Predicting Taxi Rides Based On Weather

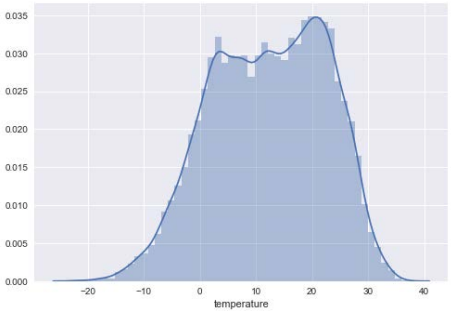
Negah Moharrami

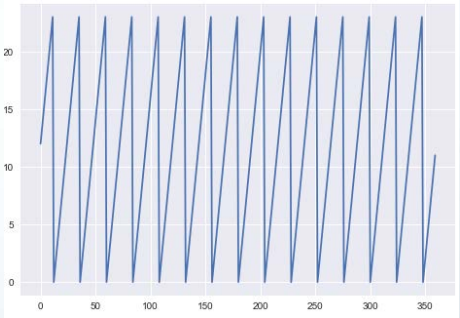
CEBD 1260 BD ANALYTICS

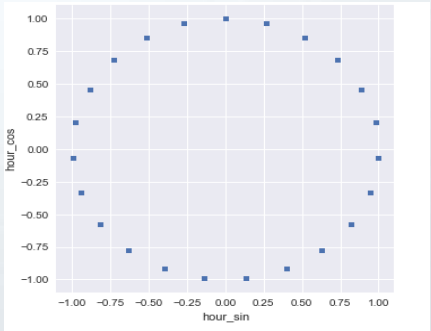
PREPARING THE DATA

	datetime	Vancouver	Portland	San Francisco	Seattle	Los Angeles	San Diego	Las Vegas	
4177	2015-07-05 21:00:00	303.02	306.130	298.930000	302.430000	299.240	295.740	306.410000	3

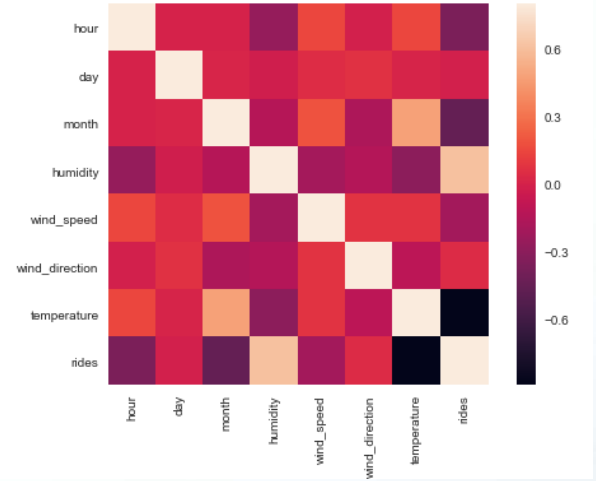
	datetime	temperature	humidity	pressure	wind_speed	wind_direction	rides	date	hour	month	day	year	hour_sin	hour_cos	day_sin	day_cc
33	2012-10-02 21:00:00	21.700000	43.0	1012.0	1.0	269.0	27	2012-10-02	21	10	2	2012	-5.195840e-01	0.854419	0.034422	0.9994(

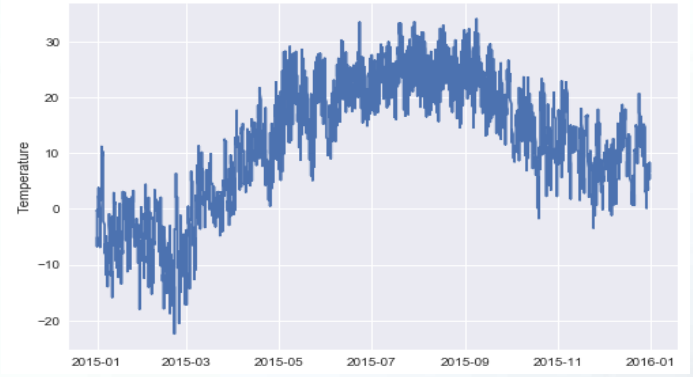


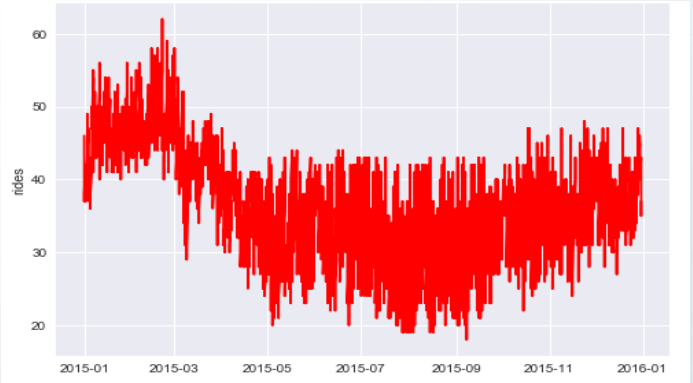




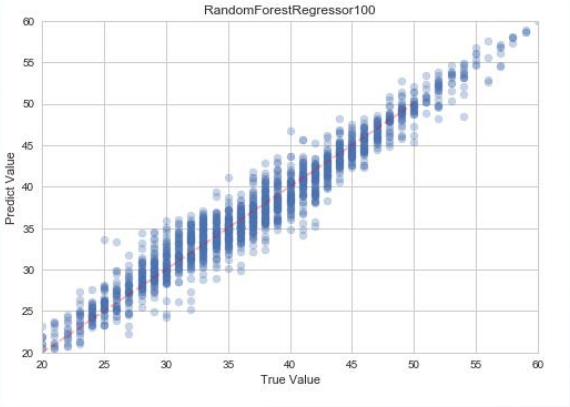
Correlation

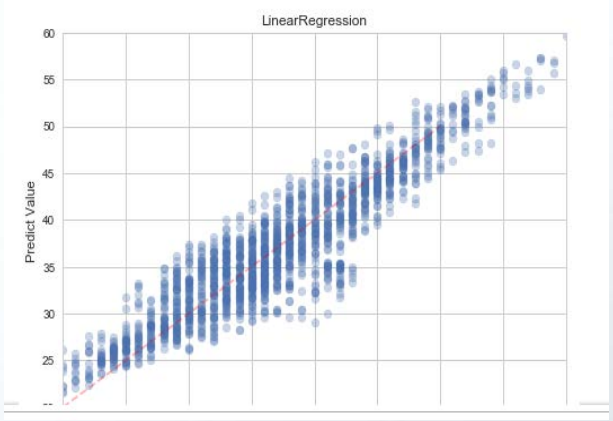






Regression



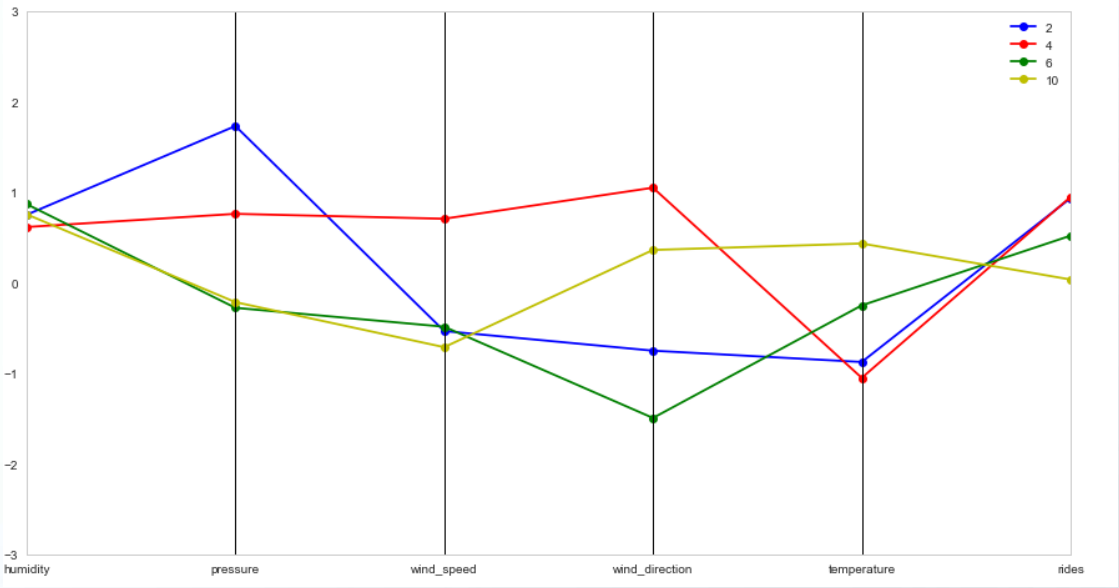


MODEL LinearRegression - encoded feature MAE 2.30558745190091 RMSE 2.9557718061243214

MODEL LinearRegression - unencoded feature MAE 5.322087434266247 RMSE 6.357275189694565

We achieved an improvement of 53% in our MSE

K_MEANS


3
2
1
0
−1
−2
−3
humidity
pressure
wind_speed
wind_direction
temperature
rides
2
4
6
10

Thanks!

Any questions?

You can findmore info and files :

- Github :

  https://github.com/Negah20/CEBD1260-ANALITICAL