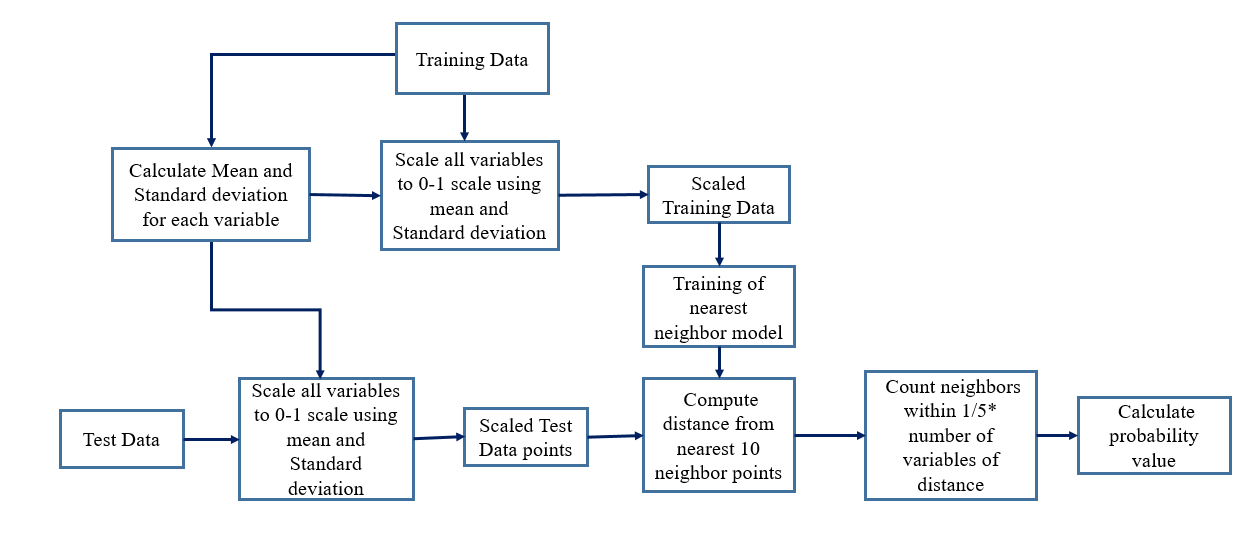
All Machine learning model provides the prediction value even when test data point is out of range from training data distribution. This article helps to determine the reliability index with prediction value. This index access the chances/probability of test data point different than training dataset.

Imagine that, we have the credit history data for default prediction, in the sample data all the customers have salary below 20 lakhs, work experience below 10 years and these features have high influence on default. We have built a classification model to predict credit default, and the model appears accurate in the training data as well as in validation data.

Suppose we get new customer with 30 lakhs salary and 5 years’ experience, the model might be inaccurate when used in practice for prediction. The reliability index will help to provide user an indication of inaccurate prediction as new data is slightly different from train data.

Reliability index is calculated using nearest neighbor approach. In this article, we expect at least 10 surrounding data points with Euclidean distance less than 1/5 of number of input variables around new test data for reliable prediction. The neighboring points and distance can be modified according to critically of use case. All the data points standardized using z-score as Euclidean distance calculation affected by scale of variables. This distance helps to form circle with equal distance to find the density of training data points around the test data point. If within defined distance number of neighbor around test point are greater than or equal to 10, then index value will be 1. If less than 10, we calculate index value by dividing number of available training data points within defined distance by 10. The overall approach is shown in below diagram:



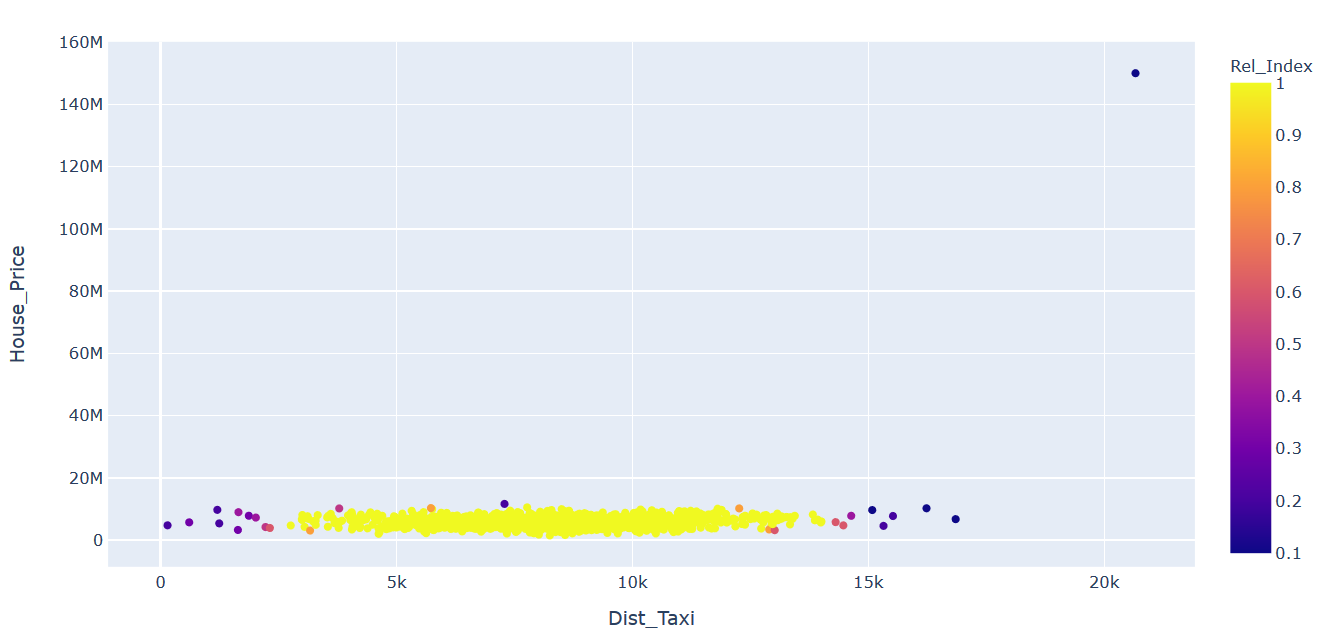
Note: To standardize test data, we use training data mean and standard deviation.

To understand, let’s take an example of house price prediction with

For simplicity example graph is shown with one input and one output variable but the concept will be applicable with multivariate data.

For example, we are predicting house price using distance to taxi (Dist\_Taxi), scatter plot is shown in below graph. The points near to yellow pixel are similar to training data and prediction will be reliable if the model is built properly but the prediction value for points near to blue pixel range will be less reliable as this are slightly different from training data distribution.

Reliability index is given in the form of probability value will range from 0 to 1 and scale independent from dataset. If the Reliability index is 0.9, then 90% chance that prediction value will be reliable. If the Reliability index is 0.2, then 20% chances that predicted value will be reliable. Similarly can be interpreted for other values.



This index is helpful when prediction process is deployed and retraining is not planned. In such cases, it is important to raise alert when the new test data points are out of training data distribution as less guarantee that predicted value will be reliable. The alert can be passed to expert/manual check before final usage of prediction value. If more frequently goes for expert check the we can thing of updating the model with new training data. The concept is applicable for both supervised and unsupervised problems.

It is not necessary if the reliability index value is 1 prediction error will be always less, it is just an indication that in the test data input points are different from training data.

Also if we pass unimportant variable and values differs from training data distribution reliability index will be less than 1, so we should run reliability analysis for important variables.

Also it is not necessary that that if reliability index is low prediction error will be always high,