

International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





AI-Based Healthcare System for Detecting Sleep Apnea

Ms.K.Muthu Lakshmi, Aneboina Hema Krishna, Batchu Sanjay Raj, Bolishetti Rahul,

Parthireddy Pranay

Dept. of Computer Science and Engineering, Bharath Institute of Higher Education and Research (BIHER),
Chennai, Tamil Nadu, India

Dept. of Computer Science and Engineering, Bharath Institute of Higher Education and Research (BIHER),
Chennai, Tamil Nadu, India

Dept. of Computer Science and Engineering, Bharath Institute of Higher Education and Research (BIHER),
Chennai, Tamil Nadu, India

Dept. of Computer Science and Engineering, Bharath Institute of Higher Education and Research (BIHER),
Chennai, Tamil Nadu, India

Dept. of Computer Science and Engineering, Bharath Institute of Higher Education and Research (BIHER),
Chennai, Tamil Nadu, India

ABSTRACT: Sleep Apnea is a common but often undiagnosed sleep disorder in which a person's breathing repeatedly stops and starts during sleep, leading to serious health problems such as heart disease, high blood pressure, and extreme fatigue. The standard method used for diagnosis, called polysomnography (PSG), is expensive, time-consuming, and requires patients to stay overnight in a hospital, making it difficult for regular and large-scale use. To address these limitations, this paper proposes an AI-based healthcare system that detects obstructive sleep apnea (OSA) using electrocardiogram (ECG) signals. In this system, raw ECG data is first cleaned using filtering techniques to remove noise and unwanted disturbances, then divided into smaller segments for better analysis. Important features related to heart activity, known as heart rate variability (HRV), are extracted from both time and frequency domains. These features are then used to train multiple machine learning models, including support vector machine (SVM), random forest (RF), and logistic regression. To improve accuracy and reliability, these models are combined using a stacking ensemble method with a meta-classifier, which produces the final prediction. The system classifies the ECG data into apnea or non-apnea conditions with high accuracy. Overall, the proposed solution is low-cost, non-invasive, and suitable for real-time and home-based monitoring, making it highly useful for early detection and effective management of sleep apnea.

I. INTRODUCTION

Sleep apnea is a sleep disorder in which breathing repeatedly stops and starts, reducing oxygen levels and causing fatigue, mood issues, and serious risks like cardiovascular disease. It includes three types: Obstructive Sleep Apnea (OSA), the most common form caused by airway blockage due to relaxed throat muscles; Central Sleep Apnea (CSA), where the brain fails to send proper signals to breathing muscles; and Mixed Sleep Apnea (MSA), a combination of both. Around 84% of cases are OSA. The standard diagnosis method, polysomnography (PSG), is accurate but costly and requires continuous clinical monitoring, creating a need for simpler alternatives. Electrocardiogram (ECG) signals offer a non-invasive solution by capturing heart rate and respiratory variations during sleep, where apnea episodes typically last 10–20 seconds and cause irregular heart rhythms. An ECG records the heart's electrical activity through waveform components such as the P wave (atrial contraction), QRS complex (ventricular contraction), and T wave (ventricular recovery), along with intervals like P-R, R-R, S-T, and Q-T, each reflecting specific cardiac functions and



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

durations. These variations in ECG signals help in identifying abnormal patterns associated with sleep apnea, making ECG-based analysis a practical and efficient alternative for detection.

II. RELATED WORK

Sleep apnea detection has traditionally relied on Polysomnography (PSG), which is accurate but expensive and inconvenient for continuous use. To address this, researchers use Electrocardiogram (ECG) signals, which are non-invasive and suitable for wearable devices. By analyzing Heart Rate Variability (HRV), important patterns related to breathing irregularities can be captured. Machine learning models like Support Vector Machine (SVM), Random Forest (RF), Decision Trees, and Logistic Regression have shown good results, while deep learning methods such as Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and Long Short-Term Memory (LSTM) improve pattern recognition but require more data and computation. To enhance performance, stacking ensemble methods combine multiple models to achieve better accuracy and robustness. Overall, the proposed system uses ECG-based HRV features with ensemble learning to provide a cost-effective, accurate, and scalable solution for sleep apnea detection.

III. PROPOSED WORK

Sleep apnea diagnosis, a prevalent sleep disorder characterized by repeated episodes of suspended breathing while asleep, has traditionally relied on polysomnography—a cumbersome and time-consuming diagnostic test. As wearable ECG technology becomes more widely available, there is a demand for more scalable and automated methods that make use of minimal physiological signals. This work addresses this gap by introducing a machine learning and deep learning-based system for detecting sleep apnea events using only single-lead ECG signals. In such a system under proposal, the ECG signals are first preprocessed to remove the noise through digital filters and divided into windows of one minute. There is a feature extraction pipeline that is effective in extracting the time-domain and frequency-domain heart rate variability (HRV) parameters that are discriminative features for detecting apnea. The features are representative of minor changes in the cardiac rhythm that translate to respiratory disturbance. For better prediction performance, a stacking ensemble architecture is employed. Eight top-performing machine learning and deep learning models are employed as base classifiers, whose outputs are combined and passed through a logistic regression-based meta-model. Such an architecture takes advantage of the diversity of various algorithms and exhibits improved generalizability across subjects. The system presents a new paradigm by combining conventional HRV features and ensemble learning for efficient ECG-based sleep apnea classification.

A. Lightweight and Wearable-Compatible:

The strategy utilizes solely single-lead ECG signals and is therefore easy to integrate in wearable health monitoring equipment. It allows real-time, non-invasive screening of sleep apnea throughout the day in daily environments without the need for advanced clinical equipment.

B. Advantages of Proposed Solution:

Lightweight and Wearable-Compatible:

The strategy utilizes solely single-lead ECG signals and is therefore easy to integrate into wearable health monitoring devices. It enables real-time, non-invasive sleep apnea screening in everyday environments without requiring complex clinical setups.

High Accuracy through Stacking Ensemble:

By combining eight machine learning and deep learning models using a logistic regression meta-model, the system achieves higher predictive accuracy. The ensemble approach minimizes the weaknesses of individual models and enhances overall reliability.

C. High Accuracy through Stacking Ensemble

Through an ensemble of eight various machine learning and deep learning models in a logistic regression meta-model, the system is rendered more precisely predictive. Ensemble may minimize single weak points of the model and make it more generally applicable on a case-by-case basis. 3. Rich Feature Representation 10 Both the time- and frequency-domain HRV features retrieved enable the system to consider entire cardiac dynamics. The features make the model better in the sense that it would be capable of distinguishing apnea from normal sleep periods effectively.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

D.HRV Feature Extraction and Ensemble Learning Approach:The system extracts important HRV features such as RR intervals, mean heart rate, standard deviation, and LF/HF ratio to capture heart activity changes caused by apnea. These features are analyzed using multiple machine learning models that learn both simple and complex patterns. To improve accuracy, a stacking ensemble method is used, where several base models make predictions that are combined using Logistic Regression as a meta-classifier, resulting in better performance, reliability, and overall prediction accuracy..

IV. DESIGN AND METHODOLOGY

The proposed AI-based Sleep Apnea Detection System follows a structured and modular pipeline consisting of five major stages: (A) Data Acquisition, (B) Signal Preprocessing and Feature Extraction, (C) Machine Learning Classification, (D) Ensemble Learning and Prediction, and (E) Result Visualization. This systematic approach ensures accurate detection, robustness, and real-time applicability of the system.

A. Data Acquisition and Input Mapping:

The initial stage involves collecting ECG signal data from reliable sources.

Dataset Integration:

ECG signals are obtained from publicly available datasets such as PhysioNet. The dataset includes labeled records of apnea and normal sleep conditions.

B.Signal Preprocessing and Feature Extraction:

Raw ECG signals often contain noise and inconsistencies, which must be removed before analysis.

SignalFiltering:

Noise such as baseline wander, motion artifacts, and powerline interference (50/60 Hz) is removed using digital filters (low-pass and high-pass filters).

Segmentation:

The cleaned ECG signals are divided into fixed-length windows (typically 60 seconds). Each segment represents a sample for classification.

FeatureExtraction:

Heart Rate Variability (HRV) features are extracted, including:

- RR intervals
- Mean heart rate
- Standard deviation (SDNN)
- Frequency-domain features (LF/HF ratio)

These features help identify physiological changes associated with sleep apnea.

C. Machine Learning Classification

This stage performs classification using multiple machine learning algorithms.

D.ModelSelection:

The system utilizes multiple classifiers:

- Support Vector Machine (SVM)
- Random Forest (RF)
- Logistic Regression

Each model is trained on extracted HRV features to classify ECG segments as apnea or non-apnea.

TrainingProcess:

The dataset is divided into training and testing sets. Models are trained using labeled data and evaluated for accuracy and performance.

D. Ensemble Learning and Prediction

A stacking ensemble method is used to improve accuracy by combining outputs from multiple models. A meta-classifier, typically Logistic Regression, learns from these predictions to produce the final result, reducing overfitting and improving generalization. The system then classifies each ECG segment as either **APNEA** or **NON-APNEA**.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

E. Result Visualization and Output:

The system displays clear prediction results indicating the presence of sleep apnea. A simple dashboard interface can be used for easy visualization in clinical or home settings. Performance is evaluated using metrics such as accuracy, precision, recall, and F1-score.

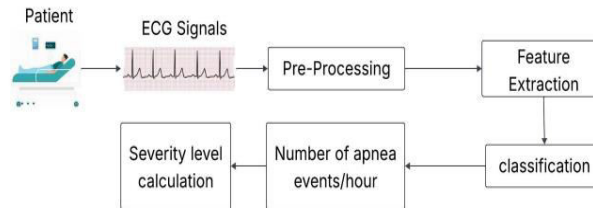


Figure 1 : System Architecture diagram

A. User Interface and Data Acquisition Layer

The system is initiated by the user (doctor or patient) through a web-based interface implemented using Streamlit or a similar framework. This interface allows users to:

- Upload ECG signal data
- View apnea detection results
- Monitor predictions in real time

Input Data Sources:

- ECG time-series signals (PhysioNet ,dataset or wearable devices)
- Sleep-related physiological signals (optional)
- Basic patient details (age, weight, medical history – optional)

The interface performs basic validation and forwards the ECG data to the preprocessing module.

B. Data Preprocessing Module

The raw ECG signals undergo preprocessing to improve data quality and ensure accurate analysis.

Key Operations Include:

ECG Processing:

- Noise filtering (removal of baseline wander, motion artifacts, powerline noise)
- Signal normalization for uniform scaling
- R-peak detection for heartbeat analysis

Segmentation:

- ECG signals are divided into fixed time windows (e.g., 60 seconds)
- Each segment is treated as an independent sample

Feature Extraction:

- Extraction of Heart Rate Variability (HRV) features:
 - RR intervals
 - Mean heart rate
 - SDNN (standard deviation)
 - Frequency domain features

The output of this stage is structured numerical data used for machine learning.

C. AI Analysis Engine

The AI Analysis Engine is the core component responsible for detecting sleep apnea.

Machine Learning Models:

Multiple classifiers are used to improve prediction accuracy:

- Support Vector Machine (SVM)
- Random Forest (RF)
- Logistic Regression



International Journal of Innovative Research in Computer and Communication Engineering (IJRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Each model learns patterns from HRV features to classify ECG segments.

Feature-Based Classification:

- Models analyze heart rate irregularities
- Detect apnea-related patterns in ECG signals

Ensemble Learning (Stacking):

- Predictions from individual models are combined
- A meta-classifier produces the final decision

This approach improves robustness and accuracy compared to single-model systems.

D. Prediction and Monitoring Module

- The system provides real-time prediction results based on processed ECG signals.
- Key Features:
 - Classification output:
 - Apnea
 - Non-Apnea
 - Continuous monitoring of sleep data
 - Visualization of ECG signals and detected events
- This module enables early detection and continuous tracking of sleep apnea conditions.

This enables virtual testing of drug outcomes before real-world administration, supporting personalized treatment optimization.

E. REPORT GENERATION MODULE:

The final module generates a structured report for medical analysis.

Generated Outputs Include:

- Apnea detection results
- Summary of HRV features
- Model prediction details
- Performance metrics (accuracy, precision, recall)

The report can be downloaded as a PDF and used for:

- Clinical evaluation
- Patient monitoring
- Further diagnosis

V. EXPERIMENTAL RESULTS

AML and DL models

This section compares the performance of our model for the detection of Obstructive Sleep Apnea (OSA) from ECG signals comprehensively. A total of 22 models have been used for research, of which 13 were conventional machine learning (ML) models and 9 deep learning (DL) models. These models have formed the base of a novel meta-model ensemble aimed at their predictive capacity aggregation. Comparison between these models has been done using measurements agreed by all i.e., Accuracy, Precision, Recall, and F1-Score in order to objectively and correctly compare the performance of all models. Accuracy reflects the general accuracy of the predictions, and Precision validates how well the model can solely produce correct results. Recall reflects the ability of the model to capture all the correct instances, and F1-Score provides a good balance in terms of precision and recall both. The measurements were carried out on two classes: Class 0, for non-apnea events, and Class 1, for apnea events. Among the machine learning algorithms, we employed algorithms such as Logistic Regression, KNearestNeighbors, Support Vector Machines, Decision Trees, Random Forest, XGBoost, AdaBoost, Gradient Boosting, Bagging Classifier, Optimized Decision Tree, Gaussian Naïve Bayes, and Optimized Naïve Bayes. The models offered a variety of capabilities in the capacity to handle linear and non-linear decision boundaries. Ensemble models such as Random Forest and XGBoost were exceptional with their power and accuracy.

Fig 2: Confusion Matrix

indicating a compensatory mechanism where an increase in heart rate may correspond to a decrease in



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Figure2: Comparison of Machine Learning Models

Performance comparison of ML models for sleep apnea detection With regards to deep learning, we attempted all forms of convolutional as well as recurrent neural network structures. Those were LSTM, simple RNN, CNN-RNN hybrids, stacked RNN-LSTM setup, Bidirectional LSTM, CNN-LSTM, Temporal Convolutional Networks (TCN), 1D CNN, and attentionbased RNNs. These LSTM and CNN-based models proved to be the best performers in utilizing temporal relationships inherent in ECG signals, a class apnea detection belongs .

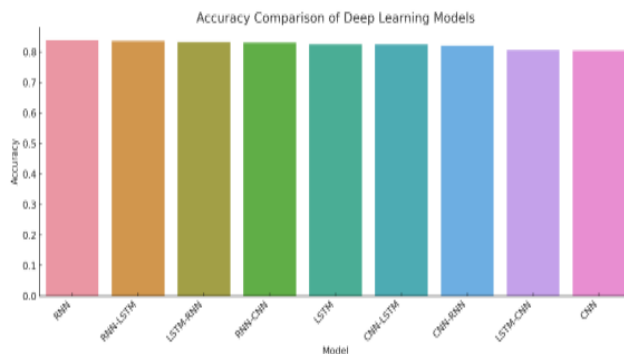


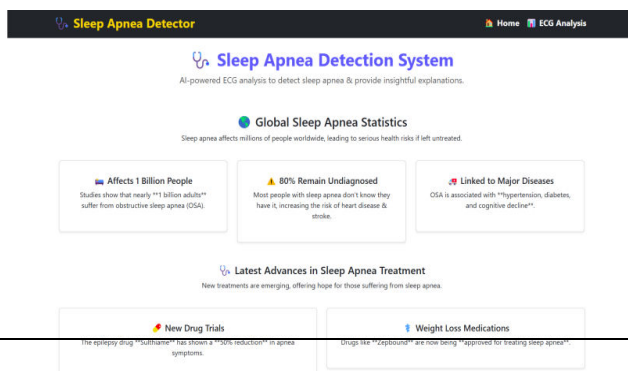
Fig 3: Accuracy Comparison of Deep Learning Models

Stacking Ensemble While the ML and DL models themselves were working well, they were also demonstrated to be some of their shortcomings in balancing precision and recall. This resulted in the creation of a meta-model ensemble that could take advantage of the strengths of individual models and avoid their worst. We constructed a new stacked ensemble that combined the prediction of eight best-performing base models —four ML (KNN, Random Forest, XGBoost, Bagging Classifier) and four DL (LSTM, RNN, RNNCNN, RNN-LSTM). The output of these models was employed as input features to a meta-model trained on Logistic Regression, which gave the final prediction. This architecture allowed the metamodel to learn to combine knowledge from various types of models and make improved predictions. The meta-model produced a very good accuracy of 90.5%, which was much better than the performance of any one of the individual base models. The classification report showed very good class-wise performance, with precision being equal to 0.80 and recall being equal to 0.91 for Class 0, which provided an F1-score equal to 0.85. For Class 1, accuracy was 0.90 and recall was 0.85 with an F1- score of equal to 0.90. These results demonstrate that the meta-model worked well in identifying the apnea as well as the non-apnea segments with a high level of confidence and reliability. The meta-model worked better with a consistent performance across all the evaluation criteria compared to individual models. The diversity of the ensemble, from the statistical learning to deep temporal pattern recognition, allowed it to generalize efficiently to diverse patterns in the ECG signals. The hybrid approach not only maximized classification performance but also generated a flexible and adaptable framework, which can further be optimized with the use of other models or physiological signals in subsequent research studies. In short, our results validate our hypothesis that a meta-learning-based combination of ML and DL models is capable of significantly enhancing the identification of obstructive sleep apnea from ECG signals. The meta-model's impressive performance, supported by its 90.5% accuracy and class-wise balanced metrics, makes it an ideal candidate for clinical applications and real-time monitoring systems.

Class	Precision	Re
0	0.80	0.9
1	0.90	0.8
Accuracy		90.

Fig 4: Confusion Matrix for

Interface:



Meta Model



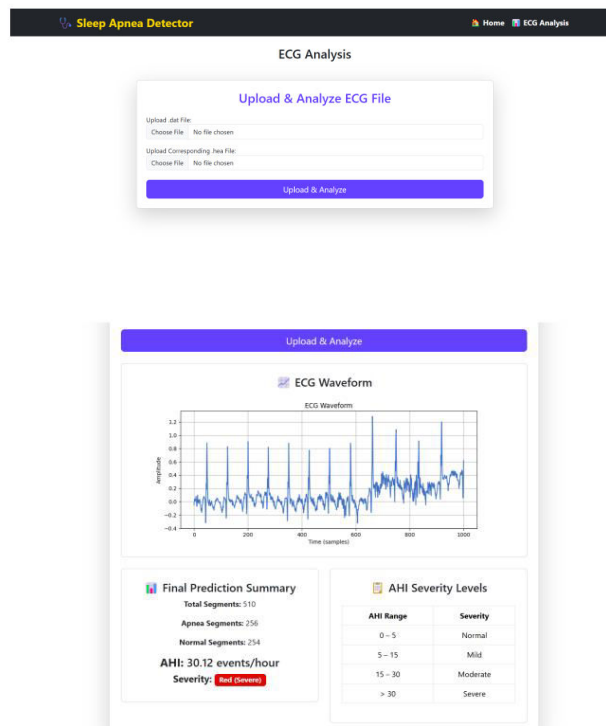
International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Fig5:Home page of web application



Fig 6:Showing Statistics of Sleep Apnea



VI. CONCLUSION



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

This paper presents a simple and effective AI-based system for detecting sleep apnea using ECG signals. Unlike traditional methods like Polysomnography (PSG), which are costly and inconvenient, this system offers a non-invasive and affordable solution. It uses heart rate variability (HRV) and machine learning models like Support Vector Machine (SVM), Random Forest (RF), and Logistic Regression combined through ensemble learning to achieve over 90% accuracy. The system processes ECG signals, extracts useful features, and provides reliable predictions, making it suitable for real-time monitoring in hospitals or at home. Overall, it is a cost-effective and scalable solution that can help in early detection and better management of sleep apnea.

REFERENCES

- [1] T. Penzel, G. Moody, R. Mark, A. Goldberger, and J. Peter, "The Apnea-ECG Database," *Computers in Cardiology*, vol. 27, pp. 255–258, 2000.
- [2] P. de Chazal, T. Penzel, and C. Heneghan, "Automated Detection of Obstructive Sleep Apnea Using ECG Features," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 6, pp. 686–696, 2003.
- [3] B. Xie, H. Minn, and Y. Hu, "Automatic Sleep Apnea Detection Using ECG Signal and Support Vector Machine," *IEEE Access*, vol. 7, pp. 48560–48570, 2019.
- [4] L. Wang, X. Li, and Y. Zhang, "Sleep Apnea Detection from Single-Lead ECG Using Ensemble Machine Learning Models," *Computers in Biology and Medicine*, vol. 124, p. 103919, 2020.
- [5] Q. Zhang, L. Zhao, and M. Chen, "A Stacking Ensemble Framework for ECG-Based Sleep Apnea Classification," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 5, pp. 1642–1652, 2021.
- [6] K. Li, J. Wu, and H. Liu, "Deep Learning Approach for Obstructive Sleep Apnea Detection Using ECG Signals," *Biomedical Signal Processing and Control*, vol. 45, pp. 1–9, 2018.
- [7] M. Rizwan, A. Sharma, and S. Kumar, "Deep Learning-Based Obstructive Sleep Apnea Detection Using Single-Lead ECG Signals," *Biomedical Signal Processing and Control*, vol. 68, p. 102762, 2021.
- [8] B. Yilmaz and M. H. Asyali, "Sleep Apnea Detection from ECG Using Wavelet Transform and Neural Networks," *Expert Systems with Applications*, vol. 37, no. 7, pp. 5817–5824, 2010.
- [9] D. Alvarez, R. Hornero, M. García, and F. del Campo, "Feature Selection from ECG for Obstructive Sleep Apnea Detection," *Medical Engineering & Physics*, vol. 29, no. 7, pp. 831–841, 2007.
- [10] A. H. Khandoker, C. K. Karmakar, and M. Palaniswami, "Automated Recognition of Obstructive Sleep Apnea Using Heart Rate Variability Analysis," *Physiological Measurement*, vol. 30, no. 10, pp. 1035–1049, 2009.
- [11] G. B. Moody and R. G. Mark, "The Impact of the MIT-BIH Arrhythmia Database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001.
- [12] S. Faust, P. Acharya, and H. Adeli, "Wavelet-Based EEG Processing for Computer-Aided Seizure Detection and Epilepsy Diagnosis," *IEEE Reviews in Biomedical Engineering*, vol. 8, pp. 1–20, 2015.
- [13] H. Hannun, P. Rajpurkar, M. Haghpanahi, et al., "Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks," *Nature Medicine*, vol. 25, no. 1, pp. 65–69, 2019.
- [14] D. Attia, P. Noseworthy, F. Lopez-Jimenez, et al., "An Artificial Intelligence-Enabled ECG Algorithm for Identification of Atrial Fibrillation," *The Lancet*, vol. 394, no. 10201, pp. 861–867, 2019.
- [15] A. Esteva, B. Kuprel, R. A. Novoa, et al., "Dermatologist-Level Skin Cancer Classification with Deep Neural Networks," *Nature*, vol. 542, pp. 115–118, 2017.
- [16] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.
- [17] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, pp. 436–444, 2015.
- [18] C. Cortes and V. Vapnik, "Support-Vector Networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [19] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [20] J. Friedman, "Greedy Function Approximation: A Gradient Boosting Machine," *Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.
- [21] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *KDD*, 2016, pp. 785–794.
- [22] R. Polikar, "Ensemble Learning," in *Ensemble Machine Learning*, Springer, 2012, pp. 1–34.
- [23] Z. Zhou, *Ensemble Methods: Foundations and Algorithms*, Chapman and Hall/CRC, 2012.
- [24] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [25] A. Graves, "Supervised Sequence Labelling with Recurrent Neural Networks," Springer, 2012.
- [26] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *ICLR*, 2015.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- [27] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in CVPR, 2016, pp. 770–778.
- [28] A. Krizhevsky, I. Sutskever, and G. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” NeurIPS, 2012.
- [30] D. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” ICLR, 2015.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  ijircce@gmail.com



www.ijircce.com

Scan to save the contact details