**Solution Sheet**

1. Which model have you used for stock price prediction? Explain your model.
2. **Task:**

The task is to predict stock price from various features given.

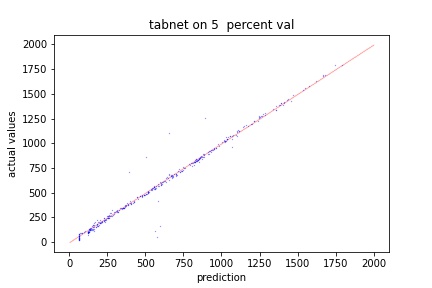
**model:**

To complete the task, a deep learning model called ‘TabNet’ is used. TabNet is a self attention model which is specially designed to deal with tabular data.

**About Data:**

Every data has problems and this is not an exception. The given data has some missing values. This is dealt by replacing the ‘Na’ with the mean of the whole column(since mean is the statistical estimate of that particular value).

The below is a graph plotted between actual and values predicted with tabnet. As we can observe clearly, the points follow a locus of y=x which means the tabnet managed to predict almost all data points correctly.



1. Which model have you used for Put-Call ratio Time series prediction? Explain your model.
2. **Task:**

Here the task is to find out the put-call ratio on august 16 from the past 6 days data.

**Model Details:**

As I said earlier, the data has some missing values which can potentially hurt the model’s performance. So, to get an overall view of the data, a self supervised way of training is adopted while training a Bi-directional RNN.

**Training Procedure:**

1.An Bi-directional RNN is built with one fully connected layer on top of it.

2.Now, this model takes sequence of length 5 of which we randomly make some elements zero(I have made 1 or 2 elements zero randomly) and a mask having 1's at the positions where we made zeros in the sequence and 0's at rest of the positions as inputs and output the a sequence of same length

3. we will optimize the above built RNN to output a sequence of same length 5, having missing values.This concludes our pretraining part which is purely inspired from masked language modelling from the field of Natural Language Processing.

4. Now, we will take the pretrained model and remove the dense layers(which are on top of RNN),remove rows which don’t have put call ratio for 15 august (since they impose large error while optimizing) and finetune this model to predict put calls ratio on august 15th from the data of august 10th,11th,12th,14th. A maximum of 8% error is observed while validating on the validation data.

Again, the below graph is plotted between actual and predicted values. Since it's almost y=x, we can say the model is working properly.

