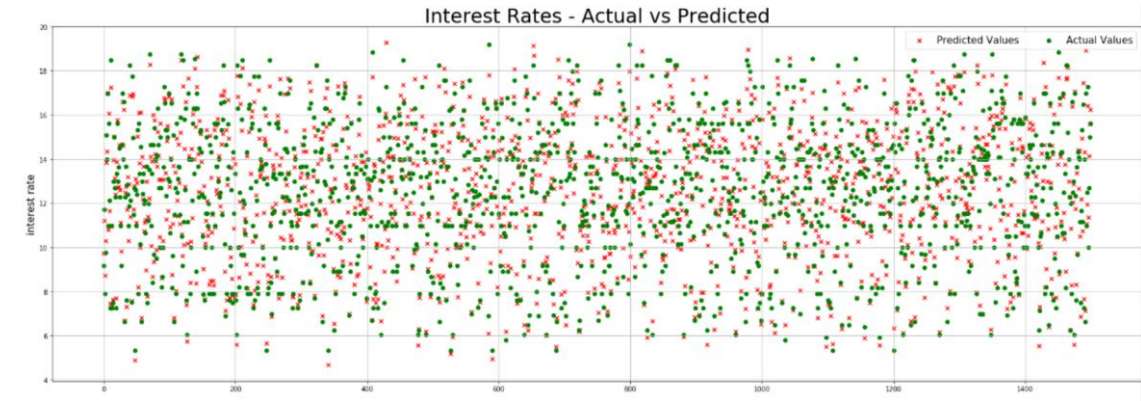


PURPOSE

To test if the model is predicting the interest rate as per the requirement of the client, who is an investor and is not willing to take risk. To make the client know that why should he use our model to predict the interest rate.

REGRESSION MODEL:



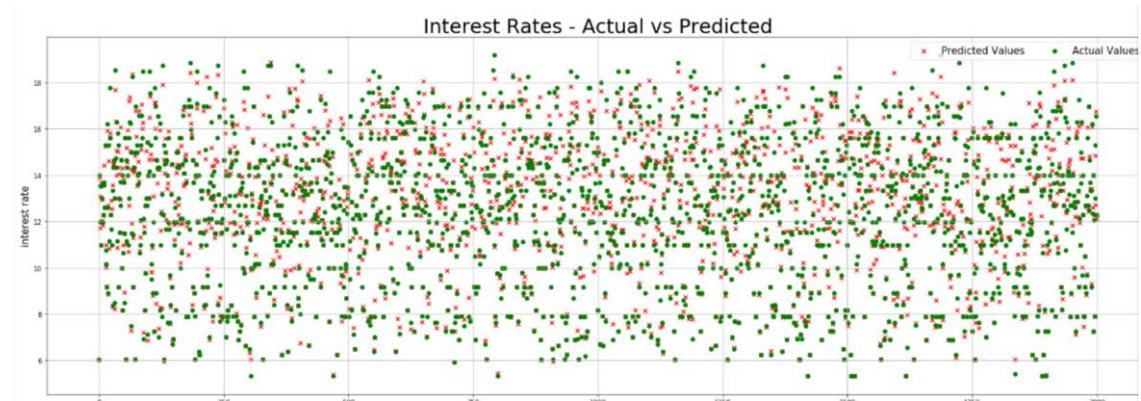
Before Cross Validation			
MODEL	Mean absolute percentage error - Test	Mean absolute percentage error - Train	Mean absolute error
Regression	4.698638583	4.698459428	0.57
After 5 fold Cross Validation			
MODEL	Mean absolute percentage error - Test	Mean absolute percentage error - Train	Mean absolute error
Regression	5017711.046	4.614891388	Mean : 448432.3228657413 ; Median : 0.8342562420368367

NEURAL NETWORK:



Before Cross Validation			
MODEL	Mean absolute percentage error - Test	Mean absolute percentage error - Train	Mean absolute error
Neural Networks	3.257460136	3.092385018	0.38
After 5 fold Cross Validation			
MODEL	Mean absolute percentage error - Test	Mean absolute percentage error - Train	Mean absolute error
Neural Networks	4.03932434	2.945391243	0.488

RANDOM FOREST:



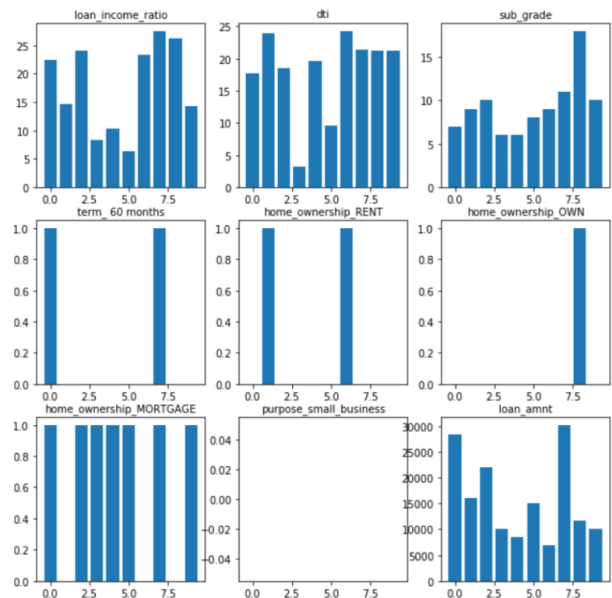
Before Cross Validation			
MODEL	Mean absolute percentage error - Test	Mean absolute percentage error - Train	Mean absolute error
Random Forest	2.638090034	1.013020849	0.341
After 5 fold Cross Validation			
MODEL	Mean absolute percentage error - Test	Mean absolute percentage error - Train	Mean absolute error
Random Forest	3.990617695	0.973973909	0.508

- As depicted in the above models, we observe that the MAPE for Random Forest is low as compared to Neural network and Regression Model. Since MAPE is a measure of prediction accuracy of a forecasting method in statistics, we can say that our client Rick should use the model to get the borrower with Grade A, B, C and D.
- So, for our model the features that the investor should have for application of the loan should be: addr\_state application\_type purpose loan\_status , verification\_status term sub\_grade home\_ownership , emp\_length , dti revol\_bal total\_pymnt total\_rec\_int total\_rec\_late\_fee , recoveries last\_pymnt\_amnt out\_prncp loan\_amnt int\_rate, annual\_inc acc\_now\_delinq delinq\_2yrs pub\_rec open\_acc,inq\_last\_6mths,collections\_12\_mths\_ex\_med revol\_util
- For the risk averse model some of the criteria being considered are : Sub Grade should be A, B , C, D ; Loan to income ratio > 0.3 and Home Ownership should be own

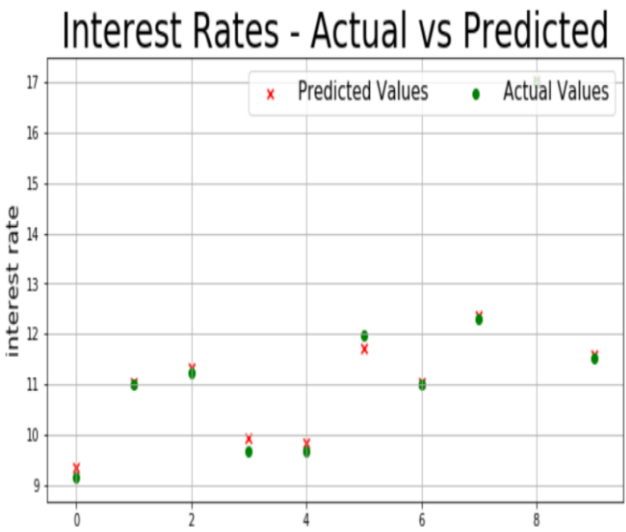
USE CASE:

Taking different values for the borrower and seeing how the model is predicting values vs the actual values. For the use case we have taken 10 randomly extracted records from the data set and seeing its value fluctuation with respect to the best model: Random Forest.

Below is the list of some of the features we have in the randomly selected data set. The pictorial representation also shows the actual and predicted values using random forest.



Profile	Actual interest rate	Predicted interest rate
Profile 1	9.17	9.355194
Profile 2	10.99	11.048431
Profile 3	11.22	11.314619
Profile 4	9.67	9.926072
Profile 5	9.67	9.821698
Profile 6	11.99	11.708390
Profile 7	10.99	11.020211
Profile 8	12.29	12.368595
Profile 9	16.99	17.031387
Profile 10	11.53	11.589946



MAPE SCORE FOR THE ABOVE DATASET: 1.1638, Since the MAPE score is very low we would suggest the client to use our model to predict interest rate of investor.