Machine Learning

Week Five Solution

Week Five

Task: Aura customer FloridaBikeRentals.com is unable to predict peaks and troughs in demand for their high-end bikes. They have approached Aura to customize a marketing tool to predict bike-sharing demand. To stabilize the demand, devise marketing strategies using the bike-sharing dataset.

Based on rented bike count, the hour of the day, the day's temperature, humidity, wind speed, rainfall, holidays, and many other factors, build a model to predict the bike count required at each hour for the stable supply of rental bikes.

1. Load the dataset
2. Check for null values in any columns and handle the missing values
3. Convert Date columns to Date format and extract day, month, day of week and weekdays/weekend from date column
4. Check correlation of features using a heatmap
5. Plot the distribution plot of Rented Bike Count
6. Plot the histogram of all numerical features
7. Plot the box plot of Rented Bike Count against all the categorical features (Hint: Categorical features on X-axis and Rented Bike Count on Y-axis)
8. Plot the Seaborn catplot of Rented Bike Count against features like Hour, Holiday, Rainfall(mm), Snowfall (cm), weekdays_weekend and give your inferences
9. Encode the categorical features into numerical features.

   (Hint: use get_dummies())
10. Identify the target variable and split the dataset into train and test with a ratio of 80:20 and random state 1
11. Perform Standard Scaling of the train dataset.

12. Perform Linear Regression, Lasso Regression and Ridge Regression for predicting the bike count required at each hour and compare the results.