

EXPLORATION TO THE MACHINE LEARNING TECHNIQUES FOR DIABETES IDENTIFICATION

Submitted in partial fulfillment of the requirements for the award of
Bachelor of Engineering degree in Computer Science and Engineering

By

T. GANESH KUMAR (Reg. No. 3511595)

S. NAVEEN (Reg. No. 3511548)



**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
SCHOOL OF COMPUTING**

SATHYABAMA

**INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)**

Accredited with Grade “A” by NAAC

JEPPIAAR NAGAR, RAJIV GANDHI SALAI, CHENNAI – 600 119

APRIL 2019



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

BONAFIDE CERTIFICATE

This is to certify that this Project Report is the bonafide work of **T.GANESH KUMAR (3511595)** who carried out the project entitled **“EXPLORATION TO THE MACHINE LEARNING TECHNIQUES FOR DIABETES IDENTIFICATION”** under our supervision from September 2018 to March 2019.

INTERNAL GUIDE

Mrs.M.S.ROOBINI M.E.,

HEAD OF THE DEPARTMENT

Dr.S.VIGNESHWARI M.E.,Ph.D.,

Submitted for Viva voce Examination held on_____

Internal Examiner

External Examiner

DECLARATION

I **T.GANESH KUMAR (3511595)** hereby declare that the Project Report entitled **“EXPLORATION TO THE MACHINE LEARNING TECHNIQUES FOR DIABETES IDENTIFICATION”** done by me under the guidance of **Mrs.M.S. ROOBINI M.E.**, at Sathyabama Institute of Science and Technology is submitted in partial fulfillment of the requirements for the award of **Bachelor of Engineering** degree in **Computer Science and Engineering**.

DATE:

PLACE: CHENNAI

SIGNATURE OF THE CANDIDATE

ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to Board of Management of **SATHYABAMA** for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

I convey my thanks to **Dr. T. SASIKALA M.E.,Ph.D., Dean**, School of Computing and **Dr.S.VIGNESHWARI M.E.,Ph.D., Head of the Department**, Department of Computer Science & Engineering for providing me necessary support and details at the right time during the progressive reviews.

I would like to express my sincere and deep sense of gratitude to my Project Guide **Mrs.M.S.ROOBINI M.E.**, Professor, Dept. of Computer Science & Engineering for his valuable guidance, suggestions and constant encouragement paved way for the successful completion of my project work.

I wish to express my thanks to all Teaching and Non-teaching staff members of the Department of **Computer Science & Engineering** who were helpful in many ways for the completion of the project.

ABSTRACT

Diabetes Mellitus known as diabetes's a metabolic issue has affected an immense number of individuals. Diabetes is considered as one of the deadliest and ceaseless illnesses which cause a development in glucose. Various complexities occur if diabetes remains untreated and unidentified. In every case, the rising in AI approaches enlightens this essential issue. Hence three AI gathering figuring to be explicit KNN (K Nearest Neighbor), SVM (Support Vector Machine) and CNN (Convolutional Neural Network) are used in this examination to recognize diabetes. The point of view of this examination is to design a model which can foresee of diabetes in patients with most prominent exactness. Examinations are performed on Pima Indians Diabetes Database (PIDD). The shows of all the three figuring are surveyed on various evaluations like exactness Accuracy. Results obtained show KNN (K Nearest Neighbor), beats with the most essential precision of 76.30% likewise different calculations. First gather the Diabetes datasets from the PIDD (Pima Indians Diabetes Database). The underlying advance is pre processing to perceive the goof and dissatisfaction in datasets and Modify it. Here we used PCA (Principle Component Analysis). Calculation to address the datasets. The second step is structure to find the capable estimation in KNN (K Nearest Neighbor), SVM (Support Vector Machine) and CNN (Convolutional Neural Network). The last development is forecast of diabetes shows that the individual have diabetes or not.

TABLE OF CONTENTS

CHAPTER	TITLE	PAGE NO
	LIST OF FIGURES	viii
1	INTRODUCTION	1
	1.1 Objectives	2
	1.2 Outline of the Report	2
2	LITERATURE SURVEY	3
	2.1. Survey 1	3
	2.2. Survey 2	4
	2.3. Survey 3	5
	2.4. Survey 4	6
	2.5. Survey 5	8
3	AIM AND SCOPE OF THE PRESENT INVESTIGATION	9
	3.1. Existing System For Diabetes	9
	3.2. Introduction to Python	9
	3.2.1. Python is Interpreted	9
	3.2.2. Python is Interactive	10
	3.2.3. Python is Object-Oriented	10
	3.2.4. The Python Framework	10
	3.2.5. The Class Library	10
	3.2.6. Languages Supported By Python	11

4	EXPERIMENTAL OR MATERIAL AND METHODS	12
	4.1. SVM Algorithm	13
	4.1.1. KNN Algorithm	14
	4.12. PCA Algorithm	15
	4.2. CNN Classification	16
	4.2.1. Convolutional Layer	16
	4.2.2. Pooling Layer	17
	4.2.3. ReLu Layer	18
	4.2.4. Classification Layer	18
	4.3. Drop Out Layer	19
5	RESULTS AND DISCUSSION	20
	5.1. Datasets Implementation	20
	5.1.1. Datasets in Graphical Form	21
	5.1.2. Presence of Diabetes or Not	22
	5.1.3. Comparative Graph	23
6	SUMMARY AND CONCLUSIONS	24
	6.1. Conclusion	24
	6.2. Future Work	24
	REFERENCES	25

LIST OF FIGURES

- 4.1. Diagram for Diabetes Prediction
- 4.2. Diagram for Admin SVM Classification
- 4.3. Diagram for Admin Distance Functions
- 4.4. Diagram for convolutional layer
- 4.5. Diagram of the Pooling
- 5.1. Datasets of Diabetes
- 5.2 Datasets in Graphical Form
- 5.3 Presence of Diabetes or Not
- 5.4 Comparative Graph