About My Code

Mayank Singh Jadon

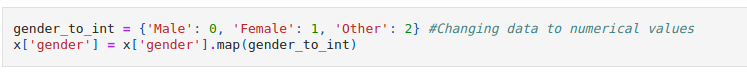
22169001 CSE Hindi

- Here, In my Code i have used Tensorflow for Neural Network, Scikit-learn for clasification report and pandas for Table.

*Pre-Processing*

-In Pre Processing, I have first copied the dataframe in ‘x’ and answers in ‘y’.

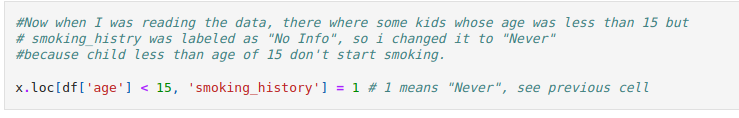
-First I changed all the String data types in int or float data types because tensorflow works with Numbers only.

 And



- Another Thing I Have done in Pre Processsing is that there were much rows where ‘No Info’ was written in rows of Column ‘smoking\_history’, but the patient was child.

So I changend it and convert ‘No Info’ to ‘Never’ Where age was less than 15, because just think will a child less than age of 15 will smoke, simply no.



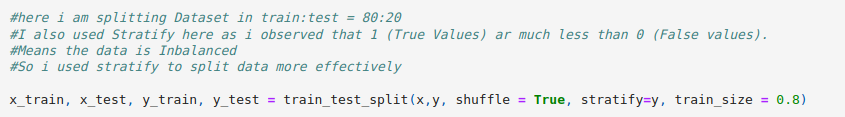
- In next line i also poped diabeties column fromx because i wanted to separate question set and answer set. Question set in ‘x’ and answer set in ‘y’.

***(Question set : x\_train and x\_test, answer set : y\_train, y\_test [will see later])***

*Training the Model*

- I have splited the data in Traing and testing set in the ratio 80:20.

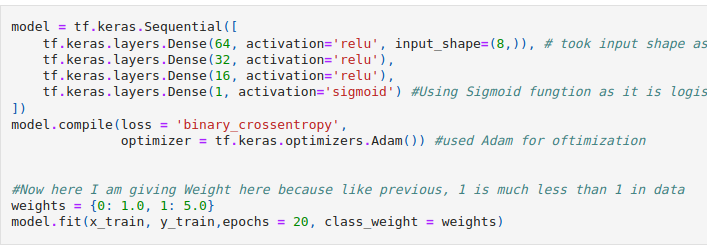
- If you look at the data, You will notice that data is **Imbalanced**. There are much patients with False diabeties report Than True Diabeties report, by which model can be trained falsely by this Imbalanced report, I have tackles this by using “*Stratify*” and Giving *Weight* to True value (1).



According to my knowledge, Stratify will distrubute sample by taking care of of 1 and 0.

*Model*

-Now let us look at our model first.



-In this Model, I have created 4 layers of neurons. And Input Shape as 8 because we only have 8 Column in Question Set (remember that diabeties clumn is poped)

- Now about Neural network architecture, I really don’t know about this, I know how to create neuron (means i know hot to use tf.keres.Sequential and tf.keras.layers.Dense but don’t now how much numbers of neurons should be placed), I just did hit and trial.

- Now i Ativaton part, i Have used ‘Sigmoid’ function because i need answer between 0 ant 1.

- As I Prevously Talked, I have given More weight to 1 because it have less occurence in Dataset.

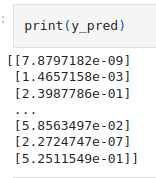
- At end, i have fitted the Model

*Prediction*

-Now prediction, i have done it like this



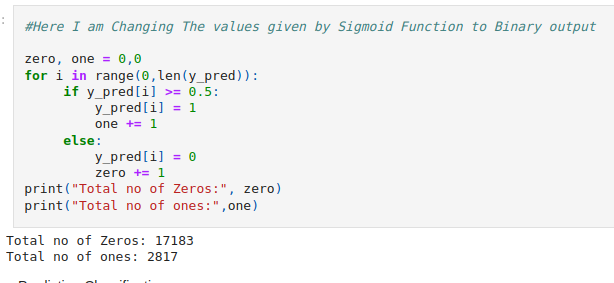
Now If i look at values of y\_pred it will look like this



But we need data in 0 or 1

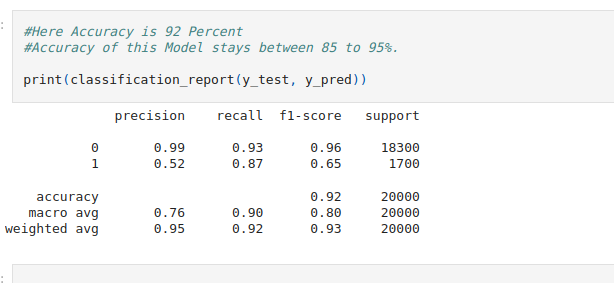
This is because in Sigmoid, the anwer comes between 0 and 1, so we neet to convert it in 0 or 1,

I have done it like this.



- In this any value greater than 0.5 is changed to 1 and other in 0.

- Now lets check acuuracy of our prediction



Here I have used *Scikit-Learns* function *classification\_report,* this gives all the details of actual and predicted values.

As seeing in the Screen shot, you can see that I gained an Accuracy of 92 percent, In Multiple run the Acuuracy was ranging between 85 to 95, I could have used Screen Shot of 95% but used this because it was close to Avg.

Thanks You

Mayank Singh Jadon

22169001 CSE Hindi