Heaven's Light is Our Guide
Computer Science & Engineering
Rajshahi University of Engineering & Technology

Course No.: CSE 4204
Course Title: Sessional based on CSE 4203

Experiment No. 2

Name of the Experiment: Design and implementation of single layer perceptron learning algorithm.

Course Outcomes: CO1

Learning Domain with Level: Cognitive (Applying, Analyzing, Evaluating & Creating)

Contents

1 Dataset

I've Selected Diabetes Dataset from Kaggle [1] .This dataset is originally from the National Institute of Diabetes and Digestive and Kidney Diseases. The objective is to predict based on diagnostic measurements whether a patient has diabetes.

1.1 Dataset Analysis

The Dataset have 8 attributes and 768 instances. The Features are :

1. **Pregnancies:** Number of times pregnant
2. **Glucose:** Plasma glucose concentration a 2 hours in an oral glucose tolerance test
3. **BloodPressure:** Diastolic blood pressure (mm Hg)
4. **Insulin:** 2-Hour serum insulin (mu U/ml)
5. **BMI:** Body mass index (weight in $kg/(heightinm)^2$)
6. **DiabetesPedigreeFunction:** Diabetes pedigree function
7. **Age:** Age (years)
8. **Outcome:** Class variable (0 or 1)

Class Distribution: class value 1 is interpreted as "tested positive for diabetes".

1.2 Training-Test Ratio

70% of data were used for training and 30% of data for validating .

2 Preprocessing

single layer perceptron learning algorithm was applied for only 3 selected features. At First the Correlation Matrix was piloted. Based on this, a 3D graph was plotted to get a idea about the Dataset. Then z-score normalization was performed to the Dataset. Normalizing the inputs makes it easier for the algorithm to find the optimal weights.

2.1 Correlation Matrix

For Feature selection a correlation matrix was printed for understating the correlation between different features . From the correlation matrix it is shown that the 'Glucose' attribute is highly correlated with the class variable. Then 'BMI' and age attributes are also more correlates with class variable than others . The correlation matrix is shown below:

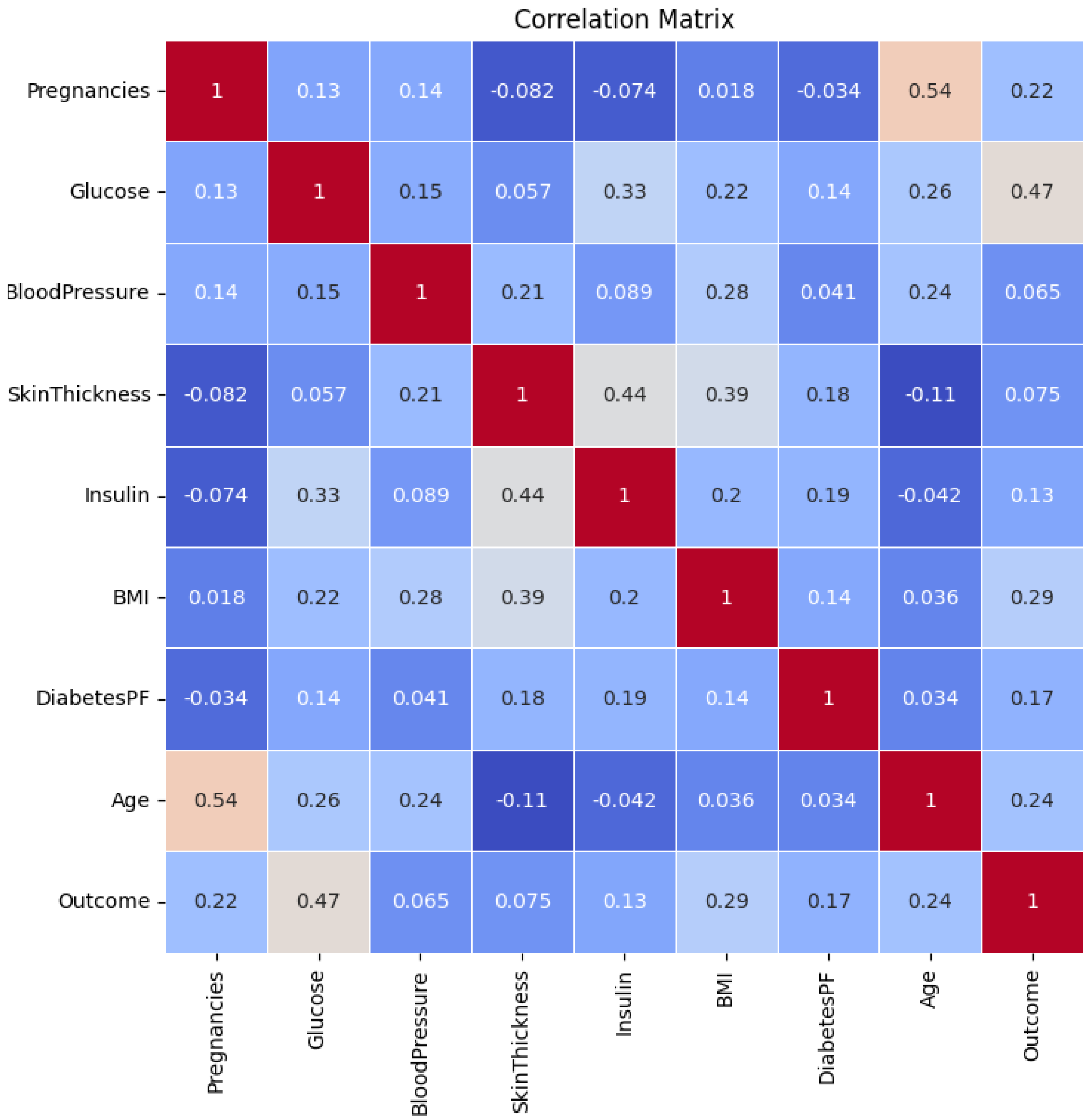

Figure 1: Correlation Matrix

Based of Correlation Matrix It was shown that "Glucose " ,"BMI" , "Age" Attributes are more correlated with output than others. So , a 3D graph was plotted(Figure 2) taking these 3 attributes. After Plotting it was clear that the Dataset was not fully linearly separable . There were overlapping and Noise in the Dataset.

2.2 Normalization

A z-score normalization was done by using this formula:

$$Z = \frac{X - \mu}{\sigma}$$

where:
Z is the is the standardized value,
X is the original value of the feature,
μ is the mean of the feature,

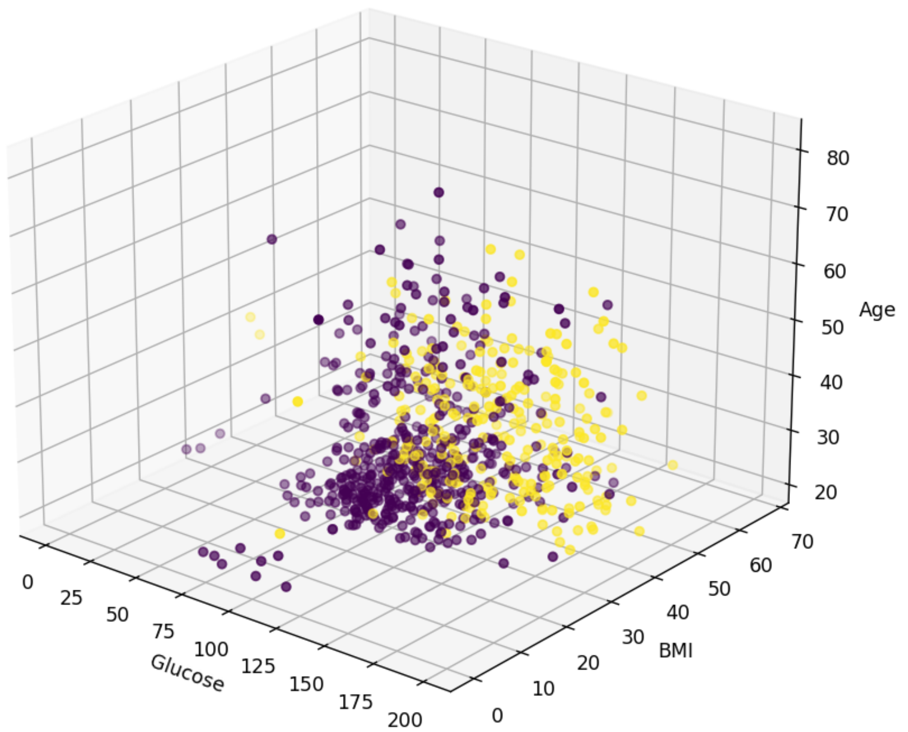

Figure 2: 3D scatter Plot of Glucose , BMI & Age

σ is the standard deviation of the feature.

3 Perceptron Learning Algorithm

The Algorithm were described as follows[2] :

1. Initialise weights and threshold:
   Define $w_i(t), (0 \leq i \leq n)$, to be the weight from input i at time t , and θ to be the threshold value in the output node.Set w_0 to be $-\theta$, the bias, and w_0 to be always 1.

   Set $w(0)$ to small random values, thus initialising all the weights and the threshold.

2. Present input and desired output Present input $x_0, x_1, x_2, \ldots, x_n$ and desired output $d(t)$

3. Calculate actual output
   $$y(t) = f_h \left[\sum_{i=0}^{n} w_i(t) x_i(t) \right]$$

4. Adapt weights

if correct

$$w_i(t+1) = w_i(t)$$

if output 0, should be 1(class A)

$$w_i(t+1) = w_i;(t) + x_i(t)$$

if output 1, should be 0 (class B)

$$w_i(t+1) = w_i(t) - x_i(t)$$

4. Adapt weights-modified version

   if correct

$$w(t+1) = w(t)$$

   if output 0, should be 1(class A)

$$w_i(t+1) = w_i(t) + \eta x_i(t)$$

   if output 1, should be 0 (class B)

$$w_i(t+1) = w_i(t) - \eta x_i(t)$$

4. Adapt weights-Widrow-Hoff delta rule

$$\Delta = d(t) - y(t)$$
$$w_i(t+1) = w_i(t) - \eta \Delta x_i(t)$$

$$d(t) = \begin{cases} 1, \text{if input from class A} \\ 0, \text{if input from class B} \end{cases}$$

where $0 \leq \eta \leq 1$, a positive gain term that controls the adaption rate.

4 Accuracy Analyzing

For Analyzing the accuracy, the algorithm using weight using 3 weight update method. Which Accuracy is given below:

- Adapt Weight : 0.7239
- Adapt Weight-Modified version : 0.7239
- Adapt weights-Widrow-Hoff delta rule : 0.7239

From the 3D plot (Figure 2) ,it was clear that the dataset was not fully linearly separable.So Adapting the weight in different method didn't make any difference .

5 Conclusion

For the chosen dataset the Single layer perceptron learning algorithm did not perform as expected because the dataset was not fully linearly separable.This algorithm can only learn and represent linearly separable functions . Also Single Layer Perceptron consists of a single layer of neurons without hidden layers. This lack of hidden layers restricts its ability to learn complex relationships in data. If the noise, outlier , overlapping regions can be removed from the dataset , the algorithm could perform better.

References

[1] Diabetes Dataset, [Online]. Available at: `https://www.kaggle.com/datasets/mathchi/diabetes-data-set`

[2] R. Beale and T. Jackson, *Neural Computing: An Introduction*, CRC Press, 1990.