

Capstone Project Submission

Seoul Bike Sharing Demand Prediction

Introduction:

Today, bike-sharing systems are blooming across more than 1000 cities around the world, particularly in big or large cities like New York City, Paris, Washington DC, London, Beijing, and Barcelona. To complete a short trip renting a bike is a faster way when compared to walking. Moreover, it is eco-friendly and comfortable compared to driving.

Due to global warming, continuous pollution, and depletion of sources of energy. Many countries have been focused on using renewable energy which doesn't harm the environment and can be reused as well. South Korea is one of the countries which has adapted to it and their most used service is rented bikes in Seoul. But in order to avoid any difficulties such as waiting time it is necessary to have an estimate of future demand. Our goal here is to build a model that can predict bike sharing demand considering all the factors which have their effects.

Problem Statement:

Currently Rental bikes are introduced in many urban cities for the enhancement of mobility comfort. It is important to make the rental bike available and accessible to the public at the right time as it lessens the waiting time. Eventually, providing the city with a stable supply of rental bikes becomes a major concern. The crucial part is the prediction of bike count required at each hour for the stable supply of rental bikes.

Approach:

Here first we imported a data set and performed EDA where we got valuable insights and further we Encoded the Categorical Columns, Feature scaling and fitting into the models. At first we tried with basic linear regression and also with Lasso regularization technique but soon realized we will need a much more complex model and so we then used a Decision tree Regressor .

Conclusion:

The analysis is done with Seoul Bike data. Four regression techniques Linear Regression, Decision Tree, are used to predict the trip duration. This statistical data analysis shows interesting outcomes in prediction methods and also in exploratory analysis.

The experimental results show that:

- Hour:- Demand for bike is mostly in morning (7 to 8) and in the evening (3to9).
- Temperature :- Temperature is positively correlated. Rented bike count is highest between 20 °C and 30 °C. So, it means temperature has an effect.
- Humidity :- Humidity is the amount of water vapor in the air. So, People preferring to borrow bike When there is less humidity.
- Wind_speed :- wind speed doesn't affect our data much.
- Visibility:- Visibility doesn't affect our results much but all we know is that it is positively correlated with bike count.
- SnowFall and Rainfall:- People don't prefer borrowing bikes, where there is rainfall or snowfall.

- Dew point: The dew point is the temperature the air needs to be cooled to (at constant pressure) in order to achieve a relative humidity. It is positively correlated with data.

Contributor

- Ajit Padole (ajitpadole1998@gmail.com)

Please paste the GitHub Repo link.
GitHub Link: https://github.com/ajitpadole/Bike-Sharing-Demand-Prediction-