Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606

Volume 09 Issue 01
2025

Research Article
https://doi.org/10.56979/901/2025

Modeling Sleep Health and Lifestyle Using Supervised Learning Algorithms

M Usman Bhatti[1*], Ali Saeed[1], Muhammad Ashir[2], Naveed Hussain[3], Mehmood Anwar[2], and Muhammad Farhat Ullah[4]

[1]Department of Software Engineering, FOIT, University of Central Punjab, Lahore, Pakistan.
[2]Department of Computer Science & IT, University of Lahore, Punjab, Pakistan.
[3]Department of Applied Computing Technologies, FOIT, University of Central Punjab, Lahore, Pakistan.
[4]School of Software, Dalian University of Technology, Dalian, Ganjingzi District, Liaoning Province, China .
Corresponding Author: M Usman Bhatti. Email: muhammad.usman1@ucp.edu.pk

Received: April 28, 2025 **Accepted**: May 31, 2025

Abstract: With the realization of the importance of sleep quality as an indicator of general well-being, this work uses the strength of machine learning to discover significant trends in information about lifestyles and health to make predictions regarding sleep health. Based on the Sleep Health and Lifestyle Dataset that consists of 373 instances (rows) and 13 features (columns), including demographic, physiological, and lifestyle-related data, one can classify the target variable Quality of Sleep as a categorical attribute, defining the task in terms of classification. Different machine learning models were implemented and compared by means of precision, recall, accuracy, and F1 score: Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Forest, Logistic Regression, Decision Tree, Gradient Boosting, Naive Bayes, and AdaBoost. Of these, the models of the ensemble type were better performers, and the Random Forest model produced the best outcomes in all measures: 98.67% accuracy, 98.74% precision, 98.67% recall, and 98.66% F1 score. The other schemes, Decision Tree and Gradient Boosting, also performed well, and SVM received the lowest scores. The results emphasize the usefulness of ensemble methods in modeling complicated and non-linear relationships in multifactor health data. The results reinforce the possibilities of machine learning in relation to data-driven, individualistic health-related recommendations and early interventions during sleep health management.

Keywords: Sleep Quality; Machine Learning; Sleep Health and Lifestyle Dataset; Ensemble Methods; Random Forest; Classification; Health Prediction

1. Introduction

Sleep is a biological phenomenon, which takes one-third of a day and has an important influence on physical and mental health [1, 2]. Moreover, sleep plays an integral role in the health and overall well-being of individuals across all age groups, including children, adolescents, and adults. Sufficient and quality sleep plays a vital role in supporting cognitive performance, mood stability, psychological health, and maintaining cardiovascular, cerebral, and metabolic functions [3]. In the past few years, research efforts have been directed toward understanding what influences sleep quality, since poor sleep is connected to serious health problems like heart disease, obesity, diabetes, and depression. [4–6]. However, the clinical study methodology has mostly been based on studying one variable in isolation, whereas data-driven methodology is required, which could simultaneously examine a multitude of variables of health and lifestyle factors to develop a more in-depth insight into their overall impact on sleep health.

The availability of comprehensive datasets has opened new avenues for researchers to analyze sleep patterns using machine learning techniques [7]. The dataset of Sleep Health and Lifestyle includes 400 records and 13 features capturing demographic details, physiological indicators, and lifestyle habits. These features range from age, occupation, and gender to physical activity levels, stress levels, BMI category, and more. Such a rich dataset allows researchers to explore the multifactorial nature of sleep health and generate predictive models that can support clinical decision-making or personal health monitoring.

In the modern era of big data, machine learning has become a powerful classification algorithm, industries and decision-making. Large volumes of data are used to train machine learning algorithms, which yield highly valuable predictions. A fundamental task in machine learning is classification, which involves assigning labels or classes to inputs and is determined by their features [8]. All fields apply machine learning, including medicine, since this field can identify trends in the information that is not quickly noticed. This study explores some machine learning models that were involved to predict sleep quality, as a categorical feature, using the features provided in the dataset. Naive Bayes, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Random Forest, Decision Tree, Logistic Regression, Gradient Boosting, and AdaBoost are the models in question. Accuracy, precision, recall, and F1 score were used to measure all the algorithms to get the effectiveness of each algorithm in capturing the relationships in the data.

Because it was explicitly stated in the experimental part, ensemble-based models like Random Forest or Gradient Boosting showed a much better result compared to the linear models or SVM. Random Forest, specifically, has gained the best accuracy (98.67%) and has been stable throughout all the assessment metrics. Conversely, SVM performed poorly, and this means that it might not suit this kind of data unless further adjusted or trans- formed. Such findings indicate that there may be complex, non-linear interactions amongst the features, which tree-based ensemble models judge better.

This research paper forms the basis that machine learning has a potential to enhance our understanding on sleep health due to predictive modelling. Predictive systems can be developed based on the information regarding different health and lifestyle aspects to assist medical professionals or citizens to identify sleep-related issues at their early onset. The study is also a contribution to the growing body of literature explaining why artificial intelligence should be implemented in preventative health to open the door to personalized recommendations and quality of life improvement.

2. Related Work

There exists a large amount of literature that follows the approach and use of machine learning practices within the scope of the healthcare sector as well as especially in regard to sleep-related data. The common purposes of these studies are to predict sleep disorders, assess the quality of sleep and identify the factors which determine the quality of sleep. As an illustration, Uezu et al. [9] examined the sleeping habits of elderly people in Okinawa Prefecture, where there is a high longevity of the population. Their contribution highlighted the importance of some lifestyle rituals in healthy sleeping, thus showing the importance of daily routines in the general health of sleep.

In the same line of thought, Dzierzewski et al. [10, 36] examined the role of lifestyle in affecting sleep at various life stages. They discovered that although routines, such as regular sleep patterns, healthy eating, and exercising are widely helpful, they may have different effects across and within demographic groups. It shows the significance of conducting population-specific sleep health approaches.

More so, strengthening the association between lifestyle and sleep, Taira et al. [11] examined the older adults and revealed that the poor sleep quality frequently co-exists with the deterioration of physical and mental health. Notably, however, they also mentioned that physiological variables, including age, blood pressure, and BMI, are important as well, as they create the impression that both behavioral and biological elements interact and should be taken into account when assessing the state of sleep health.

In conjunction with these lifestyle-centered investigations, a separate body of literature has used machine learning to promote the objectivity and productivity of sleep assessment. Machine learning algorithms were used in one such study [12, 37] to automatically classify sleep stages based on EEG spectrograms, providing an alternative that is several times faster and more accurate than manual classification. Such a direction indicates the increased promise of ML-based tools in the automation of sophisticated clinical tasks.

Extending beyond individual-level data, research has also begun to examine environmental influences on sleep. For instance, [13, 38, 39] explored the relationship between sleep health and factors exacerbated by climate change, such as temperature variation, pollution, and trauma. This study also identified vulnerable populations disproportionately affected by these conditions and highlighted the need for both adaptive strategies and further investigation into under-researched areas.

Parallel to these developments, the use of physiological biofeedback in sleep prediction has gained momentum. A recent study [14] introduced machine learning models that utilize heart rate variability and skin temperature to estimate sleep quality, based on the Pittsburgh Sleep Quality Index (PSQI). The results emphasized the effectiveness of wearable and contactless technologies in real-time, non-invasive sleep monitoring.

Taking the technological innovation further, another study [15] proposed an advanced framework that integrates Transformer-based multivariate time series modeling with ensemble learning methods. This model not only predicted sleep quality but also assessed emotional states and stress levels, outperforming traditional approaches and showcasing the power of deep learning in capturing complex temporal patterns. A review by [16] explored the current and future roles of AI in various aspects of sleep care, including screening, monitoring, prevention, prediction, diagnosis, and treatment of sleep disorders.

According to the authors, AI is a powerful and versatile tool with the potential to revolutionize clinical workflows by enhancing the efficiency of healthcare providers and improving the quality of care delivered to patients with sleep disorders. A comprehensive study conducted in China explored the multifactorial influences on insomnia, focusing on lifestyle, health conditions, and environmental factors [17, 40]. Using interviews and questionnaires, the researchers analyzed data from patients experiencing poor sleep, revealing the complex and interconnected nature of insomnia. This study underscores the importance of considering multiple influences—beyond just physiological or environmental factors—when assessing and predicting sleep disorders, particularly insomnia.

An article [18] reviews the architecture of sleep and provides evidence for its critical role in sleep health, including its impacts on mental and emotional well-being, as well as cognitive function and performance. The study also discusses strategies for improving sleep health through public health initiatives, highlighting the need to tackle sleep disparities within populations in order to improve overall health outcomes.

Despite these advancements, much of the existing literature remains limited by a narrow focus on isolated variables or single-source data, such as physiological signals alone. In contrast, the Sleep Health and Lifestyle Dataset enables a more comprehensive analysis by incorporating a diverse set of features, including demographic, behavioral, and physiological indicators. This multifaceted approach supports the growing shift toward personalized and preventive healthcare driven by artificial intelligence, enabling a deeper understanding of the factors influencing sleep quality.

3. Dataset

The Dataset of Sleep Health and Lifestyle is a structured and tabular dataset formulated to study how different demographic, physiological, and lifestyle factors affect sleep quality, which is obtained from Kaggle1. It has 373 records and 13 features (both numerical and categorical variables) presented in Table 1.

The important characteristics are gender, age, physical activity level, occupation, daily steps, stress level, heart rate, and BMI category. The target variable, quality of Sleep is numerical though in this study it has been considered as a categorical variable due to classification purpose. Also, the dataset contains a sleep disorder label and blood pressure along with a heart rate, which makes it possible to study health-related indicators in an in-depth manner. A variety of features qualifies this dataset to train machine learning models that would be used to find patterns and predict sleep health outcomes.

Table 1. Sleep Dataset Features

Feature	Type / Description
Person ID	Identifier
Gender	Categorical
Age	Numerical
Occupation	Categorical
Sleep Duration	Numerical
Physical Activity Level	Numerical
Stress Level	Numerical
BMI Category	Categorical
Blood Pressure	Textual format (e.g., "126/83")
Heart Rate	Numerical
Daily Steps	Categorical
Quality of Sleep	Target variable (numerical but treated as categorical)

3.1. Data Preprocessing

Data preprocessing is a foundational step in making the dataset appropriate for machine learning model development. The following techniques were applied:

- **Stop Words Removal:** Extraction of redundant words with low or no semantic meaning, known as stop words, was performed to reduce noise and enhance the relevance of textual features [19].
- **Replace Missing Values:** Incomplete data entries were addressed by replacing missing values using appropriate imputation techniques [20] mean is used for numerical values, and the mode is used for categorical values, ensuring the dataset's integrity and minimizing bias.
- **Standardization (Z-Score Normalization):** In order to make the data more appropriate for algorithms that are sensitive to feature distributions, this technique was used to modify features so that they had a mean of zero and a standard deviation of one [21, 22].

4. Feature Engineering

To transform unstructured attributes into formats appropriate for machine learning models, feature engineering was done. One-hot encoding was used to convert categorical variables, including gender, occupation, BMI category, and daily steps, which made it possible for the algorithms to efficiently process non-numeric data. Furthermore, the Blood Pressure information was processed by feature extraction by separating it into two separate numerical components: Systolic and Diastolic pressure. Originally, the feature was recorded as a textual value (for example, "126/83"). This change improved the model's ability to extract insights from health-related indicators and allowed for more accurate analysis. While maintaining the interpretability of the input data, these engineering features helped to improve the model's predicted accuracy.

5. Research Methodology

This study thoroughly investigates how well several machine-learning algorithms predict brain cancers using clinical medical data. ZeroR, K-Nearest Neighbours (KNN), KStar, J48 decision tree, Multilayer Perceptron (MLP), Support Vector Machine (SVM), AdaBoost, and Naïve Bayes are the eight models that are put to the test.

5.1. Naïve Bayes

Naïve Bayes is a classification model grounded in probability theory, utilizing Bayes' Theorem to estimate the likelihood of an outcome based on prior information. The algorithm is termed "naïve" because it assumes that all input features are conditionally independent given the class label, which simplifies computation [24–26, 41–43]. While this assumption may not always reflect real-world scenarios, Naïve Bayes often delivers strong performance, particularly in text classification tasks like spam filtering. It evaluates the posterior probability for each class and selects the one with the highest value for prediction. Equation 1 provides the mathematical formulation for the posterior probability of a given class c. P (cx_1, x_2, ..., x_n) is the posterior probability of class c.

$$P(c \mid x_1, x_2, \ldots, x_n) = \frac{P(c) \cdot \prod_{i=1}^{n} P(x_i \mid c)}{P(x_1, x_2, \ldots, x_n)} \tag{1}$$

5.2. Support Vector Machine (SVM)

SVM is a supervised learning technique that is applied to regression and classification [27, 44–46]. The primary goal is to uncover the ideal hyperplane that divides the data points of various classes with the largest possible margin between them. These data points are known as support vectors because they define the limits. When the data is not linearly separable, SVM maps it to a high-dimensional space where linear separability is feasible using kernel functions. The linear, polynomial, and RBF kernels represent standard choices for kernel functions in machine learning. Equation 2 provides the decision boundary.

$$W^T x + b = 0 \tag{2}$$

Where w represents weight, x represents the input, and b is the bias.

5.3. K-Nearest Neighbors

KNN is a very simple as well as instance-based algorithm that stores the entire training dataset and classifies new data points by comparing them with their nearest neighbors. When a new instance is introduced, the algorithm identifies the k closest training examples based on distance metrics like Euclidean, Manhattan, or cosine distance [28]. Among these neighbors, the majority class determines the anticipated label. When used on large datasets, KNN can become computationally costly because it doesn't require a training step, which makes it simple to build.

In the eq 3. d (x_i, x_j) is the Euclidean distance between instances x_i and x_j.

$$d(x_i, x_j) = \sqrt{\sum_{n=1}^{N} (X_{in} - X_{jn})^2} \tag{3}$$

5.4. Decision Tree (C4.5 / J48)

An enhancement on the previous ID3 method, the C4.5 decision tree technique is implemented by J48. The dataset is recursively divided according to the characteristic that provides the most information gain, with the goal of making the target labels more uniform within the resulting subsets [29] [47–49]. The procedure keeps going until a stopping condition—like a homogenous class distribution or a minimum number of instances—is satisfied. Internal nodes (feature splits) and leaf nodes (class predictions) make up the final decision tree, which creates a collection of comprehensible decision rules. Where eqs. 4 and 5 are used to determine information gain and entropy, respectively.

Information Gain:

$$IG(S, A) = \text{Entropy}(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|} \times \text{Entropy}(S_v) \tag{4}$$

Entropy:

$$H(S) = -\sum_{c \in C} P(c) \cdot \log_2 P(c) \tag{5}$$

5.5. Random Forest

A powerful ensemble learning technique called Random Forest constructs many decision trees during the training phase and uses the individual trees to provide the mean prediction for regression tasks or the mode of the classes for classification tasks [30]. Using bootstrap aggregating or bagging, each tree is trained on a randomly selected subset of features and a randomly selected part of the dataset. The model's ability to generalize is enhanced and overfitting is reduced by introducing randomness. The Random Forest model's prediction in the regression scenario is represented mathematically by Equation 6.

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} h_t(x) \qquad (6)$$

Where y^ is the final prediction, T stands for the total number of trees, and h_t (x) is the prediction of the t^{th} decision tree for the input x.

By aggregating the outputs of multiple trees, Random Forest improves accuracy and robustness [31]. It is especially effective for high-dimensional datasets and demonstrates resilience to noise and overfitting. Furthermore, Random Forest models can provide estimates of feature importance, making them valuable for both predictive performance and interpretability in various machine learning tasks.

5.6. Logistic Regression

A statistical model for binary and multi-class classification issues is called logistic regression [32, 33]. In contrast to linear regression, it uses the logistic (sigmoid) function to forecast the likelihood that an input with a particular input will belong to a particular class. The estimated parameters of a model map input features to a probability score between 0 and 1.

The probability output defines the decision boundary in the case of a binary classification problem. An input is assigned the class 1 when the estimated probability exceeds a preset threshold (typically 0.5), otherwise, it is assigned the class 0. Linear combination of the input features is given mathematically by:

$$z = w^T x + b \qquad (7)$$

Here, w stands for the weight vector, x corresponds to the input vector, and b indicates the bias component. The linear combination of these components yields z, which is then passed through the sigmoid (logistic) function to generate the predicted probability.

$$\sigma(z) = \frac{1}{1+e^{-z}} \qquad (8)$$

Thus, the predicted probability that the output y is 1 given input x is:

$$P\,(y = 1 \mid x) = \sigma\,(w^T x + b) \qquad (9)$$

The final classification decision based on a threshold τ (typically 0.5) is:

$$\hat{y} = \begin{cases} 1, & \text{if } P(y = 1 \mid x) \geq \tau \\ 0, & \text{otherwise} \end{cases} \qquad (10)$$

In Equation (10), the probability P (y = 1 ∣ x) is given by the sigmoid function applied to the linear combination of weights and input features. Logistic Regression is valued for its simplicity, interpretability, and efficiency when applied to linearly separable datasets.

5.7. Gradient Boosting

By gradually adding weak learners, typically decision trees, gradient boosting is an ensemble strategy that builds a powerful predictive model [34]. The residual errors produced by the ensemble of previously trained models are corrected by each new model. The algorithm optimizes a chosen loss function by applying gradient descent techniques to minimize the prediction error, which is where the term "gradient boosting" originates.

$$\hat{y}(x) = \sum_{m=1}^{M} \eta \cdot h_m(x) \qquad (11)$$

A new learner hm(x) is trained to approximate the negative gradient of the loss function with regard to the predictions of the existing model at each iteration m. The ensemble prediction after M iterations is given by:

Where y^(x) is the final prediction, M is the total number of boosting rounds, hm(x) is the m^{th} weak learner, and η ∈ (0, 1] is the rate of learning that is used to control the contribution of each learner.

Equation (11) represents the cumulative output of all weak learners, each scaled by the learning rate and aligned in the direction that minimizes the loss function. Gradient Boosting is highly flexible, supporting different types of loss functions, data types, and regularization methods. However, it requires careful tuning of hyperparameters to prevent overfitting and to ensure generalization.

5.8. Adaboost

The AdaBoost word stands for Adaptive Boosting. By integrating several weak models, this kind of ensemble learning technique improves classification accuracy [35]. It operates iteratively, training a new model at each step that directs more attention to the instances misclassified by previous models. These difficult cases are given more weight in subsequent rounds, while the outputs of individual models are aggregated based on their accuracy. AdaBoost enhances predictive performance and is often used with simple base classifiers like decision stumps.

$$H(x) = Sign(\sum_{t=1}^{T} \alpha_t h_t(x)) \tag{12}$$

In eq 12. h_t (x) is the weak classifier at iteration t, and alpha is a weight. T is used to show the total number of classifiers.

$$P(c \mid x_1, x_2, \ldots, x_n) = \frac{P(c) \cdot \prod_{i=1}^{n} P(x_i \mid c)}{P(x_1, x_2, \ldots, x_n)} \tag{13}$$

P (cx_1, x_2, ..., x_n) is the posterior probability of class c.

6. Model Implementation

The structured machine learning methodology used by the suggested model is described in Section 5 and is depicted in Figure 1. It starts with the input layer, which contains personal characteristics such as numerical and categorical attributes. Few of them, like person ID, are used as an identification, and blood pressure is given in text format. The data preprocessing includes Z-score standardization to match feature scales, stop word removal (if any text is included), and missing value imputation. To enhance prediction performance, feature engineering is then performed utilizing feature extraction and one-hot encoding. Different ML models, which are described in the methodology section, are trained using the prepared data. After assessing all the models, the optimal model is identified based on its performance metrics. In order to offer actionable insights into sleep quality, the model now predicts Quality of Sleep, which was initially a numerical value but is now regarded as categorical for classification.

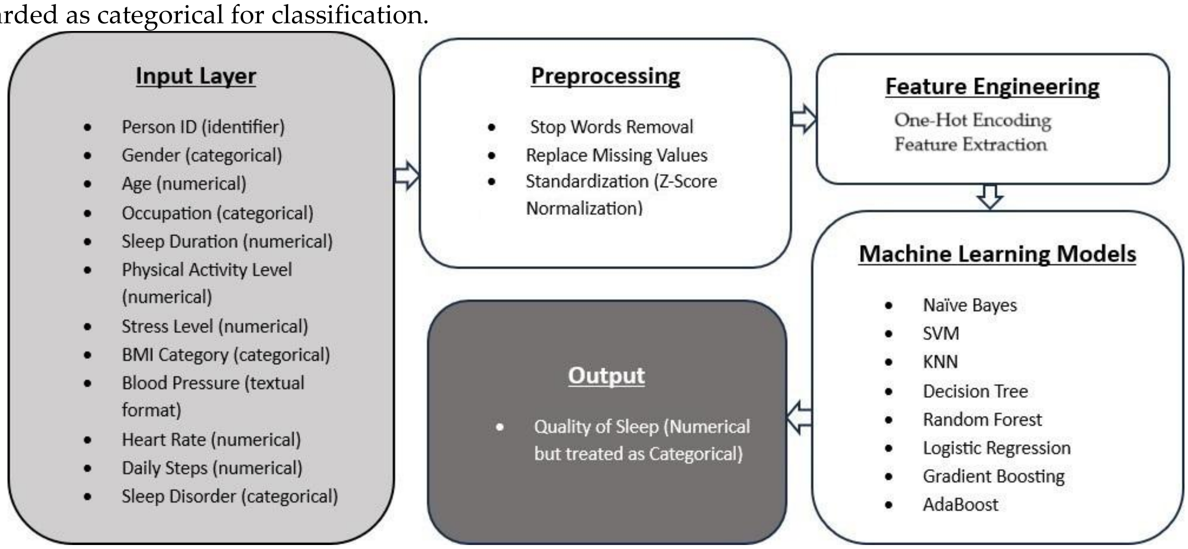

Figure 1. Architecture diagram of the Proposed Model

7. Evaluation Measures

In order to determine the effectiveness of the classification model, the confusion matrix is utilized, offering an in-depth breakdown of predicted versus actual outcomes. A confusion matrix includes four essential elements: True Positives (TP), indicating correctly identified positive cases; True Negatives (TN), representing correctly classified negative cases; False Positives (FP), where negative instances are mistakenly labeled as positive; and False Negatives (FN), where the model fails to detect actual positive cases. These elements are fundamental in deriving various performance evaluation metrics.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

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

By calculating the percentage of positive and negative predictions that are accurate out of all the forecasts produced, accuracy provides insight into the model's overall performance (Equation 1). Precision measures the fraction of instances predicted as positive that are correctly identified as positive, highlighting the metric's importance when false positives must be minimized (Equation 2). Recall—commonly referred to as sensitivity—reflects the proportion of actual positives accurately identified by the model. It is crucial when the cost of missing positive instances (false negatives) is significant (Equation 3). Finally, by calculating their harmonic mean, the F1 score offers a compromise between memory and precision.

Equation 4 represents a metric that is effective in scenarios requiring a trade-off between false positives and false negatives. When combined, these metrics offer a comprehensive assessment of the model's classification performance. By calculating the percentage of positive and negative predictions that are accurate out of all the forecasts produced, accuracy provides insight into the model's overall performance.

8. Model Performance Evaluation

Table 2 shows the performance metrics of different ML algorithms used for the prediction task. The evaluation is based on four key indicators: F1 score, precision, accuracy, and recall. Among all models, the performance metrics revealed that the Random Forest is superior because it achieves a 0.9867 score in terms of accuracy, a precision of 0.9874, a recall of 0.9867, and an F1 score of 0.9866, indicating its effectiveness across all metrics. The effectiveness of the ensemble and tree-based approach on this dataset can be proven by the good results of Naive Bayes, Decision Tree, and Gradient Boosting, which have the accuracy of 0.9733 with stable and high precision, recall, and F1 scores.

With the accuracy of 0.3733 and an F1 score of 0.3134, as illustrated in Figure 2, the results of the Support Vector Machine (SVM) are the lowest. This is an indication that the SVM would not suit this classification job well without further feature engineering or parameter tuning. Having a lower accuracy of 0.7200 and an F1 score of 0.6788, AdaBoost is quite a bad result despite the fact that it is also an ensemble method along with Random Forest and Gradient Boosting. This could imply it performs poorly in the presence of noise, and one more thing, the dataset used for this study is imbalanced data, so it needs more optimization.

K-Nearest Neighbors (KNN) and logistic regression perform moderately well. KNN achieves a solid accuracy of 0.9600 and an F1 score of 0.9495, making it a reliable and straightforward choice. Logistic Regression attains an accuracy of 0.9200 and a respectable F1 score of 0.9112, reflecting balanced precision and recall, though it lags behind tree-based models. Overall, these results highlight that ensemble methods, particularly Random Forest and Gradient Boosting, are the most effective for this task, while simpler linear models and SVM require further improvement for competitive performance. As can be shown in Figure 3, Random Forest and gradient boosting outperform all other models in terms of relative performance comparison (baseline: SVM Accuracy, because the accuracy of SVM is low compared to other models).

Table 2. Performance Metrics of Machine Learning Algorithms

Algorithm	Accuracy	Precision	Recall	F1 Measure
Naïve Bayes	0.9733	0.9748	0.9733	0.9714
Support Vector Machine	0.3733	0.7528	0.3733	0.3134
K-Nearest Neighbors	0.9600	0.9696	0.9600	0.9495
Decision Tree	0.9733	0.9760	0.9733	0.9731
Random Forest	0.9867	0.9874	0.9867	0.9866
Logistic Regression	0.9200	0.9440	0.9200	0.9112
Gradient Boosting	0.9733	0.9760	0.9733	0.9731
AdaBoost	0.7200	0.6974	0.7200	0.6788

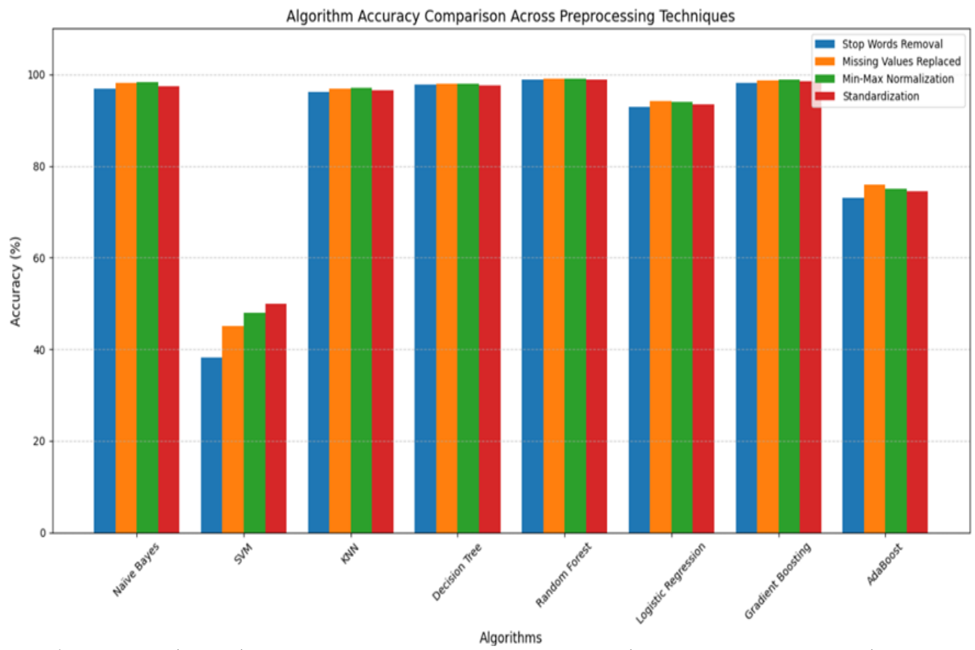

Figure 2. Algorithm accuracy comparison across the preprocessing techniques

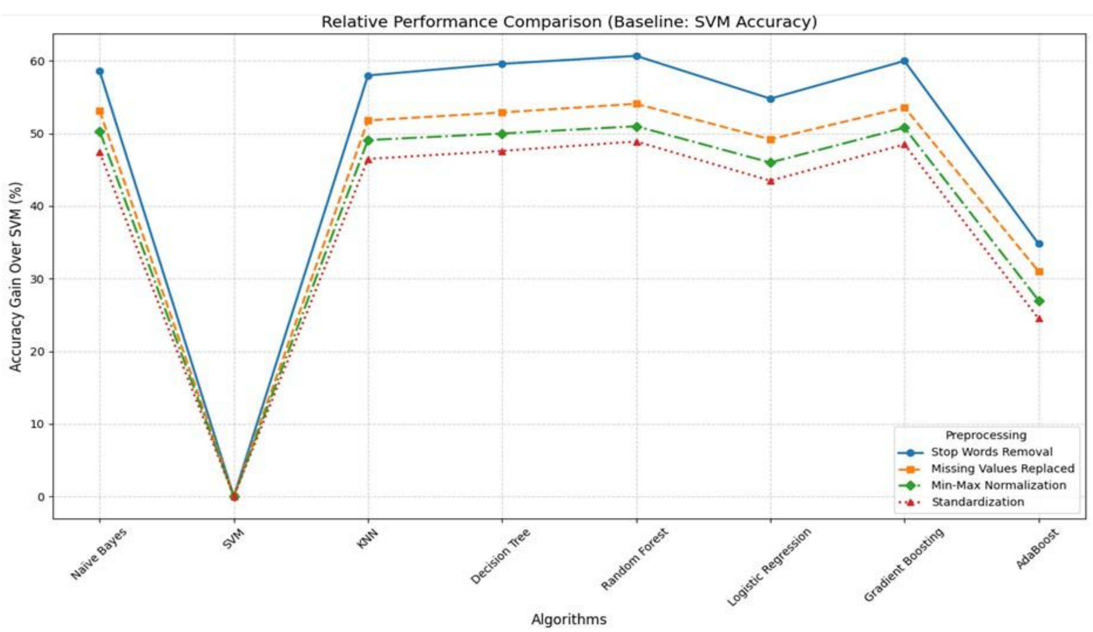

Figure 3. Relative Performance comparison (Baseline: SVM Accuracy)

9. Conclusion

The research illustrates how predictive modeling through machine learning can be applied to predict sleep health using a variety of lifestyle and demographic variables. As far as the considered models are concerned, ensemble techniques, specifically Random Forest in particular, are demonstrated to be the best as compared to the rest, since it was proven to yield the best results in all of the metrics. The strong results of Decision Tree and Gradient Boosting further reinforce the value of ensemble learning approaches. These findings suggest that integrating machine learning into health analytics can enable more accurate and personalized insights, paving the way for data-driven interventions in sleep health management and the prevention of sleep-related disorders. Additionally, the use of wearable technology to obtain real-time data may expand the influence of this in the future and produce better outcomes.

References

1. Cavallino, Vincent et al. "Antimony and sleep health outcomes: NHANES 2009-2016." Sleep health vol. 8,4 (2022): 373-379. doi:10.1016/j.sleh.2022.05.005
2. Ordway, Monica R., et al. "A systematic review of the association between sleep health and stress biomarkers in children." Sleep medicine reviews 59 (2021): 101494.
3. Ramar, Kannan, et al. "Sleep is essential to health: an American Academy of Sleep Medicine position statement." Journal of Clinical Sleep Medicine 17.10 (2021): 2115-2119.2016.
4. Kennedy, Gerard Anthony. "Social disadvantage, insufficient sleep, and cardiovascular disease." Frontiers in Sleep 4 (2025): 1500218.
5. Grech, Elizabeth, and Sarah Cuschieri. "A Good Night's Sleep in Malta in 2023: A Cross-sectional Study Exploring Sleep Quality and its Determinants via Social Media." Journal of Research in Health Sciences 24.1 (2024): e00602.
6. Boyd Jr, Kevin. Between the Environment and BMI: The Influence of Physical Activity, Sleep Quality, and Inflammation in Overweight African Americans. Diss. Howard University, 2024.
7. BV, Santhosh Krishna, et al. "Machine Learning based Human Stress Detection through Sleep Patterns." 2024 5th International Conference on Smart Electronics and Communication (ICOSEC). IEEE, 2024.
8. Baik, Shefaa, et al. "Enhancing Spare Parts Inventory Management through Machine Learning Based Classification."
9. Uezu, Eiko, et al. "Survey of sleep-health and lifestyle of the elderly in Okinawa." Psychiatry and clinical neurosciences 54.3 (2000): 311-313.
10. Dzierzewski, Joseph M., et al. "Lifestyle factors and sleep health across the lifespan." International Journal of Environmen- tal Research and Public Health 18.12 (2021): 6626.
11. Taira, Kazuhiko, et al. "Sleep health and lifestyle of elderly people in Ogimi, a village of longevity." Psychiatry and clinical neurosciences 56.3 (2002): 243-244
12. Li, Chengfan, et al. "A deep learning method approach for sleep stage classification with EEG spectrogram." International journal of environmental research and public health 19.10 (2022): 6322.
13. Wallace, Danielle A., and Dayna A. Johnson. "Climate Change, Sleep, and Mental Health." Climate Change and Mental Health Equity. Cham: Springer International Publishing, 2024. 177-203.
14. Di Credico, Andrea, et al. "Predicting Sleep Quality through Biofeedback: A Machine Learning Approach Using Heart Rate Variability and Skin Temperature." Clocks & Sleep 6.3 (2024): 322-337.
15. Kim, Jinjae, et al. "TraM: Enhancing User Sleep Prediction with Transformer-Based Multivariate Time Series Modeling and Machine Learning Ensembles." 2024 15th International Conference on Information and Communication Technology Convergence (ICTC). IEEE, 2024.
16. Verma, Ram Kishun, et al. "Artificial intelligence in sleep medicine: Present and future." World Journal of Clinical Cases 11.34 (2023): 8106.
17. Li, X., Zhou, X., Zhang, Y., Mei, R., Liu, J. (2025). Multifactor Analysis of Insomnia Influences: Predictive Ef- fects of Patients' Lifestyle and Health Status in China. American Journal of Health Research, 13(2), 109-119. https://doi.org/10.11648/j.ajhr.20251302.14
18. Hale, Lauren, Wendy Troxel, and Daniel J. Buysse. "Sleep health: an opportunity for public health to address health equity." Annual review of public health 41.1 (2020): 81-99.
19. Raulji, Jaideepsinh K., and Jatinderkumar R. Saini. "Stop-word removal algorithm and its implementation for Sanskrit language." International Journal of Computer Applications 150.2 (2016): 15-17.
20. Song, Qinbao, and Martin Shepperd. "Missing data imputation techniques." International journal of business intelligence and data mining 2.3 (2007): 261-291.
21. Jain, Y. Kumar, and Santosh Kumar Bhandare. "Min max normalization based data perturbation method for privacy pro- tection." International Journal of Computer & Communication Technology 2.8 (2011): 45-50.
22. Imron, Muhammad Ali, and Budi Prasetyo. "Improving algorithm accuracy k-nearest neighbor using z-score normalization and particle swarm optimization to predict customer churn." Journal of Soft Computing Exploration 1.1 (2020): 56-62
23. Qader, Wisam A., Musa M. Ameen, and Bilal I. Ahmed. "An overview of bag of words; importance, implementation, appli- cations, and challenges." 2019 international engineering conference (IEC). IEEE, 2019.

24. B.Seethalakshmi, “Brain tumourignorance prediction using machine learning techniques,”Brain,vol.8,no.2,pp.86–93,2024.
25. e. a. Tran Ngoc Viet, “The naïve bayes algorithm for learning data analytics,” Indian Journal of Computer Science and Engineering,vol.12,no.4,pp.1038–1043,2021.
26. Bhatti, M. U.; Saeed, A.; Farhat Ullah, M.; Sauood, M.; Ashir, M.; Hussain, N. Optimizing Brain Tumor Prediction: A Com- parative Study Of Machine Learning Algorithms. VFAST trans. softw. eng. 2024, 12, 209-219.
27. Pisner, Derek A., and David M. Schnyer. "Support vector machine." Machine learning. Academic Press, 2020. 101-121.
28. Acito, Frank. "k nearest neighbors." Predictive Analytics with KNIME: Analytics for Citizen Data Scientists. Cham: Springer Nature Switzerland, 2023. 209-227.
29. Adeyemo, O. O., T. O. Adeyeye, and D. Ogunbiyi. "Comparative study of ID3/C4. 5 decision tree and multilayer perceptron algorithms for the prediction of typhoid fever." African Journal of Computing & ICT 8.1 (2015): 103-112.
30. Rigatti, Steven J. "Random forest." Journal of Insurance Medicine 47.1 (2017): 31-39.
31. Roy, Marie-Hélène, and Denis Larocque. "Robustness of random forests for regression." Journal of Nonparametric Statis- tics 24.4 (2012): 993-1006.
32. Zabor, Emily C., et al. "Logistic regression in clinical studies." International Journal of Radiation Oncology* Biology* Physics 112.2 (2022): 271-277.
33. Das, Abhik. "Logistic regression." Encyclopedia of quality of life and well-being research. Cham: Springer International Publishing, 2024. 3985-3986.
34. Sibindi, Racheal, Ronald Waweru Mwangi, and Anthony Gichuhi Waititu. "A boosting ensemble learning based hybrid light gradient boosting machine and extreme gradient boosting model for predicting house prices." Engineering Reports 5.4 (2023): e12599.
35. Wang, Wenyang, and Dongchu Sun. "The improved AdaBoost algorithms for imbalanced data classification." Information Sciences 563 (2021): 358-374.
36. Muneer, I., Saeed, A., & Adeel Nawab, R. M. (2025). Cross-Lingual English–Urdu Semantic Word Similarity Using Sentence Transformers. The European Journal on Artificial Intelligence, 30504554241297614.
37. Ahmad, M. H., Saeed, A., Bhatti, M. U., Hussain, N., Ullah, M. F., & Anwar, M. (2025). Next Word Prediction for Urdu using Deep Learning Techniques. VFAST Transactions on Software Engineering, 13(1), 49-59.
38. https://www.kaggle.com/code/sangeethamarikkannan/brain-tumor-dataset/input Last Visited: 28-Nov-2024
39. Bhatti, M. U., Saeed, A., Ullah, M. F., Sauood, M., Ashir, M., & Hussain, N. (2024). Optimizing Brain Tumor Prediction: A Comparative Study Of Machine Learning Algorithms. VFAST Transactions on Software Engineering, 12(4), 209-219.
40. Saeed, A., Ullah, M. F., Sauood, M., Ali, S. N., & Hussain, N. (2024). Prediction of dengue cases and deaths using machine learning algorithm. Pak. J. Sci. Res, 3(2), 210-216.
41. Saeed, A., Farrukh, M. A., ul Haque, H. M., & Javaid, D. (2024). Advanced Machine Learning Algorithms for Accurate Pre- diction of Band Gaps in Rare Earth Metal Oxides Nanoparticles. ES Energy & Environment.
42. Ullah, M. F., Saeed, A., & Hussain, N. (2023). Comparison of Pre-trained vs Custom-trained Word Embedding Models for Word Sense Disambiguation. ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal, 12, e31084- e31084.
43. Ullah, M. F., Saeed, A., Li, J., Mahmood, T., & Adeel, M. (2023). BERT model for Roman Urdu fake review identification.
44. Fatima, G., Nawab, R. M. A., Khan, M. S., & Saeed, A. (2021). Developing a cross-lingual semantic word similarity corpus for English–Urdu language pair. Transactions on Asian and Low-Resource Language Information Processing, 21(2), 1-16.
45. Vardag, M. H. K., Saeed, A., Hayat, U., Ullah, M. F., & Hussain, N. (2022). Contextual Urdu text emotion detection corpus and experiments using deep learning approaches. ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal, 11(4), 489-505.
46. Hayat, U., Saeed, A., Vardag, M. H. K., Ullah, M. F., & Iqbal, N. (2022). Roman urdu fake reviews detection using stacked lstm architecture. SN Computer Science, 3(6), 470.
47. Muneer, I., Fatima, G., Khan, M. S., Adeel Nawab, R. M., & Saeed, A. (2023). Developing a large benchmark corpus for Urdu semantic word similarity. ACM Transactions on Asian and Low-Resource Language Information Processing, 22(3), 1-19.

48. Saeed, A., Nawab, R. M. A., & Stevenson, M. (2021). Investigating the feasibility of deep learning methods for urdu word sense disambiguation. Transactions on Asian and Low-Resource Language Information Processing, 21(2), 1-16.
49. Saeed, A., Nawab, R. M. A., Stevenson, M., & Rayson, P. (2019). A word sense disambiguation corpus for Urdu. Language Resources and Evaluation, 53, 397-418.
50. Saeed, A., Nawab, R. M. A., Stevenson, M., & Rayson, P. (2019). A sense annotated corpus for all-words Urdu word sense disambiguation. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(4), 1-14.