JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A MODEL FOR PREDICTING INSULIN DOSAGE FOR DIABETIC PATIENTS USING MACHINE LEARNING - Review

K.V.M. SUDHEER KUMAR
Department of Master of Computer Science
Miracle Educational Society Group of Institutions
Vizianagram– 535216 (AP) India

SARAGADAM SRIDHAR
Department of Master of Computer Science
Miracle Educational Society Group of Institutions
Vizianagram– 535216 (AP) India

Abstract

Diabetes is a chronic metabolic disorder characterized by high levels of glucose (blood sugar) in the bloodstream. The body normally regulates blood sugar levels through the hormone insulin, which is produced by the pancreas. However, in diabetes, there is either insufficient insulin production or the body's cells do not respond effectively to insulin, leading to elevated blood sugar levels. Managing diabetes involves maintaining blood sugar levels within a target range through a combination of medication, diet, regular physical activity, and monitoring blood sugar levels. Uncontrolled diabetes can lead to various complications, including cardiovascular disease, kidney damage, nerve damage, and eye problems. Gradient Boosting Classifier for predicting diabetes and the Linear Regression algorithm for predicting insulin dosage in diabetic patients. You plan to use the PIMA diabetes dataset for training the models and the UCI insulin dosage dataset for predicting insulin dosage. You have chosen the PIMA diabetes dataset for training the Gradient Boosting Classifier and the UCI insulin dosage dataset for predicting insulin dosage. Make sure you have access to these datasets and that they are properly formatted for your machine learning algorithms. Before training the models, you might need to preprocess the datasets. This may involve handling missing values, normalizing or standardizing the features, and splitting the data into training and testing sets. Use the PIMA diabetes dataset to train the Gradient Boosting Classifier. This algorithm will learn patterns and relationships in the data to predict the presence of diabetes. Once the classifier is trained, you will upload a test dataset with no class labels. Use the trained model to predict the presence of diabetes for each sample in the test dataset. For the samples predicted to have diabetes by the Gradient Boosting Classifier, you can use the UCI insulin dosage dataset to predict the insulin dosage. Preprocess the dataset as necessary and extract relevant features for insulin dosage prediction. The preprocessed UCI insulin dosage dataset to train a Linear Regression model. This model will learn the relationship between the input features and the insulin dosage. Once the Linear Regression model is trained, apply it to the samples that were predicted to have diabetes by the GBC. The model will predict the insulin dosage for each sample. Evaluate the performance of both the Gradient Boosting Classifier and the Linear Regression model. You can use metrics such as accuracy, precision, recall, MSE to assess the models' performance.

Keywords— mean squared error (MSE), Gradient Boosting Classifier (GBC).

Introduction

A group of metabolic diseases known as diabetes mellitus, more commonly referred to as diabetes, is characterized by high blood glucose concentrations due to malfunctions in insulin secretion, insulin action, or both. Diabetes pervasiveness has been rising all the more quickly in center and low-pay nations. From 108 million people in 1980 to 422 million in 2014, diabetes has increased. Diabetes symptoms can be brought on by either a lack of insulin or an inability to properly respond to insulin. To keep glucose levels under control, diabetics must take the prescribed dose of insulin. The specialist of the patient additionally needs to realize the necessary insulin portion from past records of dosages and from patient's ongoing determined glucose level There are three principal sorts of diabetes: Type 1 diabetes, which includes type 1, type 2, and gestational diabetes, is thought to be brought on by an autoimmune reaction, or the body attacking itself by mistake. Insulin production is halted by this reaction. Type 1 diabetes affects between 5 and 10% of diabetes patients. Side effects of type 1 diabetes frequently grow rapidly. Most of the time, children, teens, and young adults are diagnosed with it. You will need to take insulin every day to live if you have type 1 diabetes. There is currently no known method for preventing type 1 diabetesDiabetes Type 2: Your body doesn't use insulin well and can't keep your blood sugar levels normal in type 2. Type 2 diabetes accounts for 95% to 95% of diabetes patients. It usually affects adults but increasingly affects children, teens, and young adults over a long period of time. If you are at risk, it is essential to have your blood sugar tested even if you don't notice any symptoms. Changing a healthy lifestyle can help prevent or delay type 2 diabetes, such as:

- Getting in shape.
- Consuming nutritious food
- Being involved.

Diabetes During pregnancy, pregnant women who have never had diabetes develop gestational diabetes. If you have gestational diabetes, your child may be more likely to have health issues. After your baby is born, gestational diabetes typically disappears. However, it makes you more likely to develop type 2 diabetes in later life. Your baby has a higher risk of becoming obese as a child or teen and later developing type 2 diabetes.In the US, 96 million grown-ups — more than 1 out of 3 — have pre diabetes. Over 80% of them are unaware that they have it. Pre diabetes is a condition where blood sugar levels are higher than normal but not yet in the range for a diagnosis of type 2 diabetes. It is considered a warning sign. You are more likely to get type 2 diabetes, heart disease, and stroke if you have pre diabetes. However, there is good news. A lifestyle change program approved by the CDC can assist you in taking healthy steps to reverse pre diabetes. The forecast of glucose focuses could work with the suitable patient response in critical circumstances like hypoglycemia. As a result, advanced data-driven methods have been taken into consideration in a number of recent studies for the creation of precise predictive models of glucose metabolism. Notwithstanding the overall rules that the patient adheres to during his regular routine, a few diabetes the board frameworks have been proposed to additional help the patient in the self- administration of the sickness. One of the fundamental parts of a diabetes the board framework concerns the prescient displaying of the glucose digestion.

Proposed System

Using the PIMA diabetes dataset and UCI insulin dosage dataset, you're training two different algorithms: Gradient Boosting for predicting the presence of diabetes and Linear Regression for predicting insulin dosage if diabetes is detected by Gradient Boosting.

Dataset selection: You have chosen the PIMA diabetes dataset and UCI insulin dosage dataset for your project.

Training phase:

a. Using the PIMA diabetes dataset, you train a Gradient Boosting algorithm to predict the presence of diabetes.
b. Additionally, you train a Linear Regression model on the UCI insulin dosage dataset to predict insulin dosage if diabetes is detected.

Test phase:

a. You upload a test dataset that does not have class labels (diabetes presence) to evaluate the performance of your trained models.
b. Using the trained Gradient Boosting model, you predict the presence of diabetes for each instance in the test dataset.
c. If diabetes is detected by Gradient Boosting, you use the trained Linear Regression model to predict the insulin dosage.

By combining these two algorithms, you aim to predict the presence of diabetes and estimate the insulin dosage if diabetes is detected. Gradient Boosting provides the initial diabetes prediction, and Linear Regression helps in estimating the insulin dosage based on that prediction.

Advantages of Proposed system:

- Individual blood glucose level and insulin dosing are highly erratic and precise prediction of them is likely impractical.
- Average blood glucose level over 24 hours can be more reliably predicted and determining whether the patient's glucose level is going to be high is a more feasible task.
- It helps diabetic patients to get proper amount of insulin dosage.
- When insulin is taken in right time the body can function properly and it can save the life of the person.

Gradient boosting algorithm:

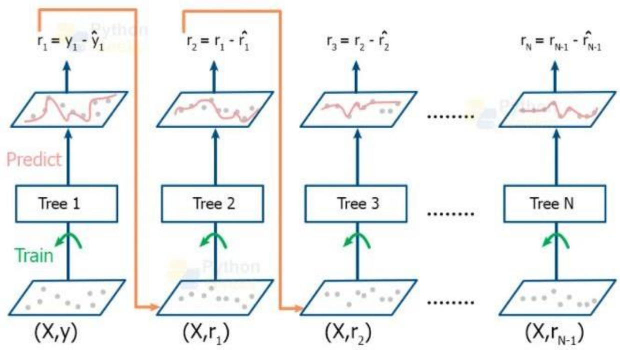
Gradient boosting is a powerful machine learning algorithm that is used to minimize both bias and variance errors in a model. While it is true that gradient boosting can help minimize bias error, it is not accurate to say that its purpose is solely to minimize bias error. Gradient boosting is a general boosting algorithm that aims to improve the performance of a model by iteratively combining weak learners (base estimators) to create a strong learner.

Unlike AdaBoost, where the base estimator can be specified by the user, the base estimator in the gradient boosting algorithm is fixed and is often a decision stump (a shallow decision tree with only one split). However, it is important to note that gradient boosting is not limited to using only decision stumps as base estimators. In fact, gradient boosting can utilize a variety of weak learners, such as decision trees, regression models, or even neural networks.

The number of iterations or estimators in the gradient boosting algorithm, often denoted as n_estimators, can be tuned to find the optimal balance between bias and variance. By increasing the number of estimators, the model can become more complex and potentially overfit the data, leading to higher variance. Conversely, using a smaller number of estimators may result in underfitting and higher bias. The default value for n_estimators in many implementations is often set to 100, but it can be adjusted based on the specific problem and data.

When used as a regressor, the cost function or loss function that is minimized during training is typically the Mean Square Error (MSE). The algorithm aims to minimize the difference between the predicted values and the actual target values by iteratively fitting new weak learners to the residuals. On the other hand, when used as a classifier, the cost function is typically the Log Loss (also known as cross- entropy loss). The algorithm minimizes the log loss by iteratively fitting new weak learners to the negative gradients of the log loss function.

In summary, gradient boosting is a versatile algorithm that can be used to minimize both bias and variance errors. It is not limited to using decision stumps as base estimators, and the number of estimators can be tuned to find the optimal trade-off between bias and variance. The choice of cost function depends on the task at hand, with MSE commonly used for regression and log loss for classification.



Gradient Boosting Algorithm

Gradient Boosting Works

The working of gradient boosting revolves around the three main elements. These are as follows:

- A loss function
- A weak learner
- An additive model

Loss Function

The objective in machine learning is to optimize the loss function. The loss function quantifies the discrepancy between the predicted values of the model and the actual target values, and the goal is to minimize this discrepancy during the training process.

As an example, we can say that regression can use the squared error & classification can use the algorithmic loss. One of the best things about gradient boosting is that with each framework a fresh boosting algorithm is not required for every loss function in question. Thus, a more generic framework would suffice.

Weak Learner

Weak learners are for the purpose of making predictions. A decision tree is basically a weak learner. Specific regression trees are used for the real output values that are used for splits. We can correct the reminders in the prediction models.

Additive Model

There are no modifications to pre-existing trees in the model, but there is the addition of a greater number of trees at a time.

At the time of adding the trees, a gradient descent procedure minimizes the losses. It minimizes the set number of parameters. In order to decrease the error, there is an updation of the weights only after

Advantages of Gradient Boosting

- Most of the time predictive accuracy of gradient boosting algorithm on higher side.
- It provides lots of flexibility and can optimize on different loss functions and provides several hyper parameter tuning options that make the function fit very flexible.
- Most of the time no data pre-processing required.
- Gradient Boosting algorithm works great with categorical and numerical data.
- Handles missing data — missing value imputation not required.

Disadvantages of Gradient Boosting

- Gradient Boosting Models will continue improving to minimize all errors. This can overemphasize outliers and cause over fitting. Must use cross-validation to neutralize.

- It is computationally very expensive — GBMs often require many trees (>1000) which can be time and memory exhaustive.

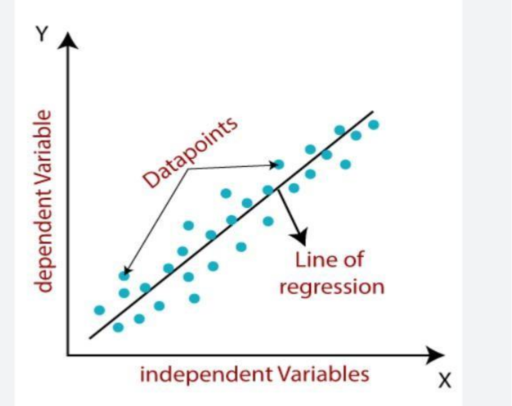
- The high flexibility results in many parameters that interact and influence heavily the behavior of the approach (number of iterations, tree depth, regularization parameters, etc.). This requires a large grid search during tuning.

Linear Regression Algorithm:

Linear Regression Algorithm is one of the easiest and most popular Machine Learning algorithms. It is a statistical method that is used for predictive analysis. Linear regression makes predictions for continuous/real or numeric variables such as sales, salary, age, product price, etc. Linear regression algorithm shows a linear relationship between a dependent (y) and one or more independent (y) variables, hence called as linear regression. Since linear regression shows the linear relationship, which means it finds how the value of the dependent variable is changing according to the value of the independent variable.

Fig Linear Regression

$y= a_0+a_1x+ \varepsilon$



Here,

Y=Dependent variable (Target Variable)

X=Independent Variable (Predictor Variable)

A0=Intercept of the line

A1=Linear Regression coefficient

E=Random error

TYPES OF LINEAR REGRESSION:

Linear regression can be further divided into two types of the algorithm:

Simple Linear Regression: Simple Linear Regression models the relationship between the independent variable (often denoted as "x") and the dependent variable (often denoted as "y") using a straight line. The objective of Simple Linear Regression is to find the best-fitting line that minimizes the difference between the predicted values and the actual values of the dependent variable.

Multiple Linear Regression:

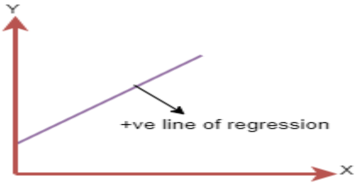
When more than one independent variable is used to predict the value of a numerical dependent variable, the linear regression algorithm is referred to as Multiple Linear Regression. Multiple Linear Regression extends the concept of Simple Linear Regression by incorporating multiple independent variables to model the relationship with the dependent variable.

Linear Regression Line:

a regression line represents the relationship between the dependent variable and one or more independent variables in a linear regression model. It is also known as a best-fit line or a line of best fit.

Positive Linear Relationship: In a positive relationship, as the values of the independent variable increase, the values of the dependent variable also tend to increase. The regression line has a positive slope, indicating that for every unit increase in the independent variable, the dependent variable is expected to increase by a certain amount. This indicates a direct or positive correlation between the variables.

a) Positive Linear Regression



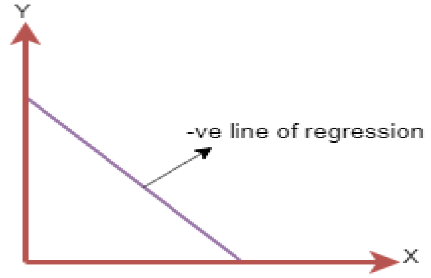
The line equation will be: $Y= a_0+a_1x$

Negative Linear Relationship:

In a negative relationship, as the values of the independent variable increase, the values of the dependent variable tend to decrease. The regression line has a negative slope, indicating that for every unit increase in the independent variable, the dependent variable is expected to decrease by a certain amount. This indicates an inverse or negative correlation between the variables Negative linear line.

Advantages

1. Linear regression is effective at determining the nature of the relationship between various variables in linear datasets.



The line of equation will be: $Y= -a_0+a_1x$

2. Both the linear regression models and the linear regression algorithms are simple to train.

3. However, dimensionality reduction methods like regularization (L1 and L2) and cross-validation can prevent Linear Regression models from overfitting.

Disadvantages

1. A significant impediment of Straight Relapse is that it accepts linearity between the reliant and free factors, which is seldom addressed in certifiable information. It makes the unlikely assumption that the dependent and independent variables have a straight line relationship.

2. Overfitting and noise are common problems. Linear regression may not be a good option for datasets where the number of observations is less than the number of attributes because it can cause overfitting. This is because the algorithm is able to begin taking into account the noise as it builds the model.

3. Because it is sensitive to outliers, Linear Regression must be applied to the data after the dataset has been pre-processed and the outliers have been removed.

4. It doesn't accept multi collinearity. Prior to applying Linear Regression, any relationship between the independent variables—also known as multicollinearity—must be eliminated using dimensionality reduction methods because the algorithm assumes that there is no such relationship.

Benefits

Benefits of Linear Regression In statistics, linear regression is common. In the field of data science, it has the following advantages:

1. Simple to Carry out A Straight Relapse AI model is computationally basic and doesn't need a lot designing above. Consequently, it is not difficult to carry out and keep up with.

2. Scalability Linear regression can be used in situations where scaling is required, such as applications that deal with large amounts of data, due to its low computational cost.

3. Ability to Interpret Linear Regression is Very Efficient to Train and Easy to Interpret. It is moderately straightforward, dissimilar to profound learning brain networks which require more information and time to prepare productively.

Proposed Method:

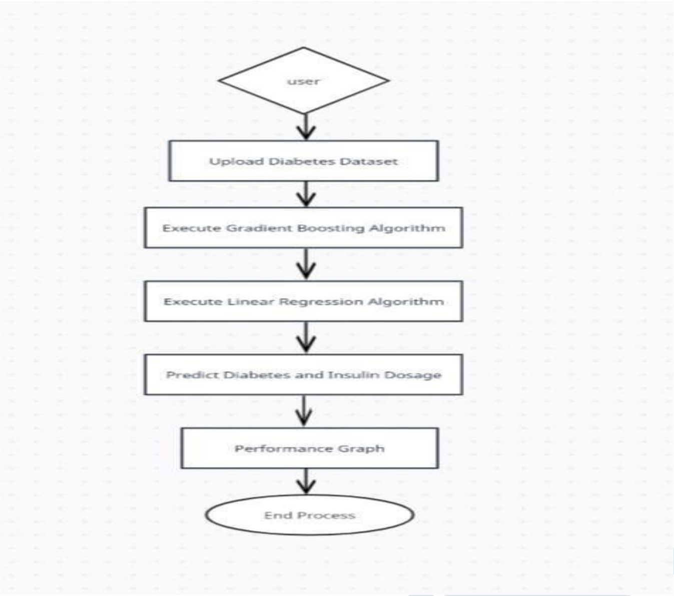

user
Upload Diabetes Dataset
Execute Gradient Boosting Algorithm
Execute Linear Regression Algorithm
Predict Diabetes and Insulin Dosage
Performance Graph
End Process

Fig – Data Flow Diagram

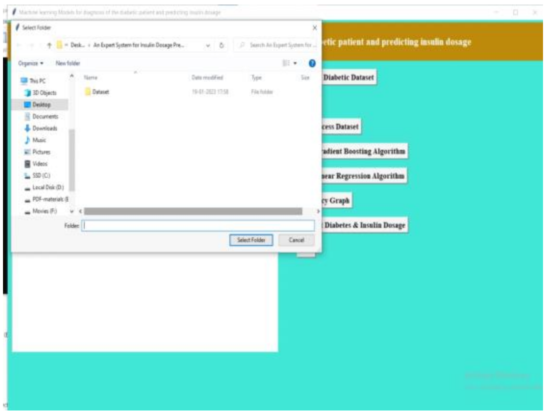
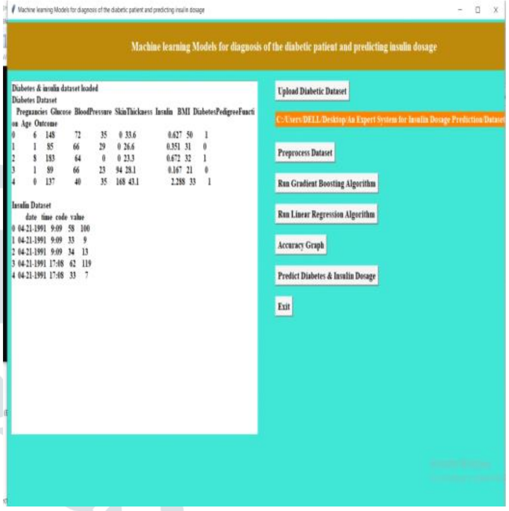


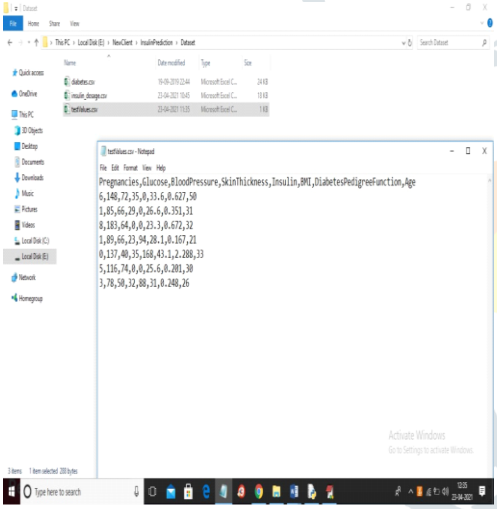
Fig- Diabetic dataset


Machine learning Models for diagnosis of the diabetic patient and predicting insulin dosage
Upload Diabetic Dataset
C:/Users/DELL/Desktop/An Expert System for Insulin Dosage Prediction/Dataset
Preprocess Dataset
Run Gradient Boosting Algorithm
Run Linear Regression Algorithm
Accuracy Graph
Predict Diabetes & Insulin Dosage
Exit

```
Diabetes & insulin dataset loaded
Diabetes Dataset
   Pregnancies  Glucose  BloodPressure  SkinThickness  Insulin   BMI  DiabetesPedigreeFuncti
on  Age  Outcome
0            6      148             72             35        0  33.6       0.627  50        1
1            1       85             66             29        0  26.6       0.351  31        0
2            8      183             64              0        0  23.3       0.672  32        1
3            1       89             66             23       94  28.1       0.167  21        0
4            0      137             40             35      168  43.1       2.288  33        1

Insulin Dataset
        date   time  code  value
0 04-21-1991   9:09    58    100
1 04-21-1991   9:09    33      9
2 04-21-1991   9:09    34     13
3 04-21-1991  17:08    62    119
4 04-21-1991  17:08    33      7
```

Dataset loaded



Name	Date modified	Type	Size
diabetes.csv	19-09-2019 22:44	Microsoft Excel C...	24 KB
insulin_dosage.csv	23-04-2021 10:45	Microsoft Excel C...	18 KB
testValues.csv	23-04-2021 11:35	Microsoft Excel C...	1 KB

```
Pregnancies,Glucose,BloodPressure,SkinThickness,Insulin,BMI,DiabetesPedigreeFunction,Age
6,148,72,35,0,33.6,0.627,50
1,85,66,29,0,26.6,0.351,31
8,183,64,0,0,23.3,0.672,32
1,89,66,23,94,28.1,0.167,21
0,137,40,35,168,43.1,2.288,33
5,116,74,0,0,25.6,0.201,30
3,78,50,32,88,31,0.248,26
```

Fig- Dataset

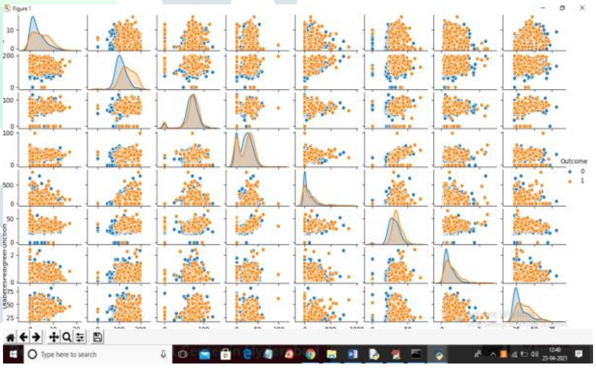


Fig- Graph

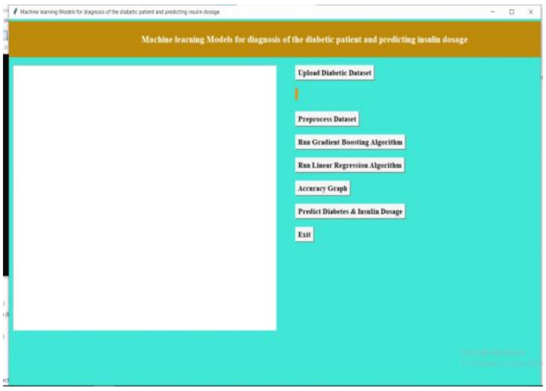
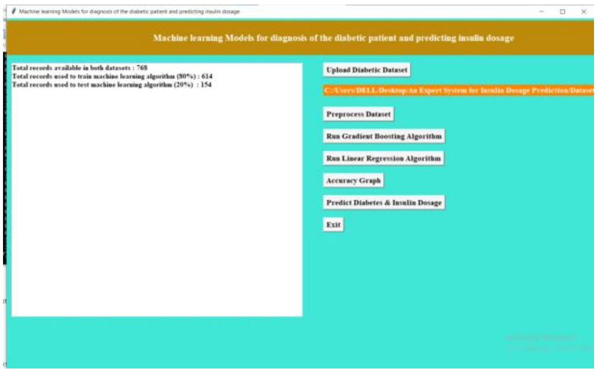
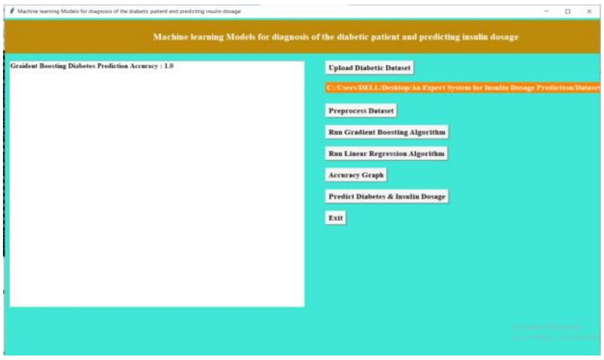

Machine learning Models for diagnosis of the diabetic patient and predicting insulin dosage
Upload Diabetic Dataset
Preprocess Dataset
Run Gradient Boosting Algorithm
Run Linear Regression Algorithm
Accuracy Graph
Predict Diabetes & Insulin Dosage
Exit

Fig- Run Code


Machine learning Models for diagnosis of the diabetic patient and predicting insulin dosage
Upload Diabetic Dataset
C:/Users/DELL/Desktop/An Expert System for Insulin Dosage Prediction/Dataset
Preprocess Dataset
Run Gradient Boosting Algorithm
Run Linear Regression Algorithm
Accuracy Graph
Predict Diabetes & Insulin Dosage
Exit

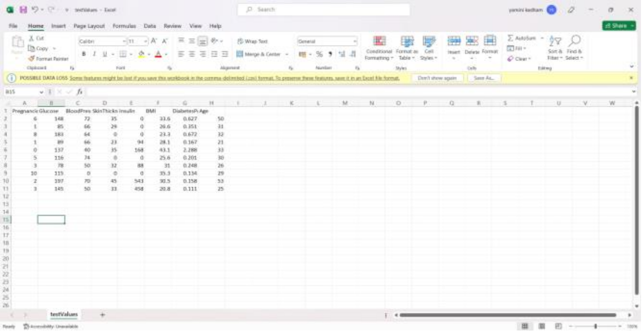
```
Total records available in both datasets : 768
Total records used to train machine learning algorithm (80%) : 614
Total records used to test machine learning algorithm (20%) : 154
```

Fig-Preprocess Dataset


Machine learning Models for diagnosis of the diabetic patient and predicting insulin dosage
Machine learning Models for diagnosis of the diabetic patient and predicting insulin dosage
Upload Diabetic Dataset
C:/Users/DELL/Desktop/An Expert System for Insulin Dosage Prediction/Dataset
Preprocess Dataset
Run Gradient Boosting Algorithm
Run Linear Regression Algorithm
Accuracy Graph
Predict Diabetes & Insulin Dosage
Exit

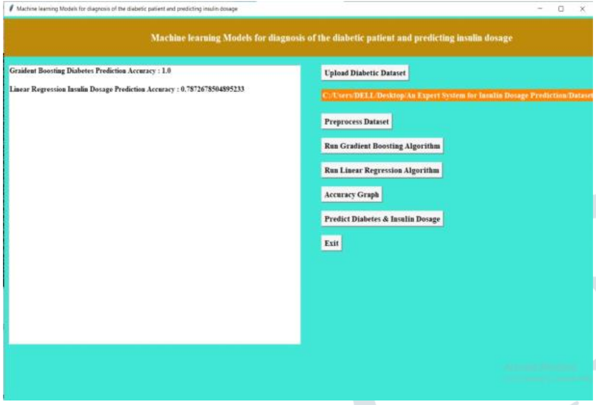
Graident Boosting Diabetes Prediction Accuracy : 1.0

Fig- Gradient Boosting Algorithm


testValues - Excel
File Home Insert Page Layout Formulas Data Review View Help
POSSIBLE DATA LOSS Some features might be lost if you save this workbook in the comma-delimited (.csv) format.
B15
testValues

Pregnancie	Glucose	BloodPres	SkinThickn	Insulin	BMI	DiabetesPi	Age
6	148	72	35	0	33.6	0.627	50
1	85	66	29	0	26.6	0.351	31
8	183	64	0	0	23.3	0.672	32
1	89	66	23	94	28.1	0.167	21
0	137	40	35	168	43.1	2.288	33
5	116	74	0	0	25.6	0.201	30
3	78	50	32	88	31	0.248	26
10	115	0	0	0	35.3	0.134	29
2	197	70	45	543	30.5	0.158	53
3	145	50	33	458	20.8	0.111	25

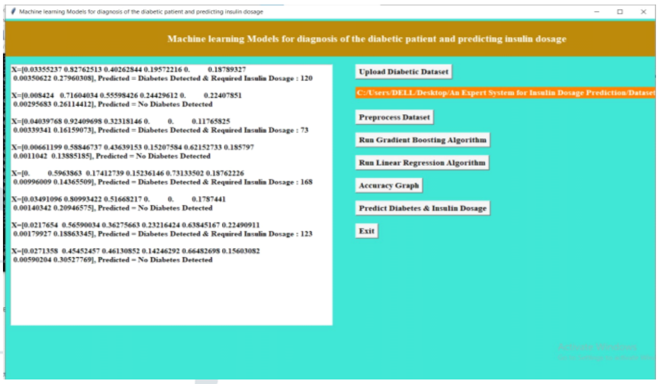
Fig-Result-1


Machine learning Models for diagnosis of the diabetic patient and predicting insulin dosage
Machine learning Models for diagnosis of the diabetic patient and predicting insulin dosage
Upload Diabetic Dataset
C:/Users/DELL/Desktop/An Expert System for Insulin Dosage Prediction/Dataset
Preprocess Dataset
Run Gradient Boosting Algorithm
Run Linear Regression Algorithm
Accuracy Graph
Predict Diabetes & Insulin Dosage
Exit

Graident Boosting Diabetes Prediction Accuracy : 1.0

Linear Regression Insulin Dosage Prediction Accuracy : 0.7872678504895233

Fig-Linear Regression Algorithm


Machine learning Models for diagnosis of the diabetic patient and predicting insulin dosage
Machine learning Models for diagnosis of the diabetic patient and predicting insulin dosage
Upload Diabetic Dataset
C:/Users/DELL/Desktop/An Expert System for Insulin Dosage Prediction/Dataset
Preprocess Dataset
Run Gradient Boosting Algorithm
Run Linear Regression Algorithm
Accuracy Graph
Predict Diabetes & Insulin Dosage
Exit

```
X=[0.03355237 0.82762513 0.40262844 0.19572216 0.         0.18789327
 0.00350622 0.27960308], Predicted = Diabetes Detected & Required Insulin Dosage : 120

X=[0.008424   0.71604034 0.55598426 0.24429612 0.         0.22407851
 0.00295683 0.26114412], Predicted = No Diabetes Detected

X=[0.04039768 0.92409698 0.32318146 0.         0.         0.11765825
 0.00339341 0.16159073], Predicted = Diabetes Detected & Required Insulin Dosage : 73

X=[0.00661199 0.58846737 0.43639153 0.15207584 0.62152733 0.185797
 0.0011042  0.13885185], Predicted = No Diabetes Detected

X=[0.         0.5963863  0.17412739 0.15236146 0.73133502 0.18762226
 0.00996009 0.14365509], Predicted = Diabetes Detected & Required Insulin Dosage : 168

X=[0.03491096 0.80993422 0.51668217 0.         0.         0.1787441
 0.00140342 0.20946575], Predicted = No Diabetes Detected

X=[0.0217654  0.56590034 0.36275663 0.23216424 0.63845167 0.22490911
 0.00179927 0.18863345], Predicted = Diabetes Detected & Required Insulin Dosage : 123

X=[0.0271358  0.45452457 0.46130852 0.14246292 0.66482698 0.15603082
 0.00590204 0.30527769], Predicted = No Diabetes Detected
```

Fig – Result-2

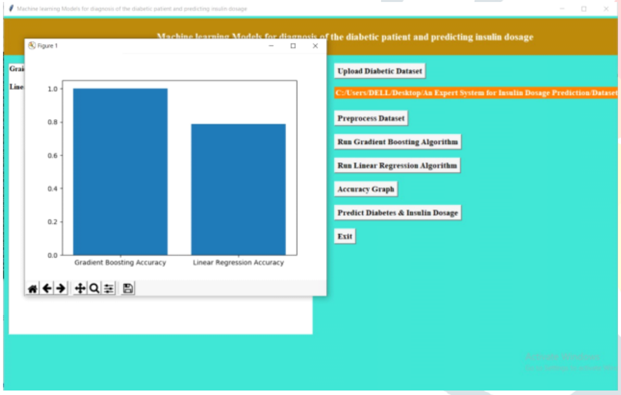


Fig: Accuracy graph

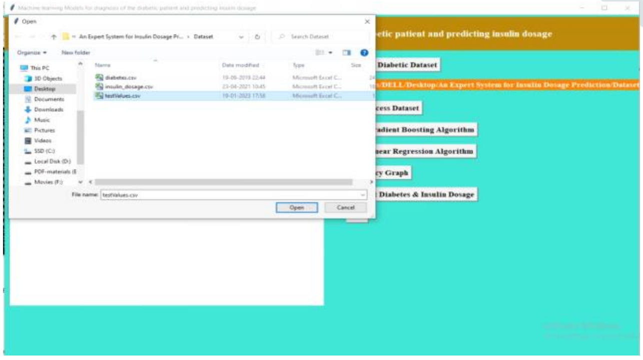

Machine learning Models for diagnosis of the diabetic patient and predicting insulin dosage
Open
An Expert System for Insulin Dosage Pr... › Dataset
Search Dataset
Organize
New folder
This PC
3D Objects
Desktop
Documents
Downloads
Music
Pictures
Videos
SSD (C:)
Local Disk (D:)
PDF-materials (E:)
Movies (F:)
Name
Date modified
Type
Size
diabetes.csv
insulin_dosage.csv
testValues.csv
16-06-2019 22:44
23-06-2021 18:45
19-01-2023 17:58
Microsoft Excel C...
File name: testValues.csv
Open
Cancel
tic patient and predicting insulin dosage
Diabetic Dataset
/DELL/Desktop/An Expert System for Insulin Dosage Prediction/Dataset
cess Dataset
adient Boosting Algorithm
near Regression Algorithm
y Graph
Diabetes & Insulin Dosage

Fig- Predict diabetes and insulin

Conclusions and Future Enhancements

The future scope of developing an expert system for insulin dosage prediction using ML (Machine Learning) holds significant potential for advancing personalized healthcare and improving the management of diabetes. Here are some key areas of future development and expansion.

ML algorithms can be further refined and optimized to improve the accuracy of insulin dosage predictions. This can be achieved by incorporating more diverse and comprehensive datasets, exploring advanced ML techniques (such as ensemble models or deep learning), and leveraging additional patient-specific factors (such as genetic information or lifestyle data) to tailor the predictions.The integration of ML-based insulin dosage prediction systems with wearable devices, continuous glucose monitoring (CGM) systems, and other Internet of Things (IoT) devices can provide real-time data inputs for more accurate predictions. This would enable patients to receive timely recommendations and adjust their insulin dosage based on current glucose levels, physical activity, and other relevant parameters.ML algorithms can be utilized to develop personalized treatment plans for individuals with diabetes. By considering various factors such as age, weight, medical history, lifestyle, and glucose patterns, ML models can generate customized insulin dosage recommendations that are tailored to the specific needs and characteristics of each patient.ML algorithms can be employed to analyze long-term glucose trends and patterns, helping healthcare providers and patients identify potential risks, optimize insulin dosages, and predict future glucose levels. This would enable proactive management of diabetes and the prevention of hyperglycemic or hypoglycemic episodes.

ML-based insulin dosage prediction systems can be integrated with electronic health record systems to provide a comprehensive view of a patient's medical history, laboratory results, and treatment plans. This integration would facilitate better decision- making by healthcare professionals and enable seamless sharing of information among different care providers.ML algorithms can be utilized to provide real-time feedback and educational resources to patients, helping them better understand their insulin dosage recommendations and make informed decisions. Interactive interfaces, mobile applications, or chatbot- based systems can be developed to deliver personalized guidance, track adherence, and offer educational materials to improve patient engagement and self-management.ML models can assist healthcare providers in making evidence-based decisions by providing insights and recommendations for insulin dosage adjustments. These models can analyze large volumes of patient data, clinical guidelines, and treatment outcomes to offer decision support tools that aid in optimizing insulin therapy and achieving better patient outcomes.

REFERENCES

[1] A. M. Syed, M. U. Akram, T. Akram, M. Muzammal, S. Khalid, and M. A. Khan, "Fundus images-based detection andgrading of macular edema using robust macula localization," IEEE Access, vol. 6, pp. 58784–58793, 2018.

[2] A. Cahn, A. Shoshan, T. Sagiv et al., "Prediction of progression from pre-diabetes to diabetes: development and validationof a machine learning model," Diabetes, vol. 36, no. 2, pp. 1–8, 2020.

[3] Veena Vijayan V. And Anjali C, Prediction and Diagnosis of Diabetes Mellitus, "A Machine Learning Approach",2015IEEE Recent Advances in Intelligent Computational Systems (RAICS) | 10- 12 December 2015 | Trivandrum.

[4] M. K. Hasan, M. A. Alam, D. Das, E. Hossain, and M. Hasan, "Diabetes prediction using ensembling of different machinelearning classifiers," IEEE Access, vol. 8, pp. 76516–76531, 2020.

[5] S. Islam Ayon, and M. Milon Islam, "Diabetes prediction: a deep learning approach," International Journal of InformationEngineering and Electronic Business, vol. 11, no. 2, pp. 21–27, 2019.

[6] M. Maniruzzaman, M. J. Rahman, B. Ahammed, and M. M. Abedin, "Classification and prediction of diabetes diseaseusing machine learning paradigm," Health Information Science and Systems, vol. 8, no. 1, pp. 1–14, 2020.

[7] P. Suresh Kumar and V. Umatejaswi, "Diagnosing Diabetes using Data Mining Techniques", International Journal ofScientific and Research Publications, Volume 7, Issue 6, June 2017 705 ISSN 2250-3153, pp. 705-709.01234567890 500 1000 1500 2000 2500 3000 3500In LunchRNN LSTM RMSE(Epoch 100) ANN RMSE0510152025300 500 1000 1500 2000 2500 3000 3500In DinnerRNN LSTM RMSE(Epoch 100) ANN RMSE www.ijcrt.org © 2022 IJCRT | Volume 10, Issue 8 August 2022 | ISSN: 2320-2882IJCRT2208305 International Journal of Creative Research Thoughts

[8] Ridam Pal, Dr. Jayanta Poray, and Mainak Sen, "Application of Machine Learning Algorithms on Diabetic Retinopathy",2017 2nd IEEE International Conference on Recent Trends in Electronics Information & Communication Technology, May19-20, 2017, India.

[9] Berina Alic, Lejla Gurbeta and Almir Badnjevic, "Machine Learning Techniques for Classification of Diabetes andCardiovascular Diseases", 2017 6th Mediterranean Conference on Embedded Computing (MECO), 11-15 JUNE 2017, BAR,MONTENEGRO.

[10] Dr. M. Renuka Devi and J. Maria Shyla, "Analysis of Various Data Mining Techniques to Predict Diabetes Mellitus",International Journal of Applied Engineering Research ISSN 0973-4562 Volume 11, Number 1 (2016) pp 727-730 © ResearchIndia Publications. http://www.ripublication.com

[11] Martinsson, John, Alexander Schliep, Bjorn Eliasson, Christian Meijner, Simon Persson, and Olof Mogren. "Automaticblood glucose prediction with confidence using recurrent neural networks." In Khd@ ijcai. 2018.

[12] Minyechil Alehegn and Rahul Joshi, "Analysis and prediction of diabetes diseases using machine learning algorithm":Ensemble approach, International Research Journal of Engineering and Technology Volume: 04 Issue: 10 | Oct -2017

[13] T. E. Idriss, A. Idri, I. Abnane and Z. Bakkoury, "Predicting Blood Glucose using an LSTM Neural Network," 2019Federated Conference on Computer Science and Information Systems (FedCSIS), 2019, pp. 35-41, doi: 10.15439/2019F159.

[14] Michael Kahn, MD, PhD, Washington University,St. Louis, MO, https: // archive.ics.uci.edu/ ml/datasets/diabetes (accessed on 10/06/2022).

Prediction of Insulin Level of Diabetes Patient Using Machine Learning Approaches. Available from: https://www.researchgate.net/publication/362904318_Prediction_of_Insulin_Level_of_Diabetes_Patient_Using_Machine_Learning_Approaches [accessed Jul 05 2023].