** Islamia College University of Peshawar**

**Mini Project Review:**

# 

#### Submitted to: Dr Naveed Abbas

**Submitted By:-**

**1. Muhammad ishaq - Roll No. 201186**

**2. Babar Ali Shah - Roll No. 201180**

**Section: B**

**Semester: 4th**

**Department: BSSE**

**Ably Guided By:**

# Abstract:

Diabetes is a chronic disease with the potential to cause a worldwide health care crisis. According to International Diabetes Federation 382 million people are living with diabetes across the whole world. By 2035, this will be doubled as 592 million. Diabetes is a disease caused due to the increase level of blood glucose. This high blood glucose produces the symptoms of frequent urination, increased thirst, and increased hunger. Diabetes is a one of the leading cause of blindness, kidney failure, amputations, heart failure and stroke. When we eat, our body turns food into sugars, or glucose. At that point, our pancreas is supposed to release insulin. Insulin serves as a key to open our cells, to allow the glucose to enter and allow us to use the glucose for energy. But with diabetes, this system does not work. Type 1 and type 2 diabetes are the most common forms of the disease, but there are also other kinds, such as gestational diabetes, which occurs during pregnancy, as well as other forms. Machine learning is an emerging scientific field in data science dealing with the ways in which machines learn from experience. The aim of this project is to develop a system which can perform early prediction of diabetes for a patient with a higher accuracy by combining the results of different machine learning techniques. The algorithms like K nearest neighbor, Logistic Regression, Random forest, Support vector machine and Decision tree are used. The accuracy of the model using each of the algorithms is calculated. Then the one with a good accuracy is taken as the model for predicting the diabetes.

# Keywords:

Machine Learning, Diabetes, Decision Tree, and K nearest neighbor, Logistic Regression, Support Vector Machine, Accuracy

# 

# INTRODUCTION

Diabetes is the fast growing disease among the people even among the youngsters. In understanding diabetes and how it develops, we need to understand what happens in the body without diabetes. Sugar (glucose) comes from the foods that we eat, specifically carbohydrate foods. Carbohydrate foods provide our body with its main energy source everybody, even those people with diabetes, needs carbohydrate. Carbohydrate foods include bread, cereal, pasta, rice, fruit, dairy products and vegetables (especially starchy vegetables). When we eat these foods, the body breaks them down into glucose. The glucose moves around the body in the bloodstream. Some of the glucose is taken to our brain to help us think clearly and function. The remainder of the glucose is taken to the cells of our body for energy and also to our liver, where it is stored as energy that is used later by the body. In order for the body to use glucose for energy, insulin is required. Insulin is a hormone that is produced by the beta cells in the pancreas. Insulin works like a key to a door. Insulin attaches itself to doors on the cell, opening the door to allow glucose to move from the blood stream, through the door, and into the cell. If the pancreas is not able to produce enough insulin (insulin deficiency) or if the body cannot use the insulin it produces (insulin resistance), glucose builds up in the bloodstream (hyperglycemia) and diabetes develops. Diabetes Mellitus means high levels of sugar (glucose) in the blood stream and in the urine.

# Types Of Diabetes. Type 1 And Type 2 Diabetes Mellitus. Insulin-Dependent Diabetes Mellitus And Non Insulin-Dependent Diabetes Mellitus. Insulin Resistance And Insufficient Insulin Production. Royalty Free SVG, Cliparts, Vectors, And Stock Illustration.

# 

# 

Type 1 diabetes means that the immune system is compromised and the cells fail to produce insulin in sufficient amounts. There are no eloquent studies that prove the causes of type 1 diabetes and there are currently no known methods of prevention.

Type 2 diabetes means that the cells produce a low quantity of insulin or the body can’t use the insulin correctly. This is the most common type of diabetes, thus affecting 90% of persons diagnosed with diabetes. It is caused by both genetic factors and the manner of living.

Gestational diabetes appears in pregnant women who suddenly develop high blood sugar. In two thirds of the cases, it will reappear during subsequent pregnancies. There is a great chance that type 1 or type 2 diabetes will occur after a pregnancy affected by gestational diabetes.

# Symptoms of Diabetes:

Frequent Urination

• Increased thirst

• Tired/Sleepiness

Frequent Urination

• Increased thirst

• Tired/Sleepiness

• Frequent Urination

• Increased thirst

• Tired/Sleepiness

• Weight loss

• Blurred vision

• Mood swings

• Confusion and difficulty concentrating

•

* Frequent Urination
* Increased thirst
* Tired/Sleepiness
* Weight loss
* Blurred vision
* Confusion and difficulty concentrating
* frequent infections

# Causes of Diabetes:

Genetic factors are the main cause of diabetes. It is caused by at least two mutant genes in the chromosome 6, the chromosome that affects the response of the body to various antigens. Viral infection may also influence the occurrence of type 1 and type 2 diabetes. Studies have shown that infection with viruses such as rubella, Coxsackievirus, mumps, hepatitis B virus, and cytomegalovirus increase the risk of developing diabetes.

## 

## PROPOSED SYSTEM

Classification is one of the most important decision making techniques in many real world problems. In this work, the main objective is to classify the data as diabetic or non-diabetic and improve the classification accuracy. For many classification problem, the higher number of samples chosen but it doesn’t leads to higher classification Accuracy. In many cases, the performance of algorithm is high in the context of speed but the accuracy of data classification is low. The main objective of our model is to achieve high accuracy. Classification accuracy can be increase if we use much of the data set for training and few data sets for testing. This survey has analyzed various classification techniques for classification of diabetic and non-diabetic data. Thus, it is observed that techniques like Support Vector Machine, Logistic Regression, KNN and Decision tree are most suitable for implementing the Diabetes prediction system.

# Proposed Model Diagram:

Performance Evaluation on various Measures

Applied Algorithms

Preprocessing Data

Dataset

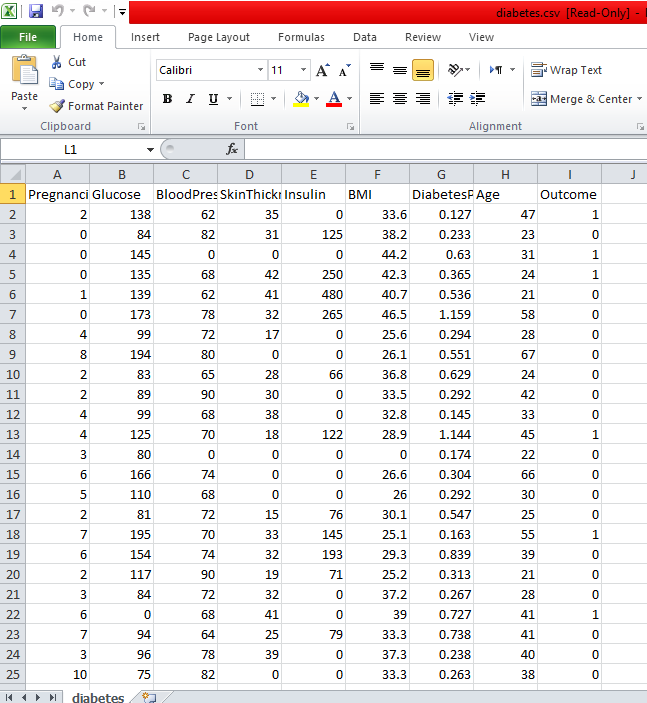
Results

Comparative Analysis Based on Accuracy

# Dataset Description:

The diabetes data set was originated from <https://www.kaggle.com/johndasilva/diabetes>.

### Diabetes dataset containing 2000 cases. The objective is to predict based on the measures to predict if the patient is diabetic or not.



The diabetes data set consists of 2000 data points, with 9 features each.

➔ “Outcome” is the feature we are going to predict, 0 means No diabetes, 1 means diabetes.

The diabetes data set consists of 2000 data points, with 9 features each.

➔ “Outcome” is the feature we are going to predict, 0 means No diabetes, 1 means diabetes.

* The diabetes data set consists of 2000 data points, with 9 features each.
* “Outcome” is the feature we are going to predict, 0 means No diabetes, 1 means diabetes.

# RESULT & DISCUSSION:

# Correlation Matrix?

## CONCLUSION AND FUTURE WORK:

One of the important real-world medical problems is the detection of diabetes at its early stage. In this study, systematic efforts are made in designing a system which results in the prediction of diabetes. During this work, five machine learning classification algorithms are studied and evaluated on various measures. Experiments are performed on john Diabetes Database. Experimental results determine the adequacy of the designed system with an achieved accuracy of 99% using Decision Tree algorithm. In future, the designed system with the used machine learning classification algorithms can be used to predict or diagnose other diseases. The work can be extended and improved for the automation of diabetes analysis including some other machine learning algorithms.

Frequent infections\