The Role of Exploratory Data Analysis and Pre-processing in the Machine Learning Predictive Model for Heart Disease

Publisher: IEEE

K. Mahalakshmi ; P. Sujatha **All Authors**

Abstract:
The heart is the organ which pumps blood to the human body. There is an important need for early disease prediction systems for heart related problems. The disease predic... **View more**

Metadata

Abstract:
The heart is the organ which pumps blood to the human body. There is an important need for early disease prediction systems for heart related problems. The disease prediction system supports the doctors to diagnose critical diseases as early as possible. The disease prediction system is built using field knowledge of the healthcare profession and machine learning (ML), artificial intelligence. This research paper aims to explore the role of Exploratory Data Analysis [EDA] and pre-processing of heart disease (HD) data for the prediction of HD. This research explores three ML classifiers, namely Random Forest (RF), Support Vector Machine (SVM) and Decision Tree (DT) with missing value imputers, feature scaling techniques, data analytics, and visualization tools using four benchmark HD datasets from the UCI repository. Imputation methods mean, median, most frequent, constant, KNN imputer and iterative imputer with scaling of features were analysed with the help of three classifiers RF, SVM and DT. Random Forest with an iterative imputer achieved 86.41% of accuracy.

Published in: 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)

More Like This

Date of Conference: 25-26 May 2023

DOI: 10.1109/ACCAI58221.2023.10199714

Date Added to IEEE *Xplore*: 04 August 2023

Publisher: IEEE

ISBN Information:

Conference Location: Chennai, India

Contents

Authors

Figures

References

Citations

Keywords

Metrics