiological and work-related factors that contribute to stress. The application of [ML] in this context serves as an innovative approach to an increasingly pertinent issue. By leveraging the power of data analytics, this model aims to organizations. Individuals we can use for self-monitoring and early intervention, while organizations can utilize them to identify high-stress environments or roles, thereby allocating resources or interventions more effectively. Our preliminary results indicate a strong correlation between the chosen features and stress levels, demonstrating the viability of using ml for stress prediction in IT professionals. This study stands as a crucial step towards a more data-driven approach to mental health condition.***

Keywords: ***Random Forest, AdaBoost Classifier, Extra Tree Classifier, Decision Tree, Stacking etc.***

I. INTRODUCTION

In a pandemic, the people's outlook of health- care constraints and lifestyles is completely switched. Since then, covid 19 had spread a lot, causing global disturbances. The administration of educational institutions has closed across the globe to prune the growth of the disease and in welfare of all people. Considering all these circumstances the people around all countries were affected by entities like food availability and medical facilities. Many surveys were conducted to study the person's stress level the stress constraints like physiological conditions. A person can be stressed out in scenarios like worrying about losing their employment, family health conditions and about the grades in examinations. Because of working with time, limited time to complete task [1]. These kinds of stressful scenarios increase the stress levels which affects the increase of heart and muscles related issues. Generally all the students with a variation of degree. So, by observing each and every student it would be a huge task to go through their profiles. This problem makes us create a new model for predicting stress automatically. For each student who is undergoing various psychological parameters of stress and proposes a solution for that. So for this to be done, some Machine Learning algorithms and Data Science techniques are used. Maintaining track records of each student's stress levels, and studying them makes us understand the degrees of stress of the students in organization .Students are categorized into 2 sub levels in regards with the stress percentage they face : i.e., over-stressed or under- stressed. And according to that, the range of stress is highlighted based on the levels. Based on prediction percentage, the authorities give advice to the students. As a results, we create a model for unlabeled data and untrained data that will determine the stress level of students using different ML and data science techniques.

II. LITERATURE REVIEW

Stress is becoming an important factor in a person's life today. The World Health Organization, stress is a type of mental illness that affects the health of citizens. There is no one person suffer from stress or depression. Everyone gets some amount of stress. Stress is a major symptom for mental health. Stress affects every aspect of a person's life such as emotions, thoughts, and behaviors. This paper presented the study on previous research on stress detection based on machine learning algorithms. presented a stress level classification framework using the Physio Bank. The evaluated results showed that the proposed model achieved accuracy (83.33%), specificity (75%), Sensitivity (75%), Positive Predictive Value (90%), Negative Predictive Value (90%), Error Rate(16.66%), F1_Score (83.33%), Recall (75%). Kumar et al. (2019) discuss the broader implications of IoT in healthcare, including its diabetic retinopathy diagnosis for remote monitoring and early intervention, emphasizing the potential for personalized treatment strategies.

Today, psychological stress is a huge problem. Mental stress, such as anxiety, over thinking, melancholy, and emotional imbalance, was common during pandemics. Pandemics are mostly caused by grouping together many measurements such as pulse rate, body temperature, heart rate, and systolic blood oxygen saturation (spo2). When a person is under stress, their bio-signals, such as thermally, electrical, impedance, acoustic, and optical, change noticeably. These bio-signals can be used to measure stress levels. Accelerometer, body temperature, respiration, blood volume pulse electrocardiogram, (BVP), electro dermal activity, and other sensor modalities (EDA). In ML Classification methods such as Kernel Support Vector Machine, K-Nearest Neighbor, Ada-Boost, Random Forest and Decision Tree methods used to evaluate and compare the classifications. The Random Forest model beat the other approaches with F1-scores of 93.77 and 70.03 for classification model and three-class classification, respectively.

III. METHODOLOGY

It will provide a detailed description of the methods utilized to finish and operate this project successfully. Many methodologies or discoveries from this subject are mostly published in journals for others to use and enhance in future research. The approach that used to attain the project's purpose of producing a faultless output. Development Life Cycle (SDLC), which consists of three primary steps: planning, implementation, and analysis.

A. Planning

Planning is done correctly and identify every piece of data, including software and hardware. Data gathering and the software requirements are the primary component of the planning process.

B. Data collection

Gather data from IT professionals, including physiological signals (heart rate, skin conductivity), self-reported stress levels (surveys or interviews),and work-related factors (e.g., workload, deadlines) Clean the data to handle missing values, outliers, and noise. Normalize or standardize features.

IV. SYSTEM ARCHITECTURE

The purpose of this proposed system is to detect stress depend on the CNN algorithm and Haar cascade algorithm using machine learning model. The system will use a camera to analyze facial expressions and detect stress it will recorde facial expressions stored in database to help users understand their stress triggers and take necessary measures.

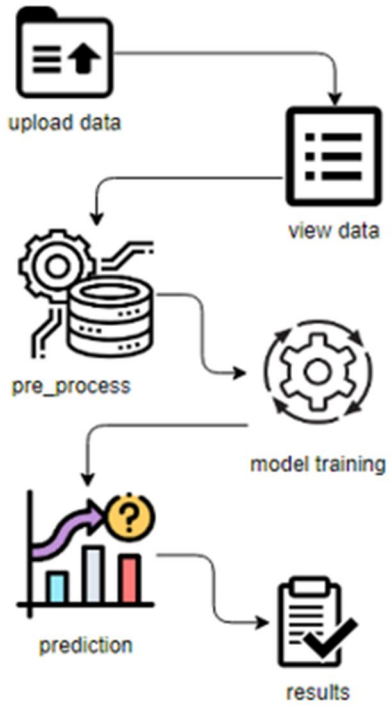


Fig 4.1 System Architecture

The live camera of the device will capture the facial expressions of the user, which will be preprocessed and then undergo feature extraction. The CNN algorithm will recognize facial features and patterns to detect stress. The Haar cascade algorithm will analyze facial expressions to detect stress-related patterns. The stress detection system will integrate the resultent of both algorithms to determine the user's stress level. The user interface will display the stress level result, along with recorded facial expressions. The recorded facial expressions will help users understand their stress triggers and take necessary measures.

V. ALGORITHM

A. *Random Forest*

A Random Forest system architecture for stress detection in IT professionals.

1) Data Collection: Gather physiological data (e.g., HRV, skin conductivity) and self-reported stress levels from IT professionals using wearable sensors or smart devices.
2) Feature Extraction: Extract relevant features from the data, including statistical measures and spectral analysis of physiological signals.
3) Random Forest Model Training: Train a Random Forest classifier using the extracted features, employing techniques like bootstrap aggregation and feature randomness.
4) Model Evaluation: Evaluate the trained Random Forest model's performance using metrics like accuracy, precision, recall, and F1-score.
5) Deployment and Integration: Deploy the trained Random Forest model in a real-world setting for continuous stress monitoring, potentially integrating it into existing IT infrastructure for seamless usage.

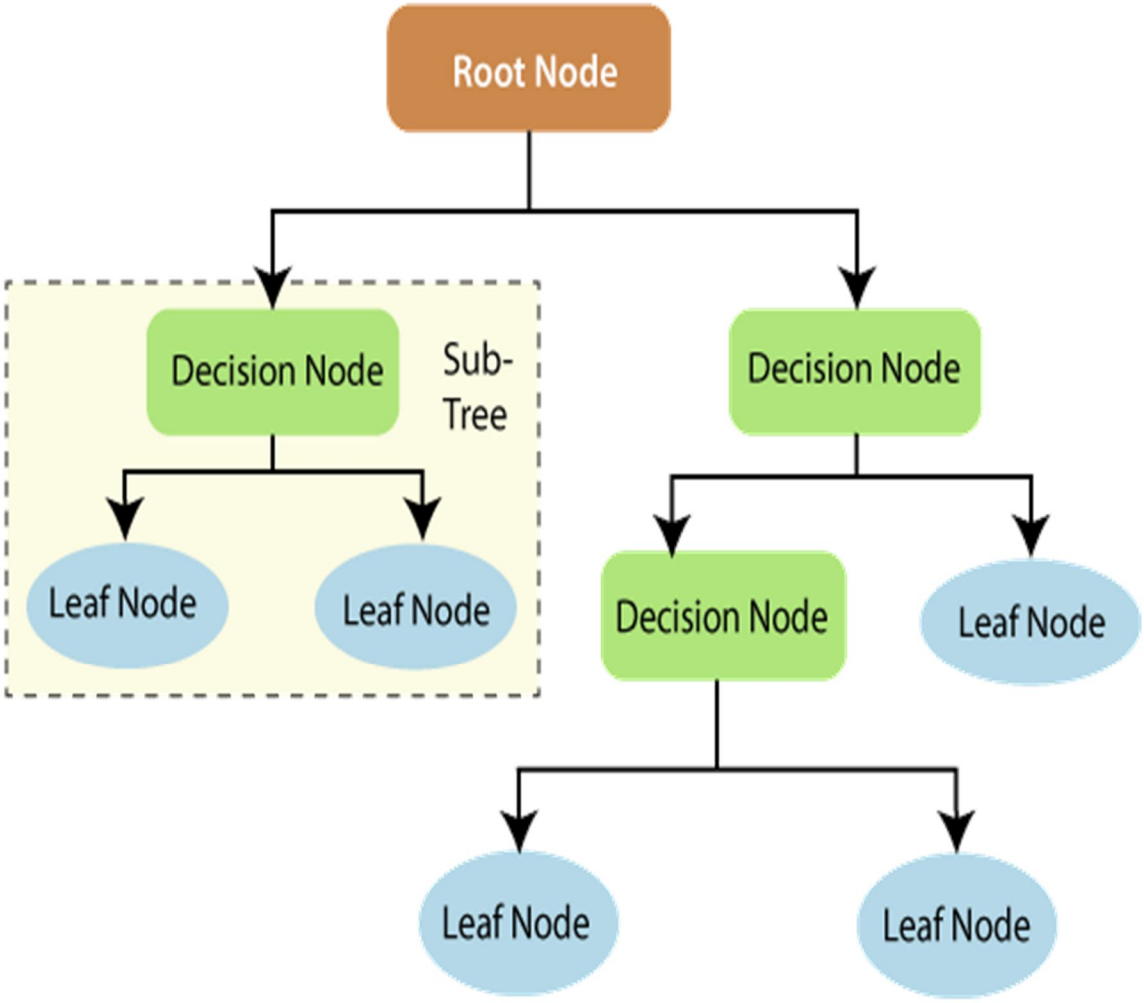


Fig 5.1 Random Forest

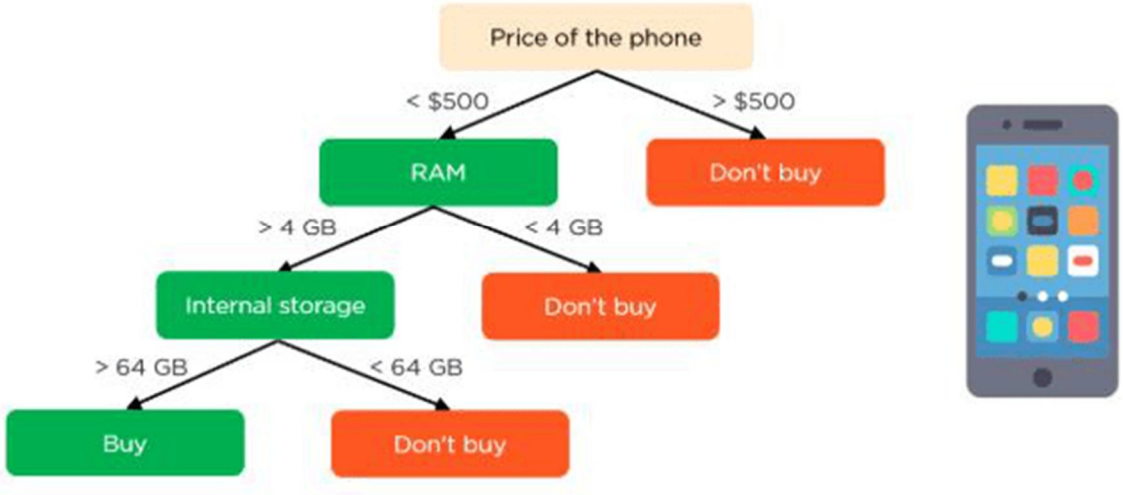


Fig 5.2 Applying decision trees in random forest

B. ADABOOST

AdaBoost, or Adaptive Boosting, is an ensemble learning technique where weak learners are sequentially trained, with each subsequent learner focusing more on instances misclassified by the previous ones. It combines their predictions through weighted majority voting, emphasizing difficult-to-classify instances, leading to a strong learner capable of robust classification performance.

VI. RESULT AND DISCUSSION

A result is the outcome of actions or occurrences, represented subjectively.

Here we can read about our project.

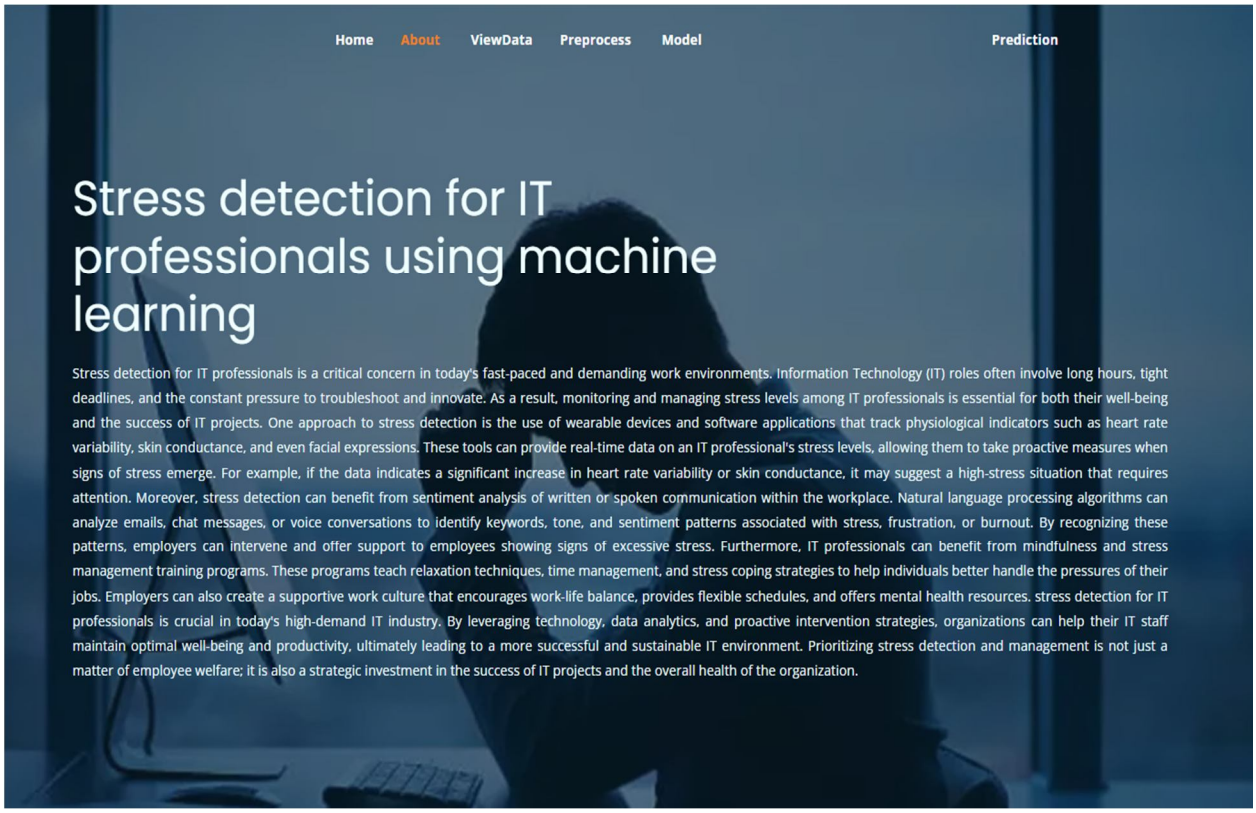


Fig 6.1: About Page

1) View Data:

In the View data page, users can view the stress dataset.

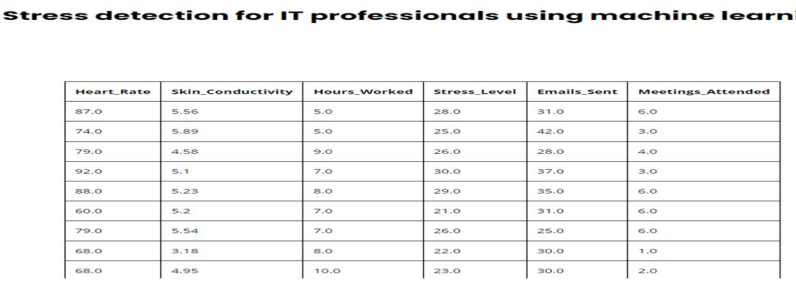


Fig 6.2: View Data Page

2) Pre-process:

Here we can pre-process and split our data into train and test.

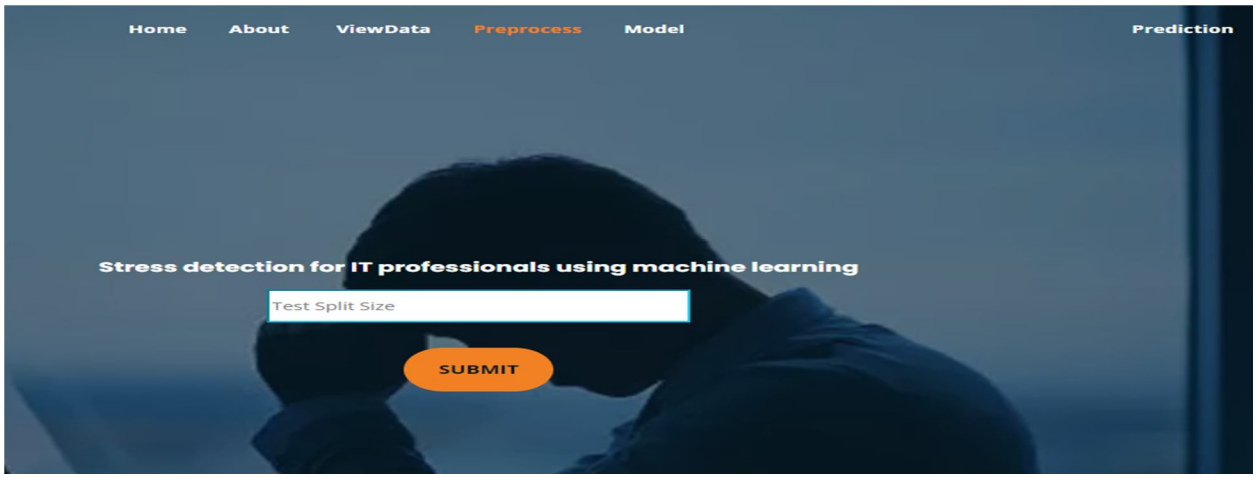


Fig 6.3: Preprocessing Page

3) Model:

Here we train our data with different ML algorithms.

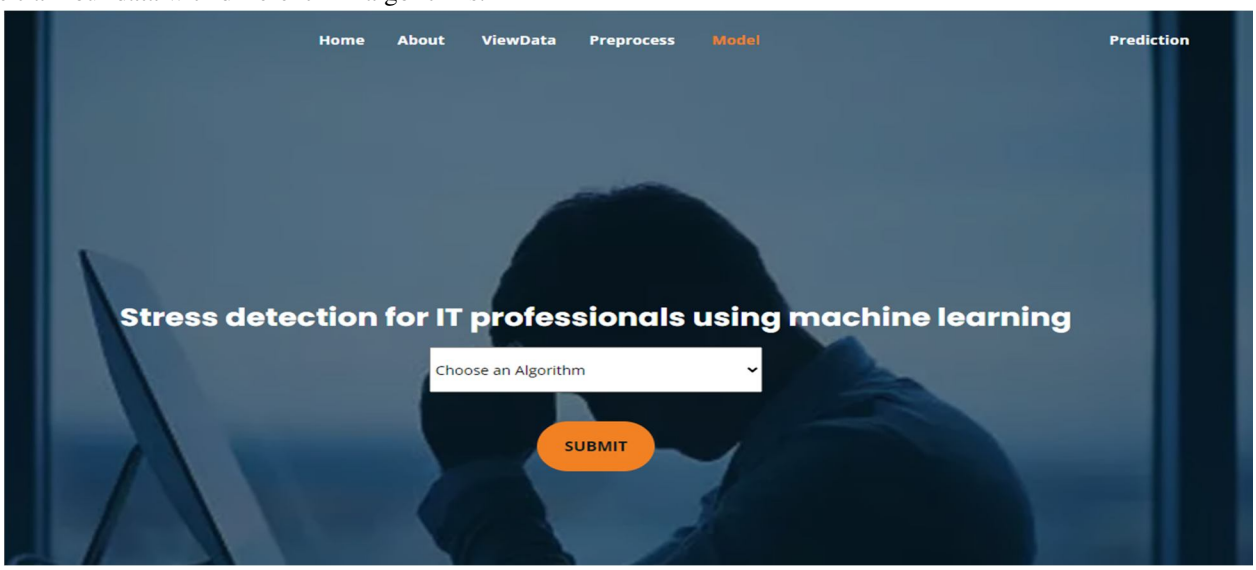


Fig 6.4: Model Page

4) Prediction:

This page show the result of theuser given input data.

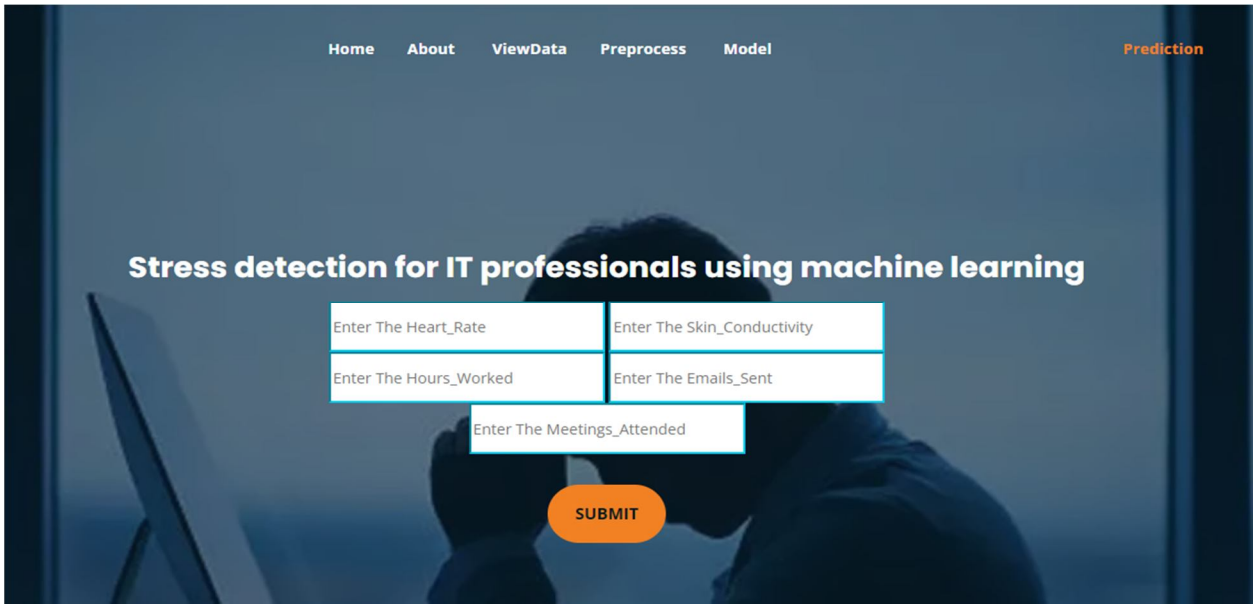


Fig 6.5: Prediction Page

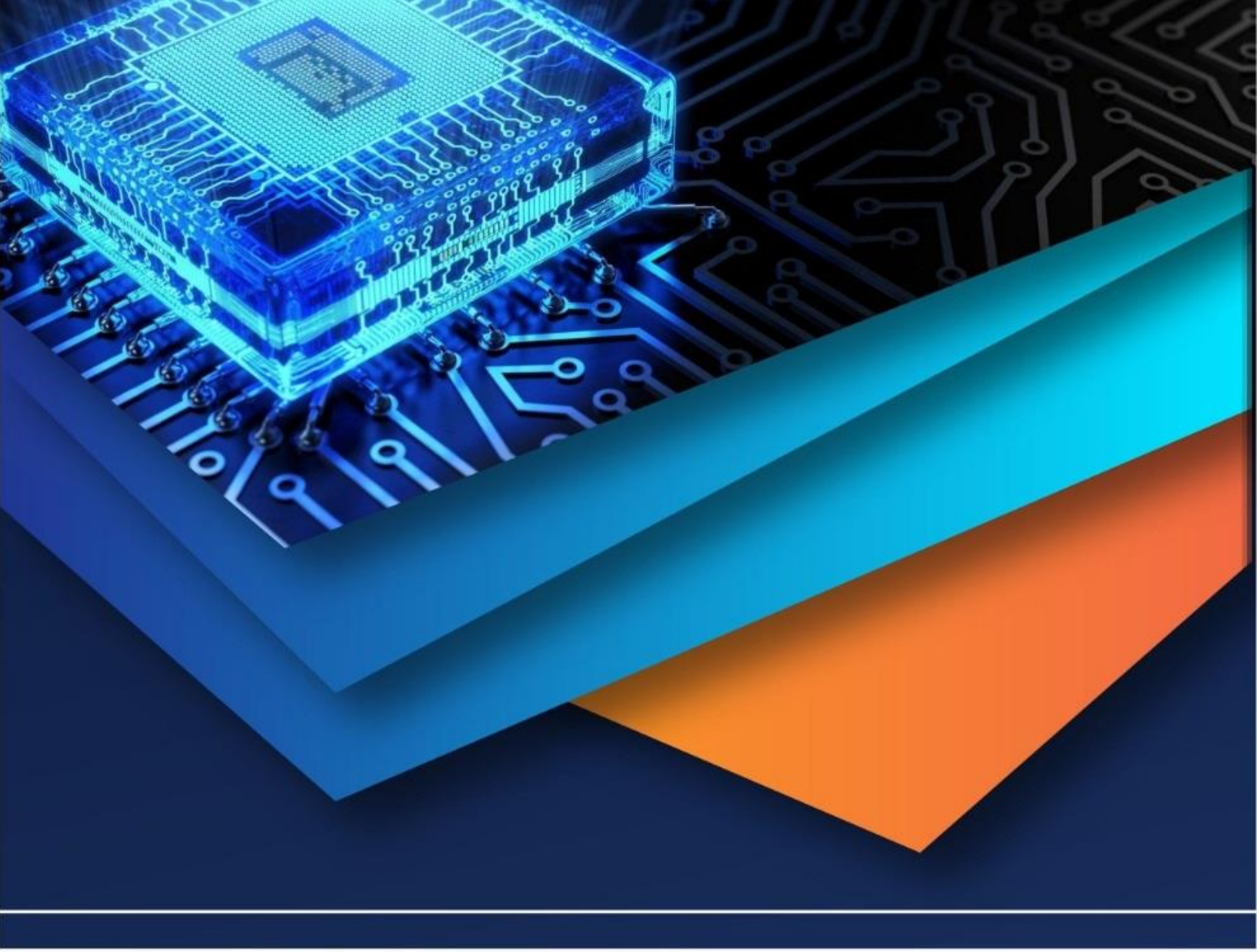
VII. CONCLUSIONS

The conclusion in the "Stress Detection in IT Professionals Using Machine Learning" lies in the comprehensive approach taken to address this critical issue on physiological data or sentiment analysis alone, our research combines both domains. By integrating variability and skin conductance with NPL techniques for sentiment analysis, we create a more holistic stress detection model. Additionally, the utilization of a diverse ensemble of regression algorithms, including RandomForestRegressor, AdaBoostRegressor, and ExtraTreeRegressor, adds robustness to the model's predictions.

REFERENCES

[1] Bakker, Mykola Pechenizkiy, Natalia Sidorova , What's your current stress level? Detection of stress patterns from GSR sensor data.Department of Computer Science Eindhoven University of Technology, 978-0-7695-4409-0/11 $26.00 © 2011 IEEE DOI 10.1109/ICDMW.2011.17

[2] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY