**A Report**

**on**

**ONLINE DIABETIES PREDICTION SYSTEM**

**Submitted for partial fulfilment of the requirements for the**

**Data Exposys Labs summer internship**

**BY**

**MOHD SAYEED SIDDIQUI**

**TABLE OF CONTENTS**

|  |  |
| --- | --- |
| **Description** | **Page No** |
| Certificate | I |
| Declaration | II |
| Acknowledgement | i |
| List of Tables | ii |
| Abstract | iii |
| Chapter 1  Introduction   * 1. Objective   Chapter 2  Literature Survey  2.1 Existing System  2.2 Proposed System  Chapter 3  Software Used  Chapter 4  Requirements  4.1 Running the jupyter    Chapter 5  Implementation  5.1 Sample Code  5.2 svm graph  Chapter 6  Conclusion  6.1 Conclusion | 6  7  8  9  10  21 |
| References | 22 |

**ii**

**LIST OF TABLES**

|  |  |  |
| --- | --- | --- |
| **Table No** | **Description** | **Page No** |
| Table 4.1 | Requirements | 14 |

**iii**

**ABSTRACT**

Diabetes is an illness caused because of high glucose level in a human body. Diabetes should not be ignored if it is untreated then Diabetes may cause some major issues in a person like: heart related problems, kidney problem, blood pressure, eye damage and it can also affect other organs of human body. Diabetes can be controlled if it is predicted earlier. To achieve this goal this project work we will do early prediction of Diabetes in a human body or a patient for a higher accuracy through applying, Various Machine Learning Techniques. Machine learning techniques Provide better result for prediction by constructing models from datasets collected from patients. In this work we will use Machine Learning Classification and ensemble techniques on a dataset to predict diabetes.

**SOFTWARE - HARDWARE REQUIREMENTS**

Software used to build:

* Visual Studio Code
* A suitable web browser (Internet explorer)
* Jupyter Notebooks
* Python frameworks
* Pandas, scikit-learn, numpy

**Hardware Required to run:**

* Windows 10,X64 bit,4gb RAM

**1.INTRODUCTION**

Diabetes is noxious diseases in the world. Diabetes caused because of obesity or high blood glucose level, and so forth. It affects the hormone insulin, resulting in abnormal metabolism of crabs and improves level of sugar in the blood. Diabetes occurs when body does not make enough insulin. According to (WHO) World Health Organization about 422 million people suffering from diabetes particularly from low- or idle-income countries. And this could be increased to 490 billion up to the year of 2030. However, prevalence of diabetes is found among various Countries like Canada, China, and India etc. Population of India is now more than 100 million so the actual number of diabetics in India is 40 million. Diabetes is major cause of death in the world. Early prediction of disease like diabetes can be controlled and save the human life. To accomplish this, this work explores prediction of diabetes by taking various attributes related to diabetes disease. For this purpose, we use the Pima Indian Diabetes Dataset, we apply various Machine Learning classification and ensemble Techniques to predict diabetes. Machine Learning Is a method that is used to train computers or machines explicitly. Various Machine Learning Techniques provide efficient result to collect Knowledge by building various classification and ensemble models from collected dataset. Such collected data can be useful to predict diabetes. Various techniques of Machine Learning can capable to do prediction, however it’s tough to choose best technique. Thus, for this purpose we apply popular classification and ensemble methods on dataset for prediction.

II. PROPOSED METHODOLOGY

Goal of the paper is to investigate for model to predict diabetes with better accuracy. We experimented with different classification and ensemble algorithms to predict diabetes. In the following, we briefly discuss the phase.

A. Dataset Description- the data is gathered from Kaggle UCI repository which is named as Diabetes Dataset. The dataset has many attributes of 520 patients. dataset has been collected using direct questionnaires from the patients of Sylhet Diabetes Hospital in Sylhet, Bangladesh, and approved by a doctor.

**OBJECTIVE**

To design a platform for our members to discuss, plan and book a slot at any Turf Ground hassle free.

Book my Turf is a website that connects the users and allows to book any ground either for playing cricket and football.

The difference between manual booking and online booking is that the online turf booking is a quick and efficient approach to book sports ground anywhere and anytime available.

The more convenient facilities we have available to us, the more we play. In turn, the revenue from renting out facilities allows turfs to develop and offer more opportunities to people in their local area.

Our aim is straightforward: we want it to be as easy as possible for people to get out and play sport in their local area.

**2. LITERATURE SURVEY**

**2.1 EXISTING SYSTEM**

**2.2 PROPOSED SYSTEM**

This Application aims to provide an interface which helps the user to Book a Ground on time and user can also pay by scan and etc

This application provides simplicity with any novice internet user being able to book the ground easily.

By considering the problems from the existing system, a more advanced system has been proposed. In the new system the user will be asked to register once at the start. The application proposed will allow users to book any sports amenity if available. Our application will then provide the user with a list nearby sports amenity. The list will also contain the information about booking availability. Based on the information the user has to decide the ground or court he/she wishes to play on. While booking a particular ground or court team/user has to pay 50% of ground fees then only booking is confirmed. If the team/user makes the booking where ground or court is available, the team will be allocated .

.

**3. SOFTWARE USED**

* **PYTHON:** Python is an interpreted, high level and general-purpose programming language. It was created by Guido van Rossum, and released in 1991. Python's design philosophy emphasizes code readability with its notable use of significant indentation. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects. Python can be used on a server to create web applications.
* **VISUAL STUDIO CODE:** Visual Studio Code is a source-code editor that can be used with a variety of programming languages, including Python, Java, C++ [JavaScript](https://en.wikipedia.org/wiki/JavaScript) and C. It is based on [Electron](https://en.wikipedia.org/wiki/Electron_(software_framework)) framework, which is used to develop [Web applications](https://en.wikipedia.org/wiki/Web_application) that run on the [Blink layout engine](https://en.wikipedia.org/wiki/Blink_layout_engine). Visual Studio Code employs the same editor component (code named "Monaco") used in [Azure DevOps](https://en.wikipedia.org/wiki/Azure_DevOps_Server) (formerly called Visual Studio Online and Visual Studio Team Services). It has a user friendly UI that makes it easy to develop applications. Simply extensions can be downloaded which makes tedious and time consuming tasks very easy.

**4.REQUIREMENTS**

The requirements to run successfully on local server are-

* Any code editor (preferably VS Code)
* Versions of Python and other libraries required:

|  |  |
| --- | --- |
| Package | Version |
| Python | 3.9 |
| Anaconda | 4.12.0 |

**5. IMPLEMENTATION**

**5.1 Sample Code-** Code runner



*#importing libraries*

**import** numpy **as** np

**import** pandas **as** pd

diabetes\_dataset= pd.read\_csv('diabetes.csv')

classifier = svm.SVC(kernel='linear')

X\_train\_prediction = classifier.predict(X\_train)

training\_data\_accuracy = accuracy\_score(X\_train\_prediction, Y\_train)

data\_accuracy **=** accuracy\_score(X\_train\_prediction, Y\_train)

input\_data = (4,110,92,0,0,37.6,0.191,30)

output:

The person is not diabetic

*#importing libraries*

**import** numpy **as** np

**import** pandas **as** pd

diabetes\_dataset= pd.read\_csv('diabetes.csv')

classifier = svm.SVC(kernel='linear')

X\_train\_prediction = classifier.predict(X\_train)

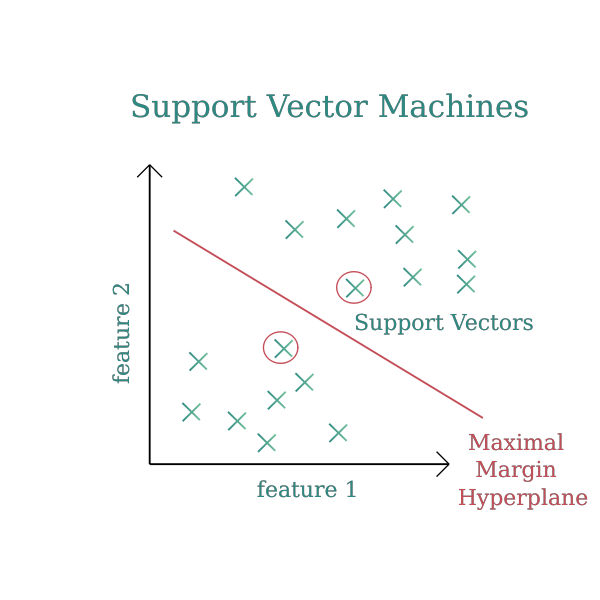
training\_data\_accuracy = accuracy\_score(X\_train\_prediction, Y\_train)

data\_accuracy **=** accuracy\_score(X\_train\_prediction, Y\_train)

input\_data = (2,197,70,45,543,30.5,0.158,53)

output:

The person is diabetic



**6.5 CONCLUSION**

The Aim of this project was to design and implement Diabetes Prediction Using Machine Learning Methods. The proposed method uses Support vector machine(svm) and has an accuracy of 98%. The Experimental results can be asst health care to take early prediction and make early decision to cure diabetes and save humans life.

**References**

* <https://codewithharry.com/videos/python-tutorials-for-absolute-beginners-127>
* https://www.youtube.com/watch?v=6mbwJ2xhgzM&list=PLu0W\_9lII9agiCUZYRsvtGTXdxkzPyItg
* https://www.youtube.com/watch?v=JxzZxdht-XY