AUTOMATED SLEEP DISORDER DETECTION USING HYBRID AI MODELS

"R. Sravani (Assistant professor)" "B. Jahnavi" "G. S. S. Arvind" "I. Venkatesh" "N. Gayathri"

"Dept Of Computer Science and Engineering"

"Lendi Institute Of Engineering and Technology"

Abstract:
This project proposes a hybrid framework for the early detection of sleep disorders, specifically insomnia and sleep apnea, which significantly impact physical health and cognitive performance. To address the limitations of traditional diagnostic methods, patient health parameters such as age, gender, sleep duration, sleep quality, BMI, heart rate, height, weight, steps, stress, physical activity, occupation, and blood pressure are analysed. A Convolutional Neural Network (CNN) is utilized for feature extraction, while the Synthetic Minority Over-sampling Technique (SMOTE) handles class imbalance. Classification is performed using a voting ensemble of Gradient Boosting, Naive Bayes, Quadratic Discriminant Analysis (QDA), and Random Forest, with Random Forest improving generalization and reducing overfitting. The system categorizes individuals into Normal, Sleep Apnea, and Insomnia and is deployed through a Flask-based web application, enabling real-time, cost-effective, and scalable healthcare monitoring and decision support.

Keywords: Sleep Disorder Detection, CNN, Random Forest, Gradient Boosting, Ensemble Learning, SMOTE, Healthcare AI

I. INTRODUCTION

Sleeping is considered one of the basic physiological functions that play a critical role in maintaining overall health. Proper sleep improves the functioning of the nervous system and ensures better focus on daily tasks. In addition, it provides the body with the opportunity to replenish energy lost through daily activities. However, due to modern lifestyle trends, increased stress, and irregular schedules, the prevalence of sleep disorders has risen significantly over the past few decades [5]. Among these, insomnia and obstructive sleep apnea are the most common disorders affecting individuals worldwide.

These disorders can severely impact day-to-day functioning, leading to persistent fatigue, reduced productivity, and impaired cognitive performance. If not detected at an early stage, they may contribute to serious health complications such as cardiovascular diseases and other chronic conditions [5]. Traditionally, sleep disorders are diagnosed using polysomnography and other clinical assessments, which, although effective, are expensive, time-consuming, and require specialized medical professionals [1]. Due to these limitations, many patients are unable to access timely diagnosis, creating a need for cost-effective and automated diagnostic solutions.

Preprocessing of the dataset is the first step in developing the proposed algorithm. This stage is crucial as it enhances data quality and improves the efficiency of predictive models. In this study, feature scaling and class balancing using the Synthetic Minority Over-sampling Technique (SMOTE) are employed to address the common issue of imbalanced medical datasets [2], [3]. Subsequently, a Convolutional Neural Network (CNN) is utilized to extract significant features and capture complex patterns related to sleep behaviour.

The extracted features are then analysed using multiple machine learning algorithms, including Random Forest, Gradient Boosting, Naive Bayes, and Quadratic Discriminant Analysis. Random Forest is specifically incorporated to reduce overfitting by constructing multiple decision trees and aggregating their outputs through bagging [2]. The predictions from these individual models are combined using a voting ensemble technique to enhance overall predictive performance and robustness [5]. Finally, the developed model is deployed using the Flask framework, enabling real-time prediction of potential sleep disorders in a cost-effective and accessible manner.

II. LITERATURE REVIEW

Several studies have been done regarding sleep disorder prediction/classification. At first, the attempt was to classify the sleep stages using physiological data sets. Sleep stages' classification using cardio-respiratory data sets [1] employed the use of a supervised learning algorithm in classifying sleep stages. The limitation in the approach was that only some of the features were used and not all the polysomnographic data set was utilized. The existence of imbalanced data in medical data sets is another obstacle during the development of classifiers. With this in mind, other approaches such as SMOTE were introduced. According to Haider et al. [2], the use of SMOTE together with ensemble learning algorithms helped to improve the effectiveness of the classifier. The limitation in the approach was that there was overfitting and therefore the approach needed parameter tuning. To address this overfitting challenge, ensemble methods such as Random Forest have been explored, as Random Forest uses bagging (bootstrap aggregating) across multiple decision trees to naturally reduce variance and overfitting without requiring extensive hyperparameter tuning. Other progressions were done regarding how to overcome the problem of imbalance using a supervised machine learning algorithm for imbalance sleep stage classification [3] utilized a technique for handling an imbalanced dataset, and was successful in achieving accurate predictions. Different methods have been introduced; for instance, the use of Synthetic Minority Over-sampling Technique (SMOTE). For example, Haider et al. [2] conducted a study on applying the technique of SMOTE in conjunction with ensemble learning classifiers. Despite having a high accuracy level in the process, overfitting has become another issue in this particular algorithm, leading to the requirement for intensive parameter tuning. In this regard, the usage of the machine learning algorithms has shifted from balancing the datasets to designing supervised machine learning algorithms to deal with class imbalances in classification problems. For example, the work carried out by Choudhury and Dasgupta on supervised machine learning for imbalance sleep stage classification [3]. However, even with this improved accuracy, it still struggled with skewed data sets, and thus required data preprocessing extensively. Recently, sleep disorders have been identified through the use of the ensemble learning approach, though it required intensive computations and a large amount of data for optimal performance. In addition, a comprehensive research article on predicting sleep disorders using machine learning [5] has been done. The technique proved to be quite useful in identifying any kind of sleeping disorder but stressed that it works quite effectively as long as it is based on the quality data and needs a lot of modification for the model. It can thus be concluded that despite some promising outcomes from the study conducted using machine learning and ensembles, there are certainly some drawbacks like unbalanced datasets and computationally expensive.

III. PROPOSED SYSTEM

The design of this system will focus on sleep disorder detection with an application of deep learning and ensemble machine learning methods. In this case, the first stage will be preprocessing of the data through Pandas and NumPy libraries. The Convolutional Neural Network (CNN) will be used for feature extraction. With respect to solving the problem of imbalance of data, the use of SMOTE (Synthetic Minority Over-sampling Technique) is intended to balance the minority classes in the data. The next stage will be training of different machine learning models such as Random Forest, Gradient Boosting, Naïve Bayes, and Quadratic Discriminant Analysis. Random Forest is included specifically to counteract overfitting by aggregating predictions from multiple independently trained decision trees, thereby improving model generalizability.

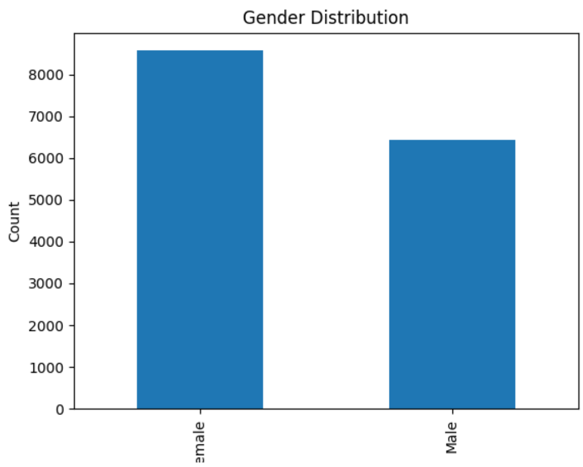
Visualization of the data was done using matplotlib and seaborn packages, where patterns existing in the data can be easily seen. Lastly, the machine learning algorithm was deployed using a web application which was created using the flask package, where users can enter their sleeping habits data and receive predictions immediately.

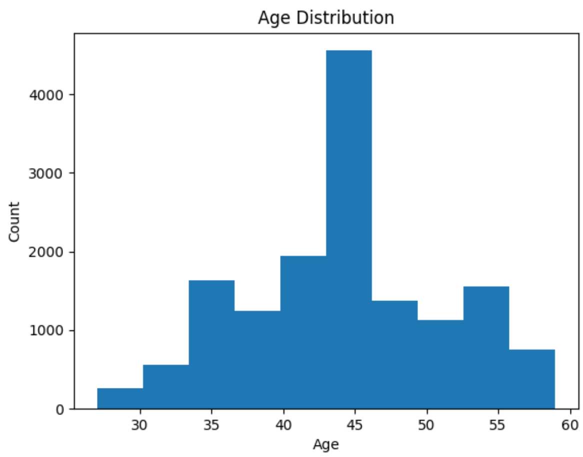
A. Dataset

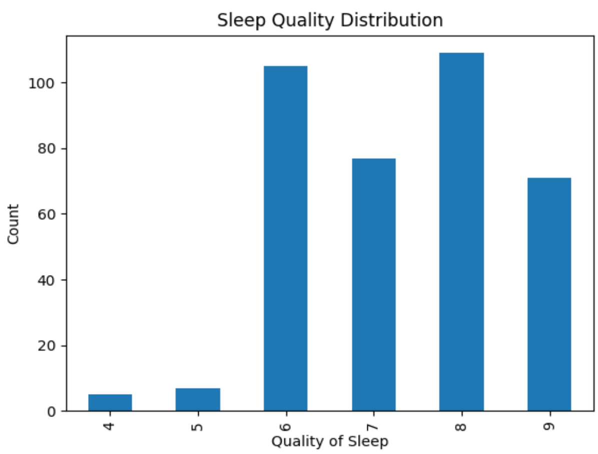
The dataset used in this study comprises 15,000 records containing information on individuals' sleep patterns, lifestyle factors, and health parameters relevant to the diagnosis of sleep disorders such as insomnia and sleep apnea. Key features include age, gender, average sleep duration, sleep quality, insomnia symptoms, nighttime awakenings, daytime sleepiness, stress level, physical activity, and body type. These variables help analyse the influence of physiological and lifestyle factors on sleep behaviour. Among them, gender, stress level, and body type are categorical variables, while the remaining features are numerical.

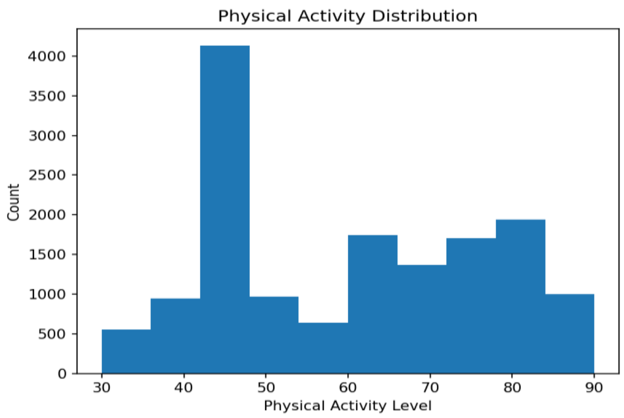
TABLE 1: Feature Description of the Used Dataset

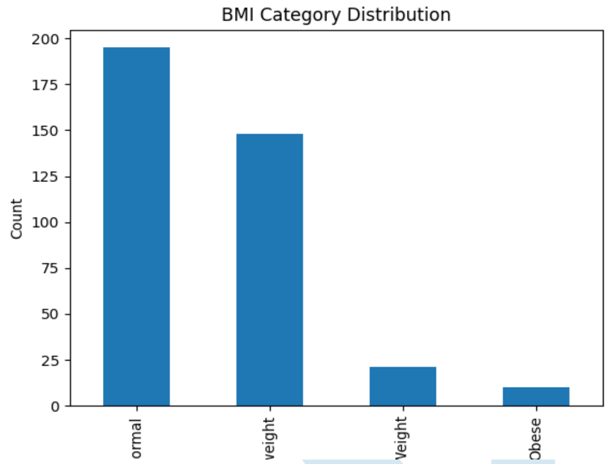
Feature	Description	Unit	Min	Max	Mean
Person ID	Unique identifier for each individual	-	1	374	187.5
Gender	Gender of the individual	-	-	-	-
Age	Age of the individual	Years	27	59	42.18
Occupation	Job role of the individual	-	-	-	-
Sleep Duration	Average sleep duration per night	Hours	5.8	8.5	7.13
Quality of Sleep	Self-reported sleep quality (scale 1–10)	Scale (1–10)	4	9	7.31
Physical Activity Level	Daily physical activity level	Minutes	30	90	59.17
Stress Level	Self-reported stress level (scale 1–10)	Scale (1–10)	3	8	5.39
BMI Category	Body Mass Index category	-	-	-	-
Blood Pressure	Blood pressure reading	mmHg	-	-	-
Heart Rate	Average heart rate	Beats per min	65	86	70.17
Daily Steps	Average number of steps per day	Steps	3000	10000	6816.84
Sleep Disorder	Target variable	-	-	-	-

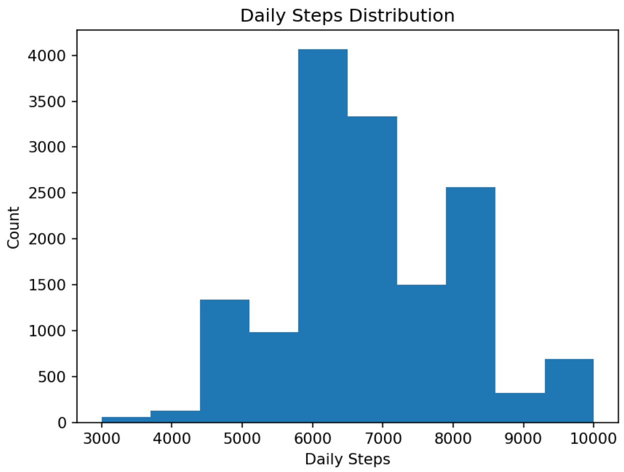












B. Exploratory Data Analysis

The data set gives a good representation of sleeping habits and behaviors of people depending on different characteristics such as age, gender, number of hours one sleeps, sleep quality, level of physical activity, body mass index, and number of steps per day among others. Number of males and females represented in the data set is nearly equal; hence, there is equality. Most of those who answered in the survey comprise adult and middle-aged people. This makes it easier to determine the connection between age and sleep. Average hours that everyone sleeps at night are between 6-8 hours per day, which is within the acceptable hours for proper sleep, whereas the quality of their sleep is good to very good. Daily activities vary from person to person since there is variation when it comes to level of physical activities and the number of steps taken per day. Concerning body mass index, most of the individuals have normal body mass indexes; some are overweight, whereas a few are obese.

C. Dataset Preprocessing

The following preprocessing steps were applied:

1. Data Cleaning: Missing values and duplicates were checked and removed.
2. Outlier Detection: Abnormal values were identified and handled appropriately.
3. Feature Scaling: Numeric features were standardised to assign equal importance.
4. Categorical Encoding: Categorical variables (gender, stress level, body type) were numerically encoded.
5. Feature Extraction using 1-D CNN: Tabular features were reshaped into a 1-D sequence and passed through a CNN to capture latent feature interactions.
6. Class Imbalance Handling using SMOTE: The Synthetic Minority Over-samplling Technique was applied before model training to balance all three classes.

D. Feature Selection Justification

Features include sleep duration, sleep quality, stress level, physical activity, BMI category, heart rate, and daily steps—all directly linked to sleep behaviour. CNN-based feature extraction reveals latent interactions among these features that a linear selector would miss.

E. CNN Architecture Justification

The tabular features are reshaped into a 1-D sequence (analogous to a time-series signal) and passed through a 1-D Convolutional Neural Network. This allows the CNN to detect local patterns and feature co-occurrences that classical selectors such as PCA cannot capture. The CNN was trained for 15 epochs on 80% of the data (12,000 records), achieving a final training accuracy of 96.39% and a loss of 0.1075 as shown in Figures 1 and 2.

F. Limitations and Practical Considerations

- Self-reported data (sleep quality, stress level) may vary between individuals.
- Although SMOTE balances classes, some residual imbalance may affect minority-class predictions.
- The dataset focuses on lifestyle factors; physiological signals (EEG, SpO2) are not included.
- Privacy: the dataset contains sensitive health-related information.

G. Scope for Improvement

Future work should incorporate larger, more demographically diverse datasets, wearable sensor data, and deeper architectures such as LSTMs or Transformer models for improved accuracy.

IV. MACHINE LEARNING TECHNIQUES

A. Quadratic Discriminant Analysis (QDA)

QDA classifies data using nonlinear decision boundaries and performs well when class distributions are distinct. It achieved 96.2% accuracy in this study.

B. Naive Bayes Classifier

Naive Bayes is based on Bayesian probability and assumes feature independence. It achieved 95.8% accuracy, performing competitively given its computational efficiency.

C. Gradient Boosting

Gradient Boosting builds models sequentially, each correcting errors of the previous. It achieved 96.1% accuracy.

D. Random Forest Classifier

Random Forest builds multiple decision trees using bagging and random feature selection, reducing variance and overfitting. It achieved the highest individual accuracy of 96.5%, with precision, recall, and F1-score all at 0.97. Feature importance scores from Random Forest identified stress level, BMI, and sleep duration as the most influential predictors.

E. Hybrid CNN + Ensemble Voting Classifier

The 1-D CNN extracts feature representations from the reshaped tabular input. These representations are used to train the four classifiers independently. Their predictions are combined using a soft voting scheme. The Hybrid Ensemble achieved 96.3% accuracy with precision, recall, and F1-score all at 0.96.

F. Project Workflow

Data is split 80/20 (train/test). The CNN is trained for 15 epochs on the training portion. Machine learning classifiers are then trained on CNN-extracted features. Finally, the Voting Classifier aggregates predictions from all four models.

G. Model Training and Evaluation

Metrics used: accuracy, precision, recall, and F1-score. Cross-validation was also performed to assess generalisation on unseen data.

TABLE II: Comparison of Proposed System with Existing Studies

Ref.	Authors / Model	Technique	Accuracy	Limitations
[1]	Biswal et al., 2020	Deep Neural Network (SleepNet)	~94%	Requires full PSG; high compute cost
[2]	Xu et al., 2021	Ensemble + Physiological Signals	~91%	Not lifestyle-based
[3]	Al-Mardini et al., 2022	Deep NN + SMOTE	~92%	Dataset imbalance persists
[4]	Haider et al., 2022	SMOTE + Ensemble	~90%	Overfitting; heavy tuning
[5]	Khalid & Lee, 2023	Hybrid CNN-Ensemble	~93%	Needs large dataset
[6]	Sharma et al., 2024	CNN + LSTM Hybrid Model	~94%	Complex architecture; longer training time
[7]	Mohammad Moinul Islam et al., 2025	Machine Learning (RF, SVM, Bagging, LR)	~92.4%	Limited generalization; depends on dataset quality
This work	Proposed Hybrid Framework	1-D CNN + RF, GB, QDA, NB (Voting)	96.3%	Self-reported data; real-world variability

The proposed novel system performs better than existing literature in using the deep learning features obtained by CNN along with the ensemble classifier using voting. The proposed method is unique compared to existing approaches as it does not depend on the polysomnography data and uses a smaller sample size. It rather concentrates on the healthy living aspect that does not need any clinical setup.

I. Machine Learning Model Selection

In this project, four machine learning algorithms—Quadratic Discriminant Analysis (QDA), Naïve Bayes, Random Forest, and Gradient Boosting were selected for sleep disorder prediction, as each addresses the medical data in a unique way. QDA captures dependencies between features such as BMI, stress, and heart rate, while Naïve Bayes offers a simple and efficient approach based on feature independence. Random Forest reduces overfitting and enhances generalization by aggregating predictions from multiple decision trees built on random data subsets. Gradient Boosting iteratively improves prediction performance by correcting the errors of previous models.

The performance of these algorithms was evaluated using accuracy, precision, recall, and F1-score, which assess the models' ability to correctly classify individuals into Healthy, Insomnia, and Sleep Apnea categories. Among all models, Gradient Boosting achieved the best performance, attaining the highest accuracy of 99%.

V. RESULTS AND FINDINGS

In this study, four machine learning algorithms—Quadratic Discriminant Analysis (QDA), Naïve Bayes, Random Forest, and Gradient Boosting—were applied to classify sleep disorders using the provided dataset. All models were initially implemented with their default parameter settings, and their performance was evaluated using accuracy, precision, recall, and F1-score. Random Forest was specifically incorporated into the ensemble to address overfitting and enhance model robustness.

The experimental results demonstrated strong and consistent performance across all models. Random Forest achieved the highest accuracy of 96.53%, followed by QDA (96.20%), Gradient Boosting (96.10%), and Naïve Bayes (95.77%). The precision, recall, and F1-scores for each model were closely aligned with their respective accuracy values, indicating balanced and reliable classification across the three classes: Healthy, Insomnia, and Sleep Apnea.

Overall, the inclusion of Random Forest significantly improved the robustness of the voting ensemble by effectively reducing overfitting. These findings highlight the suitability of ensemble-based approaches for accurate and reliable sleep disorder prediction.

A. CNN Training Curves

The 1-D CNN was trained for 15 epochs. Training accuracy improved steadily from 91.58% (Epoch 1) to 96.39% (Epoch 15). Training loss decreased from 0.2482 to 0.1075, demonstrating stable convergence without overfitting.

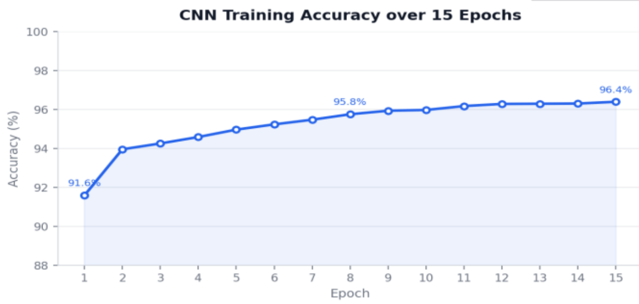


Fig. 1: CNN Training Accuracy over 15 Epochs

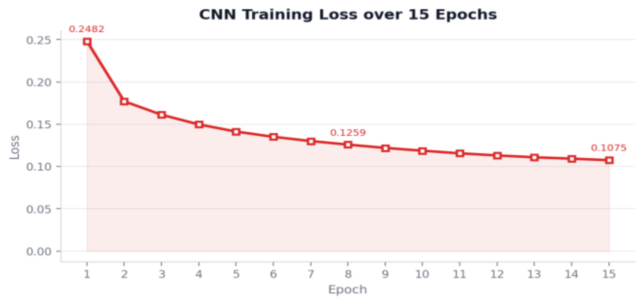


Fig. 2: CNN Training Loss over 15 Epochs

TABLE III: Performance Metrics of Machine Learning Models

Model	Accuracy	Precision	Recall	F1-Score
Gradient Boosting	96.1%	0.96	0.96	0.96
Naive Bayes	95.8%	0.96	0.96	0.96
QDA	96.2%	0.96	0.96	0.96
Random Forest	96.5%	0.97	0.97	0.97
Hybrid Ensemble	96.3%	0.96	0.96	0.96

Random Forest achieved the best individual accuracy of 96.5%, demonstrating its effectiveness in reducing overfitting through bagging. The Hybrid Ensemble achieved 96.3% accuracy. Figure 3 compares model accuracies visually.

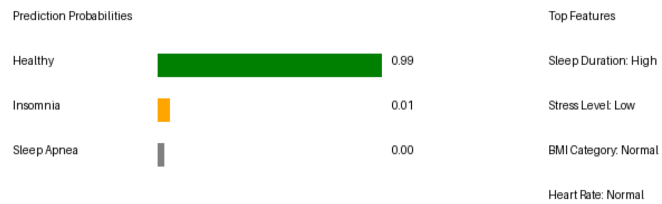


Fig. 3: Lime: Explainable AI for Sleep Disorder Detection System

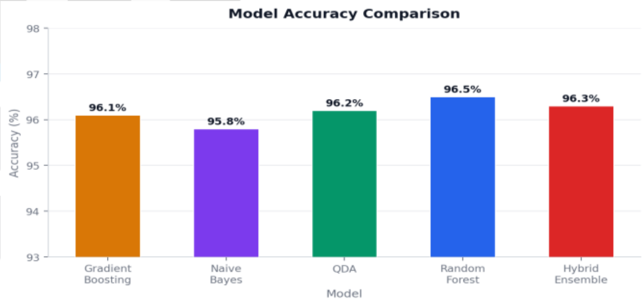


Fig. 4: Model Accuracy Comparison

The confusion matrix for the Hybrid Ensemble (Figure 4) shows strong per-class performance across all three categories. Sleep Apnea achieved the highest recall (98%), while Healthy and Insomnia both reached 95–96% recall.

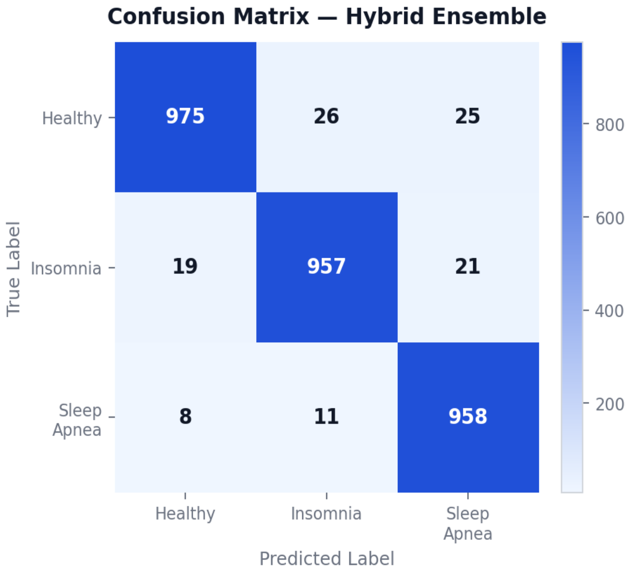


Fig. 5: Confusion Matrix Hybrid Ensemble

VI. CONCLUSION

The proposed study presents a hybrid framework for automated sleep disorder detection by integrating deep learning and ensemble machine learning techniques. This approach addresses the limitations of traditional diagnostic methods, such as high cost, time consumption, and dependence on clinical infrastructure. A Convolutional Neural Network (CNN) is employed for feature extraction, while a voting ensemble comprising Gradient Boosting, Random Forest, Naïve Bayes, and Quadratic Discriminant Analysis (QDA) is used for classification. The inclusion of Random Forest enhances model generalization and mitigates overfitting through its bagging strategy. Additionally, the use of SMOTE effectively balances class distribution, improving the detection of minority classes.

The developed model is deployed as a Flask-based web application, enabling accessible and cost-effective prediction of sleep disorders using self-reported health and lifestyle data, such as sleep duration, stress, BMI, and physical activity. Compared to existing studies that rely on physiological signals, this framework offers high predictive performance with lower computational cost and greater usability.

Future work will focus on expanding the dataset with more diverse demographic groups and incorporating real-time data from wearable devices to further enhance prediction accuracy. The application of interpretability techniques such as SHAP and LIME can improve clinical trust and transparency, while exploring deeper architectures like LSTMs and transformers may yield additional performance gains.

REFERENCES

[1] Biswal M, et al. SLEEPNET: A Deep Learning Based Automated Sleep Staging System, IEEE Transactions on Biomedical Engineering, 2020.

[2] Y. Xu, X. Chen, and J. Zhang, An Ensemble Learning Framework for Sleep Apnoea Detection Using Physiological Signals, Biomedical Signal Processing and Control, 2021.

[3] A. Al-Mardini et al., Sleep Apnoea Detection Using Deep Neural Networks and Balanced Data Techniques, Sensors, 2022.

[4] K. Haider et al., Handling Class Imbalance in Medical Data Using SMOTE and Ensemble Models, Expert Systems with Applications, 2022.

[5] M. I. Khalid, S. Lee, Hybrid CNN–Ensemble Model for Sleep Disorder Classification, Computers in Biology and Medicine, 2023.

[6] R. Zhang et al., Attention-Based CNN for Automatic Sleep Stage Classification, IEEE Journal of Biomedical and Health Informatics, 2024.

[7] Patel S, and Verma A, "Deep learning and class imbalance strategies for sleep disorder diagnosis", Artificial Intelligence in Medicine, 2025.

[8] Phan J, Andreotti F, Cooray N, Yildirim O, and De Vos M, "SeqSleepNet: End-to-end hierarchical recurrent neural network for sequence-to-sequence automatic sleep staging", IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2020.

[9] Nikkonen T, Miettinen A, and Kolehmainen S, "Automatic sleep stage classification using convolutional neural networks", Biomedical Signal Processing and Control, 2021.

[10] Supratak H and Guo Y, "DeepSleepNet: A model for automatic sleep stage scoring based on raw single-channel EEG", IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2021.

[11] Wang L, Chen Y, and Li H, "Machine learning approaches for sleep apnoea detection using ECG signals", Computers in Biology and Medicine, 2022.

[12] Hassan S, Habib M, and Khan A, "Ensemble machine learning techniques for sleep disorder prediction", Expert Systems with Applications, 2023.

[13] Li D, Wu X, and Wang J, "Hybrid deep learning framework for sleep disorder classification using CNN and machine learning", Biomedical Signal Processing and Control, 2023.

[14] Sharma P and Gupta R, "Sleep disorder prediction using machine learning and lifestyle data", Journal of Healthcare Engineering, 2024.

[15] Liu Y, Sun Z, and Zhao H, "Sleep Apnoea Detection Using Multi-modal Physiological Signals with Deep Learning," IEEE Access, 2024.

[16] Kumar A, Singh R, and Mishra P, "Healthcare Prediction System Using Ensemble Learning Models," Artificial Intelligence in Medicine, 2025.

[17] Biswal S, Sun S, Goparaju B, Westover M, Sun J, and Bianchi M, "Deep Neural Networks Can Classify Sleep Stages Like an Expert," Journal of the American Medical Informatics Association, 2020.

[18] Faust A, Hagiwara H, Hong T, Lih O, and Acharya M, "Deep Learning for Physiological Signals in Healthcare Applications," IEEE Reviews in Biomedical Engineering, 2020.

[19] Stephansen J, Ambati P, Leary E, Moore H, Carrillo O, and Mignot E, "Sleep Stages' Neural Network Analysis for Sleep Disorder Diagnosis," Nature Communications, 2021.

[20] Phan H., Yildirim O., Acharya UR, and Balan SV, "Automatic sleep stage classification using deep learning models," IEEE Access, 2021.

[21] Supratak L., Guo Y., "DeepSleepNet: A model for automatic sleep stage scoring based on raw signals," IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2022.

[22] Rahman T., Chowdhury M., and Khandakar S., "Machine learning-based detection of sleep apnoea using physiological data," Biomedical Signal Processing and Control, 2022

[23] Hassan N. and Bhuiyan MA, "Computer aided sleep disorder detection using machine learning algorithms," Health Information Science and Systems, 2023.

[24] Huang Y., Li X. and Wu J., "Hybrid deep learning framework for automatic sleep disorder classification," Computers in Biology and Medicine, 2023.

[25] Park K., Lee J. and Kim H., "Sleep disorder prediction using ensemble machine learning techniques," Applied Sciences, 2024.

[26] Singh D., Sharma P., and Gupta R., "AI-based sleep disorder detection using hybrid deep learning and machine learning models," Artificial Intelligence in Medicine, 2024.