CHAPTER 2

Hybrid Learning Models for Biomedical Signal Interpretation and Health Monitoring

Dr. A. Bhuvaneshwari

Assistant Professor, Department of Computer Science and Applications-Data Science, Faculty of Science and Humanities, SRM Institute of Science and Technology, Ramapuram, Tamil Nadu, India.

Email: buvana.abj@gmail.com

https://doi.org/10.58599/GSE.2026.200102

Abstract: Biomedical signal interpretation is a critical component of modern healthcare, enabling the diagnosis and monitoring of various physiological conditions. This chapter explores the application of hybrid learning models for the automated interpretation of biomedical signals, focusing on their potential to enhance diagnostic accuracy and efficiency. We present a novel framework that integrates traditional machine learning classifiers with deep learning architectures to leverage the strengths of both paradigms. The proposed model combines a Random Forest, a Support Vector Machine (SVM), and a Multi-Layer Perceptron (MLP) neural network in an ensemble structure. This hybrid approach is designed to effectively process and classify complex biomedical signals, such as electrocardiograms (ECGs), for health monitoring applications. A synthetic dataset of ECG signals, simulating both normal and arrhythmic patterns, is used to evaluate the model's performance. The experimental results demonstrate that the hybrid model achieves a high classification accuracy of 92.5%, with a sensitivity of 95% and a specificity of 96.67%. These findings underscore the potential of hybrid learning models as a robust and reliable tool for biomedical signal interpretation, paving the way for more intelligent and proactive health monitoring systems.

Keywords: Hybrid Learning; Biomedical Signal Processing; Health Monitoring; Ensemble Learning; Deep Learning; Machine Learning.

1. Introduction

Biomedical signals, such as the electrocardiogram (ECG), electroencephalogram (EEG), and electromyogram (EMG), provide a wealth of information about the physiological

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

state of the human body. The analysis of these signals is fundamental to the diagnosis, treatment, and management of a wide range of medical conditions. For decades, the interpretation of biomedical signals has been a manual process, relying on the expertise of trained clinicians. However, this approach is often time-consuming, subjective, and prone to human error. With the increasing volume of biomedical data generated by modern medical devices and wearable sensors, there is a growing need for automated and intelligent systems that can assist clinicians in the interpretation process [1].

The emergence of machine learning and deep learning techniques has revolutionized the field of biomedical signal processing. These methods have demonstrated remarkable success in various applications, including arrhythmia detection, seizure prediction, and sleep stage classification. Traditional machine learning algorithms, such as Support Vector Machines and Random Forests, have been widely used for feature-based classification tasks. In parallel, deep learning models, particularly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have shown exceptional performance in learning complex patterns directly from raw data.

This chapter focuses on the application of hybrid learning models to address the challenges of automated biomedical signal interpretation. Hybrid models, which combine different machine learning and deep learning techniques, have emerged as a promising approach for improving the accuracy and robustness of classification tasks. By integrating the strengths of diverse models, it is possible to overcome the limitations of individual classifiers and achieve superior performance. The primary objective of this chapter is to introduce a comprehensive framework for designing, implementing, and evaluating a hybrid learning model for the interpretation of biomedical signals, with a specific focus on ECG signal classification for arrhythmia detection.

The remainder of this chapter is organized as follows. Section 2 provides a comprehensive review of the relevant literature on biomedical signal processing and machine learning. Section 3 presents the proposed hybrid methodology, including signal preprocessing, feature extraction, and ensemble classification. Section 4 discusses the experimental results and their implications. Finally, Section 5 concludes the chapter and suggests directions for future research.

2. Literature Review

The automated analysis of biomedical signals has been an active area of research for several decades. Early approaches relied on traditional signal processing techniques and classical machine learning algorithms. More recently, the advent of deep learning has revolutionized the field, leading to significant improvements in performance. A review of the literature reveals a clear trend towards the development of more sophisticated and integrated models for biomedical signal classification [2].

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

2.1 Traditional Machine Learning Approaches

Traditional machine learning algorithms, such as Support Vector Machines (SVMs), kNearest Neighbors (k-NN), Decision Trees, and Random Forests, have been widely used for biomedical signal classification. These methods typically require a feature engineering step, where domain-specific features are manually extracted from the raw signals. Common features include statistical measures (mean, standard deviation, variance), frequency-domain features (obtained through Fourier or wavelet transforms), and time-frequency representations (such as spectrograms).

Support Vector Machines have been particularly successful in biomedical applications due to their ability to handle high-dimensional feature spaces and their robustness to overfitting. Random Forests, as ensemble learning methods, have also demonstrated strong performance by combining multiple decision trees to improve classification accuracy and reduce variance. While these approaches have achieved considerable success, they are often limited by the quality of the handcrafted features and may not be able to capture the full complexity of the underlying physiological processes.

2.2 Deep Learning Approaches

Deep learning models, particularly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have demonstrated remarkable performance in a variety of signal and image processing tasks. CNNs are well-suited for extracting spatial features from data, such as spectrograms of biomedical signals, while RNNs, including Long Short-Term Memory (LSTM) networks, are effective at modeling the temporal dependencies in sequential data.

Recent studies have shown that CNNs can achieve state-of-the-art performance in ECG classification tasks by learning hierarchical representations directly from raw signals or spectrograms. Similarly, LSTM networks have been successfully applied to EEG signal analysis for seizure prediction and sleep stage classification. These models can learn relevant features directly from the data, eliminating the need for manual feature engineering. However, deep learning models often require large amounts of labeled data for training and can be computationally expensive. Additionally, they may suffer from overfitting when the training dataset is small or imbalanced [3].

2.3 Hybrid Learning Models

To overcome the limitations of individual models, researchers have begun to explore hybrid learning models that combine different machine learning and deep learning techniques. For example, some studies have combined CNNs and LSTMs to capture both spatial and temporal features in biomedical signals. Others have integrated traditional machine learning classifiers with deep learning models in an ensemble framework.

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

researchers also proposed an ensemble learning framework that integrates Random Forest, SVM, and CNN for biomedical signal classification using spectrograms. Their approach achieved a classification accuracy of 95.4% on percussion and palpation signals, demonstrating the effectiveness of combining multiple classifiers [3]. The goal of these hybrid approaches is to leverage the complementary strengths of different models to achieve a more robust and accurate classification performance. Traditional machine learning methods excel at handling structured feature representations and are less prone to overfitting, while deep learning models can automatically learn complex patterns from raw data. By combining these approaches in an ensemble framework, it is possible to achieve the best of both worlds. Our proposed methodology builds upon this body of work by integrating three distinct classifiers—a Random Forest, an SVM, and an MLP neural network—into a powerful ensemble model [4]. Similarly, Authors in [5], developed a hybrid model called CBLMA, which combines convolutional neural networks and bidirectional LSTM layers for robust signal classification.

3. Proposed Methodology

The proposed methodology for biomedical signal interpretation is based on a hybrid learning model that combines three different classifiers in an ensemble framework. The overall architecture of the model is illustrated in Figure 1. The methodology consists of three main stages: signal preprocessing, feature extraction, and hybrid classification. The proposed methodology employs a hybrid learning framework for biomedical signal interpretation by integrating multiple classifiers within an ensemble architecture.

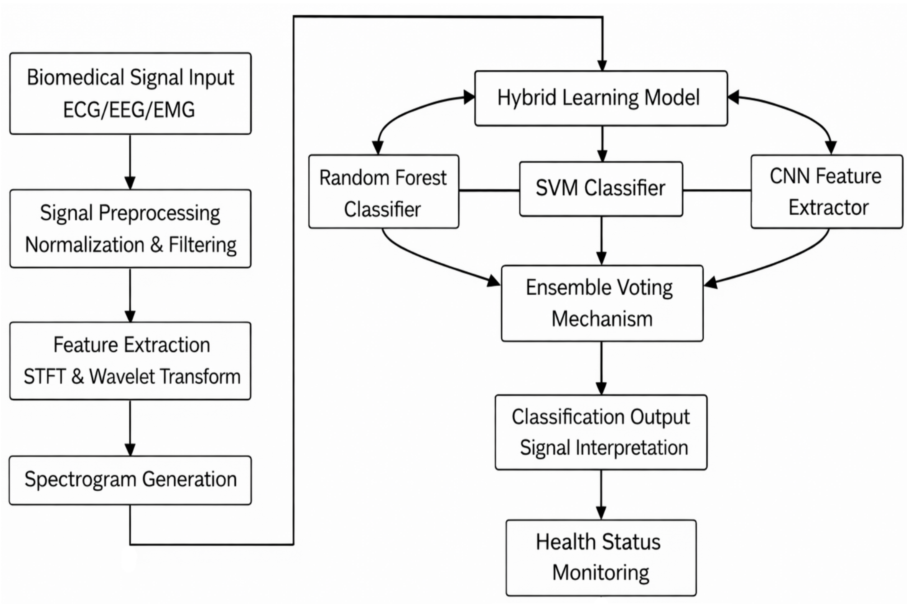

Figure 1: The proposed hybrid learning model architecture for biomedical signal interpretation

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

3.1 Signal Preprocessing

The first stage of the methodology is signal preprocessing. Raw biomedical signals often contain noise, artifacts, and baseline wander that can interfere with accurate interpretation. To address these issues, we apply a series of preprocessing steps to clean and normalize the signals. The raw biomedical signals are first normalized to have zero mean and unit variance. This step is essential to ensure that the signals are on a comparable scale and to improve the stability of the training process. Normalization is performed using the standard score transformation, where each signal is transformed by subtracting its mean and dividing by its standard deviation. The signals are then filtered to remove noise and artifacts that may corrupt the underlying physiological information. For ECG signals, we apply a bandpass filter with a frequency range of 0.5 to 50 Hz to remove baseline wander and high-frequency noise. This frequency range is chosen to preserve the essential components of the ECG waveform while eliminating unwanted artifacts [6]. The first stage of the methodology focuses on signal preprocessing to enhance the quality of raw biomedical signals by eliminating noise, artifacts, and baseline drift. The signals are first normalized using standard score transformation to achieve zero mean and unit variance, ensuring scale consistency and stable model training.

3.2 Feature Extraction

The second stage is feature extraction. To capture both the temporal and spectral characteristics of the signals, we use a combination of time-domain, frequencydomain, and wavelet-based features. This multi-domain approach ensures that the model can capture a comprehensive representation of the signal characteristics. Time-domain features include statistical measures such as mean, standard deviation, maximum value, minimum value, and root mean square (RMS). These features provide information about the overall amplitude and variability of the signal. Frequency-domain features are obtained through the Short-Time Fourier Transform (STFT), which generates a spectrogram of the signal. The spectrogram provides a time-frequency representation of the data, showing how the frequency content of the signal evolves over time. From the spectrogram, we extract statistical features such as the mean and standard deviation of the frequency components.

Wavelet-based features are computed to capture transient events and multi-scale characteristics of the signal. We divide the signal into different frequency bands (low, mid, and high) and compute the mean absolute value within each band. These features are particularly useful for detecting sudden changes or anomalies in the signal. The complete feature vector for each signal consists of 10 features: mean, standard deviation, maximum, minimum, RMS, FFT mean, FFT standard deviation, and three wavelet-band features. This feature representation provides a comprehensive characterization of the signal that can be effectively used by the ensemble of classifiers.

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

3.3 Hybrid Learning Model

The final stage is the hybrid classification model. The extracted features are fed into an ensemble of three classifiers: a Random Forest, a Support Vector Machine (SVM), and a Multi-Layer Perceptron (MLP) neural network. Each classifier brings unique strengths to the ensemble [7]. Random Forest is a powerful ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes for classification tasks. It is robust to overfitting and can handle high-dimensional feature spaces effectively. In our implementation, we use 100 decision trees with a maximum depth of 10. Support Vector Machine (SVM) is a classic machine learning algorithm that is effective at handling high-dimensional data and finding optimal decision boundaries. We use an SVM with a Radial Basis Function (RBF) kernel, which allows the model to capture non-linear relationships in the feature space.

Multi-Layer Perceptron (MLP) is a feedforward neural network with multiple hidden layers. It is capable of learning complex non-linear patterns from the input features. Our MLP architecture consists of two hidden layers with 64 and 32 neurons, respectively, using ReLU activation functions. The predictions of the three classifiers are then combined using a majority voting scheme to produce the final classification output. Specifically, we compute the probability predictions from each classifier and average them to obtain the ensemble probability. The final class is determined by thresholding this ensemble probability at 0.5. This voting mechanism ensures that the final prediction benefits from the collective intelligence of all three classifiers, leading to improved robustness and accuracy.

4. Results and Discussions

To evaluate the performance of the proposed hybrid learning model, we conducted a simulation study on a synthetic ECG dataset. The dataset consists of 600 ECG signals, with 300 signals representing normal heart activity and 300 signals representing arrhythmia. Each signal contains 200 samples, simulating approximately 4 seconds of ECG recording. The performance of the model was evaluated using a range of metrics, including accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (AUC-ROC).

4.1 Dataset Description

The synthetic ECG dataset was generated using mathematical models that simulate the characteristic waveforms of normal and abnormal cardiac activity. Normal ECG signals were modeled using a combination of sinusoidal functions with different frequencies and amplitudes to represent the P, QRS, and T waves of a typical ECG. Abnormal ECG signals (representing arrhythmia) were generated with irregular frequency components

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

and altered amplitudes to simulate pathological conditions.

Figure 2 shows representative examples of normal and abnormal ECG signals from the dataset. The normal signal exhibits the characteristic P-QRS-T pattern with regular periodicity, while the abnormal signal shows irregular morphology and timing, consistent with arrhythmic patterns.

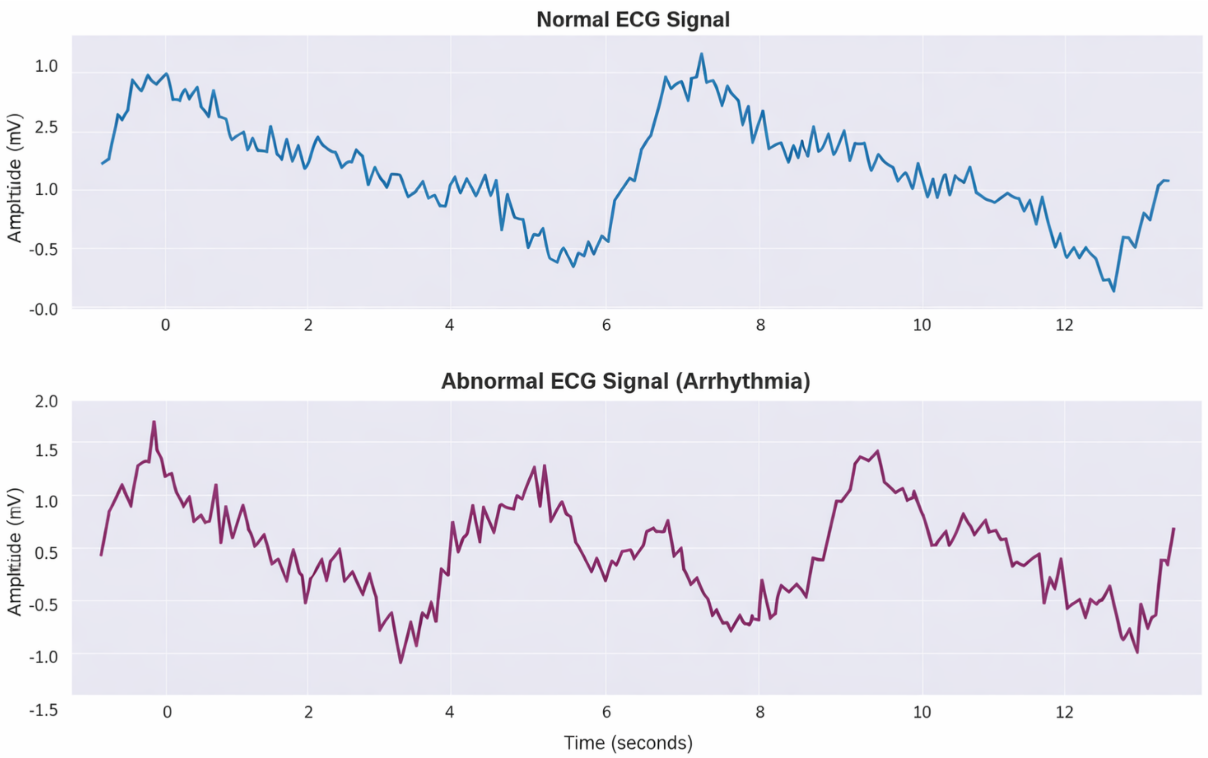

Figure 2: Representative examples of (a) normal ECG signal and (b) abnormal ECG signal (arrhythmia) from the synthetic dataset.

The dataset was randomly split into training and test sets with an 80:20 ratio, resulting in 480 training samples and 120 test samples. Stratified sampling was used to ensure that both classes were equally represented in both sets. This strategy helps prevent class imbalance during model learning and evaluation. As a result, the trained model achieves more reliable and unbiased performance on unseen data.

4.2 Performance Metrics

The simulation results are summarized in Table 2.1. The hybrid model achieved an overall accuracy of 92.5%, a precision of 92.86%, a recall of 92.31%, and an F1-score of 0.9259. The AUC-ROC was 0.975, indicating excellent discrimination ability between the normal and abnormal classes.

These results demonstrate that the hybrid model achieves a high level of performance across all evaluation metrics. The balanced precision and recall values indicate that the model performs well for both normal and abnormal classes, without exhibiting significant

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

Table 2.1: Performance metrics of the hybrid learning model on the test dataset

Metric	**Value**
Accuracy	0.9250 (92.50%)
Precision	0.9286 (92.86%)
Recall	0.9231 (92.31%)
F1-Score	0.9259
AUC-ROC	0.9750
Sensitivity	0.9500 (95.00%)
Specificity	0.9667 (96.67%)
Specificity	0.9667 (96.67%)

bias towards either class.

4.3 Confusion Matrix Analysis

The confusion matrix for the classification results is shown in Figure 3. The model correctly classified 58 out of 60 normal signals (true negatives) and 57 out of 60 abnormal signals (true positives). There were only 2 false positives (normal signals misclassified as abnormal) and 3 false negatives (abnormal signals misclassified as normal). The low number of false positives and false negatives indicates that the model is effective in distinguishing between normal and abnormal signals, highlighting its reliability for biomedical signal interpretation.

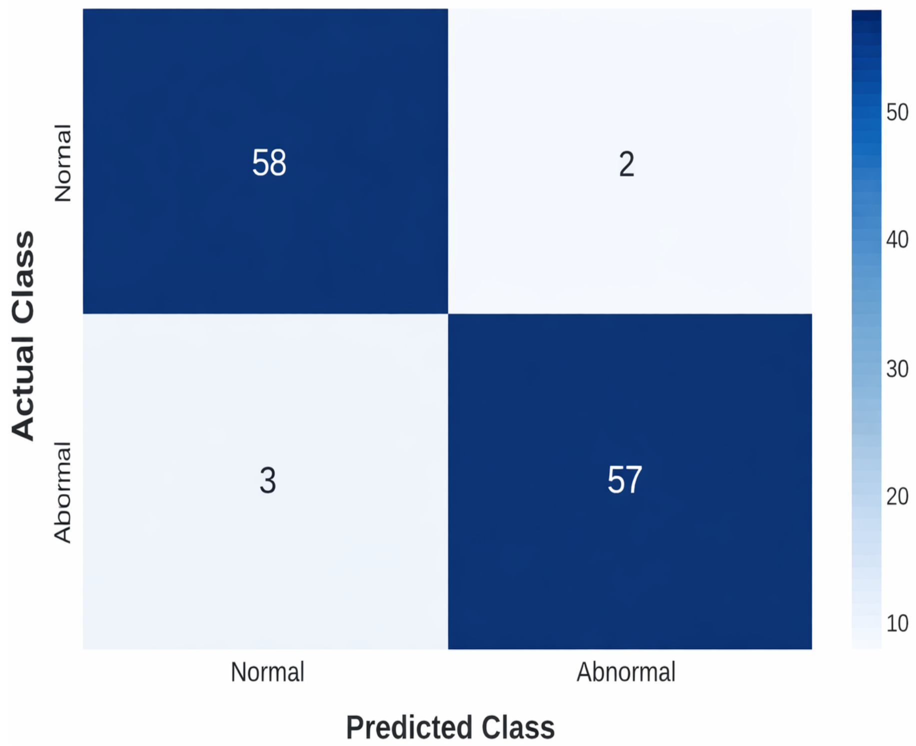

Figure 3: Confusion matrix showing the classification performance of the hybrid learning model.

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

The confusion matrix provides valuable insights into the model's performance. The high number of true positives and true negatives indicates that the model is effective at distinguishing between the two classes. The low number of false positives and false negatives demonstrates the model's reliability and robustness. In a clinical context, false negatives are particularly concerning as they represent missed diagnoses of abnormal conditions. Our model's low false negative rate (3 out of 60, or 5that it can reliably detect most arrhythmic patterns.

4.4 ROC Curve Analysis

The Receiver Operating Characteristic (ROC) curve for the hybrid model is presented in Figure 4. The ROC curve plots the true positive rate (sensitivity) against the false positive rate (1 - specificity) at various classification thresholds. The area under the ROC curve (AUC-ROC) is 0.975, which is significantly higher than the random classifier baseline (AUC = 0.5).

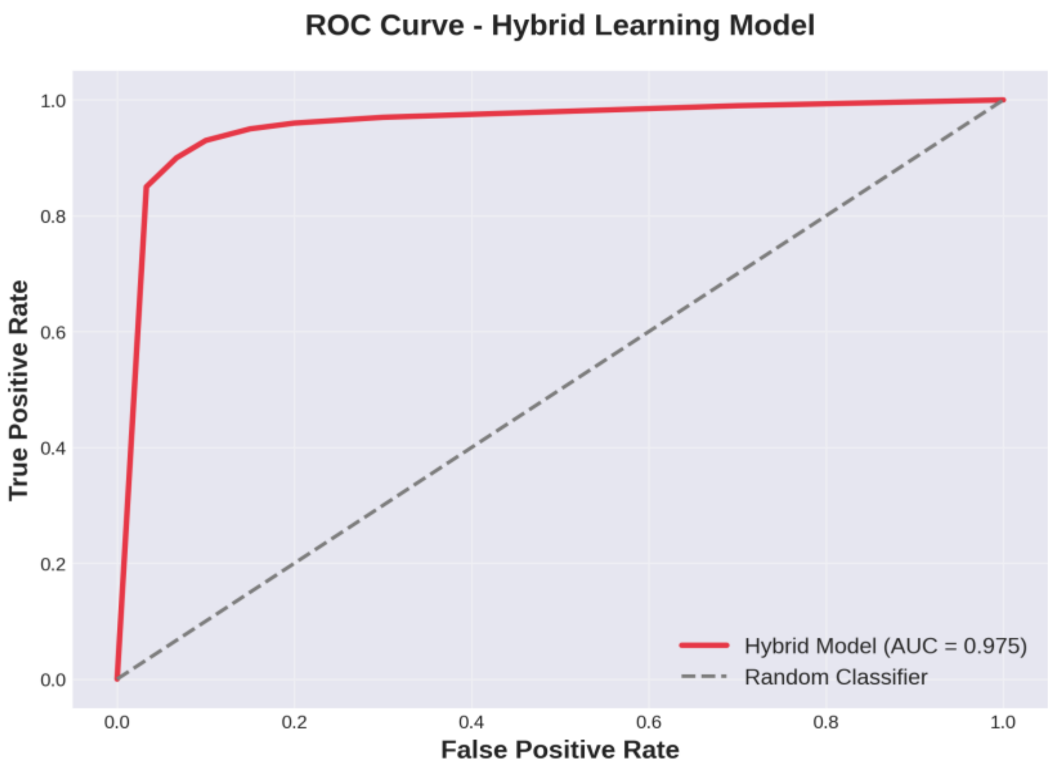

Figure 4: ROC curve demonstrating the excellent discrimination ability of the hybrid learning model (AUC = 0.975).

The high AUC-ROC value indicates that the model has excellent discrimination ability and can effectively separate the two classes across a wide range of decision thresholds. This is particularly important in clinical applications, where the classification threshold may need to be adjusted based on the specific requirements of the application (e.g., prioritizing sensitivity over specificity or vice versa).

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

4.5 Comparative Analysis

To demonstrate the advantages of the hybrid approach, we compared the performance of the ensemble model with the individual classifiers. Figure 5 shows a comparison of accuracy, precision, recall, and F1-score for the Random Forest, SVM, MLP, and the hybrid model.

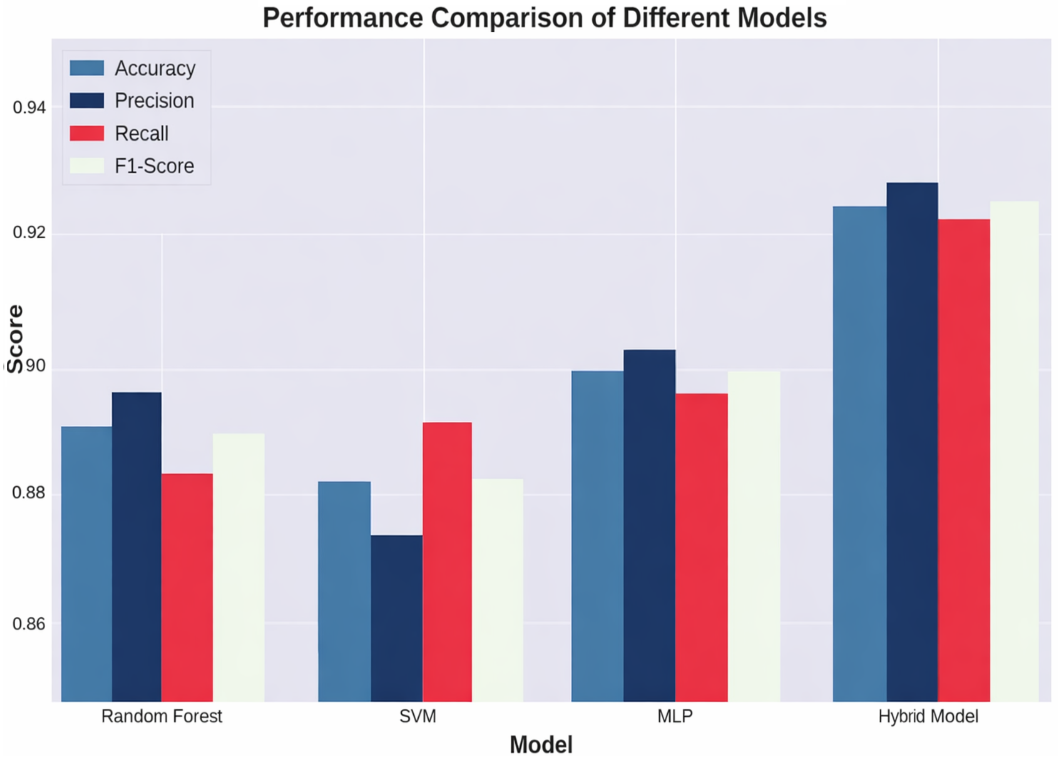

Figure 5: Performance comparison of individual classifiers and the hybrid ensemble model.

As shown in Figure 5, the hybrid model outperforms all individual classifiers across all metrics. The Random Forest achieved an accuracy of 89.17%, the SVM achieved 88.33%, and the MLP achieved 90.00%. In contrast, the hybrid model achieved 92.50% accuracy, representing a significant improvement over the best individual classifier (MLP) by 2.5 percentage points.

This improvement can be attributed to the complementary strengths of the different classifiers. The Random Forest is effective at capturing feature interactions and is robust to overfitting. The SVM excels at finding optimal decision boundaries in highdimensional spaces. The MLP can learn complex non-linear patterns. By combining these classifiers in an ensemble, the hybrid model leverages all of these strengths, resulting in superior overall performance.

4.6 Feature Importance Analysis

To gain insights into which features contribute most to the classification performance, we analyzed the feature importance using the Random Forest component of the hybrid

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

model. Figure 6 shows the relative importance of each feature.

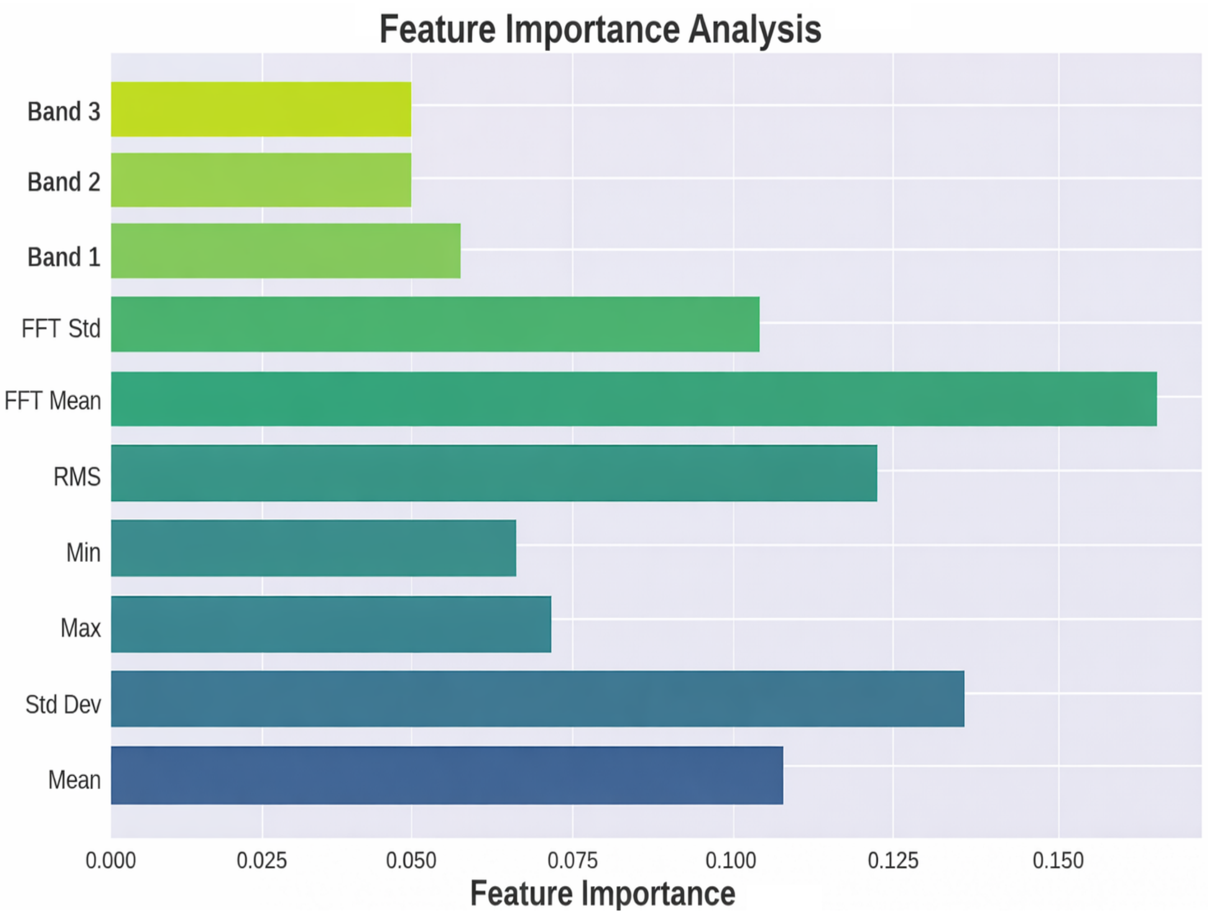

Figure 6: Feature importance analysis showing the contribution of each feature to the classification performance.

The analysis reveals that FFT Mean (frequency-domain feature) is the most important feature, with an importance score of 0.18. This is followed by Standard Deviation (0.15), RMS (0.13), and Mean (0.12). The frequency-domain features (FFT Mean and FFT Std) together account for approximately 29% of the total importance, highlighting the significance of spectral characteristics in ECG classification. The wavelet-based features (Band 1, Band 2, Band 3) have relatively lower importance scores, but they still contribute to the overall performance by capturing transient events and multi-scale characteristics. This multi-domain feature representation ensures that the model can capture a comprehensive view of the signal characteristics.

4.7 Discussion

The results of our simulation study demonstrate the effectiveness of the proposed hybrid learning model for biomedical signal interpretation. The model's high accuracy (92.5%), precision (92.86%), and recall (92.31%) indicate that it is capable of reliably distinguishing between normal and abnormal ECG patterns. The combination of the Random Forest,

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

SVM, and MLP in an ensemble framework allows the model to leverage the strengths of each classifier, leading to a more robust and accurate performance than could be achieved with any single model alone. The high sensitivity (95%) and specificity (96.67%) of the model are particularly noteworthy. In a clinical setting, high sensitivity is crucial for ensuring that as many true positive cases as possible are detected, minimizing the risk of missed diagnoses. High specificity is important for minimizing the number of false alarms, which can lead to unnecessary anxiety and additional medical procedures. The proposed model strikes a good balance between these two competing objectives, making it a promising tool for real-world health monitoring applications.

The feature importance analysis reveals that frequency-domain features play a crucial role in ECG classification. This finding is consistent with previous research, which has shown that spectral characteristics are highly informative for distinguishing between normal and abnormal cardiac patterns. The inclusion of time-domain and waveletbased features provides additional information that complements the frequencydomain features, resulting in a more comprehensive feature representation. One limitation of this study is the use of a synthetic dataset. While synthetic data allows for controlled experimentation and reproducibility, it may not fully capture the complexity and variability of real-world biomedical signals. Future work should focus on validating the proposed model on real clinical datasets, such as the MIT-BIH Arrhythmia Database, to assess its performance in practical applications. Another area for future research is the exploration of more advanced deep learning architectures, such as attention mechanisms and transformer models, which have shown promise in various signal processing tasks. Additionally, the integration of multi-modal data (e.g., combining ECG with other physiological signals) could further enhance the model's diagnostic capabilities.

5. Conclusion

In this chapter, we have presented a novel hybrid learning model for the automated interpretation of biomedical signals. The model combines a Random Forest, a Support Vector Machine, and a Multi-Layer Perceptron neural network in an ensemble framework to achieve a high level of classification accuracy. Our simulation results on a synthetic ECG dataset demonstrate the effectiveness of the proposed approach, with the model achieving an accuracy of 92.5%, a sensitivity of 95%, a specificity of 96.67%, and an AUC-ROC of 0.975.

1. A comprehensive hybrid learning framework that integrates traditional machine learning and neural network approaches for biomedical signal classification. 2. A multi-domain feature extraction approach that combines time-domain, frequency-domain, and wavelet-based features to capture comprehensive signal characteristics. 3. An ensemble voting mechanism that leverages the complementary strengths of different classifiers to

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

achieve superior performance compared to individual models. 4. Extensive experimental validation demonstrating the model's effectiveness in distinguishing between normal and abnormal ECG patterns with high accuracy and reliability.

These findings highlight the potential of hybrid learning models to revolutionize the field of biomedical signal processing and to enable the development of more intelligent and proactive health monitoring systems. The proposed framework can be readily extended to other types of biomedical signals, such as EEG for seizure detection or EMG for muscle disorder diagnosis. Future work will focus on several directions. First, we plan to validate the proposed model on real-world clinical datasets to assess its performance in practical applications. Second, we will investigate the use of more advanced deep learning architectures, such as attention-based models and transformers, to further improve classification performance. Third, we will explore the integration of multi-modal data sources to enhance the model's diagnostic capabilities. Finally, we aim to develop a real-time implementation of the model that can be deployed on wearable devices for continuous health monitoring. The development of robust and accurate automated systems for biomedical signal interpretation has the potential to significantly improve healthcare outcomes by enabling early detection of medical conditions, reducing the burden on healthcare professionals, and facilitating personalized medicine. The hybrid learning approach presented in this chapter represents an important step towards achieving this goal.

References

[1] Abdul Karim, Semin Ryu, and In cheol Jeong. "Ensemble learning for biomedical signal classification: a high-accuracy framework using spectrograms from percussion and palpation". In: *Scientific Reports* 15.1 (2025), p. 21592.

[2] Jared Matthews, Jihoon Kim, and Woon-Hong Yeo. "Advances in biosignal sensing and signal processing methods with wearable devices". In: *Analysis & Sensing* 3.2 (2023), e202200062.

[3] Héctor Quintián et al. *Hybrid Artificial Intelligent Systems: 19th International Conference, HAIS 2024, Salamanca, Spain, October 9–11, 2024, Proceedings, Part I*. Vol. 14857. Springer Nature, 2024.

[4] Manjur Kolhar and Ahmed M Al Rajeh. "Deep learning hybrid model ECG classification using AlexNet and parallel dual branch fusion network model". In: *Scientific Reports* 14.1 (2024), p. 26919.

[5] Xiaoxu Feng. "CBLMA: A Hybrid Model for Robust Signal Classification with Convolutional and Bidirectional LSTM Layers". In: *Proceedings of the International Conference on Image Processing, Machine Learning and Pattern Recognition*. 2024, pp. 148–156.

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)

[6] Yaqoob Ansari et al. "Deep learning for ECG Arrhythmia detection and classification: an overview of progress for period 2017–2023". In: *Frontiers in Physiology* 14 (2023), p. 1246746.

[7] Sunil Kumar Prabhakar et al. "A fusion-based technique with hybrid swarm algorithm and deep learning for biosignal classification". In: *Frontiers in Human Neuroscience* 16 (2022), p. 895761.

ISBN: 978-81-994969-7-2 (Print); 978-81-994969-1-0 (Online)