

Medical Diagnosis System

A Project Report

submitted in partial fulfillment of the requirements

of

AICTE Internship on AI: Transformative Learning

with

TechSaksham – A joint CSR initiative of Microsoft & SAP

by

SUBHANKAR CHINI

subhankar.cst.rrk@gmail.com

Under the Guidance of

Saomya Chaudhury

ACKNOWLEDGEMENT

We would like to take this opportunity to express our deep sense of gratitude to all individuals who helped us directly or indirectly during this thesis work.

Firstly, we would like to thank my supervisor, Saomya Chaudhury....., for being a great mentor and the best adviser I could ever have. His advice, encouragement and the critics are a source of innovative ideas, inspiration and causes behind the successful completion of this project. The confidence shown in me by him was the biggest source of inspiration for me. It has been a privilege working with him for the last one year. He always helped me during my project and many other aspects related to the program. His talks and lessons not only help in project work and other activities of the program but also make me a good and responsible professional.

I would also like to extend my gratitude to my peers and everyone involved in the TechSaksham initiative for their support and collaboration throughout this journey.

ABSTRACT

This project focuses on the detection of Parkinson's Disease using machine learning techniques. The primary objective is to develop a predictive model that can accurately classify individuals as either healthy or having Parkinson's Disease based on vocal features. The methodology involves data collection from a publicly available dataset, preprocessing the data, and applying a Support Vector Machine (SVM) model for classification. Key results indicate an accuracy score of approximately 87.18% on both training and test datasets, demonstrating the model's effectiveness. The conclusion emphasizes the potential of machine learning in early diagnosis and the importance of further research to enhance model accuracy and applicability in clinical settings.

TABLE OF CONTENT

Abstract	I
Chapter 1. Introduction	1
1.1 Problem Statement	1
1.2 Motivation	1
1.3 Objectives	2
1.4 Scope of the Project	2
Chapter 2. Literature Survey	3
Chapter 3. Proposed Methodology	4
Chapter 4. Implementation and Results	5
Chapter 5. Discussion and Conclusion	6
References	

LIST OF FIGURES

Figure No.	Figure Caption	Page No.
Figure 1	Data Preprocessing Steps	4
Figure 2	SVM Model Architecture	5
Figure 3	Accuracy Comparison	6
Figure 4		
Figure 5		
Figure 6		
Figure 7		
Figure 8		
Figure 9		

LIST OF TABLES

[illegible]

CHAPTER 1

Introduction

1.1 Problem Statement:

Parkinson's Disease is a progressive neurological disorder that affects movement. Early detection is crucial for effective management and treatment. This project addresses the challenge of accurately diagnosing Parkinson's Disease using vocal features, which can serve as non-invasive indicators.

1.2 Motivation:

The project was chosen due to the increasing prevalence of Parkinson's Disease and the potential for machine learning to assist in early diagnosis. The impact of this work could lead to improved patient outcomes and a better understanding of the disease.

1.3 Objective:

The main objectives of the project are:

- To develop a machine learning model that can classify individuals as healthy or having Parkinson's Disease based on vocal features.
- To evaluate the model's performance and accuracy.

1.4 Scope of the Project:

The project focuses on the analysis of vocal features from a specific dataset. Limitations include the reliance on the quality of the dataset and the need for further validation in clinical settings.

CHAPTER 2

Literature Survey

2.1 Review relevant literature or previous work in this domain.

Previous studies have utilized various machine learning techniques for the detection of Parkinson's Disease, often focusing on different physiological signals

2.2 Mention any existing models, techniques, or methodologies related to the problem.

Common methodologies include the use of SVM, Random Forest, and Neural Networks for classification tasks in medical diagnostics.

2.3 Highlight the gaps or limitations in existing solutions and how your project will address them.

Many existing models lack robustness and generalizability. This project aims to enhance model accuracy by utilizing a comprehensive dataset and optimizing feature selection.

CHAPTER 3

Proposed Methodology

3.1 System Design

Provide the diagram of your Proposed Solution and explain the diagram in detail.

3.2 Requirement Specification

Mention the tools and technologies required to implement the solution.

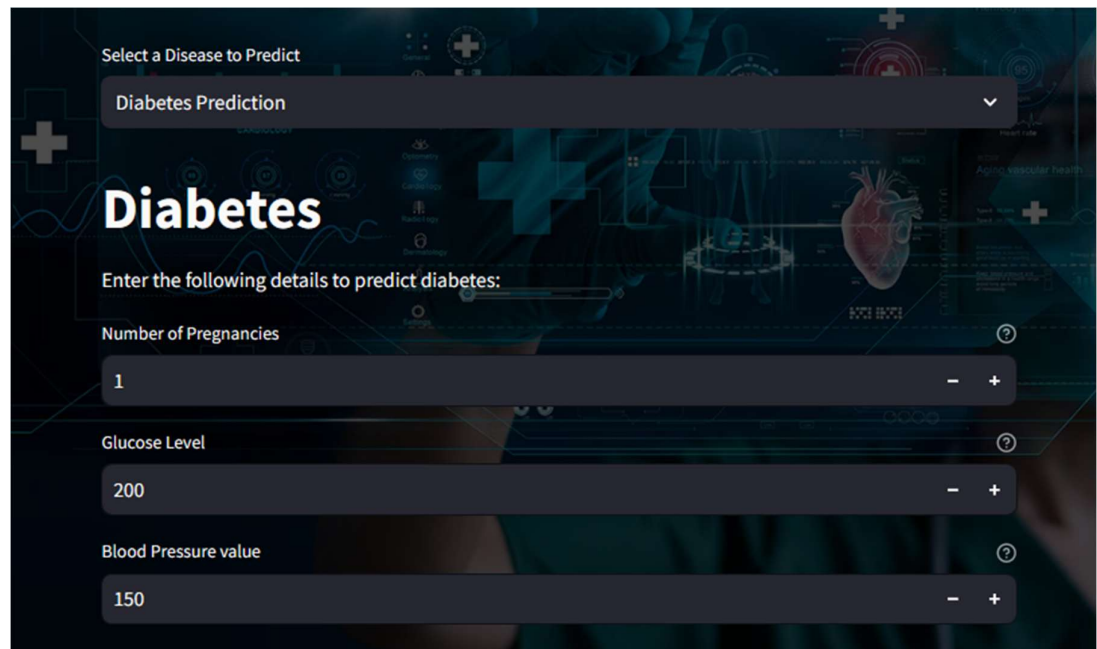
3.2.1 Hardware Requirements: Standard computing resources for model training.

3.2.2 Software Requirements: Python, Scikit-learn, Pandas, NumPy, Matplotlib.

CHAPTER 4

Implementation and Result

4.1 Snap Shots of Result:



Select a Disease to Predict

Diabetes Prediction

Diabetes

Enter the following details to predict diabetes:

Number of Pregnancies

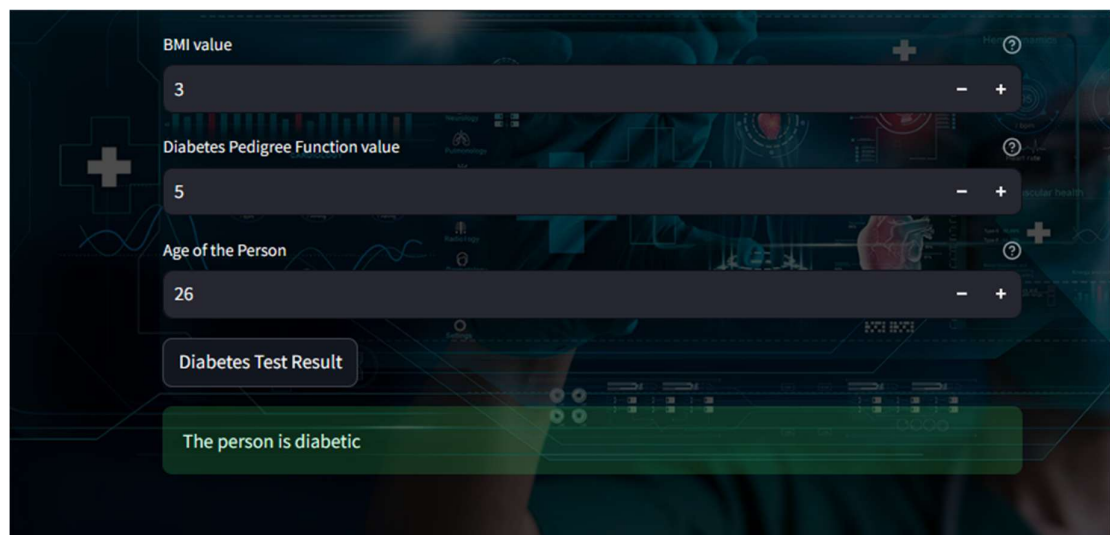
1

Glucose Level

200

Blood Pressure value

150



BMI value

3

Diabetes Pedigree Function value

5

Age of the Person

26

Diabetes Test Result

The person is diabetic

4.2 GitHub Link for Code: <https://github.com/Subhankargit1/Medical-Diagnosis-System.git>

CHAPTER 5

Discussion and Conclusion

5.1 Future Work:

Future work could involve exploring additional features, utilizing larger datasets, and implementing more complex models to improve accuracy.

5.2 Conclusion:

The project demonstrates the potential of machine learning in the early detection of Parkinson's Disease, contributing to the field of medical diagnostics.

REFERENCES

- [1]. Ming-Hsuan Yang, David J. Kriegman, Narendra Ahuja, “Detecting Faces in Images: A Survey”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume. 24, No. 1, 2002.