Machine Learning Approach for ECG-Based Heart Disease Diagnosis

1st Subalakshmi

Assistant Professor

*Department of Computer Science and Engineering*

*Dhanalakshmi Srinivasan College of Engineering*

Coimbatore, India

[subalakshmi@dsce.ac.in](mailto:subalakshmi@dsce.ac.in)

4th Dharun D

UG Scholar

*Department of Computer Science and Engineering*

*Dhanalakshmi Srinivasan College of Engineering*

Coimbatore, India

[dharund.21@dsce.ac.in](mailto:dharund.21@dsce.ac.in)

2nd Senthil vel S

UG Scholar

*Department of Computer Science and Engineering*

*Dhanalakshmi Srinivasan College of Engineering*

Coimbatore, India

[senthilvel.s21@dsce.ac.in](mailto:senthilvel.s21@dsce.ac.in)

5th Vignesh S

UG Scholar

*Department of Computer Science and Engineering*

*Dhanalakshmi Srinivasan College of Engineering*

Coimbatore, India

[vignesh.21@dsce.ac.in](mailto:vignesh.21@dsce.ac.in)

3rd Priyanka S

UG Scholar

*Department of Computer Science and Engineering*

*Dhanalakshmi Srinivasan College of Engineering*

Coimbatore, India

[priyankas.21@dsce.ac.in](mailto:priyankas.21@dsce.ac.in)

***Abstract*—** The heart performs a vital function within the human body as it enables the circulation of blood throughout all organs and tissues. Cardiac disease (CD) is considered one of the primary causes of mortality worldwide. However, accurately detecting and predicting heart diseases has always been a complex challenge in the healthcare field. Common symptoms such as rapid heartbeats, chest pain, and breathing difficulties make early detection crucial in taking necessary preventive measures before the condition worsens.

This study aims to summarize and apply different classification models to a diagnostic heart dataset and compare them to achieve the highest accuracy. The models used include K-Nearest Neighbors (KNN), Decision Tree (DT), Naïve Bayes (NB), Logistic Regression (LR), and Random Forest (RF) on five repositories of cardiac disease datasets. Additionally, a **web and mobile application** has been developed to facilitate heart disease prediction, enabling users to assess their heart health efficiently. The implementation is carried out using Anaconda Jupyter Notebook, leveraging various libraries to enhance precision and accuracy. Among all models, after ten iterations, **the Random Forest classifier achieved an accuracy of 96.5%**, outperforming previous models in the Cleveland dataset. This system provides a reliable and accessible solution for early heart disease detection, which can benefit individuals globally.

**Keywords:** Heart Disease Prediction, Machine Learning, Random Forest, Logistic Regression, Classification Models, Web Application, Mobile Application, Anaconda, Jupyter Notebook, Early Diagnosis, Healthcare AI.