International Journal on Science and Technology (IJSAT)

E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Sleep Based Stress Level Prediction Using Machine Learning

Mallarapu Venkat Sai[1], N. Musrat Sultana[2], Dr V. Subba Ramaiah [3], Dr. K. Rajitha[4]

[1] Student, Department of Computer Science and Engineering, Mahatma Gandhi Institute of Technology, Gandipet, India.

[2,3,4] Assistant Professor, Department of Computer Science and Engineering, Mahatma Gandhi Institute of Technology, Gandipet, India.

Abstract

Sleep-based Stress Level Prediction Using Machine Learning is a system that evaluates an individual's stress level by analyzing physiological sleep parameters, including snoring intensity, respiration rate, body temperature, limb movement, blood oxygen saturation, eye movement, sleep duration, and heart rate. Machine learning algorithms are applied to preprocessed data to classify stress levels with high accuracy and generalization. Unlike traditional self-reporting methods, this system continuously monitors sensor-based physiological data, enabling objective and real-time stress detection. The use of decision-tree-based classifiers ensures model interpretability for both users and healthcare professionals. The system can be integrated into wearable devices and health monitoring applications, contributing to early stress detection, preventive mental wellness, and personalized healthcare.

Keywords: Sleep-Based Stress Prediction, Machine Learning, Random Forest Classifier, Physiological Parameters, Decision Tree.

1. Introduction

Stress is the body's response to any change requiring attention or action, and it has become increasingly common in daily life, even affecting individuals during sleep. Existing systems detect stress through face-to-face interaction or social media data, which are inaccurate and ineffective during sleep. The proposed system addresses this by analyzing sleep-based physiological parameters such as heart rate, respiration rate, blood oxygen saturation, body temperature, limb movement, eye movement, snoring intensity, and sleep duration using machine learning algorithms. When this data is submitted to the web application, the integrated machine learning model processes it and predicts the individual's stress level on a scale of five, enabling early detection and better health outcomes.

1.1 Problem Definition

The major challenge is detecting stress during sleep. Existing systems identify stress through face-to-face interaction, communication, or social media data such as tweets, which are inaccurate and cannot monitor a person during sleep. These methods work only on person interaction or social networking data, not on body-based sensor data. Since continuous monitoring during sleep is not possible through traditional methods, a machine learning model trained on physiological sensor data is needed to accurately detect and predict stress levels during sleep.

1.2 Existing Applications

Self-reporting and survey-based methods are widely used to collect information about individuals' thoughts, feelings, and experiences. However, these methods suffer from several limitations such as personal biases, recall bias, and subjectivity. Responses can be influenced by emotions or the desire to present oneself favorably, leading to inaccurate results. Additionally, these methods are limited to conscious experiences and fail to capture deeper processes. Overall, self-reporting and surveys provide only a partial and sometimes distorted view of human behavior and stress levels.

1.3 Proposed Application

The proposed system detects stress during sleep by analyzing physiological parameters such as snoring range, respiration rate, body temperature, limb movement, blood oxygen, eye movement, sleep duration, and heart rate. It uses machine learning algorithms, with Random Forest achieving the highest accuracy of 98.4%, to classify stress into five levels from Low to High. The system includes data preprocessing, model training, and performance evaluation to ensure accurate and reliable stress prediction.

2. Literature Survey

"**Stress Prediction of Working Employees Using ML"** by Shanmugapriya et al. Predicts employee stress using RF, SVM, KNN with Random Forest achieving 92% accuracy[1]. A web-based application was developed for stress monitoring.

"The study Machine Learning-Driven Stress Prediction" by Migliaccio, Abate, Bisogni, and Cimmino used the Sleep Health and Lifestyle dataset to predict stress with five ML models[2]. Key predictors were sleep duration, physical activity, and heart rate. A web app was developed for real-time stress detection, showing AI's potential in personalized health monitoring and early intervention.

"Evaluating SVM and Soft Voting Ensemble Methods for Classifying Stress Levels Among University Students" by Maria Susan Anggreainy and Manik Hapsara explores stress classification using anonymous student text posts scraped from online platforms[3]. The dataset was manually labeled into three categories: no stress, mild stress, and high stress. Two approaches were tested—a standalone SVM and a Soft Voting Ensemble combining SVC, Logistic Regression, and Random Forest. Results showed that the Soft Voting Ensemble achieved better accuracy, proving more effective for handling multi-class text classification of emotionally nuanced student content.

"Stress Detection Using SVM and Virtual Assistant" by Pranav Amale, Anup Vibhute, Mahesh S. Mathpati, Ashwini V. Jatti, and Somnath B. Thigale presents a chatbot model that detects stress using simple yes/no questions[4]. The system employs the Support Vector Machine (SVM) algorithm, achieving 100% accuracy on the dataset. Built with Anaconda Navigator and Python, the chatbot offers a user-friendly interface and provides users with actionable advice or medical recommendations, making it a practical tool for stress detection and early intervention.

“Health Prediction Based on Day-to-Day Life Activity using Machine Learning Approach” by Razeena Begum Shaik, Bhimavarapu Harshini, Vasireddy Aarya Chowdary, Jonnadula Malleswari, and Tummala Jaya Vardhan proposes a comprehensive health monitoring framework using smartwatch data to analyze key indicators such as blood pressure, stress levels, and heart health[5]. The system applies K-Means and Agglomerative Clustering, with the Elbow Method used to determine optimal clusters and K-Means achieving a silhouette score of 0.85. By offering personalized suggestions based on historical data, the model supports proactive health management, promotes preventive healthcare, and empowers users with actionable insights for improved well-being.

“Analysis of Students Stress Level using Machine Learning Algorithms” by S. Dharaneedharan, S. Dhivakar, S. Sibidharan, and N. T. Renukadevi examines student stress levels using academic and personal factors, with stress classified as low, moderate, or high[6]. Machine learning models including SVM, Decision Trees, ANN, and AdaBoost were applied, with Decision Trees achieving the highest accuracy by effectively capturing complex relationships. Key predictors identified were study hours, sleep duration, and GPA, highlighting critical factors influencing student stress.

“DeepNeuroRest: Deep Learning-Assisted Stress Detection in Sleep with Smart Web App Insights” by Angelina George, Achal Baniya, Alphonsa Jose, Roshni M. Balakrishnan, and Peeta Basa Pati presents a two-phase study on stress prediction using machine learning and deep learning models[7]. In Phase 1, nine ML models and five DL models were trained, with CatBoost and Extra Trees achieving 98.77% accuracy, while the Multi-Layer Perceptron emerged as the best deep learning model with 97.85% accuracy. Fine-tuned AC-GAN was used for dataset enhancement, improving model robustness. Phase 2 involved developing an email client (using Microsoft Azure, Flask, and Automation) and a web client (using Wix, Streamlit, and ngrok) to provide real-time stress analysis through a speedometer graph. The study demonstrates the integration of AI models with practical applications for stress monitoring and management.

“Enhanced Stress Detection via Heart Rate Data: A Feature Selection and Stacking Ensemble Approach” by Kazi Ferdous Hasan, Md. Shakil Bhuiyan, Nahida Hoque, Alfe Suny, and Rafiqul Islam proposes a machine learning-based stress detection system using heart rate data[8]. The approach involves preprocessing, outlier handling, ANOVA-based feature evaluation, K-Best feature selection, and data normalization. Multiple ML models—including Random Forest, KNN, Decision Trees, and MLP—were trained, and a stacking ensemble was used for enhanced prediction. The stacking model achieved 99.95% accuracy with zero false positives or negatives, demonstrating high reliability, though with longer computation time. This method shows promise for effective stress monitoring, requiring professional validation for real-world application.

3. Design and Methodology

The design and methodology for stress level prediction using machine learning involve using Python and its key libraries. Pandas and NumPy are used for data handling, Scikit-learn provides machine learning algorithms, Streamlit enables interactive web applications, Matplotlib and Seaborn are used for visualizations, Pickle helps in saving and loading trained models, and the OS module allows file and environment management.

3.1 Technologies Used

Hardware Requirements

- RAM: 8GB DDR4
- Processor: Any Intel or AMD x86-64 processor
- Hard Disk: 10GB HDD (SSD recommended)
- Input Devices: Keyboard and Mouse
- Output Devices: Screen (Monitor)

Software Requirements

- Operating System: Windows 10/11 or Linux
- Language: Python 3.10
- IDE: Jupyter Lab or VS Code or Google Colab
- Libraries: NumPy, Pandas, Matplotlib, Seaborn, Sklearn, Streamlit

3.2 System Design

The design and methodology involve using Python and its key libraries to achieve accurate results, forming a robust framework for implementing and deploying stress prediction models.

3.2.1 System Architecture

The system architecture takes physiological input parameters including snoring rate, respiration rate, body temperature, limb movement, blood oxygen, eye movement, sleep duration, and heart rate. These are trained under five algorithms: K Neighbors Classifier, Decision Tree Classifier, Random Forest Classifier, Gradient Boosting Classifier, and XGB Classifier. The dataset is preprocessed by removing null and unnecessary values. The user enters sleep data through the frontend, which sends it to the backend via a REST API. The backend preprocesses the data, extracts features, and uses a Random Forest model to predict the stress level. The result is stored in a PostgreSQL database and sent to an AI Chatbox layer, which generates personalized relaxation or health suggestions.

International Journal on Science and Technology (IJSAT)

E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

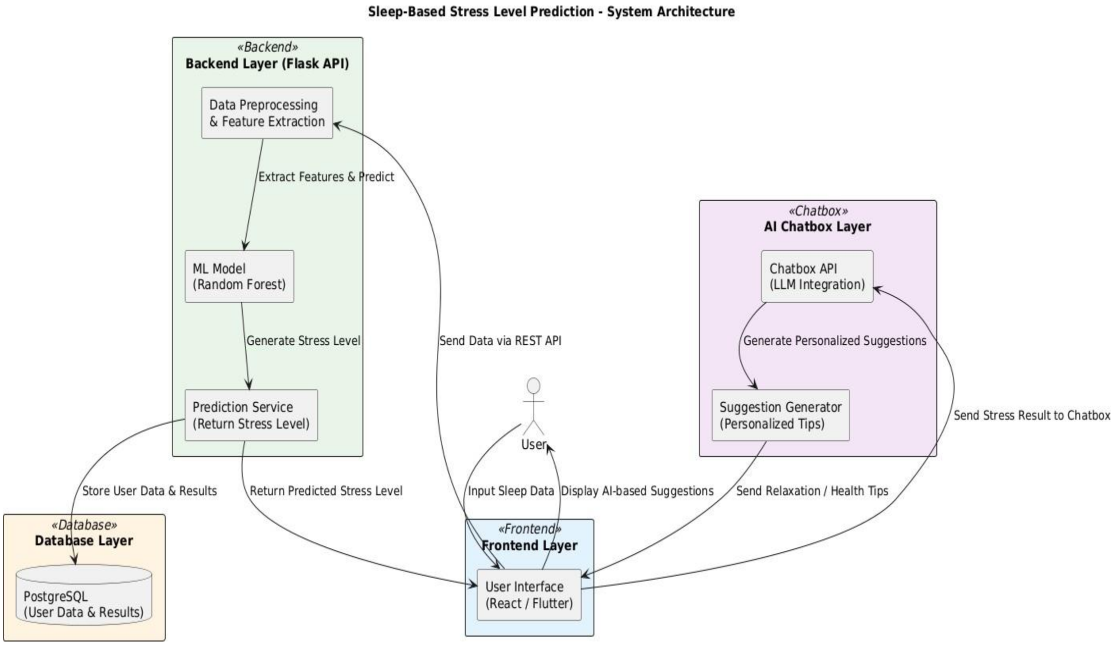


Fig 3.2.1: System Architecture of Stress Level Prediction

3.2.2 Class Diagram

It describes the structure of the system by showing classes, attributes, operations, and relationships among objects. It includes classes like User and Stress Detection System showing how individuals interact to determine the stress level.

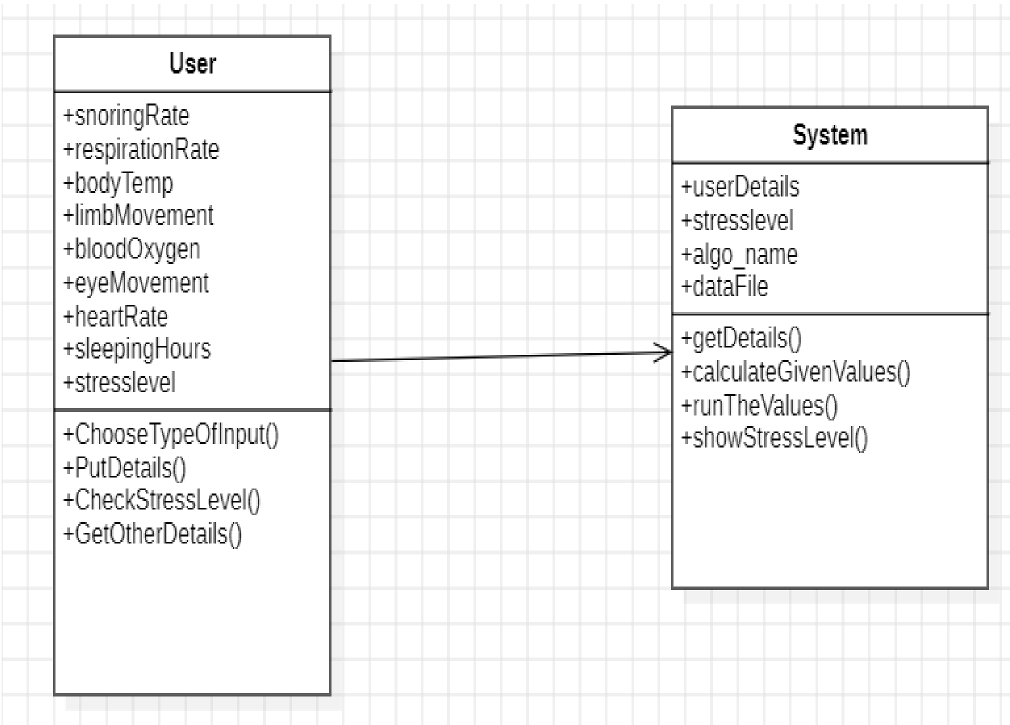


Fig 3.2.2: Class Diagram of Stress Level Prediction

3.2.3 Sequence Diagram

It shows the chronological interactions between the user and the stress prediction system including data preprocessing, dataset splitting, model training, model evaluation, and final stress prediction result.

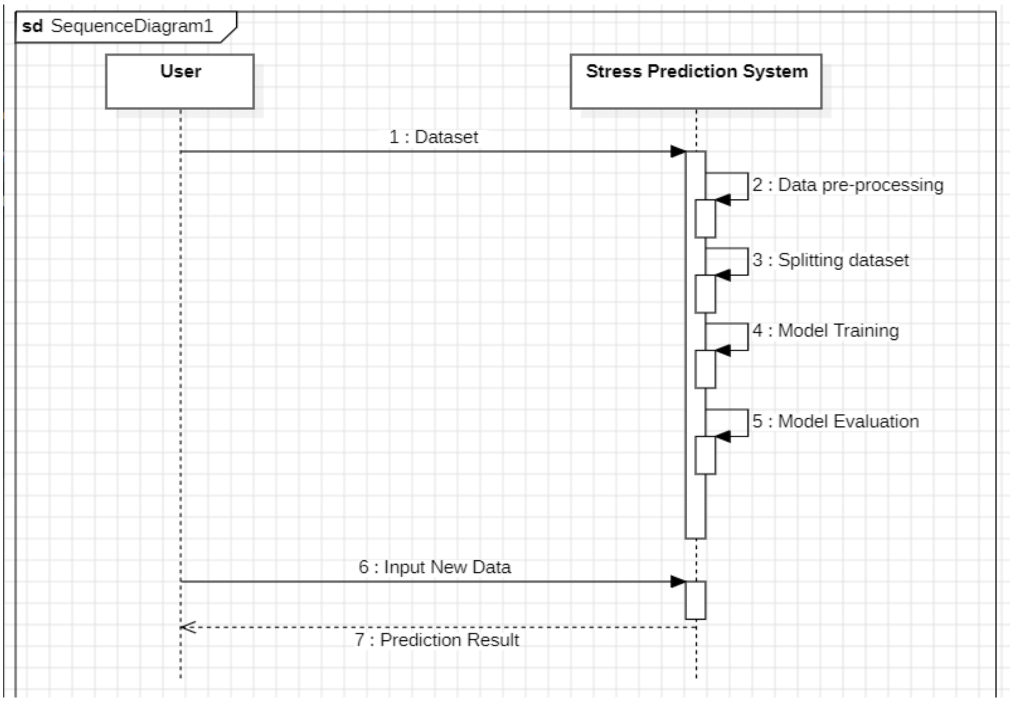


Fig 3.2.3: Sequence Diagram of Stress Level Prediction

3.2.4 Use Case Diagram

It shows the interaction between user and system where the user selects input type, enters values, runs the file, and the system calculates and displays the predicted stress level.

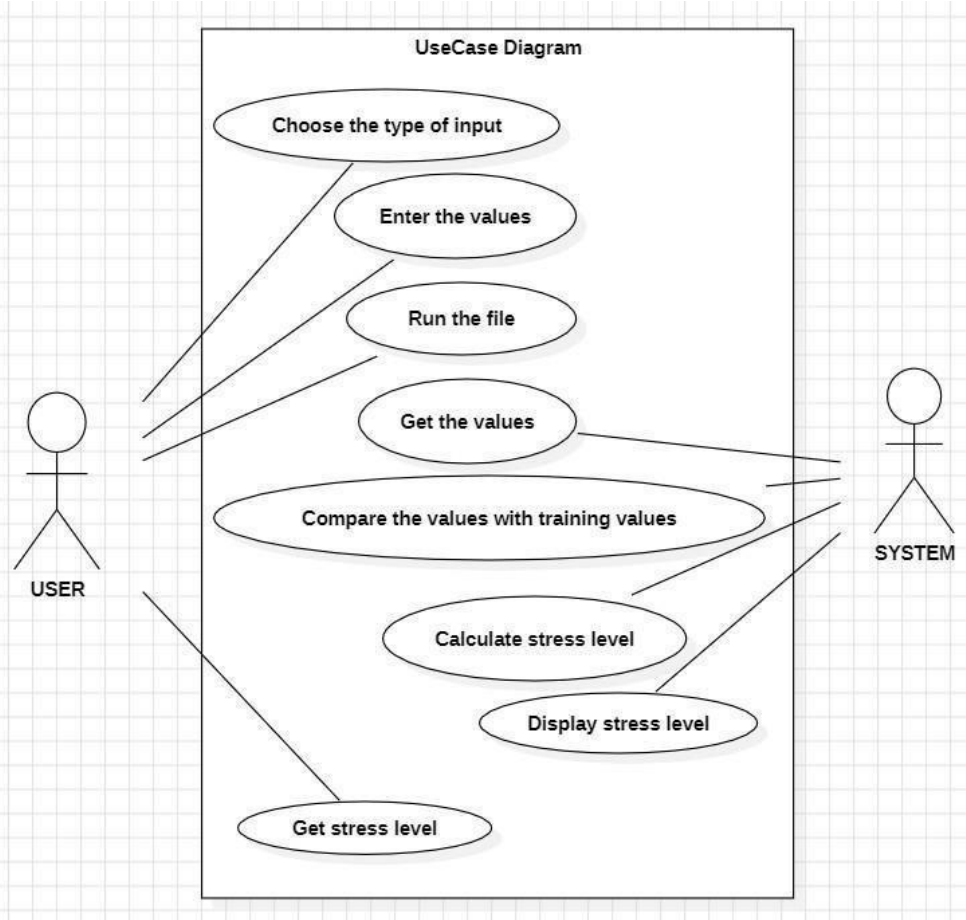


Fig 3.2.4: Use Case Diagram of Stress Level Prediction

3.2.5 Activity Diagram

It illustrates the stress level detection system which collects data from the user, estimates values, validates the range, compares with trained model values, calculates, and displays the stress level.

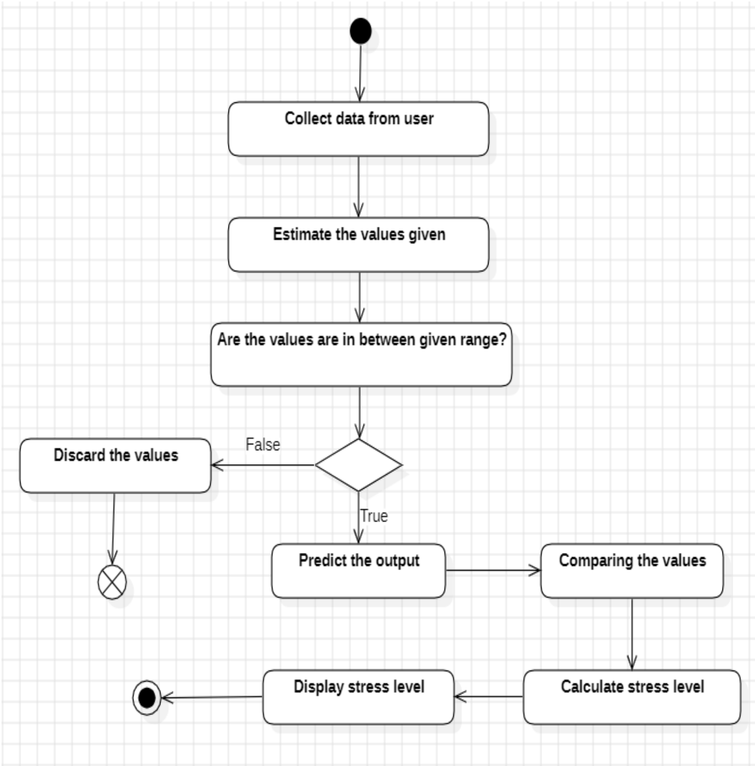


Fig 3.2.5: Use Case diagram of Stress Level Prediction

4. Implementation

The system is developed as a smart healthcare application integrating frontend, backend, machine learning models, database management, and prediction modules.

4.1 Frontend Implementation

Built using HTML, CSS, JavaScript, and REST API Integration. Provides a clean, responsive interface for users to input physiological details and view prediction results.

4.2 Backend Implementation

Built using Python 3.10, Flask, PostgreSQL, Pickle, Pandas and NumPy for handling prediction requests, data preprocessing, and model integration.

4.3 Machine Learning Implementation

Uses Scikit-learn, Jupyter Lab/Google Colab/VS Code, Matplotlib and Seaborn for model development and evaluation.

4.4 Algorithms

- Algorithm 1: Data Preprocessing and Feature Preparation
- Algorithm 2: Decision Tree Stress Prediction
- Algorithm 3: Stress Level Classification
- Algorithm 4: Personalized Suggestion Generation

5. Testing and Results

The Sleep Based Stress Level Prediction System underwent thorough testing across all modules to evaluate functionality, accuracy, and reliability. Each module was assessed for correct behaviour in real-world scenarios using physiological sleep data. Below are the major interfaces and their testing outcomes:

5.1 Testing

The system is tested using multiple testing strategies:

- **Unit Testing** - Individual components such as data preprocessing functions, ML model prediction functions, API endpoints, and database operations are tested independently.
- **Integration Testing** - Interaction between frontend, backend, database, and machine learning model is tested to ensure seamless communication and proper data flow.
- **Machine Learning Model Testing** - Trained models are tested using training and testing datasets to verify prediction accuracy, model stability, and classification performance across Logistic Regression, Random Forest, KNN, Gradient Boosting, and Decision Tree.
- **API Testing** - Backend APIs are tested using tools like Postman to verify request handling, input validation, and response generation.
- **Functional Testing** - System is tested from user perspective to validate workflows such as entering sleep parameters, generating stress predictions, storing results, and displaying personalized health suggestions.

5.2 Results

5.2.1 Home Page - Provides overview of AI Stress Prediction platform with navigation options and action buttons.

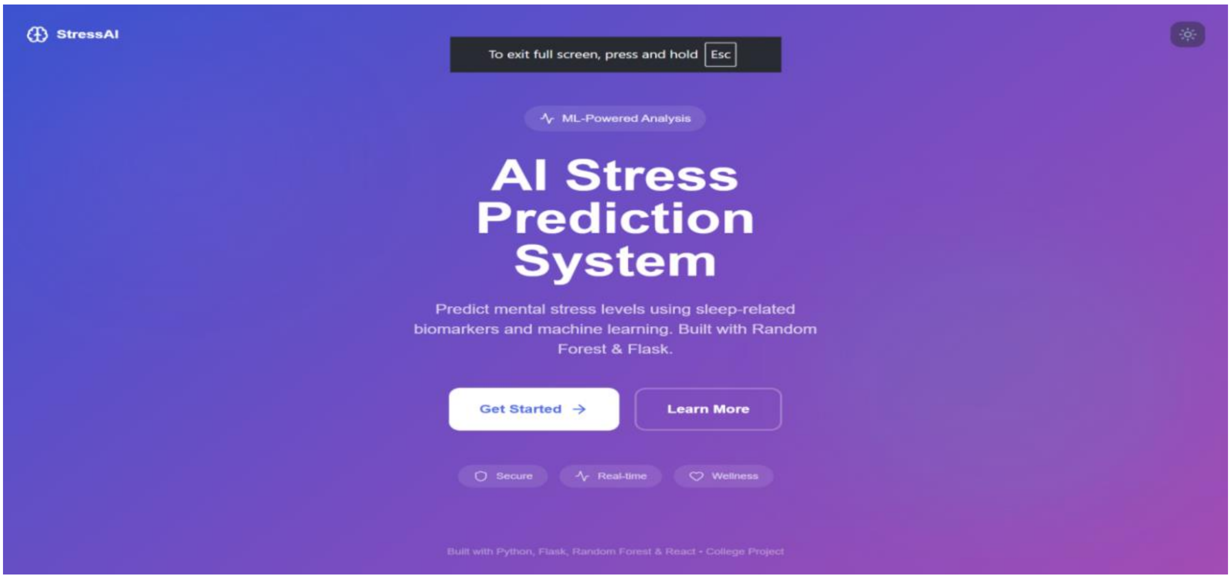
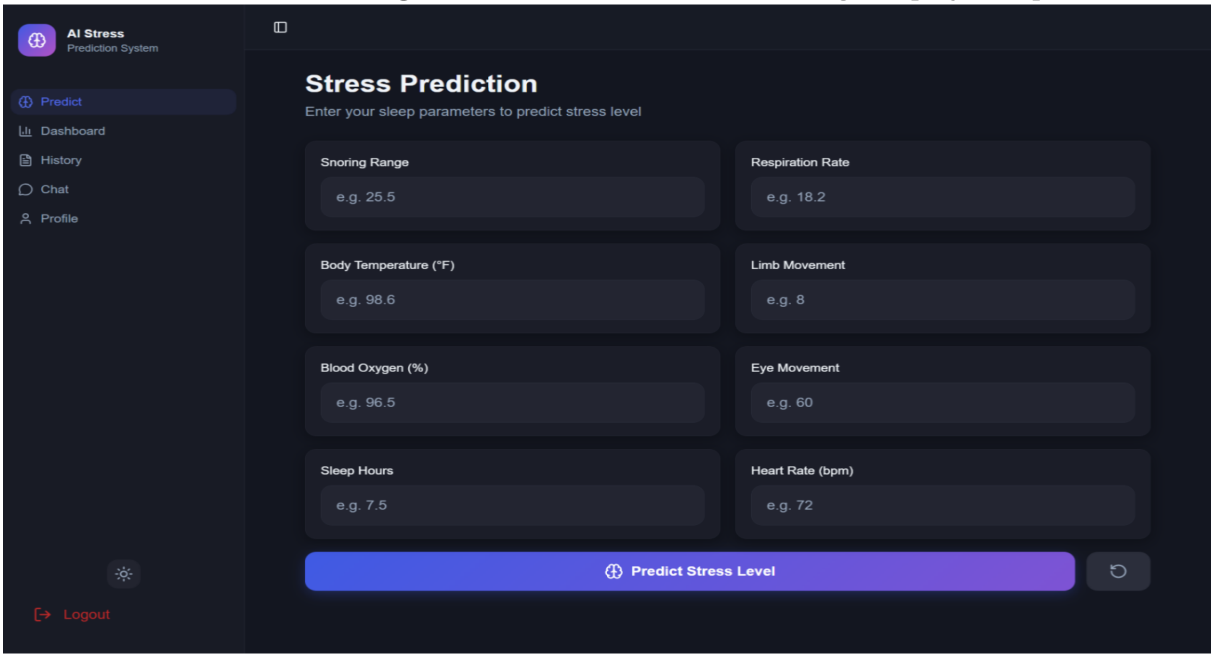


Figure 5.2.1: Home Page of AI Stress Level Prediction

5.2.2 Stress Prediction Input Page The Stress Prediction Input Page allows users to enter sleep-related physiological parameters including snoring range, respiration rate, body temperature, limb movement, blood oxygen level, eye movement, sleep hours, and heart rate. Users can click on the "Predict Stress Level" button to generate the prediction result instantly.

Figure 5.2.2: Stress Prediction Input Page of AI Stress Level Prediction

5.2.3 Stress Prediction Result Page The Stress Prediction Result Page displays the predicted stress level



such as "Low" along with a visual emoji indicator. The system also provides personalized feedback and wellness suggestions encouraging users to maintain healthy sleep habits and stay active.

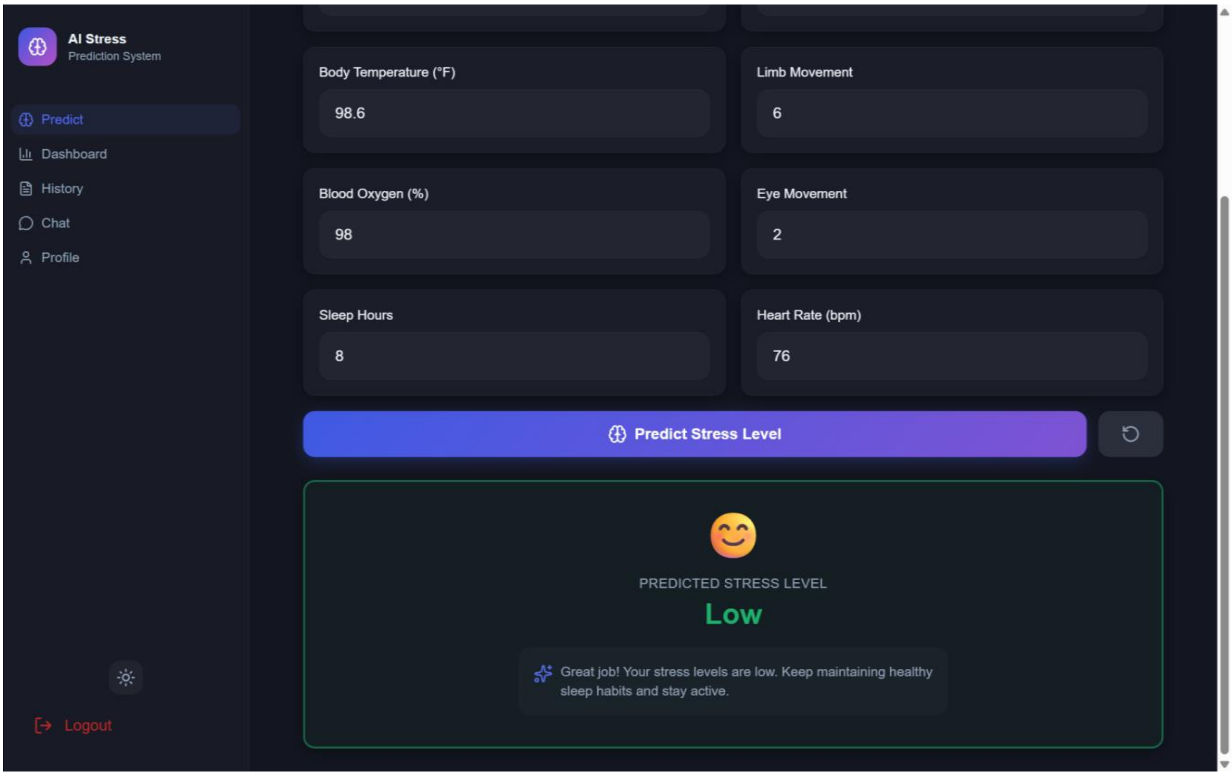


Figure 5.2.3: Stress Prediction Result Page of AI Stress Level Prediction

5.2.4 Moderate Stress Prediction Result Page Displays the output when the system detects a moderate level of stress. The predicted result is indicated as "Moderate" accompanied by a visual emoji. Personalized recommendations include improving sleep schedules, reducing screen time, and practicing deep breathing relaxation techniques.

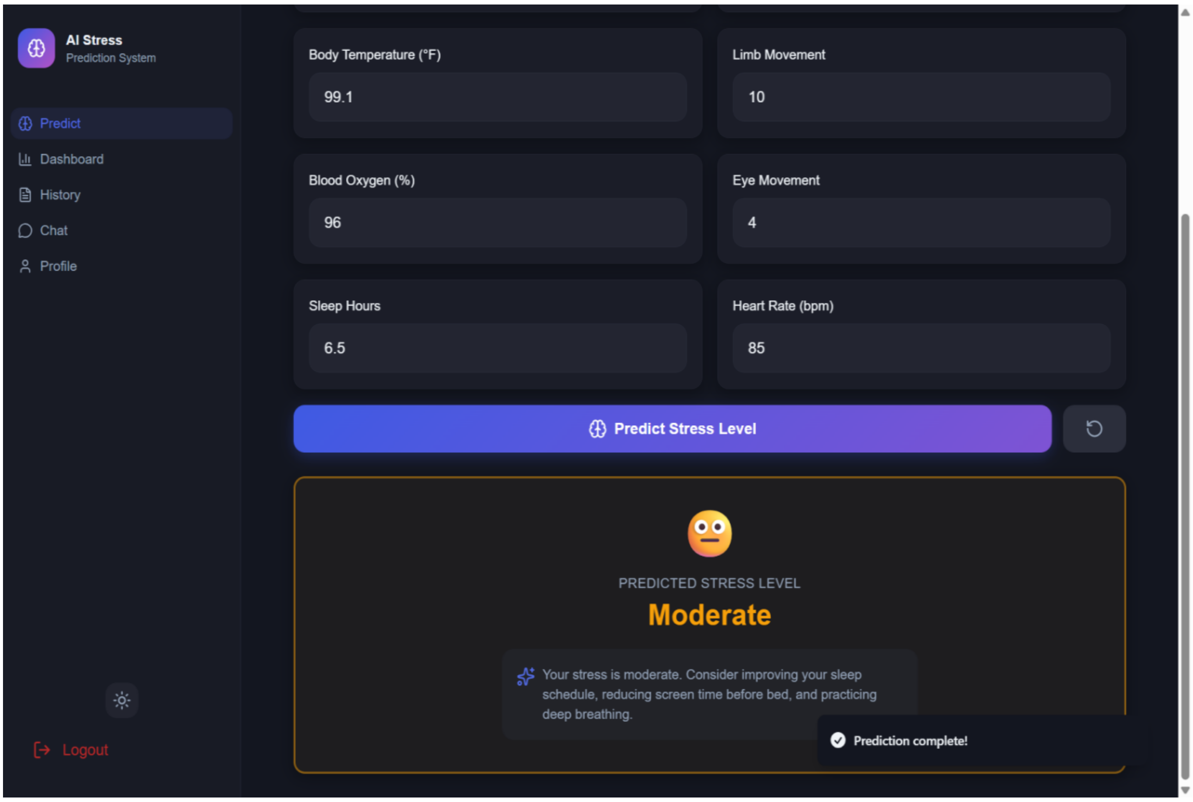


Figure 5.2.4 Moderate Stress Prediction Result Page of AI Stress Level Prediction

5.2.5 High Stress Prediction Result Page Displays the output when the system identifies a high level of stress. The predicted result is prominently displayed as "High" with a visual emoji. Recommendations include taking regular breaks, improving sleep quality, practicing relaxation techniques, and seeking professional guidance if necessary.

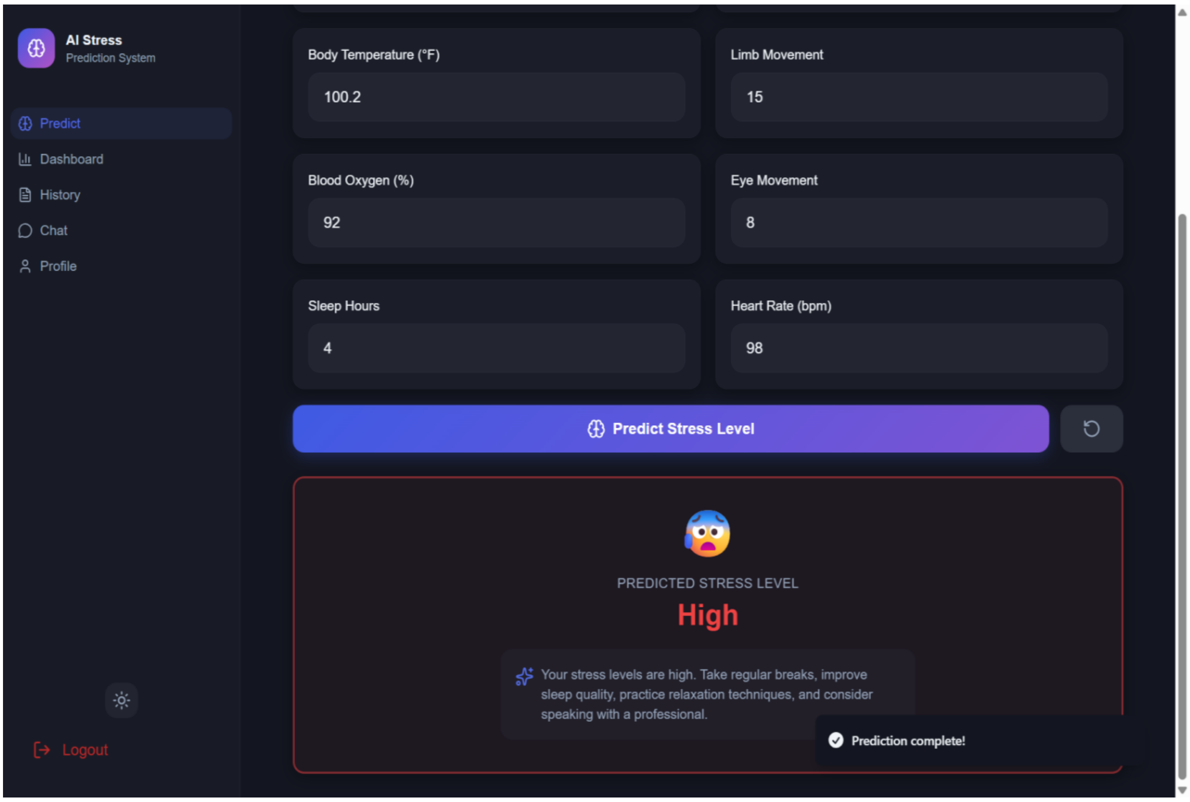


Figure 5.2.5 High Stress Prediction Result Page of AI Stress Level Prediction

5.2.6 Dashboard Page - Shows total predictions, latest stress level, stress distribution chart, and stress trend graph.

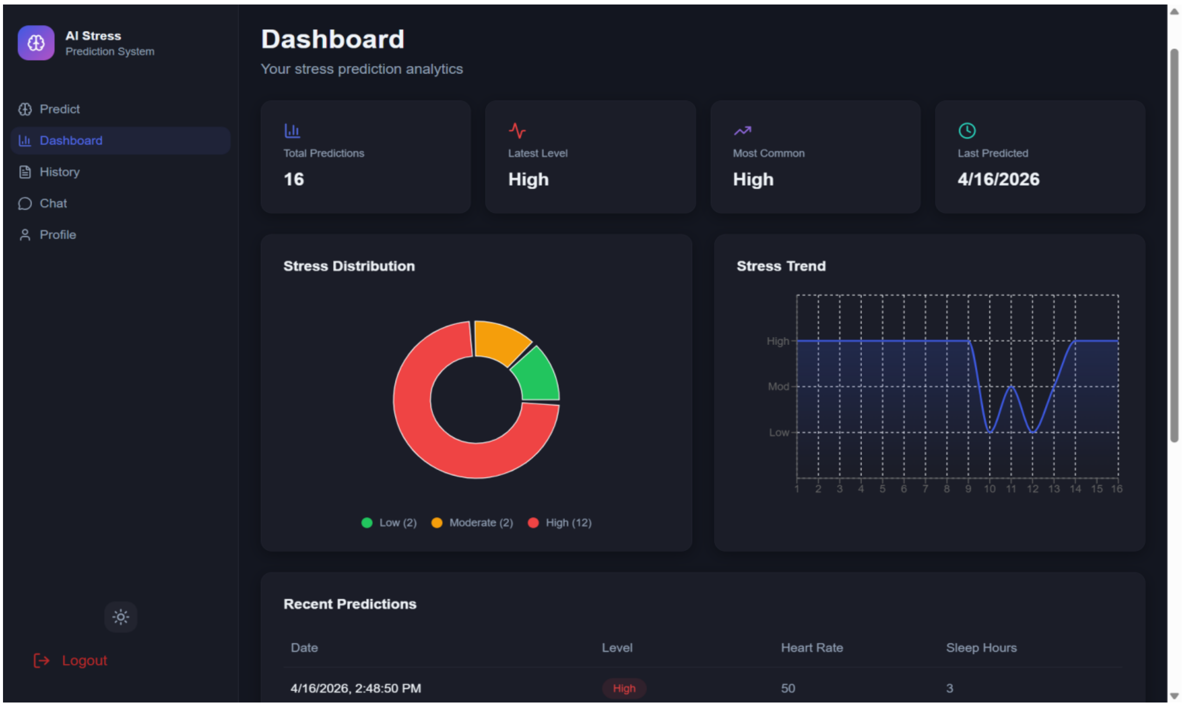


Figure 5.2.6 Dashboard Page of AI Stress Level Prediction

5.2.7 Prediction History Page - Displays chronological list of all previous predictions with export and clear options.

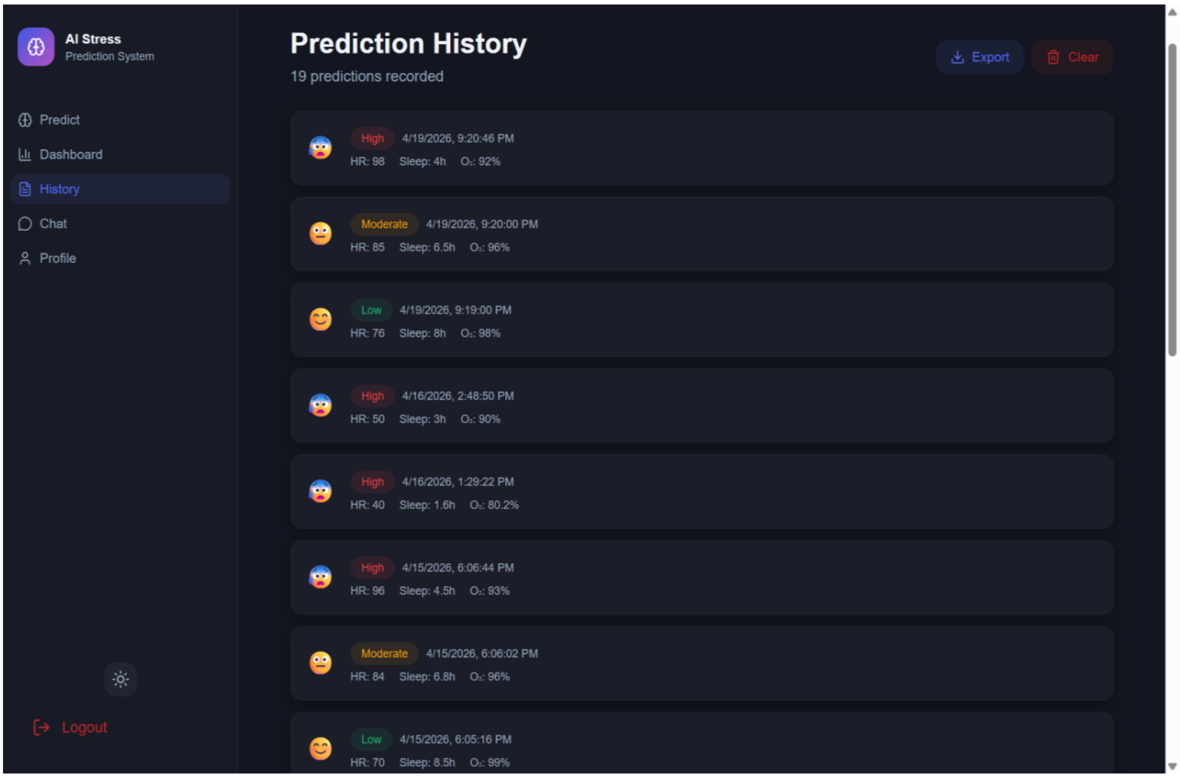


Figure 5.2.7 Dashboard Page of AI Stress Level Prediction

5.2.8 AI Assistant Chat Page - Interactive chatbot providing stress management and wellness recommendations.

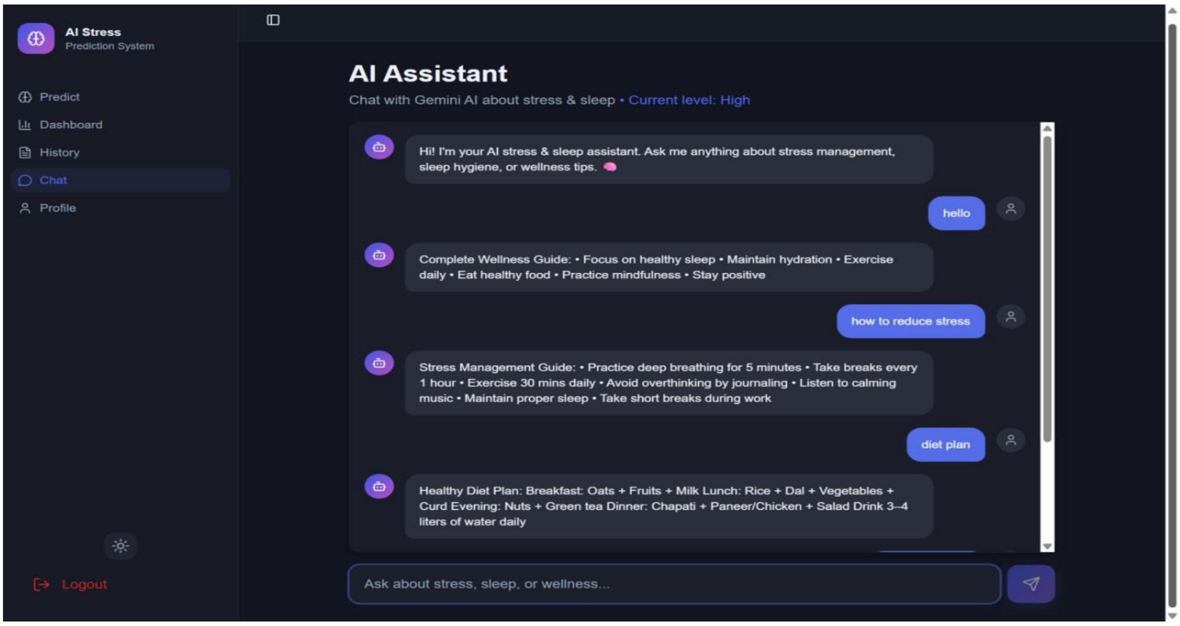


Figure 5.2.8 AI Assistant Chat Page of AI Stress Level Prediction

5.2.9 Profile Page - Displays user details, total predictions, and information about the application.

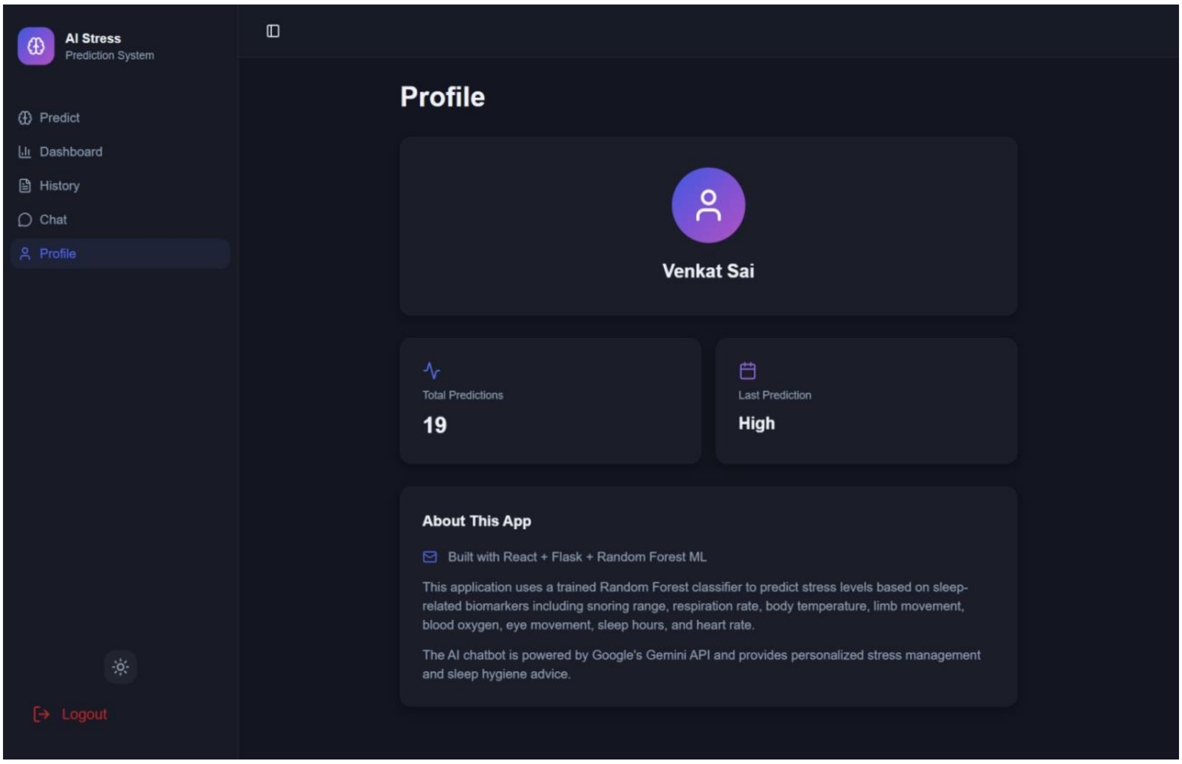


Figure 5.2.9 Profile Page of AI Stress Level Prediction

5.3 Performance Analysis

5.3.1 Model Accuracy - Random Forest achieved highest accuracy of 98.4%, followed by XGBoost 97.8%, Gradient Boosting 97.1%, Decision Tree 96.7%, and KNN 85.2%.

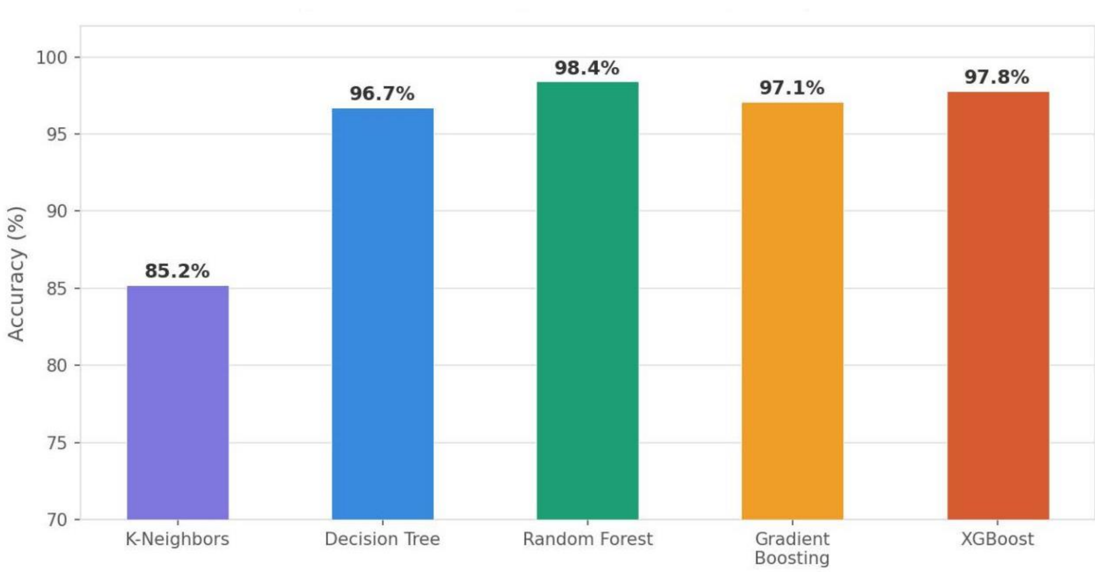


Figure 5.3.1 Machine Learning Model Accuracy Comparison

5.3.2 Feature Importance - Heart rate (0.23) and Blood oxygen (0.20) are the most influential features, followed by Sleep hours (0.18) and Respiration rate (0.13).

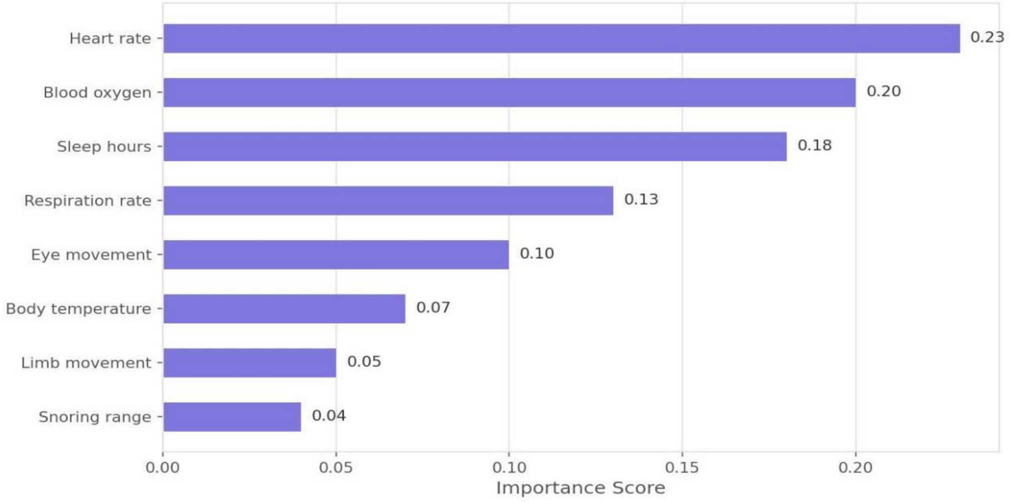


Figure 5.3.2 Feature Importance Analysis

5.3.3 Classification Performance - Random Forest achieves Precision, Recall, and F1-Score ranging between 0.97 and 0.99 across all stress classes.

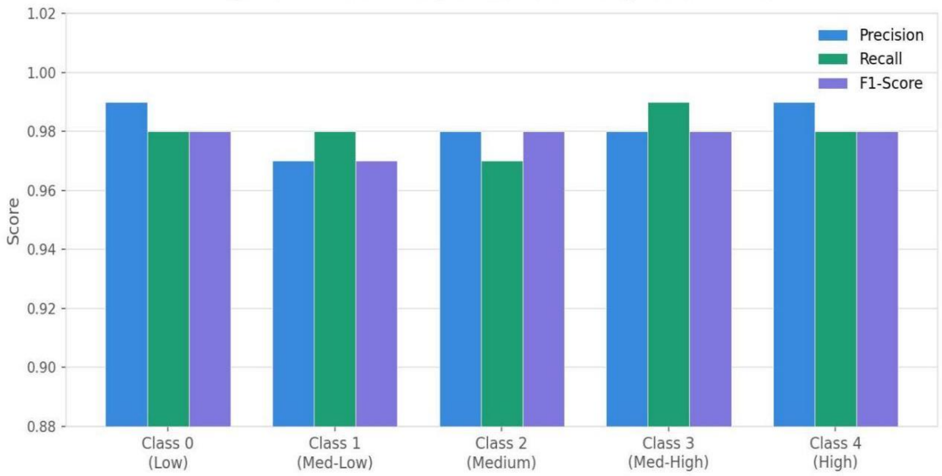


Figure 5.3.3 Precision, Recall, and F1-Score Analysis

5.3.4 Train vs Test Accuracy - Training accuracy stabilizes at 98.5% and testing accuracy at 98.2%, confirming no significant overfitting.

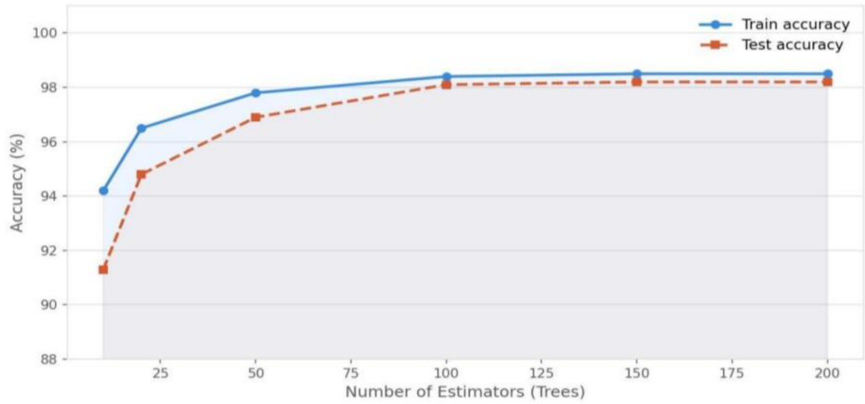


Figure 5.3.4 Train vs Test Accuracy Analysis

6. Conclusion and Future Scope

6.1 Conclusion

The proposed Sleep Based Stress Level Prediction System presents an efficient and intelligent solution for predicting human stress levels using machine learning techniques and sleep-related physiological parameters. It effectively addresses the limitations of traditional stress assessment methods such as delayed diagnosis, manual observation, and lack of continuous monitoring.

The system analyzes important health parameters such as snoring range, respiration rate, body temperature, limb movement, blood oxygen level, eye movement, sleep duration, and heart rate to accurately predict stress levels. Among multiple machine learning models tested, the Random Forest Classifier achieved the highest accuracy and was selected as the final model for deployment due to its strong performance, better generalization, and reliable prediction capability.

The system also provides personalized health suggestions based on predicted stress levels, helping users take preventive actions for better mental wellness. The integration of a user-friendly frontend, Flask backend, PostgreSQL database, and real-time API communication ensures smooth system interaction and effective healthcare monitoring. Overall, the system achieves improved prediction accuracy, faster response time, and better healthcare support, making it a reliable solution for modern stress management and preventive healthcare systems.

6.2 Future Scope

The proposed system can be further enhanced by implementing several improvements to support real-world deployment and scalability:

- **Deep Learning Integration** - Implementing advanced deep learning models such as Artificial Neural Networks (ANN), Long Short-Term Memory (LSTM), and hybrid machine learning approaches to further improve prediction accuracy and model performance.
- **Wearable Device Integration** - Real-time wearable device integration can be added to collect live physiological data automatically, reducing manual input and improving monitoring efficiency.
- **Mobile Application Development** - The system can be extended to support mobile application development for better accessibility and continuous health tracking.
- **Hospital Management Integration** - Integration with hospital management systems and electronic health records can improve clinical usage and professional healthcare monitoring.

- **Additional Features** - Doctor consultation support, mental health recommendations, and personalized therapy suggestions can further enhance the usefulness of the system.
- **Security and Scalability** - Stronger security mechanisms, cloud deployment, and compliance with healthcare standards can improve reliability and support large-scale practical implementation in smart healthcare environments.

References

1. S. P, P. Balasubramanie, C. R. Dhivyaa, D. S. Raia and P. Dhivya, "Stress Prediction of Working Employees in Different Environments Using Machine Learning Techniques," 2025 6th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI), Goathgaun, Nepal, 2025, pp. 1409-1414.
DOI: https://doi.org/10.1109/ICMCSI64620.2025.10883523
2. M. Migliaccio, A. Abate, C. Bisogni and L. Cimmino, "Machine Learning-Driven Stress Prediction: A Comparative Analysis and Web Application Using the Sleep Health and Lifestyle Dataset," 2025 6th International Conference on Bio-engineering for Smart Technologies (BioSMART), Paris, France, 2025, pp. 1-4.
DOI: https://doi.org/10.1109/BioSMART66413.2025.11046172
3. M. S. Anggreainy and M. Hapsara, "Evaluating SVM and Soft Voting Ensemble Methods for Classifying Stress Levels Among University Students," 2025 IEEE International Conference on Artificial Intelligence for Learning and Optimization (ICoAILO), Bali, Indonesia, 2025, pp. 7-12. DOI: https://doi.org/10.1109/ICoAILO66760.2025.11156062
4. P. Amale, A. Vibhute, M. S. Mathpati, A. V. Jatti and S. B. Thigale, "Stress Detection Using SVM and Virtual Assistant," 2025 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), Bhopal, India, 2025.
DOI: https://doi.org/10.1109/SCEECS64059.2025.10940886
5. R. B. Shaik, B. Harshini, V. A. Chowdary, J. Malleswari and T. J. Vardhan, "Health Prediction Based on Day-to-Day Life Activity using Machine Learning Approach," 2025 International Conference on Electronics, Computing, Communication and Control Technology (ICECCC), Bengaluru, India, 2025, pp. 1-8.
DOI: https://doi.org/10.1109/ICECCC65144.2025.11063746
6. U. Harsha Vardhan, S. Dhivakar, S. Sibidharan and N. T. Renukadevi, "Analysis of Students Stress Level using Machine Learning Algorithms," 2025 5th International Conference on Expert Clouds and Applications (ICOECA), Bengaluru, India, 2025, pp. 916-922.
DOI: https://doi.org/10.1109/ICOECA66273.2025.00160
7. A. George, A. Baniya, A. Jose, R. M. Balakrishnan and P. B. Pati, "Human Stress Detection in and Through Sleep using Artificial Intelligence," 2023 4th International Conference on Electronics and Sustainable Communication Systems (ICESC), 2023.
DOI: https://doi.org/10.1109/ICESC57686.2023.10193562
8. K. F. Hasan, M. S. Bhuiyan, N. Hoque, A. Suny and R. Islam, "Enhanced Stress Detection via Heart Rate Data: A Feature Selection and Stacking Ensemble Approach," 2025 2nd International Conference on Next-Generation Computing, IoT and Machine Learning (NCIM), Gazipur, Bangladesh, 2025, pp. 1-6.
DOI: https://doi.org/10.1109/NCIM65934.2025.11160016