PROJECT REPORT ON

**DIABETES PREDICTION USING**

**SUPPORT VECTOR MACHINE**

Submitted by**:**

Samrajya Chand

Trinity International College

03/22/2024

**ABSTRACT**

“Diabetes prediction using support vector machine” is the implementation of algorithm support vector machine to correctly detect if a patient has diabetes or not. We used python to implement the system using Google colabs as the IDE. By using python libraries like Numpy and Pandas we worked with the dataset and array of dataset provided. It can predict if a person has diabetes or not by using medical profiles such as Pregnancies, Glucose, Skin thickness, Age, Insulin , BMI and Blood Pressure .

**Keywords:** Support Vector Machine, Pandas, Numpy, Prediction System.

TABLE OF CONTENT

CHAPTER 1

INTRODUCTION

* 1. Introduction
  2. Problem Statement
  3. Objectives
  4. Scope and Limitation

1.4.1 Scope

1.4.2 Limitation

CHAPTER 2

LITERATURE REVIEW AND RESEARCH METHODOLOGY

2.1 Literature Review

2.2 Framework of the Model

2.3 Methodology

**CHAPTER 1**

**INTRODUCTION**

* 1. Introduction

Diabetes is a chronic disease that occurs either when the pancreas does not produce enough insulin or when the body cannot effectively use the insulin it produces. According to a report from WHO, diabetes and kidney disease due to diabetes caused an estimated 2 million deaths [1]. According to data from the IDF, around 7.1% of adults aged 20-79 in Nepal had diabetes in 2019 [2]. Nepal has a considerable burden of diabetes, with both type 1 and type 2 diabetes affecting a significant portion of the population. The integration of ICT (Information Communication & Technology) might be helpful for a developing country like Nepal. It helps in analyzing the current situation of Health in Nepal and act for necessary solutions.

“Diabetes Prediction System” is the system curated so the user can dictate if he/she is suffering from diabetes or not. The parameters to detect if a person is suffering from Diabetes or not are: [Pregnancies, Glucose, Blood Pressure, Skin Thickness, Insulin, BMI, DiabetesPedigreeFunction, Age .The prediction is done using computer aided diagnosis methods. This is a manual process and also dependent on each case as the system detects about ongoing pattern and model. By using this system user can directly detect if they have diabetes or not, which saves them cost for going to a health care institution and spending more amount of money. It would remove the uncertainty and give results based on the data fed to it.

The types of diabetes that are found normally are:

1. Type 1 Diabetes:

Type 1 diabetes is thought to be caused by an autoimmune reaction (the body attacks itself by mistake). This reaction stops your body from making insulin. Approximately 5-10% of the people who have diabetes have type 1. Type 1 diabetes can be diagnosed at any age, and symptoms often develop quickly.

1. Type 2 Diabetes:

With type 2 diabetes, your body doesn’t use insulin well and can’t keep blood sugar at normal levels. About 90-95% of people with diabetes have type 2. It develops over many years and is usually diagnosed in adults (but more and more in children, teens, and young adults).

1. Gestational Diabetes

Gestational diabetes develops in pregnant women who have never had diabetes. If you have gestational diabetes, your baby could be at higher risk for health problems. Gestational diabetes usually goes away after your baby is born. However, it increases your risk for type 2 diabetes later in life. Your baby is more likely to have obesity as a child or teen and develop type 2 diabetes later in life.

* 1. Problem Statement

Diabetes is a chronic disease that unless treated and detected early might lead to severe consequences later. Diabetes is a major cause of blindness, kidney failure, heart attacks, stroke and lower limb amputation. Since diabetes is a chronic disease people might ignore the simple symptoms of diabetes as just a consequences of dehydration. The only possibility of detecting diabetes is by blood test of our glucose (sugar) level.

In the early stages, diabetes may not present noticeable symptoms, leading individuals to underestimate its seriousness and delay seeking medical attention and hospitals to not seeing necessity for blood test .As diabetes symptoms vary from person to person the health officials too might misinterpret or misdiagnosis it, the doctors might just diagnosis different disease instead of preparing diagnosis based on the data and patterns stored in the database.

Hence, to develop an intelligent and cost-effective computer-aided diagnosis system for predicting the risk of diabetes in individuals using the Support Vector Machine (SVM) algorithm, utilizing relevant features such as Pregnancies, Glucose Level, Blood Pressure, Skin Thickness, Insulin Level, BMI, Diabetes Pedigree Function, and Age this system was made.

* 1. Objectives

The objectives of this system are too:

1. Develop a machine learning model for diabetes prediction.
2. Utilizing relevant features for accurate prediction.
3. Evaluating the algorithm’s performance.
4. To provide basic suggestions according to disease.
5. Providing instant result to patients.
   1. Scope and Limitations
      1. Scope

The primary scope of the project is to develop a machine learning model, specifically using the SVM algorithm, for predicting the risk of diabetes in individuals. The project aims to create a user-friendly interface, such as a web application or desktop application that allows users to input their personal data and receive predictions from the trained SVM model.

* + 1. Limitations

The project is limited to using the specified features or parameters (Pregnancies, Glucose Level, Blood Pressure, Skin Thickness, Insulin Level, BMI, Diabetes Pedigree Function, and Age). Other potentially relevant features may not be considered. The accuracy of the predictions relies on the accuracy of the user inputs. Incorrect or incomplete information provided by the user may result in inaccurate predictions.

CHAPTER 2

LITERATURE REVIEW AND RESEARCH METHODOLOGY

2.1 Literature Review

The Decision Support in Diabetes Prediction System, a developed system, utilizes data mining modeling techniques to uncover relationships within healthcare data. This system serves as a training tool for nurses and medical students to diagnose patients with diabetes.

The neural network is divided into two parts: training data and testing data. Eight parameters, including pregnancies, glucose levels, blood pressure, skin thickness, insulin, BMI, diabetes pedigree, and age, are considered. When doctors input unknown data, the system compares it with the trained data to generate outcomes. This project, also referred to as an Intelligent System, employs data mining techniques such as Support Vector Machine, Logistic Regression, Naïve Bayesian algorithm, K-Nearest Neighbors algorithm, Decision Tree algorithm, and Random Forest algorithm. These algorithms are implemented as web-based applications that compare user inputs with the trained dataset. This assists healthcare providers in making informed clinical decisions.

The research paper proposes a weighted fuzzy rule-based Clinical Decision Support System for diabetes diagnosis. It extracts knowledge from patients' clinical data and employs a computerized approach to generate weighted fuzzy rules. Consequently, a fuzzy rule-based decision support system accurately assesses the likelihood of having diabetes.

2.2 Framework of the Model

Input Attribute of Patient

Determine its Accuracy

Congratulations! No Detection

Is Disease

Diagnosed?

Fig 1: Framework of diabetes production system.

2.3 Methodology

**REFERENCES**

# [1] World Health Organization, "Global report on diabetes," WHO Press, Geneva, Switzerland, 2016, pp. 1-88.

# [2] International Diabetes Federation, "IDF Diabetes Atlas, 9th ed.," Brussels, Belgium, 2019.

# 