Enhancing Sleep Disorder Diagnosis with a Machine Learning Approach Using Ensemble Neural Networks

Sleep Disorder Diagnosis with Ensemble Neural Networks

Md Samiul Alom
International University of Business Agriculture and Technology
Dhaka, Bangladesh
samiulalom090@gmail.com

Samiha Maisha Jeba
International University of Business Agriculture and Technology
Dhaka, Bangladesh
jebam615@gmail.com

Angkon Debnath
International University of Business Agriculture and Technology
Dhaka, Bangladesh
angkonnath1@gmail.com

Tanjim Taharat Aurpa
Bangabandhu Sheikh Mujibur Rahman Digital University
Gazipur, Bangladesh
aurpa0001@bdu.ac.bd

Rifat Siddiqua
Govt. College of Applied Human Science
Dhaka, Bangladesh
siddiquarifat49@gmail.com

Abstract

Sleep disorders, including conditions such as obstructive sleep apnea, significantly impact the quality of life and pose serious health risks. Early detection and diagnosis are crucial for effective management. This study explores various machine learning techniques to predict sleep disorders, utilizing a comprehensive dataset to enhance diagnostic accuracy. We implement multiple Artificial Neural Network architectures, including Bagging, Boosting, and Weighted Average ensembles, to identify the most effective model. Our results indicate that both the ANN Bagging Ensemble and ANN Boosting Ensemble achieved the highest accuracy of 94.7%, followed closely by the ANN Weighted Average Ensemble with an accuracy of 92.9%. This research underscores the potential of applying sophisticated ensemble methods in clinical settings for the timely diagnosis of sleep disorders, paving the way for improved patient outcomes.

Keywords

Sleep Disorders, Artificial Neural Networks , Obstructive Sleep Apnea, Health Monitoring Systems, Sleep Quality Assessment

ACM Reference Format:
Md Samiul Alom, Samiha Maisha Jeba, Angkon Debnath, Tanjim Taharat Aurpa, and Rifat Siddiqua. 2024. Enhancing Sleep Disorder Diagnosis with a Machine Learning Approach Using Ensemble Neural Networks: Sleep Disorder Diagnosis with Ensemble Neural Networks. In *11th International Conference on Networking, Systems, and Security (NSysS '24), December 19–21, 2024, Khulna, Bangladesh.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3704522.3704533

NSysS '24, December 19–21, 2024, Khulna, Bangladesh

ACM ISBN 979-8-4007-1158-9/24/12
https://doi.org/10.1145/3704522.3704533

1 Introduction

Sleep is a distinct state of awareness in which there is reduced engagement with the surroundings, and the body's sensory functions and voluntary muscle movements are temporarily halted. The intricate mechanisms occurring in the brain and body during sleep are still being studied extensively. Sleep is essential and restorative, as it energizes us and enables us to manage the challenges of daily living [11].

The 2005 update of the International Classification of Sleep Disorders (ICSD-2) improved upon previous versions by incorporating both symptom-based and physiological approaches to categorize sleep disorders. It organizes 81 categories of sleep disorders, enabling more accurate diagnoses and better communication among healthcare professionals. The classification differentiates between various types of insomnia, breathing disorders during sleep, excessive sleepiness, circadian rhythm disturbances, abnormal behaviors during sleep, and movement disorders.[17].

People typically spend around one-third of their lives sleeping or resting. Although this is a common expectation, the reality is that 1 out of every 3 adults does not get sufficient sleep, with many individuals averaging only 6.8 hours per night. Sleep is essential for people of all ages, as it allows the body's systems to recalibrate and repair themselves. During sleep, the brain, nervous system, immune system, heart, and digestive system undergo important restorative processes. Without sufficient sleep, the body's cells can age faster and become more susceptible to infection and chronic diseases, leading to problems like brain fog, mood issues, and other health concerns. [16],[2]. Sleep deprivation can have serious consequences on an individual's physical and mental health, including an increased risk of obesity, diabetes, cardiovascular disease, and mental health issues.

We developed and evaluated predictive models for sleep disorders using Artificial Neural Networks (ANN). The goal is to explore the potential of ANN for forecasting the presence or absence of sleep disorders by using extensive datasets that include lifestyle choices, sleep habits, BMI, and various health indicators. Combining lifestyle and sleep health data offers a comprehensive approach to understanding the complex interactions between different factors that impact sleep quality and patterns. By using ANN algorithms, the study can reveal hidden trends and connections within these datasets, thereby improving the accuracy of predicting and diagnosing sleep disorders.

In this research work, we aim to explore the use of Artificial Neural Networks (ANN) for predicting sleep disorders. The main objectives are given below:

- Implementing ANN techniques to forecast the presence or absence of sleep disorders.
- Utilizing datasets that include lifestyle choices, sleep habits, BMI, and various health indicators.
- Analyzing the interactions between various factors affecting sleep quality.
- Evaluating the performance of ANN models for accurate prediction and diagnosis of sleep disorders.
- Identifying hidden trends within the data to improve sleep disorder prediction.

2 Related Work

Recent advancements in artificial intelligence (AI) have significantly enhanced the prediction and classification of sleep disorders. Several studies have employed various machine learning and deep learning techniques to address this complex health issue. For instance, one research utilized a Long Short-Term Memory (LSTM) network with a Stacked LSTM approach to classify sleep stages and predict disorders like sleep apnea and insomnia, achieving notable accuracy [12]. Another study employed a boosting-based ensemble learning algorithm to predict mild sleep-disordered breathing, reporting an accuracy around 70%, which considers multiple factors such as body mass index, age, and snoring characteristics [9]. Additionally, research has demonstrated the effectiveness of Gradient Boosting classifiers in analyzing sleep health and lifestyle datasets, identifying key attributes contributing to sleep issues. Logistic Regression and Support Vector Machine (SVM) models were also evaluated, with a polynomial SVM model achieving 91.44% accuracy [11]. One significant study [18] proposed four distinct architectures: 1D-CNN, ConvLSTM, 1D-CNN-LSTM, and 2D-CNN-LSTM, designed to predict OSA events every 30 seconds using 90-second segments of respiratory signals. Among these models, the 1D-CNN-LSTM emerged as the most effective, achieving an accuracy exceeding 83%, alongside a sensitivity of 81% and a specificity of 85%. Other studies [15] have focused on specific populations, such as a study that utilized machine learning and deep learning models to predict sleep disorders in asthma patients using the Taiwan NHIRM dataset.This study evaluated KNN, SVM, and Random Forest (RF) models, with RF achieving the highest accuracy of 0.813. This study evaluates five machine learning models KNN, SVM, DT, RF, ANN to identify which performs best in classifying sleep disorders. Among these models, the ANN emerged as the top performer, achieving a classification accuracy of 92.92% and F1-score of 91.93%, surpassing the other models tested [4].This research focused on predicting obstructive sleep apnea (OSA) using transfer learning with the AlexNet framework, trained on ECG data from the MIT-BIH polysomnographic database. The Stochastic Gradient Descent with Momentum (SGDM) algorithm achieved an accuracy of 86.63%, with sensitivity at 92.20% and precision at 90.55%. The study highlights deep learning's potential in enhancing non-invasive OSA diagnostics [14]. In another work, [1] has shown how the problematic use of mobile phones affects sleep using different feature selection methods and machine learning.

Artificial Neural Networks (ANN) have emerged as powerful tools in various predictive modeling applications, including healthcare. A study focused on diabetes prediction employed several machine learning models, including LGBM, XGBoost, and Random Forest, with a weighted average ensemble achieving an impressive accuracy of 96.10%. The integration of ensemble techniques has proven beneficial in enhancing predictive performance across different domains [8]. Another notable application of ANN involves traffic flow prediction, where Ensemble Empirical Mode Decomposition (EEMD) was used to preprocess data [10]. The paper [13] illustrates the application of machine learning in the petroleum industry, emphasizing that Artificial Neural Networks (ANN) are the most frequently utilized model. Furthermore, a hybrid model combining SVM and ANN for diabetes prediction achieved an accuracy of 94.87%, illustrating the effectiveness of combining different algorithms to improve model performance [3]. Overall, the literature highlights the versatility and efficacy of ANN in various applications, paving the way for future research to explore innovative approaches in predictive analytics across different fields.

3 Proposed Methodology

In this section, we will examine the key concepts that form the foundation of the proposed approach.

3.1 Data Preprocessing

In this paper, we have used the Sleep Health and Lifestyle Dataset[1] from Kaggle. Generally, The dataset may contain missing values and extreme data points that can disrupt the data categories. To address this, we remove the observations with null values and the outliers, or the extreme values. Additionally, we encode the variables to optimize the performance of the model. Finally, we normalize the data to ensure that the model is more effective in predicting the outcome. We first convert the dataset into numerical values. Then, we remove the occupation column as it has a limited correlation with the outcome column. The outcome column represents sleep disorder, which has three classes: none, insomnia, and sleep apnea. These are assigned the values 0, 1, and 2, respectively [7].

3.2 Machine Learning Models

Artificial Neural Network (ANN): Artificial Neural Networks (ANN) are computational systems designed to simulate the learning process of the human brain. Composed of interconnected neurons arranged in layers, ANNs process input data through a sequence of transformations to produce an output. Typically, an ANN is structured with the following types of layers, as illustrated in Figure 1:

[1]https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset

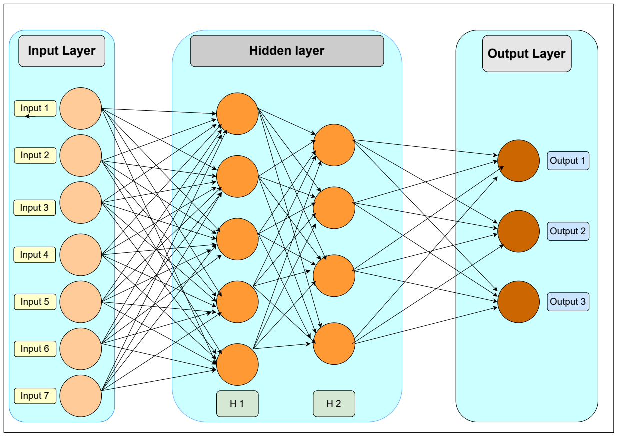

Figure 1: ANN Architecture

- **Input Layer:** The layer where input data is fed into the network, with each node corresponding to a feature in the dataset.
- **Hidden Layer(s):** Intermediate layers that transform the input into higher-level features. Each hidden layer neuron applies a weighted sum of its inputs, followed by an activation function.
- **Output Layer:** The final layer that outputs predictions based on the transformed input from the hidden layers.

In each neuron, the weighted sum of inputs is calculated, as shown in Equation 1:

$$z = W^T \cdot X + b \tag{1}$$

Where:

- W is the weight vector,
- X is the input feature vector,
- b is the bias term,
- z is the weighted sum of the inputs.

The neuron's output is computed by applying an activation function $g(z)$ to this weighted sum, as shown in Equation 2:

$$y = g(z) \tag{2}$$

Some commonly used activation functions include:

- **ReLU (Rectified Linear Unit):** Popular for deep neural networks, ReLU is defined as:

$$g(z) = \max(0, z)$$

It effectively introduces non-linearity without saturating gradients, which can speed up training.

- **Softmax Function:** Used for multi-class classification, the softmax function normalizes the output of the neurons in the output layer into a probability distribution. It is given by Equation 3:

$$g(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{N} e^{z_j}} \tag{3}$$

Where:
– z_i is the weighted input for class i,
– N is the number of classes,
– $g(z_i)$ is the probability that the input belongs to class i.
The softmax ensures that the output values are positive and sum to 1, making it ideal for multi-class problems.

The process of training an ANN involves minimizing the difference between the predicted output and the actual output by adjusting the weights through the backpropagation algorithm and gradient descent optimization. The weight update rule is described by Equation 4:

$$W_{\text{new}} = W_{\text{old}} - \eta \cdot \frac{\partial J(W)}{\partial W} \tag{4}$$

Where:

- W_{new} is the updated weight,
- W_{old} is the previous weight,
- η is the learning rate that controls the magnitude of the weight updates,
- $\frac{\partial J(W)}{\partial W}$ is the gradient of the loss function $J(W)$ with respect to W.

ANNs are highly flexible and powerful tools for capturing complex patterns in data 1, making them ideal for applications such as image classification, speech recognition, and natural language processing. However, their performance often requires large datasets and considerable computational resources.

Artificial Neural Network (ANN) Bagging Ensemble: Bagging, or Bootstrap Aggregating, is an ensemble learning technique that aims to reduce variance and mitigate overfitting in high-variance models like ANNs. It involves training multiple ANN models on different random subsets of the training data through a process called bootstrap resampling. Each model learns unique patterns, contributing to a more robust ensemble. After training, the predictions of all models are combined using majority voting for classification tasks or averaged for regression tasks. This aggregation improves generalization, enabling the ensemble to perform better on unseen data.

Artificial Neural Network (ANN) Boosting Ensemble: Boosting is an iterative ensemble technique designed to reduce bias and enhance accuracy. In ANN Boosting, multiple neural networks are trained sequentially, with each model focusing on the misclassified examples from the previous models. This is accomplished by adjusting the weights of the training samples, giving more emphasis to difficult cases. The final predictions are made using weighted voting, where the contribution of each model is based on its accuracy. This method progressively corrects errors, leading to better performance, especially in complex classification scenarios.

Weighted Average Ensemble with Artificial Neural Networks (ANN): The Weighted Average Ensemble is a method that combines predictions from multiple ANN models trained on the same task, assigning different weights based on each model's performance. After independent training on resampled datasets, the predictions are evaluated, and each model is assigned a weight reflecting its accuracy. The final prediction is computed as a weighted average of the individual model predictions, ensuring that more accurate models exert a greater influence on the outcome. This approach enhances overall accuracy and generalization by leveraging the strengths of various models, making it particularly effective in ensemble learning.

3.2.1 Proposed Framework: Here, we have discussed our methodology, which is illustrated in Figure 2. It shows the workflow of our paper, and we will explain it below:

- The workflow begins with input features (Feature 1 through Feature N), which undergo data preprocessing, including normalization and handling missing values.
- The preprocessed data is divided into ANN and ensemble ANN model (e.g., ANN 1 to ANN 50), each consisting of an input layer, hidden layers for learning complex patterns, and an output layer for making predictions.
- Predictions from the individual ANN models are combined using ensemble techniques, such as bagging, which trains multiple models on random subsets of the data to average their predictions, and a weighted average method that assigns varying weights based on model performance.
- The final output presents the predicted outcomes derived from the combined predictions of the individual ANN models.

3.3 Model Evaluation

To evaluate the effectiveness and reliability of the proposed classifier, we utilize confusion matrices after the training process[5]. Using the confusion matrix that our classifier produced, we created a heatmap to display this data, as seen in Figure 3 4 5 6. This detailed breakdown enables the calculation of key performance metrics that assess the classifier's overall effectiveness. The key elements of the confusion matrix are:

- **True Positive (TP)**: The number of instances correctly classified as belonging to the positive class.
- **True Negative (TN)**: The number of instances correctly classified as belonging to the negative class.
- **False Positive (FP)**: The number of instances wrongly classified as positive when they actually belong to the negative class. This error is also referred to as a Type I error.
- **False Negative (FN)**: The number of instances wrongly classified as negative when they actually belong to the positive class. This error is also referred to as a Type II error.

From the confusion matrix, we derive key evaluation metrics such as Accuracy, Precision, Recall, and F1-Score, which are used to quantify the classifier's performance.

Accuracy: Accuracy is a commonly used metric in classification that calculates the proportion of correct predictions out of the total number of predictions. It gives an overall measure of how well the classifier distinguishes between both positive and negative instances. The accuracy is defined as the ratio of correctly classified instances (both true positives and true negatives) to the total number of predictions:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \tag{5}$$

Precision Precision is a metric that evaluates the classifier's capability to identify positive instances correctly. It measures the ratio of true positive predictions to the total number of positive predictions, including false positives. Precision is particularly relevant when false positives are costly, such as in domains like medical diagnoses or fraud detection. Precision is computed as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\% \tag{6}$$

A higher precision score suggests that the classifier commits fewer Type I errors (i.e., false positives).

Recall Recall also referred to as Sensitivity or True Positive Rate (TPR), measures the classifier's ability to correctly identify all actual positive instances. It is calculated as the ratio of true positives to the total number of actual positive cases (true positives and false negatives). Recall is especially important in situations where missing positive instances (Type II errors) can have severe consequences. Recall is given by:

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\% \tag{7}$$

A high recall score indicates that the model successfully identifies most positive cases, which is vital in fields like disease detection.

F1-Score The F1-Score is the harmonic mean of precision and recall, providing a balanced view of the classifier's performance when both false positives and false negatives are significant. This score combines both precision and recall, offering a more comprehensive evaluation. The F1-Score is particularly valuable in cases of imbalanced datasets where focusing on just precision or recall might give a misleading picture. The F1-Score is calculated as follows:

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{8}$$

A higher F1-Score suggests that the model strikes a good balance between precision and recall, reducing both false positives and false negatives. This metric is especially useful when both types of classification errors have significant implications.

The metrics derived from the confusion matrix—accuracy, precision, recall, and F1-Score—offer a comprehensive evaluation of the classifier's performance[6]. While accuracy provides an overall assessment, precision and recall focus on specific error types, and the F1-Score combines both to give a holistic performance measure. These metrics are crucial for assessing models in tasks where the cost of classification errors is high, especially in imbalanced datasets. **Table 1** present the accuracy, precision, recall, and F1-Score of the bagging ensemble and other models, respectively.

4 Results and Discussion

We employed four distinct modeling approaches to analyze the data: Artificial Neural Network (ANN), ANN Bagging Ensemble, ANN Boosting Ensemble, and ANN Weighted Average Ensemble. We achieve better performance than other researchers as illustrated in **Table 2**. Besides the fact that we use four different ensmble models in our research, that is another factor that makes this study superior to others.

The ANN model is a fundamental artificial intelligence algorithm that mimics the structure and function of the human brain, allowing it to learn and make predictions from complex datasets. The ANN

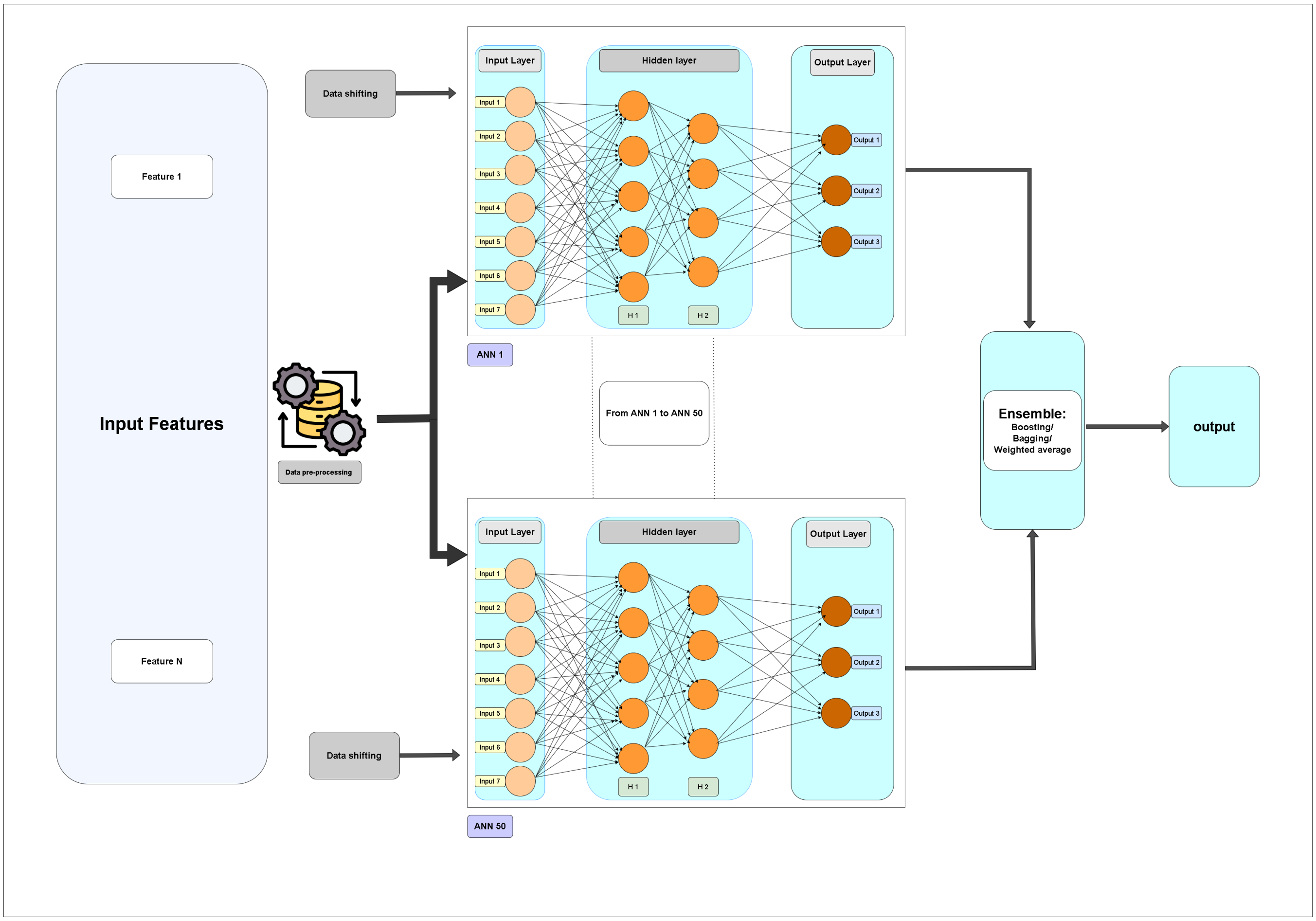

Figure 2: System Diagram

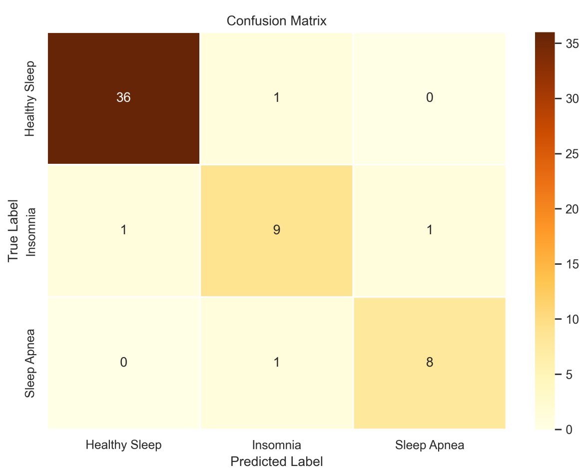

Figure 3: Confusion Matrix of ANN model

Figure 4: Confusion Matrix of Ensemble ANN Bagging

Model	Metric	Healthy Sleep	Insomnia	Sleep Apnea	Accuracy
ANN	Precision	0.97	0.89	0.80	0.93
	Recall	1.00	0.73	0.89	0.93
	F1-Score	0.99	0.80	0.84	0.93
	Macro Avg	0.89	0.87	0.88	0.93
	Weighted Avg	0.93	0.93	0.93	0.93
ANN Bagging	Precision	0.97	1.00	0.99	0.95
	Recall	0.90	0.82	0.86	0.95
	F1-Score	0.89	0.89	0.89	0.95
	Macro Avg	0.92	0.90	0.91	0.95
	Weighted Avg	0.95	0.95	0.95	0.95
ANN Boosting	Precision	0.97	0.97	0.97	0.93
	Recall	0.82	0.82	0.82	0.93
	F1-Score	0.89	0.89	0.89	0.93
	Macro Avg	0.89	0.89	0.89	0.93
	Weighted Avg	0.93	0.93	0.93	0.93
ANN Weighted Avg Ensemble	Precision	0.97	0.97	0.97	0.93
	Recall	0.82	0.82	0.82	0.93
	F1-Score	0.89	0.89	0.89	0.93
	Macro Avg	0.89	0.89	0.89	0.93
	Weighted Avg	0.93	0.93	0.93	0.93

Table 1: Comparison of Precision, Recall, F1-Score, Accuracy, Macro Average, and Weighted Average for different models: ANN, ANN Bagging, ANN Boosting, and ANN Weighted Average Ensemble.

Paper Title	Model Used	Best Model	Matrices	Accuracy/Precision
A Novel Approach to Prediction of Mild Obstructive Sleep Disordered Breathing in a Population-Based Sample[7]	Ensemble model on Sleep Heart Health data	The Boosting algorithm	Sensitivity, Specificity	Accuracy: 70%
Utilizing Multi-Class Classification Methods for Automated Sleep Disorder Prediction[9]	Support Vector Machine (SVM), Logistic Regression (LR)	2-degree polynomial SVM	Accuracy, Precision	Accuracy: 91.44%
Applying Machine Learning Algorithms for the Classification of Sleep Disorders[3]	KNN, Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), Artificial Neural Network (ANN), Genetic Algorithm	Artificial Neural Network (ANN)	Accuracy, Precision	Accuracy: 92.92%, Precision: 92.01%
Sleep Disorder Classification Using Convolutional Neural Networks[12]	AlexNet deep learning framework	AlexNet with (SGDM)	Accuracy, Sensitivity, Precision	Accuracy: 86.63%
Proposed Framework	**ANN, Ensemble Techniques: Bagging, Boosting, Weighted**	**ANN Bagging, ANN Boosting**	**Accuracy**	**Accuracy: 94.7%**

Table 2: Overview of models, best-performing algorithms, evaluation metrics, and their corresponding accuracies/precisions for various sleep disorder prediction studies

Bagging Ensemble combines multiple ANN models, each trained on a different subset of the data, to create a more robust and accurate prediction. The ANN Boosting Ensemble, on the other hand, sequentially trains weak learners (in this case, ANNs) to correct the errors of the previous model, resulting in a stronger overall prediction. Finally, the ANN Weighted Average Ensemble combines the outputs of multiple ANN models, with each model's contribution weighted based on its performance.

To evaluate the performance of these models, we determined the p-value of each feature in Figure 7 with Random forest Classifier, which indicates the statistical significance of the relationship between the feature and the target variable. Additionally, we compared the F1 scores, accuracy, across the different models in **Table**

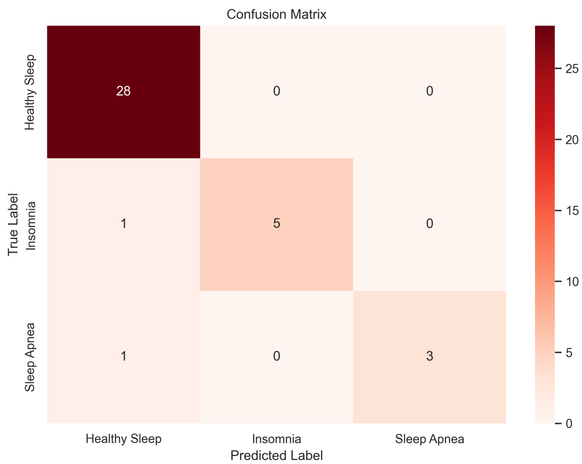

Figure 5: Confusion Matrix of Ensemble ANN Boosting

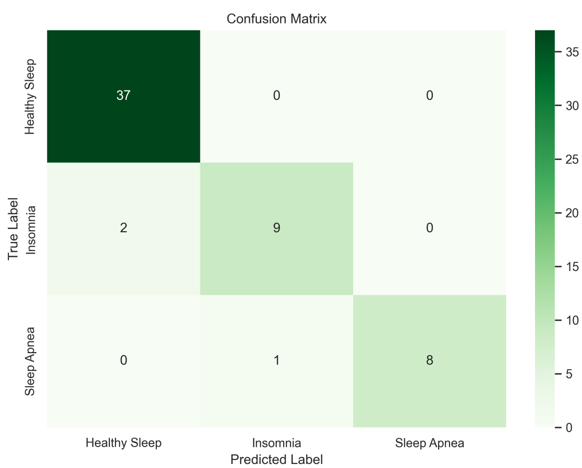

Figure 6: Confusion Matrix of Ensemble ANN Weighted Average.

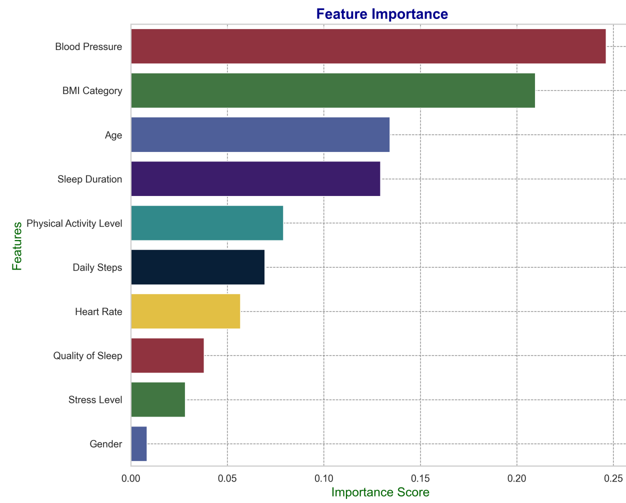

Figure 7: Feature Importance Scores for Model Variables

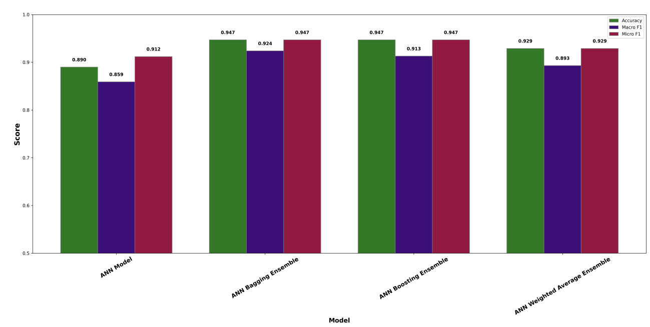

Figure 8: Evaluation metrics of different Models

3. The F1 score comprehensively evaluates the model's ability to classify data accurately.

In Figure 8, we explored the performance of various Artificial Neural Network (ANN) models and their ensemble configurations in classifying sleep disorders, specifically focusing on healthy sleep, insomnia, and sleep apnea. The results of our analysis, summarized in the table below, highlight the effectiveness of the models across several evaluation metrics, including accuracy, Macro F1, Micro F1, and F1 Score.

The standalone ANN model, as shown in Table 3, achieved an accuracy of 89.0%, with a Macro F1 score of 0.859. The model demonstrated robust classification capabilities, as evidenced by its Micro F1 score of **0.912** and an F1 Score of **0.911**. These metrics indicate a strong performance in distinguishing between the various sleep disorder categories.To enhance classification accuracy, we applied ensemble methods, including bagging and boosting techniques. Both the **ANN Bagging Ensemble** and the **ANN Boosting Ensemble** reached impressive accuracies of **94.7%**, showcasing significant improvements over the baseline ANN model. The Macro F1 scores for these ensembles were **0.924** and **0.913**, respectively, indicating their ability to manage class imbalances effectively while maintaining high precision. Additionally, both ensembles recorded F1 Scores of **0.946** and **0.945**, reflecting their balanced performance in classification.The **ANN Weighted Average Ensemble** also demonstrated commendable performance, achieving an accuracy of **92.9%**. Its Macro F1 score of **0.893** and Micro F1 score of **0.929** confirmed its reliability in classifying sleep disorders, although it slightly trailed behind the bagging and boosting ensembles.Interestingly, we noted a decrease in performance metrics following fine-tuning, with the accuracy dropping to **92.1%** and a Macro F1 score of **0.816**. This observation suggests that while fine-tuning can often improve model performance, it may not always yield better results across all metrics.

5 Conclusion

Sleep disorders, such as insomnia and sleep apnea, pose significant health risks, affecting both physical and mental well-being.

Model	Accuracy	Macro F1	Micro F1	F1 Score
ANN Model	0.890	0.859	0.912	0.911
ANN Bagging Ensemble	0.947	0.924	0.947	0.946
ANN Boosting Ensemble	0.947	0.913	0.947	0.945
ANN Weighted Average Ensemble	0.929	0.893	0.929	0.929
After Fine Tuning	0.921	0.816	0.868	0.872

Table 3: Comparison of Accuracy, Macro F1, Micro F1, and F1 Score among different ANN Models and Ensembles

They are often influenced by a variety of factors, including lifestyle choices, sleep habits, BMI, and other health indicators, making accurate diagnosis and treatment essential. Our study explored the potential of Artificial Neural Networks (ANN) to predict sleep disorders using these diverse datasets. The baseline ANN model achieved strong results, with an accuracy of 89.0%, but the application of ensemble methods like bagging and boosting significantly improved performance, raising the accuracy to 94.7%. These techniques demonstrated enhanced classification capabilities, particularly in managing class imbalances. Interestingly, fine-tuning led to a slight drop in performance, underscoring the importance of careful optimization. Overall, the combination of ANN and ensemble methods offers a promising approach to improving the prediction, diagnosis, and early intervention of sleep disorders. Future work should focus on expanding datasets with more health indicators and exploring advanced machine learning techniques, like deep learning, to enhance predictive accuracy for sleep disorders. Developing personalized models based on individual health profiles and integrating real-time data from wearable devices could enable continuous monitoring and timely interventions.

References

[1] Fairuz Tanzim Achal, Mosammat Suraiya Ahmmed, and Tanjim Taharat Aurpa. 2023. Severity Detection of Problematic Smartphone Usage (PSU) and its Effect on Human Lifestyle using Machine Learning. In *2023 IEEE 8th International Conference for Convergence in Technology (I2CT)*. IEEE, 1–6.

[2] Md Shoaib Ahmed, Tanjim Taharat Aurpa, and Md Musfique Anwar. 2020. Online topical clusters detection for top-k trending topics in twitter. In *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. IEEE, 573–577.

[3] Usama Ahmed, Ghassan F. Issa, Muhammad Adnan Khan, Shabib Aftab, Muhammad Farhan Khan, Raed A. T. Said, Taher M. Ghazal, and Munir Ahmad. 2022. Prediction of Diabetes Empowered With Fused Machine Learning. *IEEE Access* 10 (2022), 8529–8538. https://doi.org/10.1109/ACCESS.2022.3142097

[4] Talal Sarheed Alshammari. 2024. Applying Machine Learning Algorithms for the Classification of Sleep Disorders. *IEEE Access* 12 (2024), 36110–36121. https://doi.org/10.1109/ACCESS.2024.3374408

[5] Tanjim Taharat Aurpa and Md Shoaib Ahmed. 2024. An ensemble novel architecture for Bangla Mathematical Entity Recognition (MER) using transformer based learning. *Heliyon* 10, 3 (2024).

[6] Tanjim Taharat Aurpa, Md Shoaib Ahmed, Richita Khandakar Rifat, Md Musfique Anwar, and ABM Shawkat Ali. 2023. UDDIPOK: A reading comprehension based question answering dataset in Bangla language. *Data in Brief* 47 (2023), 108933.

[7] Tanjim Taharat Aurpa, Md Shoaib Ahmed, Rifat Sadik, Sabbir Anwar, Md Abdul Mazid Adnan, and Md Musfique Anwar. 2021. Progressive guidance categorization using transformer-based deep neural network architecture. In *International Conference on Hybrid Intelligent Systems*. Springer, 344–353.

[8] Tanjim Taharat Aurpa, Samiha Maisha Jeba, and Shoaib Ullah Rasel. 2023. Ensemble Methods of Machine Learning Algorithms for Early Diabetic Detection in Comparison. In *2023 1st International Conference on Circuits, Power and Intelligent Systems (CCPIS)*. 1–6. https://doi.org/10.1109/CCPIS59145.2023.10291566

[9] Brian Caffo, Marie Diener-West, Naresh M. Punjabi, and Jonathan Samet. 2010. A Novel Approach to Prediction of Mild Obstructive Sleep Disordered Breathing in a Population-Based Sample: The Sleep Heart Health Study. *Sleep* 33, 12 (12 2010), 1641–1648. https://doi.org/10.1093/sleep.33.12.1641 arXiv:https://academic.oup.com/sleep/article-pdf/33/12/1641/7303897/sleep-33-12-1641.pdf

[10] Xinqiang Chen, Jinquan Lu, Jiansen Zhao, Zhijian Qu, Yongsheng Yang, and Jiangfeng Xian. 2020. Traffic flow prediction at varied time scales via ensemble empirical mode decomposition and artificial neural network. *Sustainability* 12, 9 (2020), 3678.

[11] Elias Dritsas and Maria Trigka. 2024. Utilizing Multi-Class Classification Methods for Automated Sleep Disorder Prediction. *Information* 15, 8 (2024), 426.

[12] S. Gokulan, S. Narmadha, M. Pavithra, R. Rajmohan, and T. Ananthkumar. 2020. Determination of Various Deep Learning Parameter for Sleep Disorder. In *2020 International Conference on System, Computation, Automation and Networking (ICSCAN)*. 1–6. https://doi.org/10.1109/ICSCAN49426.2020.9262331

[13] Daniel Asante Otchere, Tarek Omar Arbi Ganat, Raoof Gholami, and Syahrir Ridha. 2021. Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models. *Journal of Petroleum Science and Engineering* 200 (2021), 108182.

[14] Chun-Cheng Peng and Chu-Yun Kou. 2023. Sleep Disorder Classification Using Convolutional Neural Networks. *IFIP International Journal on Artificial Intelligence Applications and Innovations* (2023), 539–548.

[15] Dinh-Van Phan, Nan-Ping Yang, Ching-Yen Kuo, and Chien-Lung Chan. 2021. Deep learning approaches for sleep disorder prediction in an asthma cohort. *Journal of Asthma* 58, 7 (2021), 903–911.

[16] Carrie Solomon. 2022. Health Benefits of Sleep: WHY IS Getting Enough Rest SO IMPORTANT. *Alternative Medicine* 66 (2022), 26–29.

[17] Michael J Thorpy. 2012. Classification of sleep disorders. *Neurotherapeutics* 9, 4 (2012), 687–701.

[18] Eileen Wang, Irena Koprinska, and Bryn Jeffries. 2023. Sleep Apnea Prediction Using Deep Learning. *IEEE Journal of Biomedical and Health Informatics* 27, 11 (2023), 5644–5654. https://doi.org/10.1109/JBHI.2023.3305980