

Predicting Diabetes using Neural Network in Python

We have used "diabetes.csv" containing 600 instances and 8 features for training.

We have used "diabetes_test.csv" containing 168 instances for testing.

Have used learnrate for learning rate, epoch for number of iterations, weights1 for input layer and hidden layer, weights 2 for hidden layer and output layer.

Then Calculating value of Sigmoid :

Also derivative of Sigmoid : $x(1-x)$

Followed by cost.

Then iteration 1 moving forward from input layer to output layer, calculating using dot function with weights for both input and hidden layer.

Then Back propagation with derivative of cost function through hidden layer and back to input layer.

Update weights after that complete cycle/iteration.

Same for training it will repeat for the number of iterations/epochs mentioned (2000 in our case)

For adding data to the list, we have used values 3 bifurcated values between 0 to 1 namely for each of the species.

There were many missing values/zero values in the data, so for that we have handled them using mean normalization mentioned in findMean() function. The program does give accurate results without using mean to replace zero values, But not as accurate comparatively.

Then used Scikit learn library StandardScaler to Standardize/Normalize data.

We have used diabetes_test.csv for test data, which does not have the respective output of whether Diabetic or non-Diabetic(1 or 0 respectively) associated with each instance.

It will read each feature of the 8 features value for each of the instance up to the last instance i.e length of test data and then predict. We have printed a matrix type formation for labelling the output that will range from 0 to 1.

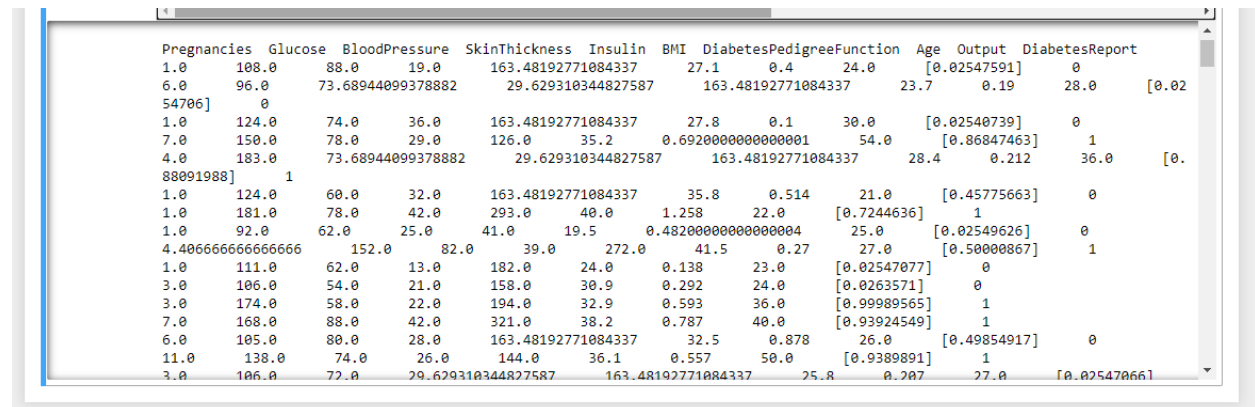
Based on the Output value and diabetic type will be labelled to the instance whether 1(person will have Diabetes) or 0(Person would be non-Diabetic).

Attached below is a part of the output set :

Expected Output –

Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
1	108	88	19	0	27.1	0.4	24	0
6	96	0	0	0	23.7	0.19	28	0
1	124	74	36	0	27.8	0.1	30	0
7	150	78	29	126	35.2	0.692	54	1
4	183	0	0	0	28.4	0.212	36	1
1	124	60	32	0	35.8	0.514	21	0
1	181	78	42	293	40	1.258	22	1
1	92	62	25	41	19.5	0.482	25	0
0	152	82	39	272	41.5	0.27	27	0
1	111	62	13	182	24	0.138	23	0
3	106	54	21	158	30.9	0.292	24	0
3	174	58	22	194	32.9	0.593	36	1
7	168	88	42	321	38.2	0.787	40	1
6	105	80	28	0	32.5	0.878	26	0
11	138	74	26	144	36.1	0.557	50	1

Predicted Output –



Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Output	DiabetesReport
1.0	108.0	88.0	19.0	163.48192771084337	27.1	0.4	24.0	[0.02547591]	0
6.0	96.0	73.68944099378882	29.629310344827587	163.48192771084337	23.7	0.19	28.0	[0.02547591]	0
1.0	124.0	74.0	36.0	163.48192771084337	27.8	0.1	30.0	[0.02540739]	0
7.0	150.0	78.0	29.0	126.0	35.2	0.6920000000000001	54.0	[0.86847463]	1
4.0	183.0	73.68944099378882	29.629310344827587	163.48192771084337	28.4	0.212	36.0	[0.02549626]	0
1.0	124.0	60.0	32.0	163.48192771084337	35.8	0.514	21.0	[0.45775663]	0
1.0	181.0	78.0	42.0	293.0	40.0	1.258	22.0	[0.7244636]	1
1.0	92.0	62.0	25.0	41.0	19.5	0.48200000000000004	25.0	[0.02549626]	0
0.0	152.0	82.0	39.0	272.0	41.5	0.27	27.0	[0.5000867]	1
1.0	111.0	62.0	13.0	182.0	24.0	0.138	23.0	[0.02547077]	0
3.0	106.0	54.0	21.0	158.0	30.9	0.292	24.0	[0.0263571]	0
3.0	174.0	58.0	22.0	194.0	32.9	0.593	36.0	[0.99989565]	1
7.0	168.0	88.0	42.0	321.0	38.2	0.787	40.0	[0.93924549]	1
6.0	105.0	80.0	28.0	163.48192771084337	32.5	0.878	26.0	[0.49854917]	0
11.0	138.0	74.0	26.0	144.0	36.1	0.557	50.0	[0.9389891]	1
3.0	106.0	72.0	29.629310344827587	163.48192771084337	25.8	0.207	27.0	[0.02547066]	0