VIT® BHOPAL
www.vitbhopal.ac.in

DSN 3039
Community Project 2020-21

Supervisor: Dr. Kanchan Lata Kashyap

Date: 25th May 2021

Diabetes Predictor

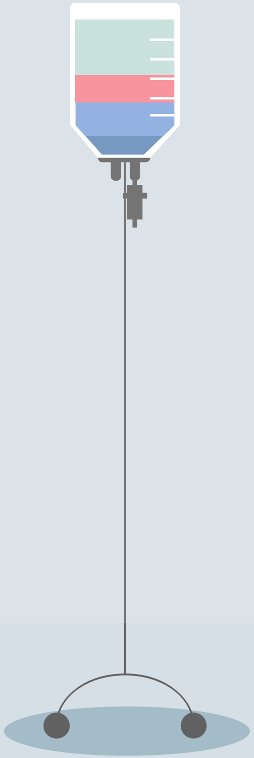
TEAM

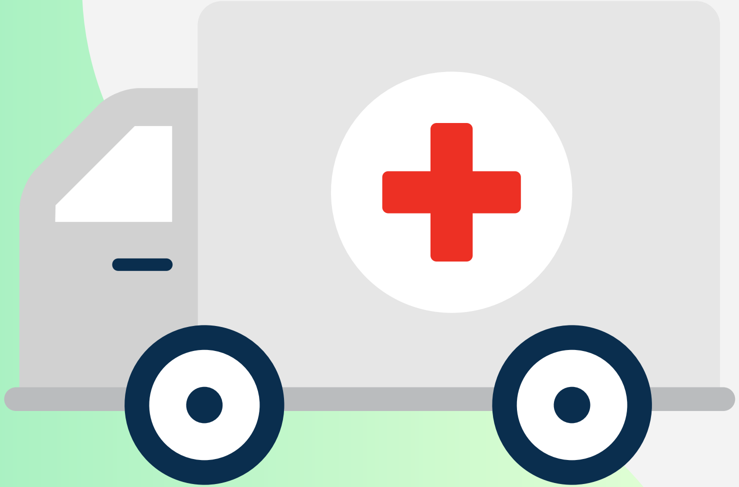
Akash
Anant Jaiswal
Avanish Sandilya

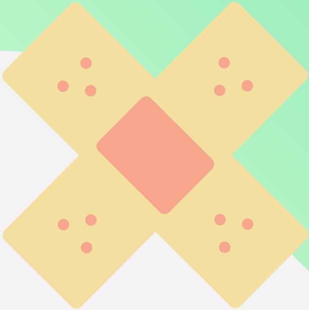
Dishant Mehta
Manasi Muchrikar
Nishq Poorav Desai

Rahul Khandebharad
Utkarsha
Umang Mahesh Rane

INDEX



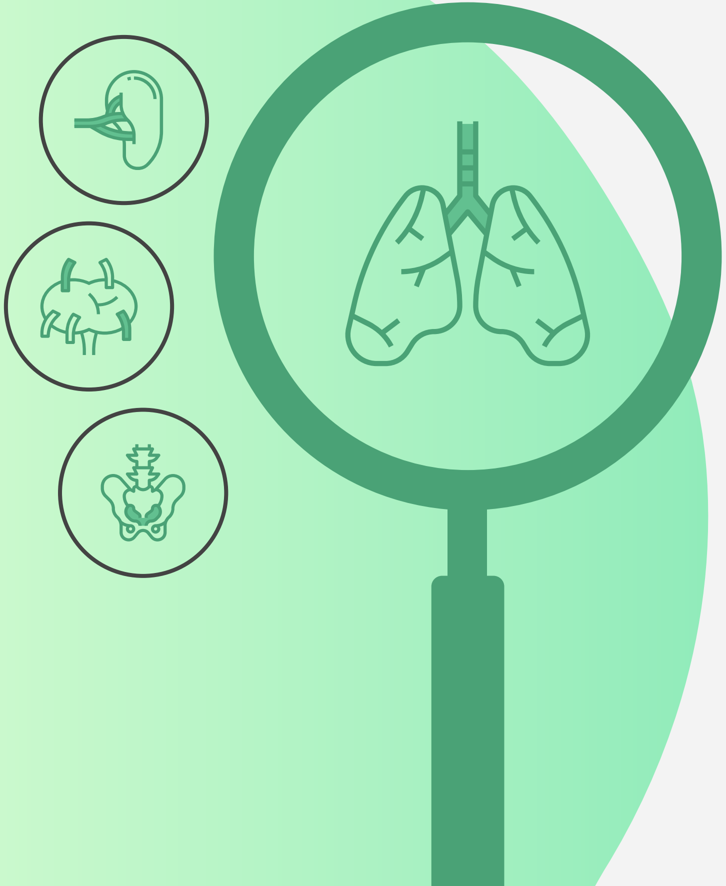




01

Objective

Diabetes is a health condition which causes a financial burden on people with no stable background. Diabetes is not a dangerous health condition, but due to lack of proper guidance ,it gets converted into a severe health issue. A basic check through a questionnaire can help people with better advice to deal with diabetes.



Here we are proposing a solution which will be a fusion of software and hardware elements that will be able to process the factor indicating diabetes-like blood sugar level, insulin level, and blood pressure. These factors then fed to the system will help analyze if the person is suffering from Diabetes or not. Later, recommending the diet for Diabetes free life.

02

Existing Project

Identifying Existing Projects

Calisir et al.

Calisir et al. developed an automatic diabetes diagnosis system based on the Linear Discriminant Analysis (LDA) technique.The Total 89.74% classification accuracy is achieved by applying the Morlet Wavelet Support vector machine classifier.

Zou et al.

Zou et al. have applied decision trees, random forest, and neural network techniques for the prediction of diabetes mellitus. The highest 80.84% accuracy is achieved by using principal component analysis and minimum redundancy maximum relevance (mRMR) for dimensionality reduction.

Tigga et al.

Tigga et al. used Logistic regression, K-Nearest Neighbour, Support vector machine, Naïve Bayes and Decision tree, and Random forest for the classification of diabetic and non-diabetic. The highest 90% accuracy is achieved by applying the Random Forest classifier.

Some More Existing Projects

Sisodia et al.

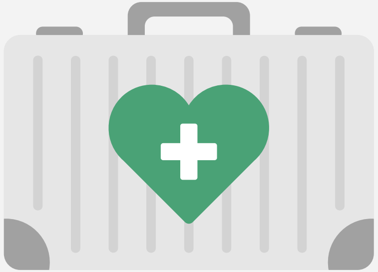
Sisodia et al. applied the decision tree, naïve Bayes, and SVM for the prediction of diabetes, and the highest 76.30% accuracy is achieved by using Naïve Bayes classifier.

Wu et al.

Wu et al. obtained 95.42% accuracy by using improved k-NN and logistic regression techniques to predict Type 2 diabetes mellitus.

Meng et al.

Meng et al. achieved 73.23% and 77.87% classification accuracy by applying artificial neural network and decision tree(C.5) model, respectively.



Choubey et al.

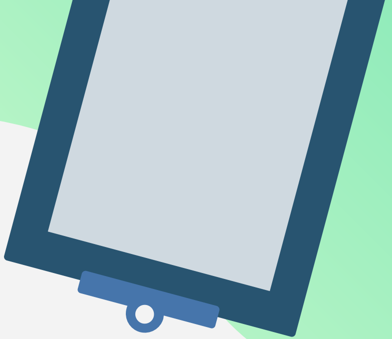
Choubey et al. used Genetic algorithms and radial basis function-based neural network techniques for feature selection and diabetes classification. The highest 76.087% classification accuracy is achieved on Pima Indian Diabetes Dataset (PIDD).

Haung et al.

Haung et al. obtained the highest 95% accuracy on Ulster Community and Hospitals Trust (UCHT) data set by applying Naïve Bayes, IB1, and decision tree classifier for diabetes prediction.

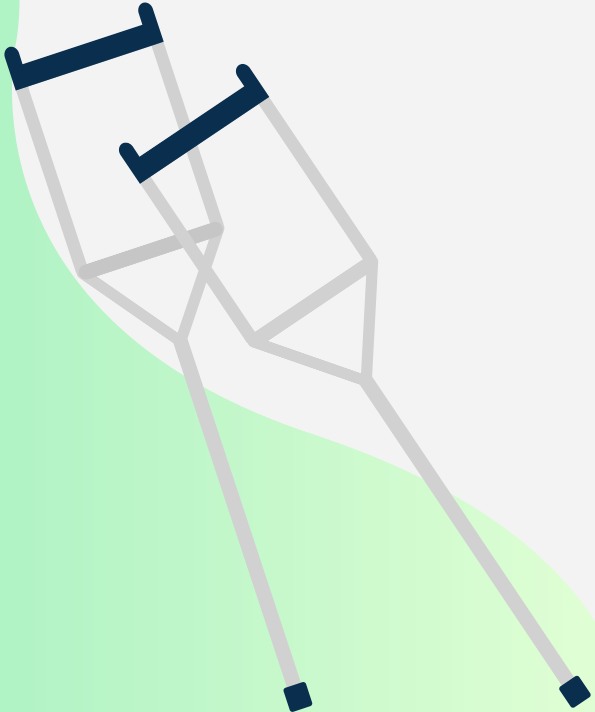
Perveen et al.

Perveen et al. applied Naïve Bayes and decision tree machine learning techniques for diabetes prediction achieved 81% and 80% true positive rate, respectively.



03

Project Modules



Identifying Information

Dataset

The PIMA diabetes dataset has been collected originally from National Institute of Diabetes , Digestive and Kidney Diseases that are specified in the Credits and Acknowledgement Section. In this , a survey was conducted on 768 people.

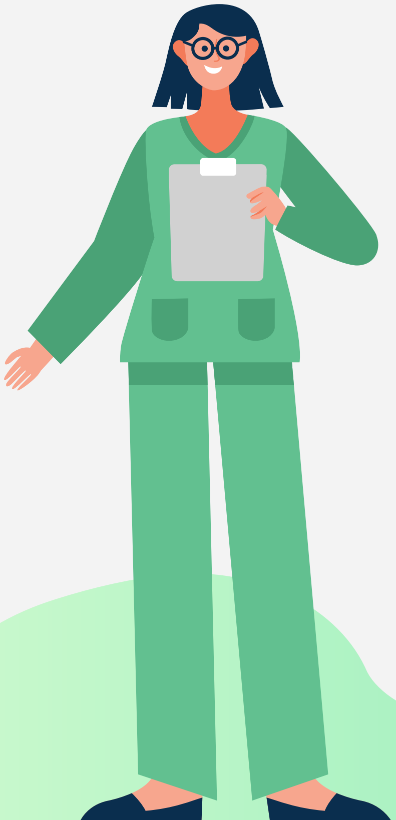
Feature Extraction

In the dataset 8 independent variables have been identified as features including Number of pregnancies, Plasma glucose concentration in 2 hours in an oral glucose tolerance test, Diastolic blood pressure (mm Hg), Triceps skin fold thickness (mm), 2-Hour serum insulin (mu U/ml), Body mass index (weight in kg/(height in m)ˆ2), Diabetes pedigree function and age.

Identifying Information

Pre-Processing

Data Standardization is a process of tuning and re-scaling features in such a manner that the resulting attribute has 0 average and a standard deviation of 1. The dataset is normalized by using the Z-score method to ensure its uniformness.



Hardware Measurement

To formulate the values of attributes, the hardware design is created to precisely measure the blood sugar and blood pressure, to input in the model.

DEEP LEARNING MODEL

It is the basic sequential model that can be defined as a network of Dense Layers, which is used for deep learning using Keras. It helps us in making a model consisting of many layers and sustaining the balance of all the layers. A Sequential model is appropriate for a plain stack of layers where each layer has exactly one input tensor and one output tensor.

The basic sequential model which is defined as a network of Dense Layers is used for training of deep learning framework. Keras deep learning framework is used in this work for prediction of diabetes. This model comprises many layers and sustains the balance of all the layers.

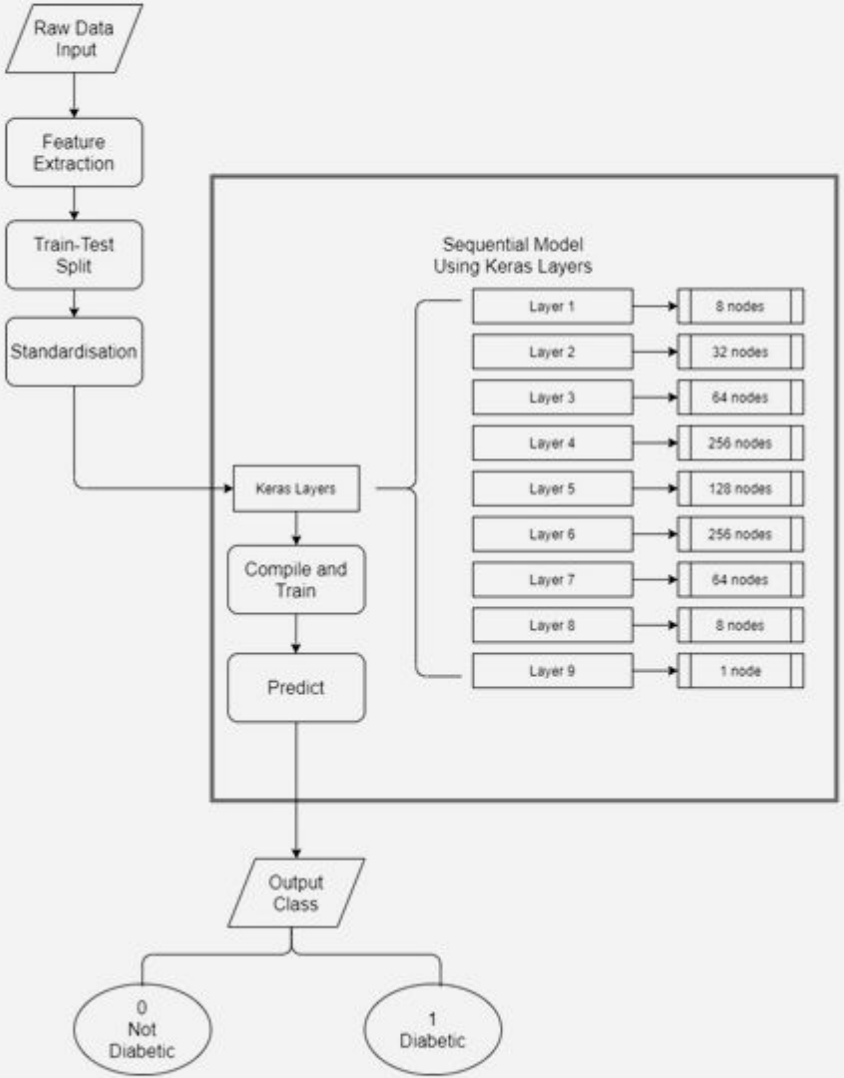
The Keras functional API is a way to create models that are more flexible than the Keras Sequential API. The functional API can handle models with non-linear topology, shared layers, and even multiple inputs or outputs. The functional API is a way to build graphs of layers. In the functional API, models are created by specifying their inputs and outputs in a graph of layers. That means that a single graph of layers can be used to generate multiple models. The functional API makes it easy to manipulate multiple inputs and outputs.

Keras Sequential Model

Keras Functional Model

WEB DEPLOYMENT

The applied Algorithmic model is embedded in a fully-functional website build via flask framework which is deployed using Heroku Hosting service, which ensures the user-availability and portability of the project.


Raw Data Input
Feature Extraction
Train-Test Split
Standardisation
Sequential Model Using Keras Layers
Keras Layers
Compile and Train
Predict
Layer 1
Layer 2
Layer 3
Layer 4
Layer 5
Layer 6
Layer 7
Layer 8
Layer 9
8 nodes
32 nodes
64 nodes
256 nodes
128 nodes
256 nodes
64 nodes
8 nodes
1 node
Output Class
0 Not Diabetic
1 Diabetic

04

Outcome

We have proposed a solution which includes a fusion of software and hardware elements that will be able to process the factor indicating diabetes-like blood sugar level, insulin level, and blood pressure and predict if the person is suffering from Diabetes or not.

The project also estimates the likelihood for a simulated input, to track the chances of future or dummy conditions.

Through the successful deployment of the project, an endeavor with a diabetes-free life is ensured which can proved to prevent pecuniary disaster and loss of loved ones.

05

Contribution

SOFTWARE TEAM

Nishq Poorav Desai
Avanish Sandilya

Would engage in constructing the algorithm for analyzing the parameters to determine if the person is suffering from Diabetes.

SOFTWARE PLANNING TEAM

Rahul Khandebharad

Would gather the Information required, Designing Engaging Presentations and assist the Documentation.

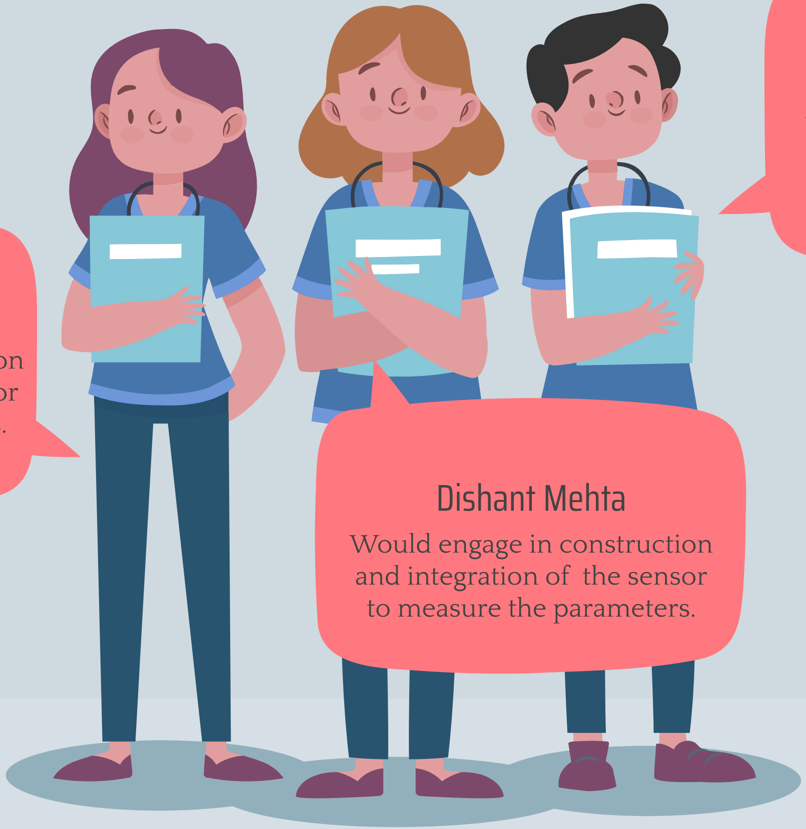
Utkarsha

Would engage in Designing Presentation and Front-End/ UI Design of the website.

Umang Mahesh Rane

Would engage in Project Diary making and Front- End designing of the website.

HARDWARE TEAM



Manasi Muchrikar
Would engage in construction and integration of the sensor to measure the parameters.

Dishant Mehta
Would engage in construction and integration of the sensor to measure the parameters.

Akash Anant Jaiswal
Would engage in construction and integration of the sensor to measure the parameters.

06.References

1. Marco Righettoni and Antonio Tricoli: Toward portable breath acetone analysis for diabetes detection. Journal of Breath Research, Volume 5, Number 3 Published 9 August 2011.

2. Kao, K.-W.; Hsu, M.-C.; Chang, Y.-H.; Gwo, S.; Yeh, J.A. A Sub-ppm Acetone Gas Sensor for Diabetes Detection Using 10 nm Thick Ultrathin InN FETs. Sensors 2012, 12, 7157-7168.

3. Eric A Orzeck, M.D., Joseph H Mooney, M.A. and John A Owen Jr, M.D.: Diabetes Detection with a Comparison of Screening Methods. Diabetes 1971 Feb;

THANKS

Any Questions?