

# STATISTICS AND DATA SCIENCE



**\*EDITORS\***

**Prin. Dr. M. M. Rajmane  
Mrs. S. V. Mahajan  
Dr. Mrs. S. P. Patil**

**Prarup Publication, Kolhapur**

**STUDY OF CLASSIFICATION TECHNIQUES (LOGISTIC REGRESSION,  
SUPPORT VECTOR MACHINE AND LINEAR DISCRIMINANT ANALYSIS) IN  
PREDICTION OF PREVALENCE OF HEART DISEASE**

**<sup>1</sup>R. D. Patil , <sup>2</sup>Dr.R.R. Kumbhar, <sup>3</sup>Dr. S.V. Kakade**

<sup>1</sup>Assistant Professor, S.G.M. College Karad,

<sup>2</sup>Principal, Vivekananda college Kolhapur,

<sup>3</sup>Professor, Krishna Institute of Medical Science Demmed to be an University karad

**Email: <sup>1</sup>patilrd1996@gmail.com ,**

**<sup>2</sup>rrkumbhar@yahoo.co.in ,**

**<sup>3</sup>satishvkakade@yahoo.co.in**

---

**Abstract:**

In this study we have used the publicly available Cleveland Dataset and compare it through various classification techniques such as Logistic regression, Support vector machine, Linear Discriminant analysis for prediction of the prevalence of heart diseases. From these different models we evaluated their accuracies in predicting a heart disease. We have claimed that logistic regression and Linear Discriminant Analysis give more accurate results than support vector machines for predicting heart disease. We also used F1 score, AUC curves, precision and recall as evaluative measures. Here our aim is to provide a benchmark and improve earlier ones in the field of heart disease diagnostics with the help of different classification techniques.

**Key words:** Heart disease, Logistic regression(LR),Linear Discriminant Analysis(LDA),classification, Cleveland Heart Disease.

**Introduction:**

In this study we focused on heart disease as it leads to death. Heart disease is very complex to determine due to many problems of health such as cholesterol, chest pain, and