**Machine Learning assignment for Internship**

I have used Keras to train the model. The model has:

* 4 Hidden layers starting with 300 neurons at the 0th(Input) layer and moving forward with 200, 150 and 100 neurons.
* The output layer has 1 neuron.
* Hidden layers use ReLu as activation function because in its formula **a=max(0,Z)**, for a lot of Z, slope is close to 1. So **learning becomes fast**.
* Output layer has Sigmoid activation because in its formula **sigma =1/(1+e^(-z))**, sigma would be either 1 (z>0.5) or 0 (z<0.5). So it’s **good for binary classification and output layer.**

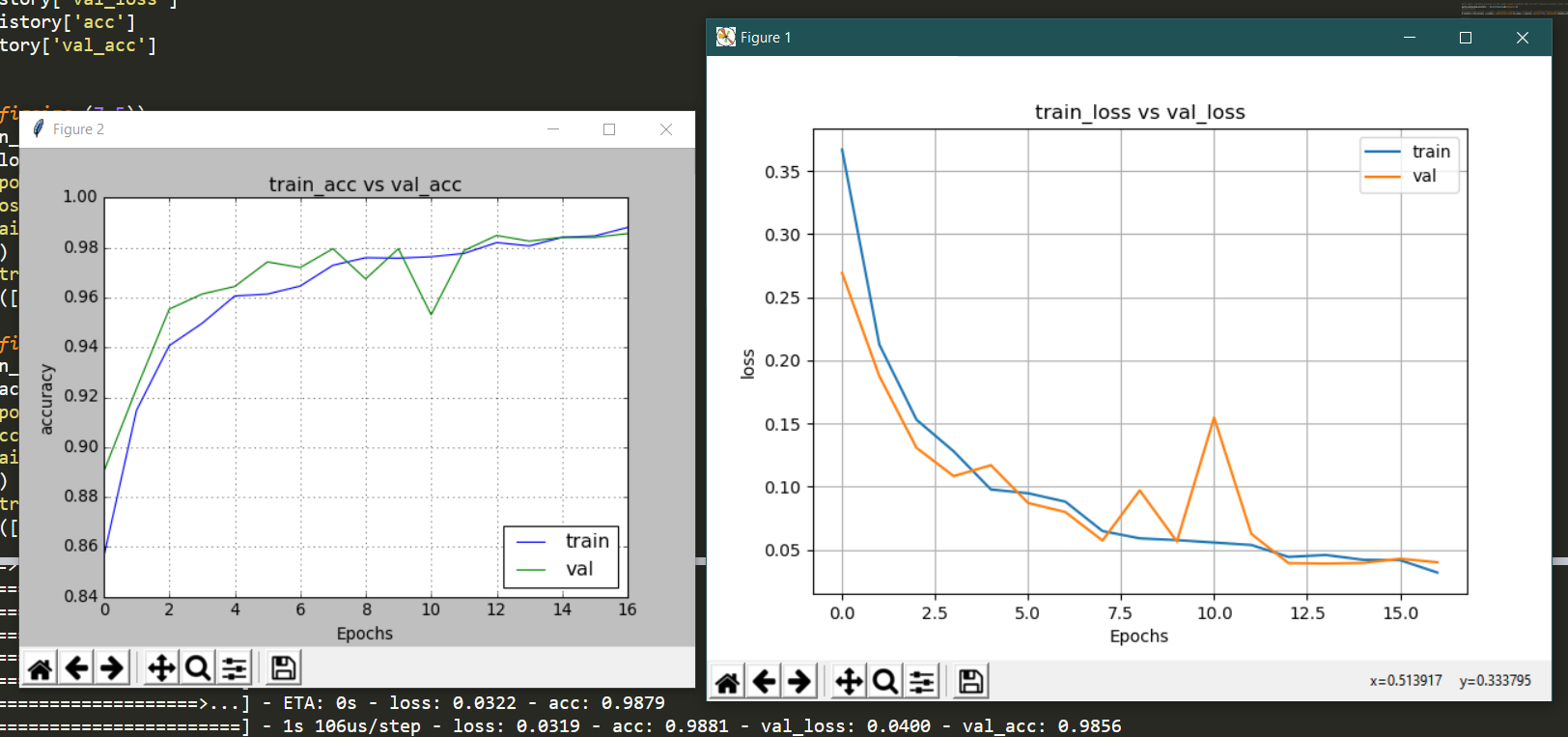
**Preprocessing**

* Dropped unnecessary coloumns like ['ID', 'molecule\_name', 'conformation\_name']
* Shuffled the dataset
* Normalized the dataset

**Post processing**

* Precision, recall and F1 score
* Plotting graphs

**Graphs plotted**



**(**I think I did some mistake in calculating Precision, recall and F1 score and I don’t have time to debug as the assignment has to be submitted before 11:59pm. But the code runs fine**)**

**Validation accuracy is 98%**

**Precision: 0.41 Recall: 0.5**

**F1 Score: 0.45**