

SMT. PARVATIBAI CHOWGULE  
OF ARTS AND SCIENCE  
GOA-403601, GOA(INDIA)  
COMPUTER SCIENCE AND ENGINEERING

---

# Clustered Machine Learning for Predicting Diabetes

---

*Submitted By:*

Alexander Roque  
Rodrigues  
SU170331  
TYBSC

*Submitted To:*

Mrs. Ashweta Fondekar  
Asst. Professor  
Dept. of CSE

June-February  
2019-2020  
Graduation Dissertation

# 1 Acknowledgements

# Contents

<b>1</b>	<b>Acknowledgements</b>	<b>1</b>
<b>2</b>	<b>Introduction</b>	<b>3</b>
<b>3</b>	<b>System Specifications and Requirements</b>	<b>5</b>
3.1	Functionalities for Doctors . . . . .	5
3.2	Functionalities for Patients . . . . .	5
<b>4</b>	<b>Finding the Right Algorithm</b>	<b>6</b>
<b>5</b>	<b>Components and Frameworks</b>	<b>6</b>
5.1	f . . . . .	7

## Abstract

Your abstract.

## 2 Introduction

In recent days, there has been a sharp increase in the cases of diabetes mellitus. Diabetes mellitus is on the rise amongst many people and the rate of contracting this lifestyle disease could be reduced significantly if proper measures and precautions were to be instilled amongst people the number of people can be reduced.

Machine learning is a growing field in computer science. With the development and introduction of many algorithms the prediction and accuracy of the predictions itself has improved substantially. Machine learning and healthcare systems are also becoming increasingly popular in the healthcare sector.

The project encompasses the qualities of Remote Patient Monitoring (RPM) and Clinical Decision Support (CDS). RPM provides medical facilities that have the ability to transmit patient data to healthcare professionals who might very well be halfway around the world. RPM can monitor blood glucose levels and blood pressure. It is particularly helpful for patients with chronic conditions such as type 2 diabetes, hypertension, or cardiac disease. Data collected and transmitted via PRM can be used by a healthcare professional or a healthcare team to detect medical events such as stroke or heart attack that require immediate and aggressive medical intervention. Data collected may be used as part of a research project or health study. RPM is a life-saving system for patients in remote areas who cannot access face-to-face health care. CDS analyzes data from clinical and administrative systems. The aim is to assist healthcare providers in making informed clinical decisions. Data available can provide information to medical professions who are preparing diagnoses or predicting medical conditions like drug interactions and reactions. CDS tools filter information to assist healthcare professionals in caring for individual clients.

The objective of this project is to create a system that is able to use the machine learning algorithms and predict the outcome of the parameters entered into the algorithm and help the patient draw a conclusion whether or not he/she has the same traits exhibited by similar patients that have diabetes. Also the system should have a UI that is capable of displaying the

data of the patients to the doctor and to the patients themselves for further interpretation.

### 3 System Specifications and Requirements

The main function of this project is to enable the doctors to advise their patients with the help of the prediction software. The system should be accessible to the patient as well as the doctor, therefore it should have a web interface for the two parties to interact with.

#### 3.1 Functionalities for Doctors

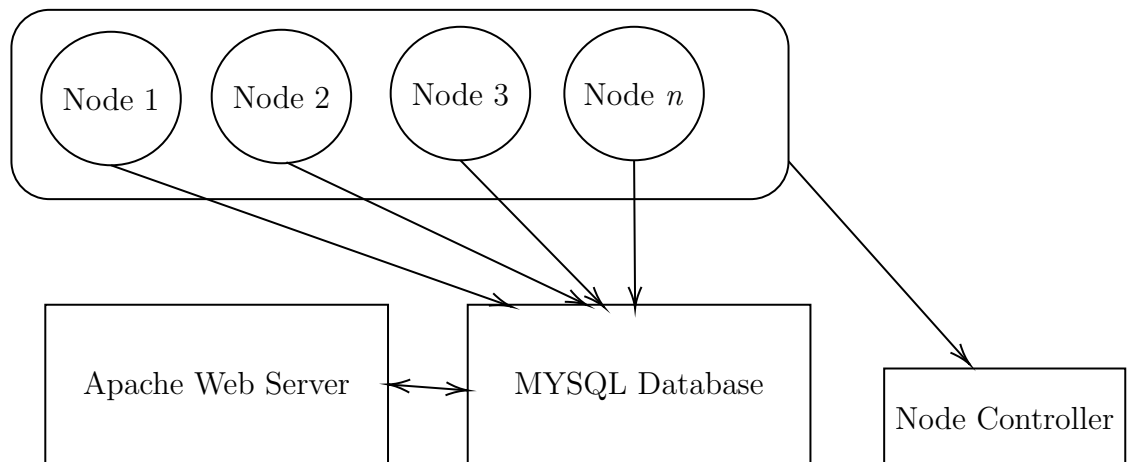
As an owner of a doctors account a doctor should be able to:

- Add new patients.
- Add new observations for the machine learning algorithm to predict.
- Analyse the patients previous records.
- Leave notes for patients to act on.

#### 3.2 Functionalities for Patients

As the patient, one should be able to:

- Should be able to see the predicted risk of developing diabetes.
- Should be



## 4 Finding the Right Algorithm

For our predictions we need to use algorithms that give us the maximum accuracy and once we find that algorithm we will further need to tune the selected algorithm from the algorithms we selected to optimise the performance of the predictions generated by the machine learning algorithm. For this purpose we turn to the EDA or exploratory data analysis stage of any machine learning project.

## 5 Components and Frameworks