Machine Learning Classification for Sleep Disorder Detection: A Comparative Analysis of Five Algorithms

Tejas Pagare
S20230010172

Krishna Suryawanshi
S20230010126

Yashwardhan Singh Chouhan
S20230010268

Vedant Kasar
S20230010118

Vasudev Dilware
S20230010253

***Abstract*—Sleep disorders represent a significant public health concern, affecting millions worldwide. This paper presents a comprehensive comparative study of five standard machine learning classifiers: K-Nearest Neighbors (KNN), Decision Tree, Random Forest, Support Vector Machine with One-vs-Rest (LinearSVM_OvR), and Logistic Regression on a sleep health and lifestyle dataset. The classification task involves predicting three sleep disorder outcomes: None, Insomnia, and Sleep Apnea. A standardized experimental framework was implemented using Python and Scikit-learn, with consistent preprocessing including feature scaling and one-hot encoding. Results demonstrate that Logistic Regression achieved the highest test accuracy (0.877), with Random Forest and LinearSVM_OvR also performing strongly (0.849). Random Forest shows superior F1-score performance on the Insomnia class (0.90). The permutation feature importance analysis for Random Forest reveals that daily steps and BMI category are the most influential predictors of sleep disorders.**

***Index Terms*—Sleep Disorders, Machine Learning, Classification, KNN, Decision Tree, Random Forest, Support Vector Machine, Logistic Regression**

I. Introduction

Sleep disorders are increasingly recognized as critical health conditions affecting various populations globally. Approximately one-third of adults experience symptoms of insomnia, while sleep apnea affects millions, often remaining undiagnosed. Early detection and accurate classification of sleep disorders can significantly improve patient outcomes and quality of life through timely intervention and lifestyle modifications.

Machine learning has emerged as a powerful tool in healthcare diagnostics, enabling the analysis of complex relationships between physiological, behavioral, and lifestyle factors. This paper investigates five foundational machine learning classification algorithms: K-Nearest Neighbors (KNN), Decision Tree, Random Forest, Support Vector Machine with One-vs-Rest (LinearSVM_OvR), and Logistic Regression.

The sleep health dataset contains 374 samples with 13 features capturing demographic information, health metrics, and lifestyle factors. The target variable is Sleep Disorder, which has three classes: None, Insomnia, and Sleep Apnea.

II. Methodology

A. Dataset Description

The dataset comprises 374 samples with the following features:

1) Demographic Features: Gender (Male, Female), Age (years), Occupation (various professions)

2) Health Metrics: BMI Category (Normal, Overweight, Obese), Blood Pressure, Heart Rate (bpm)

3) Sleep and Lifestyle Features: Sleep Duration (hours), Quality of Sleep (1-10 scale), Physical Activity Level, Stress Level (1-8), Daily Steps

4) Target Variable: Sleep Disorder: None (43 samples), Insomnia (15 samples), Sleep Apnea (15 samples)

B. Preprocessing Pipeline

Categorical features were transformed using one-hot encoding. Numeric features were standardized using StandardScaler. The dataset was split using an 80-20 train-test split.

C. Classification Models

KNN: Instance-based learner using $k = 5$ and Euclidean distance.

Decision Tree: Recursive feature partitioning with Gini impurity.

Random Forest: Ensemble of 100 decision trees with majority voting.

LinearSVM_OvR: One-vs-Rest linear kernel SVM for multi-class classification.

Logistic Regression: Linear model using the one-vs-rest approach for multi-class classification.

D. Evaluation Metrics

All models were evaluated using Accuracy, Precision, Recall, and F1-Score metrics.

III. Results and Discussion

A. Overall Model Performance

TABLE I
Overall Model Performance Comparison on Test Set

Model	Acc.	Prec.	Recall	F1	Time (s)
KNN	0.808	0.82	0.86	0.81	˜0.00
DT	0.836	0.89	0.91	0.83	0.0700
RF	0.849	0.85	0.91	0.85	0.4400
SVM	0.849	0.85	0.91	0.85	0.0700
LR	**0.877**	**0.90**	**0.93**	**0.88**	0.0800

1) Key Findings:

1) **Top Performer**: **Logistic Regression (LR) achieved the highest test accuracy (0.877)** and the best overall F1-score (0.88).
2) **Strong Contenders**: Random Forest (RF) and LinearSVM_OvR (SVM) achieved identical and strong test accuracy (0.849).
3) **KNN Performance**: KNN achieved the lowest accuracy (0.8082) but still maintained reasonable performance.

B. Test Accuracy Comparison

Fig. 1 illustrates the test accuracy comparison across all models, clearly showing the superior performance of Logistic Regression.

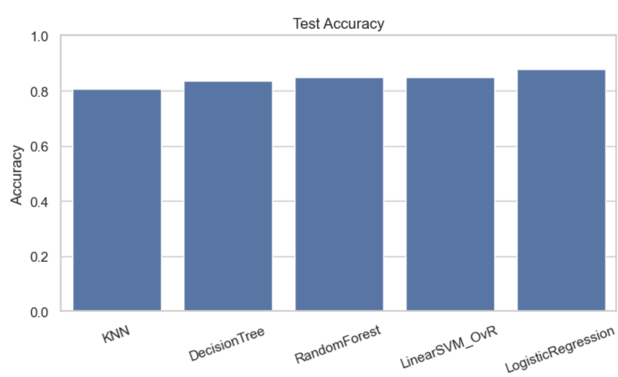

Fig. 1. Model comparison: Test accuracy across all five classifiers

C. Training Time Analysis

Model training efficiency is an important practical consideration for deployment. As shown in Fig. 2, Random Forest requires the longest training time (0.44 seconds) due to its ensemble nature. Logistic Regression is highly efficient, training in approximately 0.08 seconds, similar to Decision Tree and LinearSVM_OvR.

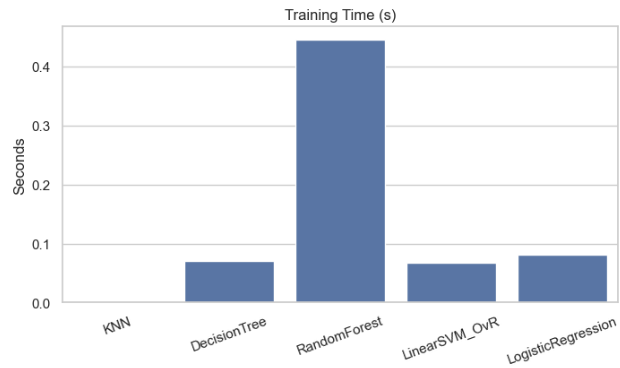

Fig. 2. Model comparison: Training time (seconds) for each classifier

D. Confusion Matrix Analysis

The confusion matrices in Fig. 3 show distinct classification patterns. Logistic Regression correctly classified 39 of 43 “None” cases, **13 of 15 “Insomnia” cases**, and **12 of 15 “Sleep Apnea” cases**. Random Forest also performed well, correctly classifying 39 of 43 “None” cases, 13 of 15 “Insomnia” cases, and 10 of 15 “Sleep Apnea” cases.

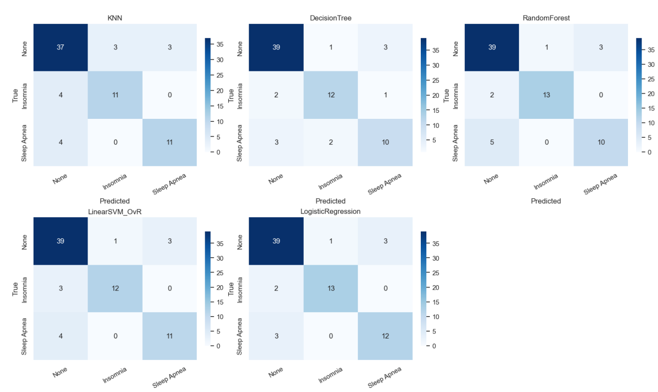

Fig. 3. Confusion matrices for all five classifiers showing true vs. predicted labels

E. Per-Class F1 Score Analysis

Fig. 4 presents per-class F1 scores. All models achieve a high F1-score for the ”None” class. Random Forest shows the strongest F1-score for the **Insomnia** class (0.90), while Logistic Regression and LinearSVM_OvR perform best on the **Sleep Apnea** class (F1-score ≈ 0.76).

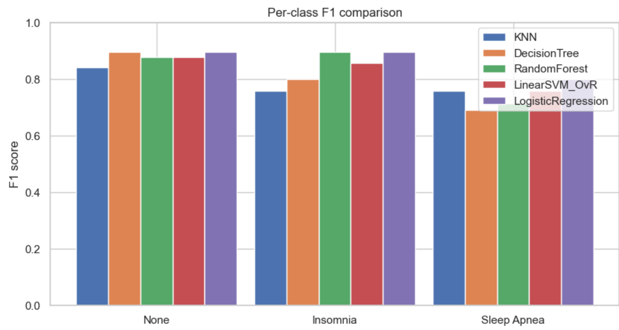

Fig. 4. Per-class F1 comparison across all five models for each sleep disorder category

F. Feature Importance Analysis

Fig. 5 visualizes the top features influencing sleep disorder predictions based on the Random Forest model. Daily steps and BMI category emerge as the strongest predictors, suggesting lifestyle factors are critically associated with sleep health.

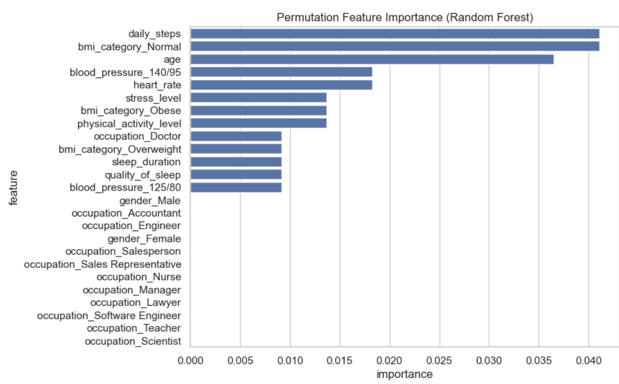

Fig. 5. Permutation Feature Importance for Random Forest showing top predictive features

TABLE II
TOP 10 FEATURE IMPORTANCE RANKINGS

Rank	Feature	Importance
1	Daily Steps	0.0411
2	BMI Category (Normal)	0.0411
3	Age	0.0365
4	Blood Pressure (140/95)	0.0183
5	Heart Rate	0.0183
6	Stress Level	0.0137
7	BMI Category (Obese)	0.0137
8	Physical Activity Level	0.0137
9	Occupation (Doctor)	0.0091
10	BMI Category (Overweight)	0.0091

1) Key Insights: Daily steps and BMI category are the strongest predictors (0.0411 each), suggesting physical activity and weight management are critically associated with sleep health. Age ranks third in importance (0.0365). Blood pressure and heart rate show moderate importance, while occupation-related features show minimal importance.

G. Model Comparison Summary

TABLE III
MODEL COMPARISON SUMMARY AND PRACTICAL RECOMMENDATIONS

Performance Metric	Best Performer
Overall Accuracy	Logistic Regression (0.877)
Insomnia F1-score	Random Forest (0.90)
Sleep Apnea F1-score	LR / SVM (≈ 0.76)
Training Speed	KNN ($<$0.01 s)
Generalization	LR / RF / SVM
Interpretability	Decision Tree / Logistic Regression

IV. CONCLUSION

This comparative study systematically evaluated five machine learning classifiers on a sleep health and lifestyle dataset for multi-class sleep disorder detection.

A. Key Conclusions

1) **Linear Model Dominance**: Logistic Regression achieved the highest overall test accuracy (87.7%), proving to be the most effective classifier for this specific dataset.
2) **Ensemble Performance**: Random Forest and LinearSVM_OvR achieved strong and competitive performance (84.9% accuracy), demonstrating robust generalization.
3) **Feature-Driven Insights**: Daily steps, BMI category, and age emerge as the most predictive features, emphasizing the importance of objective lifestyle factors in sleep health assessment.
4) **Class Imbalance Challenge**: While overall accuracy is high, all models struggle comparatively with the minority Sleep Apnea class, suggesting future work should address this class imbalance.
5) **Trade-offs in Model Selection**:
   - *Maximum Accuracy and Speed*: Logistic Regression offers the best combination of top performance and fast training.
   - *Maximum Robustness*: Random Forest provides a stable, high-performance option.
   - *Interpretability*: Decision Tree and Logistic Regression provide more explicit decision rules/feature weights.
6) **Clinical Implications**: The best-performing models can be used as effective decision support tools in conjunction with clinical evaluation by sleep medicine specialists.

B. Future Work

1) Hyperparameter optimization using grid search or Bayesian optimization
2) Advanced ensemble methods (XGBoost, LightGBM, Stacking)
3) Addressing class imbalance via SMOTE and weighted loss functions
4) Feature engineering and interaction terms

5) Clinical validation with sleep specialists
6) Model explainability using SHAP or LIME

Acknowledgment

The authors acknowledge the availability of the sleep health and lifestyle dataset and the support of the institution. We thank the open-source Python community for Scikit-learn, NumPy, and Pandas.

References

[1] D. H. Wolpert and W. B. Macready, "No free lunch theorems for optimization," *IEEE Trans. Evol. Comput.*, vol. 1, no. 1, pp. 67–82, 1997.

[2] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, 1967.

[3] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[4] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[5] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, C. Cui, G. Corrado, S. Thrun, and J. Dean, "A guide to deep learning in healthcare," *Nat. Med.*, vol. 25, no. 1, pp. 24–29, 2019.

[6] A. Rajkomar, J. Dean, and I. Kohane, "Machine learning in medicine," *New Engl. J. Med.*, vol. 380, no. 14, pp. 1347–1358, 2019.

[7] S. Dreiseitl and L. Ohno-Machado, "Logistic regression and artificial neural network classification models: a methodology review," *J. Biomed. Inform.*, vol. 35, no. 5–6, pp. 352–359, 2002.

[8] M. S. Redline and S. L. Kirchner, "Machine learning in sleep medicine," *J. Clin. Sleep Med.*, vol. 16, no. 1, pp. 1–10, 2020.

[9] V. Vapnik, *The Nature of Statistical Learning Theory*. New York: Springer-Verlag, 1995.

[10] J. R. Quinlan, "Induction of decision trees," *Mach. Learn.*, vol. 1, no. 1, pp. 81–106, 1986.