

Changes – 2: 28/04/2018

I've removed the graphs and just displaying the results as discussed. MSE, R2 & Cross Validation scores (scoring = 'neg_mean_squared_error') For the initial page, I am going with value of K as 5.

An Interactive Explorer for Machine Learning datasets

Select the features to be included in the Random Forest Model

Prepared by **Adil Khan**.

[More information on the parameters](#)

Target feature is ['churn']

Top 5 features according to Select K Best (Chi2) : ['avg_dist', 'trips_in_first_30_days', 'luxury_car_user', 'city_KingsLanding', 'phone_Android']

Features

avg_dist
avg_rating_by_driver
avg_rating_of_driver
avg_surge

No of folds: 5

Fit_intercept:
True

Normalize:
False

Copy_X:
True

Results:

Mean Squared Error: 0.19
Variance score: 0.19
Cross Validation score: -0.19

After selecting the features:

An Interactive Explorer for Machine Learning datasets

Select the features to be included in the Random Forest Model

Prepared by **Adil Khan**.

[More information on the parameters](#)

Target feature is ['churn']

Selected features : ['avg_rating_by_driver', 'avg_rating_of_driver', 'avg_surge', 'surge_pct']

Features

avg_dist
avg_rating_by_driver
avg_rating_of_driver
avg_surge

Results:

Mean Squared Error: 0.23
Variance score: 0.00
Cross Validation score: -0.23

No of folds: 5



Fit_intercept:

True

Normalize:

False

Copy_X:

True