ADVANCEMENTS IN AUTOMATED SLEEP STAGE SCORING: A STUDY USING K-FOLD VALIDATION FOR MACHINE LEARNING TECHNIQUES

Tanmay Rathod[1][0009-0002-6991-054X] and Santosh Kumar Satapathy [1111-2222-3333-4444]

[1] Princeton University, Princeton NJ 08544, USA
[2] Springer Heidelberg, Tiergartenstr. 17, 69121 Heidelberg, Germany
lncs@springer.com

Abstract. Sleep stage scoring, automated in polysomnography, has evolved significantly due to the advancement in machine learning approaches. The performance of these models was determined on polysomnographic data, using the models of Random Forest, Ensemble Learning, Gradient Boosted Classifier, and SGD in association with K-Fold validation. Among them, the Ensemble Learning model achieved maximum accuracy of 95% with precision and F1 score of 98.23% and 98.93%, respectively, by the Wake model. Random Forest achieved an accuracy of 88% while the N2 stage showed a precision of 89.74% and an F1-score of 91.36%. Gradient Boosted Classifier and SGD delivered 87%, and 80% accuracy correspondingly, whose performance was not quite consistent at various stages. Here, our proposed methodology includes a pipeline of features extracted from power bands of EEG data for better presentation. These results show the immense potential of Machine Learning in arriving at robust interpretable automated Sleep Stage Scoring.

Keywords: Automated sleep stage scoring (ASSS) , Machine Learning (ML) , Electroencephalogram (EEG) , Feature Engineering (FE) , Electrooculogram (EOG) ,

1 Introduction

Sleep is essential for human health, influencing various physiological functions necessary for daily activities, mental well-being, and overall health. Disruptions in sleep can lead to conditions such as restless legs syndrome, sleep apnea, and insomnia, which are becoming increasingly common due to modern lifestyles and stress. These sleep-related disorders can be monitored and managed through bioelectric signals, particularly through the use of electroencephalogram (EEG) recordings. In recent years, advancements in electrical and signal processing have facilitated the automation of sleep data analysis. Traditional methods, like polysomnography and EEG recordings, are time-consuming, but machine learning and deep learning advancements offer promising automation solutions. This discussion will focus on the benefits and challenges of these modern techniques, which could soon be integrated into consumer devices. Specifically, it will explore how wearable technology can be utilized to extract robust features and how machine learning algorithms can improve accuracy to levels comparable to medical professionals.

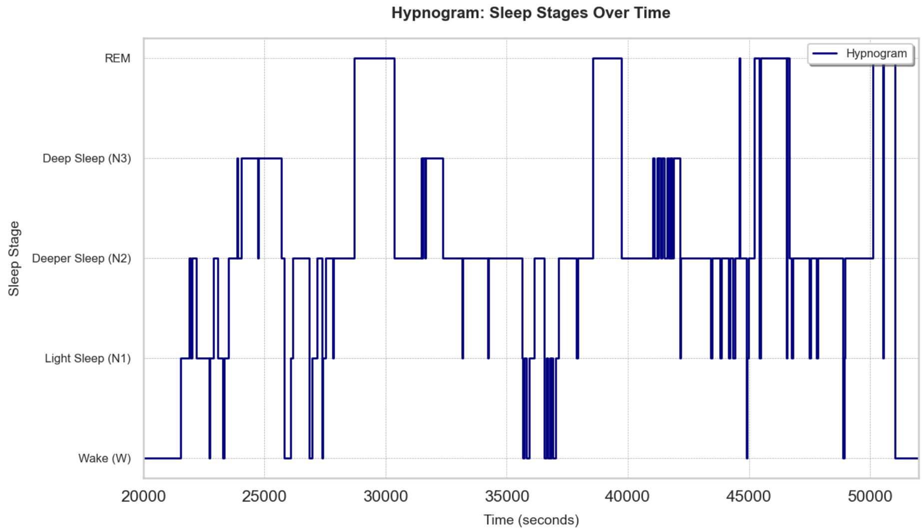

Fig. 1. Hypnogram Sleep Stages Over Time

Despite progress in sleep scoring methods, significant gaps remain. Sleep, a state of reduced physical and mental activity, is essential for processes such as dreaming, brain activity changes, neuron regeneration, muscle recovery, and overall health maintenance. It also helps reduce the risk of chronic conditions like type 2 diabetes and heart disease, while promoting mental health and immunity. Sleep plays a crucial role in the growth, development, and cognitive function of younger individuals, enhancing memory and learning. Bio signals, which reflect physiological activities such as breathing, heartbeats, and brain signals, are central to sleep analysis. Techniques like EEG, electrocardiogram (ECG), and polysomnography (PSG) allow us to capture these bio signals. Additionally, other physiological signals include electromyogram (EMG), electrogastrogram (EGG), electroretinogram (ERG), electrooculogram (EOG), and phonocardiogram (PCG). In sleep studies, EEG, EOG, and EMG are commonly used signals to analyze the sleep cycle, which consists of three stages: non-rapid eye movement (NREM) and rapid eye movement (REM). Although signal processing and computational methods are rapidly advancing, there is still a gap between machine learning techniques and fully automated sleep staging, which calls for greater collaboration between academia, research, and industry to manage large datasets for effective sleep analysis.

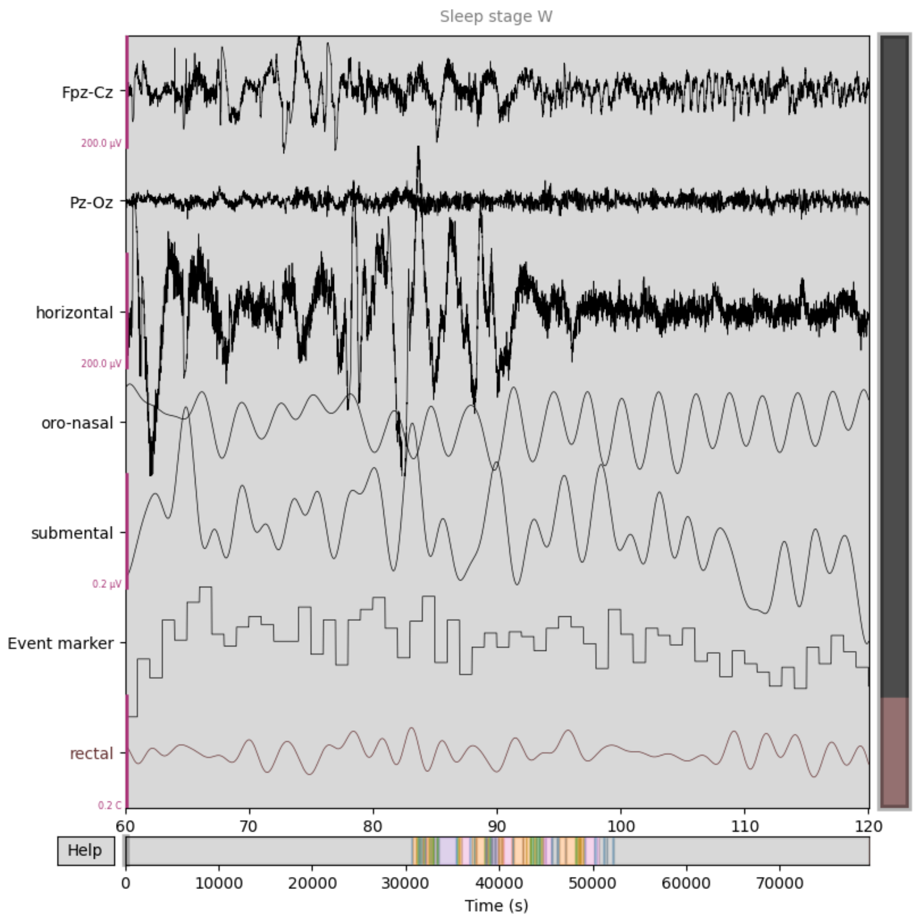

Fig. 2. Bio Signals Chanels

In sleep studies, several important signals are used to classify different sleep stages. The electroencephalogram (EEG) records the brain's electrical activity and helps identify sleep stages based on specific wave frequencies. The electrooculogram (EOG) monitors eye movements, which is especially useful for detecting REM sleep, characterized by rapid eye movements. The electromyogram (EMG) measures muscle activity, assisting in distinguishing between sleep stages, with a particular focus on recognizing the muscle atonia typical of REM sleep. The sleep cycle consists of NREM and REM stages, with NREM further divided into three phases: Stage 1, which marks the transition from wakefulness to light sleep and is characterized by theta waves; Stage 2, which represents deeper sleep with distinctive K-complexes and sleep spindles; and Stage 3, or slow-wave sleep, marked by delta waves. EEG data is typically recorded across multiple channels to capture spatial information from various brain regions. The data is then preprocessed, regularized, and labeled according to standardized methods to facilitate accurate classification. Feature extraction plays a crucial role in this process, where features such as power spectral density, amplitude, frequency of brain waves, and entropy measures are extracted to aid in the classification and analysis of the sleep stages. These features are subsequently used in conjunction with machine learning algorithms to enhance the accuracy of sleep stage detection.

1.1 Related Work

This paper examines the impact of physiological arousals during sleep, particularly their effects on cognitive impairments and heart rate variations. Interruptions to sleep continuity caused by arousals can reduce the quality of rest by impairing cognitive function and destabilizing heart rates. To address this issue, the researchers utilized EEG signals and machine learning algorithms, successfully identifying arousal patterns with a model that achieved an intensity level and sensitivity of 82.68%, specificity of 95.68%, and an AUROC (Area Under the Receiver Operating Characteristic curve) of 96.30%. This model shows strong potential for identifying physiological disruptions during various sleep stages.

The paper also presents new approaches using time and frequency domains to address misclassified data, leveraging ensemble learning strategies. This multi-model machine learning method helps improve reliability by increasing model dependency. Feature extraction for sleep stage identification was conducted according to the American Academy of Sleep Medicine’s guidelines. Additionally, class-balanced random sampling techniques were employed to enhance model performance, resulting in higher accuracy and mean F1 scores. This approach demonstrated robustness in classifying sleep stages and highlighted the potential of ensemble learning in improving sleep analysis.

Another study utilized polysomnographic signals, including EEG, EMG, EOG, and ECG, to classify sleep stages by capturing the interdependencies and nonlinear complexities within the signals. Using a small dataset, the model was able to distinguish between sleep states with 74% accuracy. This study reinforced the importance of using multiple bioelectrical signals to accurately interpret sleep patterns when data is limited.

A separate study focused on automating the traditionally manual process of polysomnography using deep learning models, specifically RCNN. This model was trained to automate sleep stage classification, sleep apnea detection, and limb movement identification, tasks usually performed by human experts. The results demonstrated an accuracy of 87.6% for sleep staging, 88.2% for sleep apnea detection, and 84.7% for limb movement detection. These findings highlight the potential of RCNNs to replicate human intervention, enabling fast and efficient sleep analysis without compromising accuracy.

Technological advancements in sleep stage scoring, driven by adaptive and parallel computing models such as artificial neural networks (ANN), have revolutionized the field. These models enable fast and reliable classification of sleep stages, processing

large datasets at high speeds without sacrificing performance. This integration of automation and computational power offers promising solutions for rapid, accurate insights into patients' sleep patterns and is likely to be the future of sleep analysis.
Another study utilized EEG data and ensemble learning for sleep stage classification, following standardized AASM guidelines for feature extraction. This approach demonstrated consistent performance with strong F1 scores and accuracies across various sleep stages. The importance of adhering to standardized guidelines in sleep research was emphasized, ensuring the clinical relevance and consistency of results.
Lastly, a comparative study analyzed machine learning techniques for feature extraction from bioelectric signals, aiming to bridge the gap between advanced machine learning models and traditional signal processing methods. The study highlighted the need for clinically useful and robust models for automating sleep staging. Despite the promise of machine learning in sleep stage classification, the research concluded that more diverse and accurately developed datasets are necessary for reliable clinical application. It was noted that datasets from a variety of sources are essential for improving model generalization across different populations and sleep conditions.

2 Contribution

In this work, we have proposed a straightforward k-fold cross-validation machine learning stack for classifying EEG signals into five distinct sleep stages. This method is computationally efficient and simple to implement, making it an ideal choice for environments with limited resources. Despite its simplicity, the approach yields promising results, demonstrating that machine learning models can be highly effective for sleep stage classification. The technique leverages basic yet powerful models, achieving impressive classification performance while minimizing computational costs. As shown in the results, our method achieves high accuracy and robust performance, offering an accessible solution for accurate sleep stage classification.

3 Methodology

in this study employed EEG data to classify sleep stages using machine learning techniques. It used raw EEG signals and the hypnograms corresponding to those signals. Data were preprocessed by attaching annotations to the EEG signals, cropping the signals for the relevant sleep cycles, and mapping the sleep stage labels to their corresponding numeric values. The data was then segmented into epochs of 30 seconds. Features from EEG epochs in terms of PSD are calculated to provide the frequency characteristics of interest. Features are therefore aggregated into defined bands, that is, delta, theta, alpha, sigma, and beta. A Random Forest classifier was implemented with a machine learning pipeline to automate the classification process. To validate the model, a 5-fold cross-validation strategy was adopted, ensuring robust evaluation by alternating training and test data splits. Performance metrics, including accuracy, precision, recall, and F1-score, were calculated to evaluate the model's classification performance. Furthermore, a confusion matrix was used to check difference between the predicted and actual labels and was used to illustrate model behavior during different stages of sleep. Finally, heatmaps of the confusion matrix were plotted to demonstrate classification performance both in terms of sample counts and percentages. All these steps collectively allowed for the determination of strengths and limitations in the prediction of sleep stages using EEG data and Random Forests.

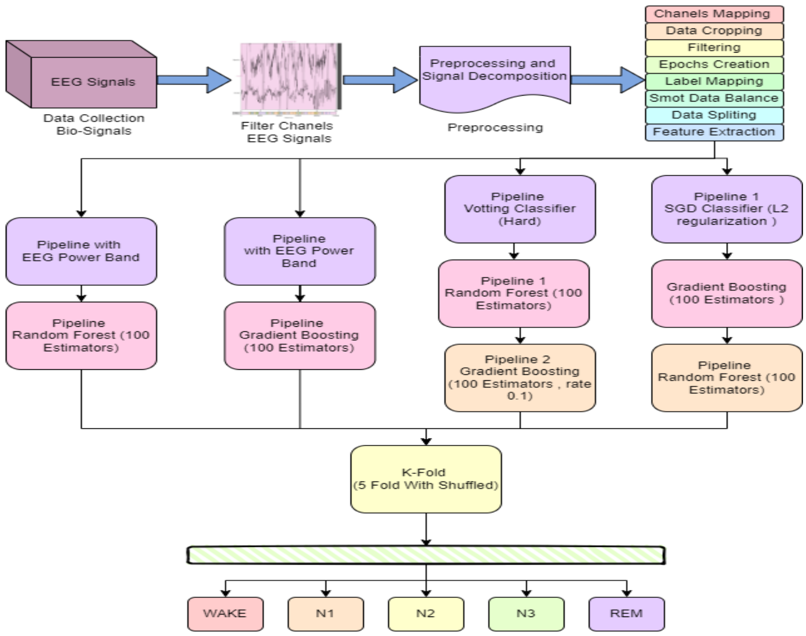

Fig. 3. K-Fold Cross Validation Diagram

3.1 Bio Signals Collection

The collection of bio-signals, such as EEG (electroencephalogram) data, involves recording physiological signals that reflect the electrical activity of the brain. This process requires specialized equipment, including electrodes strategically placed on the scalp to capture neural oscillations across various frequency bands. The signals are recorded over time and often synchronized with corresponding annotations, such as hypnograms, to label states like sleep stages. Proper preprocessing is critical to ensure data quality, including removing noise, aligning signals with annotations, and segmenting recordings into meaningful intervals for analysis. These high-quality recordings serve as the foundation for computational analysis and model training.

We have not collected our own EEG signal data; instead, we have utilized publicly available EEG datasets for this study. These datasets, such as the Sleep-EDF dataset, include pre-recorded EEG signals and their corresponding annotations, which detail various sleep stages. The data was collected from multiple subjects under controlled conditions and made publicly accessible for research purposes. By leveraging these existing datasets, we ensure a standardized approach for analysis while avoiding the complexities of new data collection, focusing on extracting meaningful insights from the provided data to train and evaluate our models.

3.2 Filter EEG Signals Chanels

EEG signal preprocessing is a crucial step for analyzing brain activity, as raw EEG data can contain noise and artifacts that hinder effective interpretation. One of the common methods for filtering EEG channels involves the use of band-pass filters to isolate specific frequency ranges of interest. For example, in the case of sleep stage classification, EEG signals are often filtered into distinct frequency bands such as delta, theta, alpha, beta, and sigma waves. These bands correspond to different states of brain activity and are critical for sleep analysis. In our study, the frequency bands are carefully chosen to cover the typical range of brain activity observed during sleep stages, from 0.5 Hz to 30 Hz, ensuring that only the relevant data is used for further analysis.

Additionally, when filtering EEG channels, it is essential to minimize noise from various sources, such as electrical interference or muscle activity. To achieve this, we apply a power spectral density (PSD) analysis, which provides a frequency-domain representation of the EEG signal. This allows us to compute the power within specific frequency bands for each EEG channel. By averaging the power across these frequency ranges, we obtain features that are more reliable and meaningful for classification tasks. This filtered and processed data is then used to train machine learning models.

3.3 Pre-Processing Signals

In this project, preprocessing begins by loading the raw EEG data from EDF files, which includes both the EEG signals and their associated annotations (hypnograms). The raw EEG data is read, and annotations are assigned to mark different sleep stages (such as Wake, N1, N2, N3, and REM). After that, the annotations are cropped to ensure that only relevant data around the marked sleep stages is used. This involves trimming a portion of the data before and after each annotated segment, ensuring we focus only on the periods of interest. The event markers are then extracted, and the data is segmented into 30-second epochs. These epochs are aligned with the sleep stages so that each segment corresponds to a specific stage of sleep, which is crucial for the classification task.

Next, the EEG data undergoes a transformation to extract meaningful features. One common method is to calculate the power spectral density (PSD) within certain frequency bands (e.g., delta, theta, alpha, sigma, beta). These frequency bands are indicative of different brain activities associated with various sleep stages. By focusing on these bands, the EEG signal is reduced to a more compact and informative representation, capturing essential brain wave activity. The PSD for each frequency band is averaged over the time window of each epoch, and these averages are used as features for training machine learning models. This process ensures that only the most relevant information from the raw EEG data is retained, making it suitable for analysis and classification.

3.4 Machine Learning Pipe-Lines Building

3.4.1 Random Forest

Random Forest is a powerful ensemble learning algorithm that combines multiple decision trees to improve classification accuracy and reduce overfitting by averaging predictions across many trees. Each tree is trained on a random subset of the data, which helps to increase the model's robustness and generalization capabilities.

3.4.2 Gradient Boosting

Gradient Boosting is an ensemble learning technique that builds a model by combining multiple weak learners, typically decision trees, in a sequential manner. In each step, the model tries to correct the errors made by the previous ones by focusing more on the incorrectly predicted data points. The algorithm minimizes the loss function using gradient descent, thus "boosting" the accuracy of the model. It’s particularly effective for both classification and regression tasks, and has advantages such as handling various data types, providing high predictive accuracy, and being less prone to overfitting when tuned properly.

3.4.3 Ensemble Learning 1

The transformed data is then passed to the GradientBoostingClassifier, which uses the ensemble learning technique to build a model by iteratively correcting errors made by previous models. The classifier operates with 100 estimators (trees) and a fixed random state for reproducibility. This setup allows the model to progressively improve its predictions on EEG data, making it well-suited for complex tasks like sleep stage classification.

3.4.4 Ensemble Learning 2

In this three separate machine learning pipelines, each implementing a different classifier for EEG data classification. The first pipeline, sgd_pipe, uses Stochastic Gradient Descent (SGD) as the classifier, with a log loss function and L2 regularization. It first applies the eeg_power_band function to extract features from the EEG signal, then scales the features using StandardScaler to ensure that the model training process is not influenced by the scale of the data. The second pipeline, gb_pipe, employs Gradient Boosting, an ensemble method known for its ability to iteratively improve predictions. Like the previous pipeline, it uses feature extraction and scaling before passing the data to the classifier. The final pipeline, rf_pipe, incorporates a Random Forest classifier, which is also an ensemble model but uses decision trees to make predictions. It applies the same feature extraction and scaling steps but adds the class_weight='balanced' parameter to address class imbalances during training. All pipelines ensure that data preprocessing is consistent, making them suitable for testing various machine learning models on the EEG data.

3.5 K-Fold Cross Validation

K-Fold Cross Validation is a robust method used for evaluating the performance of machine learning models. It works by dividing the dataset into K subsets or "folds". The model is trained on K-1 folds and tested on the remaining fold, and this process is repeated K times, each time with a different fold serving as the test set. The results are then averaged to provide a more reliable estimate of model performance. This technique helps to mitigate overfitting and ensures that the model generalizes well to unseen data. In our methodology, we have used **5-fold cross-validation**, meaning the data is divided into five parts, and the model undergoes five rounds of training and testing. This approach offers a good balance between computational efficiency and model evaluation, making it suitable for our EEG data classification task.

3.6 Classification & Plots

For classification purposes, we have used a **confusion matrix**, which provides a detailed view of the performance of the classification model by showing the true positive, true negative, false positive, and false negative values for each class. It helps in understanding how well the model is distinguishing between different classes and where misclassifications occur. By examining the confusion matrix, we can identify patterns of misclassification, such as whether certain classes are being confused with others. This detailed breakdown is essential for evaluating the model's accuracy, precision, recall, and overall effectiveness in classifying the EEG data into different sleep stages. The confusion matrix serves as a valuable tool for improving model performance by highlighting areas for further refinement.

4 Results & Discussion

4.1 Random Forest

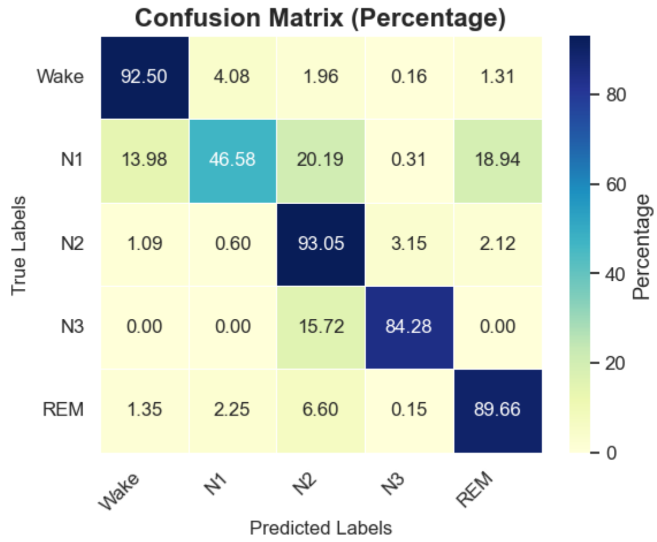

Fig. 4. Random Forest Confusion Matrix

Class	Precision	Recall	F1 Score
Wake	0.88	0.92	0.9
N1	0.75	0.47	0.57
N2	0.9	0.93	0.91
N3	0.87	0.84	0.86
REM	0.85	0.9	0.87

Table 1. Random Forest Results

4.2 Gradient Boosting Classifier

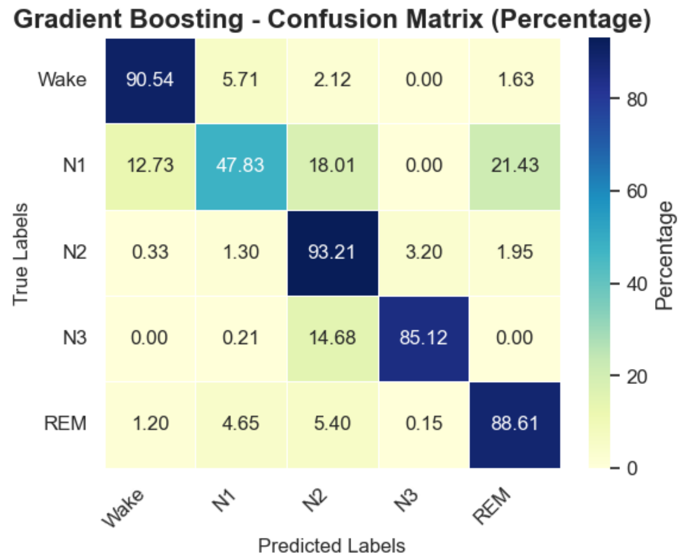

Fig. 5. Gradient Boosting Confusion Matrix

Class	Preci-sion	Recall	F1 Score
Wake	0.91	0.9	0.91
N1	0.63	0.49	0.55
N2	0.91	0.93	0.92
N3	0.87	0.85	0.86
REM	0.84	0.89	0.86

Table 2. Gradient Boosting Results

4.3 Ensemble Learning Stack 1 (G.B. + R.F.)

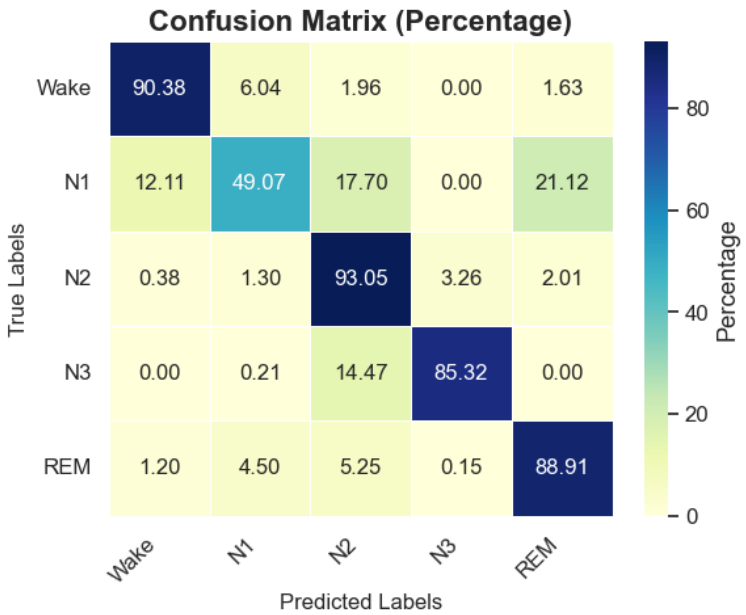

Fig. 6. Ensemble Stack Random Forest & Gradient Boosting Confusion Matrix

Class	Precision	Recall	F1 Score
Wake	0.98	1.0	0.99
N1	0.65	0.39	0.48
N2	0.89	0.93	0.91
N3	0.89	0.8	0.84
REM	0.87	0.84	0.86

Table 3. Ensemble Learning 1 Results

4.4 Ensemble Learning Stack 2 (SGD Pipeline)

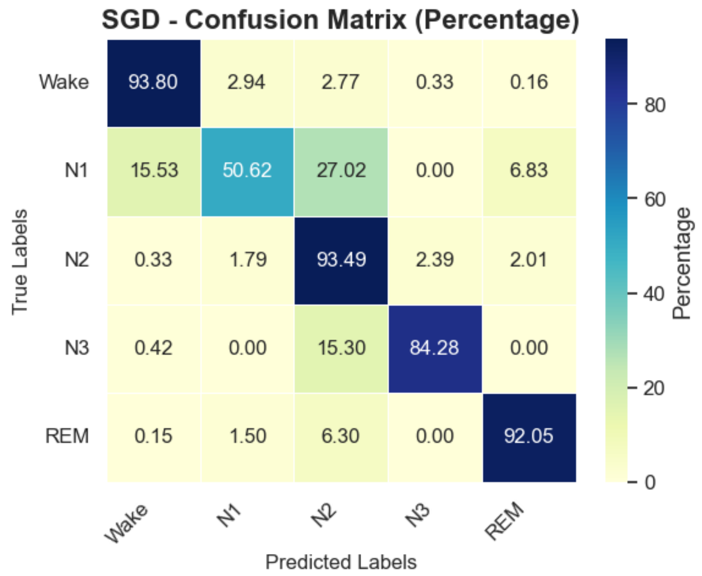

Fig. 7. SGD Classifier Pipe Line Ensemble Learning Confusion Matrix

Class	Preci-sion	Recall	F1 Score
Wake	0.79	0.82	0.8
N1	0.43	0.22	0.29
N2	0.86	0.91	0.88
N3	0.86	0.78	0.82
REM	0.7	0.76	0.73

Table 4. Ensemble Learning 2 Results

4.5 Comparison

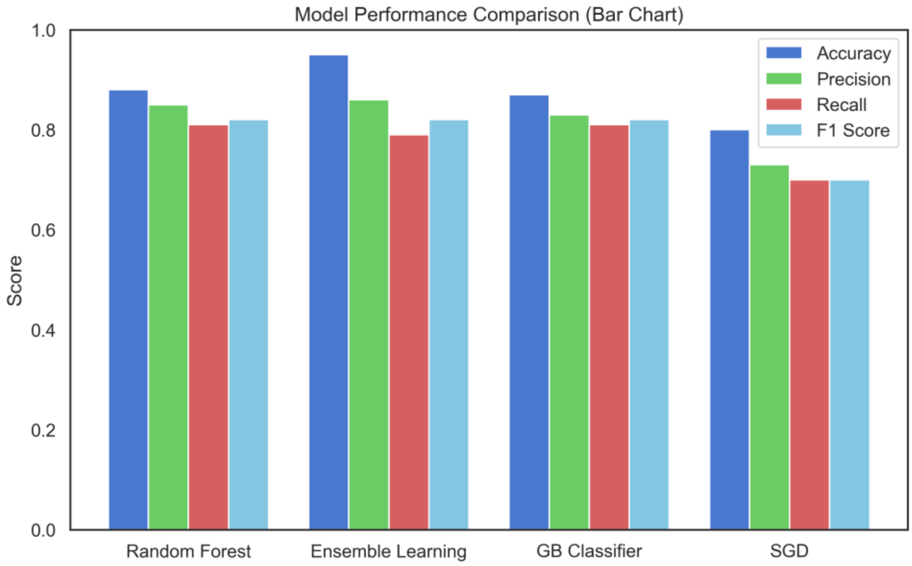

Fig. 8. comparisons Model Bar Chart

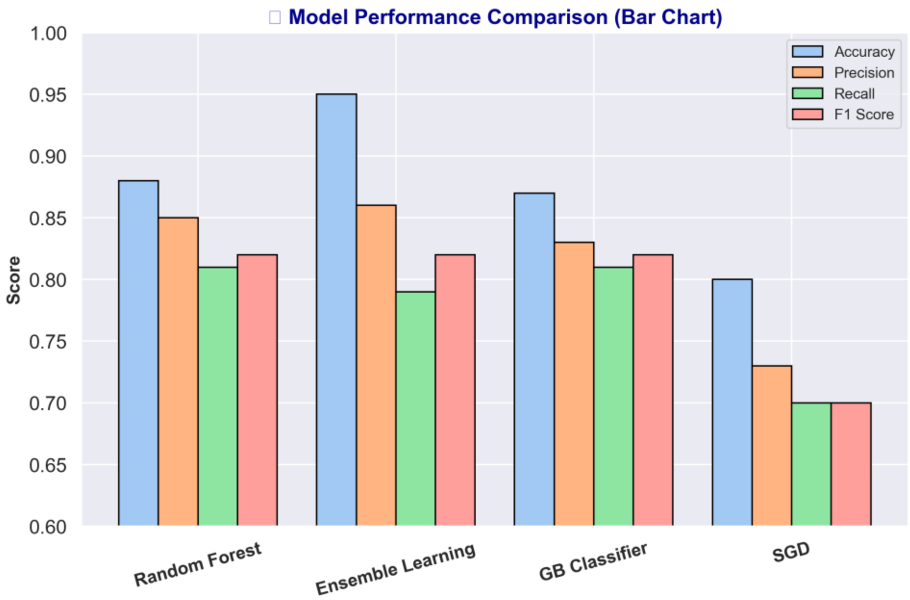

Fig. 9. Model Performance Comparison Bar Chart

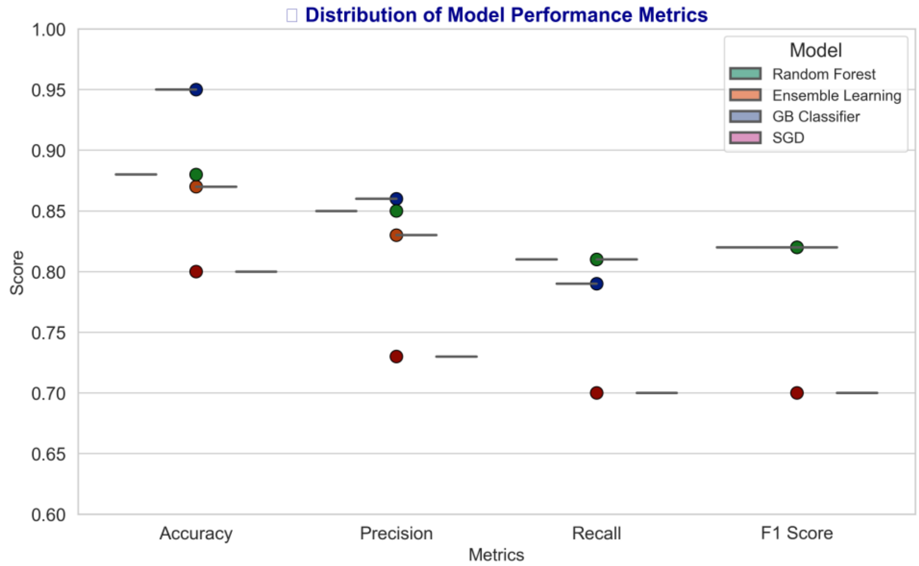

Fig. 10. Distribution Model Performance Metrics

5 Conclusion

The results show that **Ensemble Learning** achieved the highest overall accuracy of 0.95, with exceptional performance in classifying the "Wake" stage (f1-score = 0.99), though it struggled with the "N1" stage (f1-score = 0.48). **Random Forest** also performed well, attaining an accuracy of 0.88, with strong results for the "Wake" and "N2" stages (f1-scores around 0.90), but its performance on "N1" was weaker (f1-score = 0.57). The **GB Classifier** showed similar results to Random Forest, with an accuracy of 0.87. **SGD** had the lowest accuracy of 0.80, particularly underperforming in classifying the "N1" stage. These findings highlight the effectiveness of simple k-fold classification pipelines using machine learning techniques for EEG signal classification, offering a computationally efficient way to achieve high accuracy, particularly in distinguishing critical sleep stages such as "Wake" and "N2."

6 Reference

1. Alsolai, H., Qureshi, S., Iqbal, S. M. Z., Vanichayobon, S., Henesey, L. E., Lindley, C., & Karrila, S. (2022). A Systematic Review of Literature on Automated Sleep Scoring. In IEEE Access (Vol. 10, pp. 79419–79443). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2022.3194145

2. Faust, O., Razaghi, H., Barika, R., Ciaccio, E. J., & Acharya, U. R. (2019). A review of automated sleep stage scoring based on physiological signals for the new millennia. In Computer Methods and Programs in Biomedicine (Vol. 176, pp. 81–91). Elsevier Ireland Ltd. https://doi.org/10.1016/j.cmpb.2019.04.032

3. Tsinalis, O., Matthews, P. M., & Guo, Y. (2016). Automatic Sleep Stage Scoring Using Time-Frequency Analysis and Stacked Sparse Autoencoders. Annals of Biomedical Engineering, 44(5), 1587–1597. https://doi.org/10.1007/s10439-015-1444-y

4. Ronzhina, M., Janoušek, O., Kolářová, J., Nováková, M., Honzík, P., & Provazník, I. (2012). Sleep scoring using artificial neural networks. In Sleep Medicine Reviews (Vol. 16, Issue 3, pp. 251–263). https://doi.org/10.1016/j.smrv.2011.06.003

5. Biswal, S., Sun, H., Goparaju, B., Brandon Westover, M., Sun, J., & Bianchi, M. T. (2018). Expert-level sleep scoring with deep neural networks. Journal of the American Medical Informatics Association, 25(12), 1643–1650. https://doi.org/10.1093/jamia/ocy131

6. Dutt, M., Redhu, S., Goodwin, M., & Omlin, C. W. (2023). SleepXAI: An explainable deep learning approach for multi-class sleep stage identification. Applied Intelligence, 53(13), 16830–16843. https://doi.org/10.1007/s10489-022-04357-8

7. Han, H., Seong, M. J., Hyeon, J., Joo, E., & Oh, J. (2024). Classification and automatic scoring of arousal intensity during sleep stages using machine learning. Scientific Reports, 14(1). https://doi.org/10.1038/s41598-023-50653-9

8. Gaiduk, M., Serrano Alarcón, Á., Seepold, R., & Martínez Madrid, N. (2023). Current status and prospects of automatic sleep stages scoring: Review. In Biomedical Engineering Letters (Vol. 13, Issue 3, pp. 247–272). Springer Verlag. https://doi.org/10.1007/s13534-023-00299-3

9. Mousavi, S., Afghah, F., & Rajendra Acharya, U. (2019). Sleepeegnet: Automated sleep stage scoring with sequence to sequence deep learning approach. PLoS ONE, 14(5). https://doi.org/10.1371/JOURNAL.PONE.0216456

10. Zhao, C., Wu, W., Zhang, H., Zhang, R., Zheng, X., & Kong, X. (2024). Sleep Stage Classification Via Multi-View Based Self-Supervised Contrastive Learning of

EEG. IEEE Journal of Biomedical and Health Informatics, 1–9. https://doi.org/10.1109/JBHI.2024.3432633

11. Chen, Y., Lv, Y., Sun, X., Poluektov, M., Zhang, Y., & Penzel, T. (2024). ESSN: An Efficient Sleep Sequence Network for Automatic Sleep Staging. IEEE Journal of Biomedical and Health Informatics. https://doi.org/10.1109/JBHI.2024.3443340

12. Jiang, C., Xie, W., Zheng, J., Yan, B., Luo, J., & Zhang, J. (2024). MLS-Net: An Automatic Sleep Stage Classifier Utilizing Multimodal Physiological Signals in Mice. Biosensors, 14(8). https://doi.org/10.3390/bios14080406

13. He, M., Tang, M., Meng, L., & Liang, Z. (2024). TBSTSleepNet: Three-branch spectro-temporal bidirectional LSTM based attention model for EEG sleep staging. Biomedical Signal Processing and Control, 97. https://doi.org/10.1016/j.bspc.2024.106695

14. Zhang, W., Zhang, S., Wang, Y., Li, C., Peng, H., & Chen, X. (2024). A CNN-Transformer-ConvLSTM-CRF Hybrid Network for Sleep Stage Classification. IEEE Sensors Journal. https://doi.org/10.1109/JSEN.2024.3434404

15. Ito, A., & Tanaka, T. (2024). SleepSatelightFTC: A Lightweight and Interpretable Deep Learning Model for Single-Channel EEG-Based Sleep Stage Classification. https://doi.org/10.1101/2024.08.02.606301

16. el Hadiri, A., Bahatti, L., el Magri, A., & Lajouad, R. (2024). Sleep stages detection based on analysis and optimisation of non-linear brain signal parameters. Results in Engineering, 23. https://doi.org/10.1016/j.rineng.2024.102664

17. Ma, J., Lin, Q., Jia, Z., & Feng, M. (2024). ST-USleepNet: A Spatial-Temporal Coupling Prominence Network for Multi-Channel Sleep Staging. http://arxiv.org/abs/2408.11884

18. Tsinalis, O., Matthews, P. M., & Guo, Y. (2016). Automatic Sleep Stage Scoring Using Time-Frequency Analysis and Stacked Sparse Autoencoders. Annals of Biomedical Engineering, 44(5), 1587–1597. https://doi.org/10.1007/s10439-015-1444-y

19. Fiorillo, L., Puiatti, A., Papandrea, M., Ratti, P. L., Favaro, P., Roth, C., Bargiotas, P., Bassetti, C. L., & Faraci, F. D. (2019). Automated sleep scoring: A review of the latest approaches. In Sleep Medicine Reviews (Vol. 48). W.B. Saunders Ltd. https://doi.org/10.1016/j.smrv.2019.07.007

20. Chriskos, P., Frantzidis, C. A., Nday, C. M., Gkivogkli, P. T., Bamidis, P. D., & Kourtidou-Papadeli, C. (2021). A review on current trends in automatic sleep staging through bio-signal recordings and future challenges. In Sleep Medicine Reviews (Vol. 55). W.B. Saunders Ltd. https://doi.org/10.1016/j.smrv.2020.101377

21. Satapathy, S.K., Loganathan, D. Automated classification of multi-class sleep stages classification using polysomnography signals: a nine-layer 1D-convolution neural network approach. *Multimed Tools Appl* (2022). https://doi.org/10.1007/s10420-022-13195-2

22. Satapathy, S.K., Loganathan, D. Multimodal multiclass machine learning model for automated sleep staging based on time series data. *SN COMPUT. SCI.* 3, 276 (2022). https://doi.org/10.1007/s42979-022-01156-3

23. Satapathy, S.K., Loganathan, D. Prognosis of automated sleep staging based on two-layer ensemble learning stacking model using single-channel EEG signal. *Soft Comput* 25, 15445–15462 (2021). https://doi.org/10.1007/s00500-021-06218-x

24. Satapathy, S.K., Bhoi, A.K., Loganathan, D., Khandelwal, B., Barsocchi, P. Machine learning with ensemble stacking model for automated sleep staging using dual-channel EEG signal. *Biomed Signal Process Control* 69, 102898 (2021). https://doi.org/10.1016/j.bspc.2021.102898

25. Satapathy, S.K., Loganathan, D., Kondaveeti, H.K., et al. Performance analysis of machine learning algorithms on automated sleep staging feature sets. *CAAI Trans. Intell. Technol.* 6(2), 155–174 (2021). https://doi.org/10.1049/cit2.12042

26. Satapathy, S.K., Loganathan, D. Machine learning approaches with heterogeneous ensemble learning stacking model for automated sleep staging. *Int. J. Comput. Digit. Syst.* (2022). https://doi.org/10.12785/ijcds/100109

27. Satapathy, S.K., Loganathan, D. A study of human sleep stage classification based on dual channels of EEG signal using machine learning techniques. *SN COMPUT. SCI.* 2, 157 (2021). https://doi.org/10.1007/s42979-021-00528-5

28. Satapathy, S.K., Loganathan, D. Automated classification of sleep stages using single-channel EEG: A machine learning-based method. *Int. J. Inf. Retr. Res.* 12(2), 1–19 (2022). https://doi.org/10.4018/IJIRR.299941