REPORT

Sindhura Uppu Anmisha Reddy Ramidi

UBIT Name: suppu UBIT Name: anmishar

Person No: 50206730 Person No: 50208673

* We have developed a linear regression model using RShiny on Bike Sharing Demand dataset so as to predict the number of users that may use bikes in a particular scenario.
* The product is available for analysis at <https://sindhurauppu.shinyapps.io/bikesharing/>
* The dataset is available [here](https://www.kaggle.com/c/bike-sharing-demand/data)
* There are various interesting features in this data such as:
  + Datetime
  + Season 1-spring, 2-summer,3-fall,4-winter.
  + Holiday 0-Yes,1-No.
  + Workingday 0-Yes, 1-No.
  + Weather 1-Clear, 2-Misty, 3-Light snow/rain, 4-Thunderstorm
  + Temp
  + Atemp
  + Humidity
  + Windspeed
  + Casual
  + Registered
  + Count
* The datetime variable is further analyzed to identify the count of users based on year, month, hours and day.
* We developed the UI using RShiny and gave suitable input elements to each of the features.
* Once the features are selected, we use them as test data, generate 2 models to predict registered and casual users based on the training data using randomForest and predict the registered and casual users for the test data using this model.
* Various levels of ntree (decision trees) were specified from the range 100-350 and the 6 different predictions are taken into a dataframe and which is rendered in the UI.
* We have also plotted the accuracy graphs by considering logarithmic root mean square error of the predicted values and actual values of registered, casual and count of the users in the test set.
* It should be noted that we are using many features while generating models using randomForest, due to which there might be a slight delay in generating the plots and the table in UI.
* We have also plotted regression plots for training and test data set in the next tab.