Bike Sharing Demand Forecasting

February 17, 2016

1 Problem Statement

Consider the data in the file *BikeDemandDaily.csv*. The data shows the demand for bike rides for a bike sharing program. There are two types of customer segments: (1) Casual customers who are not registered for the program but share bike ride on a spot payment basis, and (2) Registered customers who are registered for the program and share bike rides by paying in advance and have a member card.

2 Variable Description

The data in the file has the following information (variables):

1. **year**: Year index. There are two years in the sample.
2. **month**: Month index. Jan=1, Feb=2, ... , Dec=12.
3. **day**: Day index for a month and year.
4. **season**: Season index indicating four seasons.
5. **holiday**: Holiday index. 1 indicates holiday and 0 indicates not a holiday.
6. **workingday**: Workingday index. 1 indicates working day and 0 indicates not a working day.
7. **meanatemp**: Average daily temperature (degree celsius).
8. **maxatemp**: Maximum daily temperature (degree celsius).
9. **minatemp**: Minimum daily temperature (degree celsius).
10. **sdatemp**: Sandard deviation of the hourly temperature during a day.
11. **meanhumidity**: Average humidity for a day (percentage).
12. **maxhumidity**: Maximum humidity for a day (percentage).

13. **minhumidity**: Minimum humidity for a day (percentage).
14. **sdhumidity**: Standard deviation of humidity for a day.
15. **meanwindspeed**: Average wind speed in kmph.
16. **maxwindspeed**: Maximum wind speed in kmph.
17. **minwindspeed**: Minimum wind speed in kmph.
18. **sdwindspeed**: Standard deviation of wind speed.
19. **Casual**: Number of casual customers using the bike during the day.
20. **Registered**: Number of registered customers using the bike sharing program.
21. **Total**: Total number of daily customers.

3 Objective of Analysis

The following list illustrates some of the expected analysis outcome:

1. Understand the pattern of bike demand for causal, registered and total number of bikes demanded.
2. Plot relevant graphs to understand the demand pattern.
3. Plot variable graphs to understand important predictors.
4. Visualization of data.
5. Use moving average method for forecasting.
6. Use regression analysis for forecsting.
7. Use regression to decide which are the important variables.
8. Use some machine learning methods for better prediction.
9. Management report preparation and decision framework analysis.
10. Simulate data to understand loss in profit at various levels of inventory of bikes.