

GlucSense - Diabetes Monitoring System for Pregnant Women

Niya Francis¹, Philip Thomas¹, Shalini Das¹, Thomas U Anthrayose¹, Er. Thomas Joseph²

¹*Department of Computer Science and Engineering, Saintgits College of Engineering,
APJ Abdul Kalam Technological University.*

²*Assistant Professor, Department of Computer Science and Engineering,
Saintgits College of Engineering,
APJ Abdul Kalam Technological University.*

Abstract

Diabetes in pregnant women is a higher case of concern. Poorly treated diabetes during pregnancy causes risk for both maternal and child health. Proper diabetic control during pregnancy stage could reduce this risk to a higher extent. This paper proposes a system in which diabetes level of pregnant women is predicted using certain features and depending on results several diet plans and exercise suggestions are given to control diabetes. Object detection is also incorporated in the system so as to detect the food and find the glucose content in the food item before consuming it. The dataset is PIMA Indians dataset obtained from Kaggle repository.

Keywords: *Machine Learning, Diabetes prediction, Object Detection, Linear Regression*

1. Introduction

Diabetes is a chronic disease condition affecting human body due to abnormal increase in blood sugar level. Onset of diabetes causes different types of diseases such as heart attack, blindness, kidney disease, etc. Diabetes in pregnant women is even more crucial, as in this period body undergoes dramatic changes. With respect to pregnant ladies, diabetes is of mainly two types: pre-existing diabetes and gestational diabetes.

Pre-existing diabetes also known as pre-gestational diabetes is mostly having type1 (insulin-dependent or juvenile) diabetes and some may have type2 (insulin-independent or maturity-onset) diabetes. A pregnant woman with diabetes is having higher risk than a normal woman since it affects both the mother and the child, which include birth defect to child, jaundice, low blood glucose level, premature birth etc. Miscarriage and high blood pressure are other severe causes.

Gestational diabetes is a medical condition of being diabetic during pregnancy which is given by the degree of glucose variations with onset or first recognition during pregnancy. So it is very important to have a system by which one can monitor diabetes by providing certain basic information about the individual.

The proposed system is specifically a web application in which one can provide these information, which forms the factors used to predict the diabetes level. Then the internal processing occurs. The final result based on analysis is displayed to the user along with the diet plan and the system also provides the functionality to determine the amount of glucose content in specified food consumed by user so that the user could control the diabetic level in an easy manner.

2. Literature Review

Many approaches had made to predict the age, disease prediction related to diabetes. This section explains some of those approaches.

[1] Mustafa S Kadhmi et al

They proposed a diabetes prediction system where initially K-nearest neighbor algorithm is used for eliminating the unwanted data and classification method used is Decision Tree. The system has achieved an accuracy of 98.7% compared to other systems based on Pima Indians Diabetes dataset.

[2] Kav kiotis et al

They compared three different algorithms using k fold cross validation with a fold of 10 as evaluation criterion, the algorithms included are Naïve Bayes, SVM and logistic regression, where better performance and accuracy of 84 % is shown by SVM than other algorithms.

[3] Zheng et al

They tried to predict diabetes mellitus at early stage by using various algorithms Random Forest, KNN, SVM, Naive Bayes, decision tree and logistic regression where for further improvement filtering criteria can be improved.

[4] Balakrishnan Valliyot et al

They studied the risk factors associated with type 2 diabetes mellitus (T2DM). The prime focus was on finding the means, binary logistics regression and multiple binary logistics regression to assess the significance of the observation.

3. Machine Learning Algorithm and Technique used for Diabetes Prediction

Now a days Machine learning methods are widely used to predict the diabetes. Machine Learning enables a machine to learn automatically from experience and to predict correctly when new instances occur. Machine learning algorithm used in the proposed system is Linear Regression. There are several techniques that can be chosen to detect objects. One such popular technique is YOLO [5]. These techniques are applied on web application. Algorithms and techniques like Linear Regression Algorithm, YOLO are used.

A. Linear Regression

Linear Regression is a statistical model used to predict the relationship between independent and dependent variable. Mainly by examine two types of variables, firstly which variable in particular are significant prediction of the outcome variable and second is how significant is the regression line to make prediction with highest possible accuracy. Linear Regression is mainly based on drawing a line through data. The simplest form of simple linear equation with one dependent and one independent variable is represented by [7],

$$Y=M*X+C$$

Y=dependent variable

X=independent variable

M=slope

C=co-efficient of line

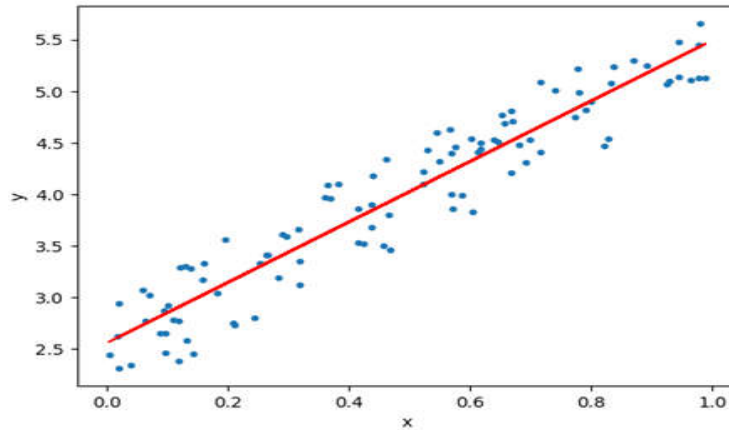


Figure [1]: Linear Regression

B. YOLO

YOLO is used for real-time object detection. The algorithm applies the entire image to a single neural network, then divides the image into regions and predicts bounding boxes and probabilities for each region. Those bounding boxes are weighted by the probabilities predicted.

YOLO is famous because it achieves high precision, while being able to run in real time as well. The algorithm "only looks at the image once" in the sense that it requires only one forward spread to pass through the neural network to make predictions.

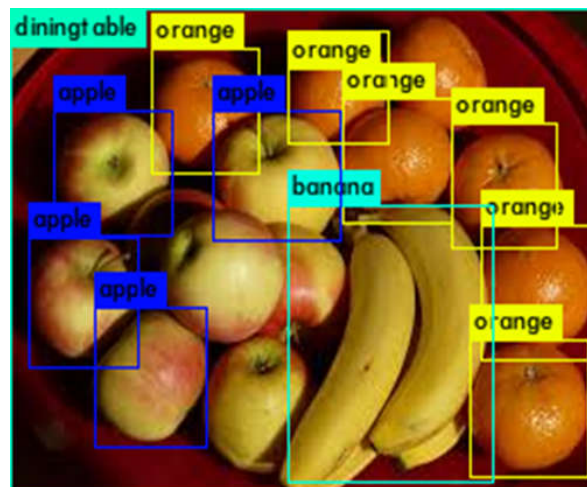


Figure [2]: YOLO Object Detection

4. Proposed Model

The proposed system gives a clear solution for poorly treated diabetes pregnancy risks. The system includes a web interface that helps the user to see the predicted diabetes level along with the diet plan. We incorporated 4 modules in this system.

1. Prediction Module
2. Object Detection Module
3. Warning Module
4. Useful Measures Module

Prediction Module: In this module, diabetes level of pregnant women is predicted using features like number of pregnancies had, insulin level, BMI, blood pressure, skin thickness etc. the predicted output will be shown as a result. For this we use Linear Regression algorithm.

Object Detection Module: In this module we will capture the image of the food that a pregnant woman going to consume and the system will identify the object. After identification of the object, the system will classify the object according to the given calorie measure. For object Detection we use YOLO and SSD algorithms.

Warning Module: In this module we will compare the current diabetes level amount and identified object calorie measure amount. If the sum of both is higher than safer limit it will give a caution message.

Useful Measures Module: In this module the system will provide useful diet plans and exercise tips for a healthy pregnancy period. It will be beneficial for both mother and child.

The data flow for the system is shown below.

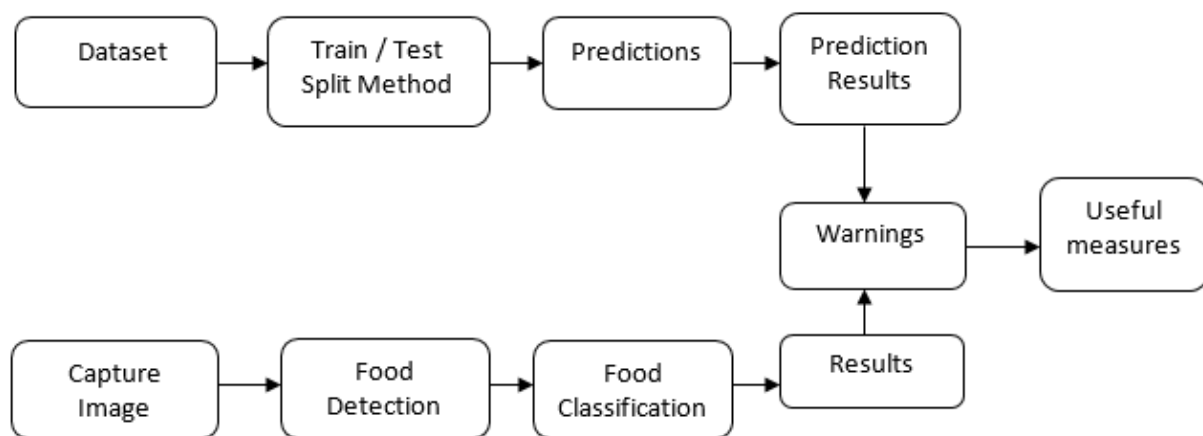


Figure [3]: Dataflow diagram

Datasets are an important part of the field of machine learning. Data munging is the initial process of refining raw data into content for consumption by downstream systems and users. In Train/Test Split, the data we use is usually split into training data and test data. The training set take more time than testing and contains a known output and the model learns on this data in order to be generalized to other data later on. Precisely in our model, prediction is used to make the person understand what he has to do according to his/her situation. After that the image of the food item will be captured and then identify what food item is that. After identification it will classify the food item. Then the system will predict the calorie value of the identified food. Both the calculated result of predicted glucose level value and food calorie value will sum up. If the new calculated value ie, the sum up value is higher than the limiting point it will move to the warning module and that module will give corresponding measures to neutralize the effect, reciprocal otherwise

5. Experiment and Result

The experimental data was taken from the PIMA Indians dataset obtained from Kaggle repository. The dataset contains information of females having at least 21 years of age and a total of 768 instances. Each tuple contains medical information like number of pregnancies had, blood pressure, skin thickness, insulin, bmi, glucose level and age. Here glucose is taken as the target variable and the rest forms the dependent variables. The dataset is used to train a Linear Regression model. Linear Regression is used as the inputs are continuous and the predictions are multi class. The model is then linked to the webpage using Flask. So the model can make predictions on the glucose level from the values received from the webpage.

A glucose level between 90-150 mg/dL is considered as normal, 150-240 mg/dL is considered as high, 240-300 mg/dL is considered as very high and greater than 300 mg/dL is considered as out of control in this system. With the limited dataset, the results of predictions got an R^2 Score of 0.21466796422482926.

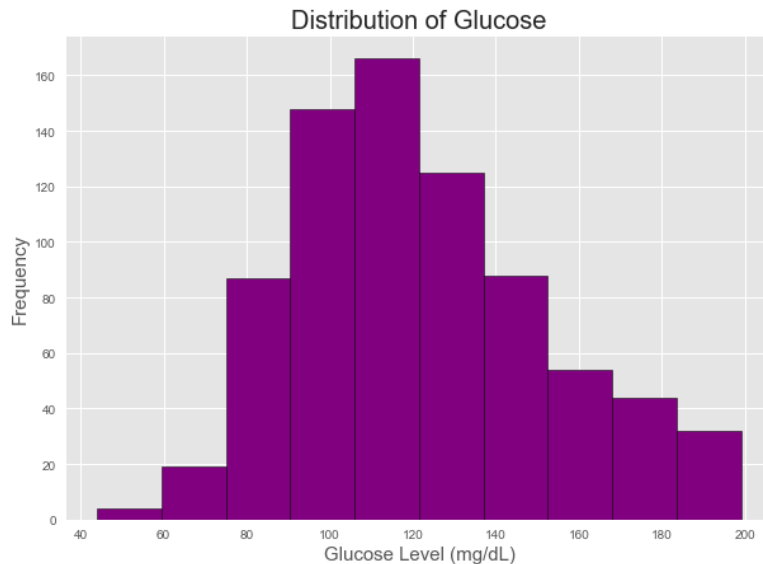


Figure [4]: Distribution of Glucose levels

GlucoSense

Please fill up the following information

Enter your age:

No. of pregnancies:

Blood Pressure (diastolic):

Skin Thickness:

Insulin Level:

BMI (weight(kg)/height(m)):

Figure [5]: Prediction Module Interface

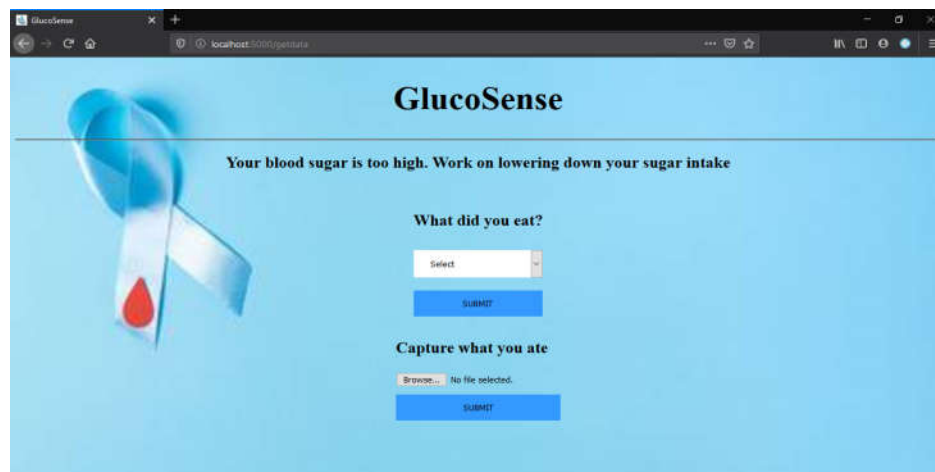


Figure [6]: Warning and Object Detection Interface 1

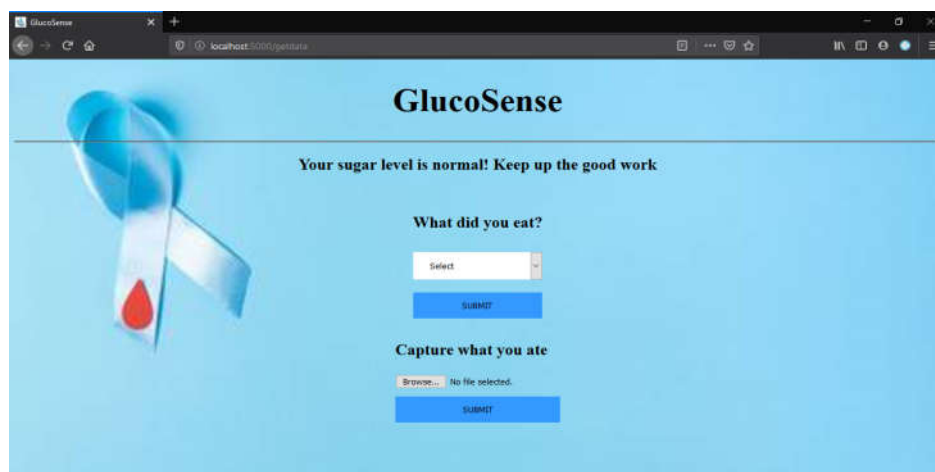


Figure [7]: Warning and Object Detection Interface 2

After making the predictions, appropriate warning messages are shown to the user. Some useful measure to help them maintain a descent level of glucose is also displayed.

The user is also given an option to click a picture of what she is going to eat. The picture is taken for object detection using YoloV3 and the algorithm detects the food. The calorie value of the food is calculated and compared with daily permissible calorie consumption limits. A pie-chart is displayed showing the daily calorie consumption of the user. If that also exceeds permissible limits, appropriate warning messages are shown.

6. Conclusion

GlucoSense is exclusively a system for pregnant women who suffers from non-periodic changes in sugar level during pregnancy period. Our system helps to predict the current diabetes level by entering some factors. Object detection is also incorporated in the system so as to detect the calorie content in each food item before consuming it.

As a result we have developed a model where diabetes is predicted for pregnant women and ways to indulge a method or other means where that person can follow a strict plan so as to reduce their chances of acquiring this deadly disease or a chance to reduce the impact of this disease.

References

- [1] Kadhm, Mustafa, Ghindawi, Ikhlas & Enteesha, Duaa, “An Accurate Diabetes Prediction System Based on K-means Clustering and Proposed Classification Approach”, International Journal of Applied Engineering Research, 2018.
- [2] Kavakiotis, Ioannis, Olga Tsave, AthanasiosSalifoglou, NicosMaglaveras, IoannisVlahavas, and IoannaChouvarda. "Machine Learning and Data Mining Methods in Diabetes Research." Computational and Structural Biotechnology Journal (2017).
- [3] Zheng, Tao et al. "A machine learning-based framework to identify type 2 diabetes through electronic health records", International journal of medical informatics 97 (2017): 120- 127.
- [4] Balakrishnan valliyot et al., “Risk factors of type 2 diabetes mellitus in the rural population of north kerala, india: a case control study”, Journal of Diabetologia croatica, Vol. 4, 2013.
- [5] Rachel Huang, Jonathan Pedoeem, Cuixian Chen, “YOLO-LITE: A RealTime Object Detection Algorithm Optimized for Non-GPU Computers”, 2018 IEEE International Conference on Big Data (Big Data), DOI: 10.1109/BigData.2018.8621865.
- [6] <https://www.kaggle.com/uciml/pima-indians-diabetes-database>
- [7] Simplilearn <https://www.youtube.com/watch?v=NUXdtN1W1FE>