

FRAUD DETECTION USING MACHINE LEARNING ALGORITHMS: A COMPARATIVE ANALYSIS

A PROJECT REPORT

Submitted by

A.AKHIL KUMAR	314126510002
A.TEJA	314126510001
MD.KAMALUDDIN	314126510063
M.SAI CHAND	314126510058

*in partial fulfillment for the award of the degree
of*
**BACHELOR OF TECHNOLOGY
IN
COMPUTER SCIENCE ENGINEERING**



Under esteemed guidance of
Mrs.S.V.S.S.LAKSHMI, M.Tech
(Assistant Professor)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
ANIL NEERUKONDA INSTITUTE OF TECHNOLOGY AND SCIENCES
(Affiliated to Andhra University)

SANGIVALASA, VISAKHAPATNAM - 531162
2017-2018

FRAUD DETECTION USING MACHINE LEARNING ALGORITHMS: A COMPARATIVE ANALYSIS

A PROJECT REPORT

Submitted by

A.AKHIL KUMAR	314126510002
A.TEJA	314126510001
MD.KAMALUDDIN	314126510063
M.SAI CHAND	314126510058

*in partial fulfillment for the award of the degree
of*
**BACHELOR OF TECHNOLOGY
IN
COMPUTER SCIENCE ENGINEERING**



Under esteemed guidance of
Mrs.S.V.S.S.LAKSHMI, M.Tech
(Assistant Professor)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
ANIL NEERUKONDA INSTITUTE OF TECHNOLOGY AND SCIENCES
(Affiliated to Andhra University)

SANGIVALASA, VISAKHAPATNAM - 531162
2017-2018

**ANIL NEERUKONDA INSTITUTE OF
TECHNOLOGY AND SCIENCES
(Affiliated to Andhra University)
SANGIVALASA, VISAKHAPATNAM-531162**

BONAFIDE CERTIFICATE

Certified that this project report “**FRAUD DETECTION USING MACHINE LEARNING ALGORITHMS: A COMPARITIVE ANALYSIS**” is the bonafide work of “A.Akhil Kumar(314126510002), A.Teja(314126510001), Md.Kamaluddin(314126510063), M.Sai Chand(314126510058)” who carried out the project work under my supervision.

Prof.R.Sivaranjani, M.Tech., Ph.D.

HEAD OF THE DEPARTMENT

Professor

Department of Computer Science

and Engineering

ANITS

Mrs.S.V.S.S.Lakshmi, M.Tech.

PROJECT GUIDE

Assistant Professor

Department of Computer Science

and Engineering

ANITS

DECLARATION

We, **A.AKHIL KUMAR, A.TEJA, MD.KAMALUDDIN, M.SAI CHAND**, of final semester B.Tech., in the department of Computer Science Engineering from ANITS, Visakhapatnam, hereby declare that the project work entitled **FRAUD DETECTION USING MACHINE LEARNING ALGORITHMS: A COMPARATIVE ANALYSIS** is carried out by us and submitted in partial fulfillment of the requirements for the award of **Bachelor of Technology in Computer Science Engineering** , under Anil Neerukonda Institute of Technology & Sciences during the academic year 2017-18 and has not been submitted to any other university for the award of any kind of degree.

A.AKHIL KUMAR	314126510002
A.TEJA	314126510001
MD.KAMALUDDIN	314126510063
M.SAI CHAND	314126510058

ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany the successful completion of task would be incomplete without the mention of the people who made it possible, whose constant guidance and encouragement always boosted the morale. We take a great pleasure in presenting a project, which is the result of a studied blend of both research and knowledge.

We first take the privilege to thank the Head of our Department, **Dr.R.Sivaranjani**, for permitting us in laying the first stone of success and providing the lab facilities, we would also like to thank the other staff in our department and lab assistants who directly or indirectly helped us in successful completion of the project.

We feel great to thank **Mrs.S.V.S.S.Lakshmi**, who is our project guide and who shared her valuable knowledge with us and made us understand the real essence of the topic and created interest in us to work day and night for the project; we also thank our B.Tech project coordinator **Mrs.T.Kranthi**, for her support and encouragement.

We also like to thank **Mr.P.Sainath Vital** for providing us the lab resources; we also thank our friends and college staff who extended their part of support in the successful completion of the project.

A.AKHIL KUMAR	314126510002
A.TEJA	314126510001
MD.KAMALUDDIN	314126510063
M.SAI CHAND	314126510058

ABSTRACT

The rapid growth in E-commerce industry has lead to an exponential increase in the use of credit cards for online purchases and consequently there has been a surge in the fraud related to it . Detecting these fraud transactions has become burdensome because, the profiles of normal and fraudulent transactions are ever-changing with time. For predicting these fraudulent transactions, banks make use of various data mining methodologies : they collect past data and devise new features for enhancing the predictive power. The performance of fraud detection in credit card transactions is greatly affected by the sampling approach on data-set, selection of variables and detection technique(s) used. We chose to compare the algorithmic approach of five classification models: Logistic Regression, Naive Bayes, Random Forest, Decision Trees, Neural Networks, to determine which algorithm suits best for our data set. The performance of the techniques is evaluated based on confusion metrics.

TABLE OF CONTENTS

TITLE	Page No.
Abstract	i
Keywords	i
1. Introduction	1
1.1 Problem statement	1
1.2 Motivation	2
1.3 Contribution	2
1.4 Research Methodology	3
2. Literature Survey	4
2.1 Introduction to Credit Card Fraud	4
2.2 Recent Developments in Fraud Detection	4
2.3 Machine Learning	6
2.4 Algorithms Implemented	9
2.5 Logistic Regression	9
2.6 Naïve Bayes	11
2.7 Decision Tree	14
2.8 Neural Network	16
2.9 Random Forest	16
3. System Software Requirements	18
3.1 Software Requirements	18
3.2 Hardware Requirements	18

3.2.1 User Interface	18
3.2.2 Hardware Interface	18
3.2.3 Software Interface	18
4. Software Methodology	19
4.1 Introduction to R	19
4.1.1 R library	19
4.1.2 R Packages	20
4.1.3 Advantage of R	21
4.2 R Studio	22
5. Proposed System	23
5.1 Architecture	23
5.2 Processing Steps	24
6. Design	25
6.1 UML Diagrams	25
6.1.1 Sequence Diagram	26
6.1.2 State chart Diagram	27
7. Implementation	28
7.1 Code for User Interface	28
7.2 Accuracy for Naïve Bayes	31
7.3 Accuracy for Neural Network	34
7.4 Accuracy for Logistic Regression	36
7.5 Accuracy for Decision Tree	39
7.6 Accuracy for Random Forest	41

8.Results	43
8.1 Data Set	43
8.2 Feature Extraction	44
8.3 Accuracy For Different Classifiers	45
8.4 User Interface	59
9. Testing	62
9.1Types of Testing	62
9.2 Hold-Out Method	65
9.3 Test Cases	66
10. Conclusion	68
References	69