

CE 263 Midterm

Bike Sharing

In this midterm you will build a predictive model for the total number of hourly bike rentals in the bike sharing network of Washington, D.C.

The data.

Training set (midterm_train.csv): The dataset provided to you is hourly rental data spanning two years with an exception of 2 months used as a testing set for predictions. You must predict the total count of bikes rented during each hour for the months of August and December 2012.



The following variables are given to you:

Inputs:

- date: date string
- year: year (0: 2011, 1:2012)
- month: month (1 to 12)
- hour: hour (0 to 23)
- holiday: if the day is holiday or not
- weekday: day of the week, 0 (Sunday) to 6 (Saturday)
- working: 1 if a day is neither weekend nor holiday, 0 otherwise.
- weather_type :
 - 1: Clear, Few clouds, Partly cloudy
 - 2: Mist, Cloudy
 - 3: Light Snow, Thunderstorm
 - 4: Heavy Rain, Ice Pellets, Thunderstorm, Mist, Snow, Fog, Apocalypse ☺
- temp: temperature in Celsius.
- feels_like: “feels like” temperature in Celsius (accounting for wind chill and humidity)
- humidity: humidity
- windspeed: Normalized wind speed. The values are divided to 67 (max)

Output:

- count: count of total rented bikes

Testing set (midterm_test.csv): The months of August and December of 2012. The same set of input variables is provided to predict the hourly counts of bike rentals.

The Score

Your submissions will be evaluated by the Root Mean Squared Logarithmic Error (RMSLE). The RMSLE is calculated as:

$$\sqrt{\frac{1}{n} \sum_{i=1}^n (\log(p_i + 1) - \log(a_i + 1))^2}$$

where:

- n is the number of hours in the test set
- p_i is your predicted count
- a_i is the actual count
- $\log(x)$ is the natural logarithm.

Submission requirements

Your submission must contain:

- 1 page description of your approach (preferably as PDF)
- an implementation of your algorithm (preferably in Python)
- a file with predictions for August and December 2012

Prediction file has to be in the following format:

```
date,hour,count  
8/1/12,0,13  
8/1/12,1,23  
8/1/12,2,10  
.....
```



Good luck!