

## **Dude - Where's My Bike?**

*Background information from Fanaee-T, Hadi and Gamma, Joao. "Event labeling combining ensemble detectors and background knowledge", Progress in Artificial Intelligence (2013); pp.1-15, Springer Berlin Heidelberg*

*Original source of data: <http://capitalbikeshare.com/system-data>*

Bike sharing systems are a new generation of traditional bike rentals where the whole process from membership, the rental and the return has become automatic. Through these systems, a user is able to easily rent a bike from a particular position and return it at another position. Currently, there are over 500 bike-sharing programs around the world which contain over 500,000 bicycles. There is great interest in understanding the use of these systems due to their important role in traffic, environmental, and health issues. In contrast to most other transport services such as bus or subway, the duration of travel, departure and arrival position is explicitly recorded in these types of systems. (Note that these types of automated systems are starting to be installed on buses - e.g. the Pittsburgh Port Transit Authority.) One potential analysis of this type of data treats the bike sharing system as a virtual sensor network that can be used to identify changes in mobility in the city. Hence, it is expected that many important events in the city could be detected via analyzing this data. Our primary concern, however, is understanding and characterizing how these bike sharing systems are used.

Your research group has been given the daily data for 2011, 2012 with the following variables:

*ID*: identification number

*Date*: date of observation

*Season*: 1 = Winter; 2 = Spring; 3 = Summer; 4 = Fall

*Year*: year of observation (2011 or 2012)

*Month*: 1 = January, ..., 12 = December

*DayOfWeek*: 0 = Sunday, ..., 6 = Saturday

*Holiday*: whether the day was a holiday (1 = Yes, 0 = No)

*WorkDay*: whether the day was a work day (i.e., not a weekend or holiday) (1 = Yes; 0 = No)

*Weather*: type of weather

1 = clear, few clouds, partly cloudy

2 = mist & cloudy, mist & broken clouds, mist & few clouds, mist

3 = light snow, light rain & Thunderstorm & scattered clouds, light rain & scattered clouds

4 = heavy rain & ice pellets & thunderstorm & mist, snow & fog

*Temp*: Temperature (in degrees Celsius)

*TempFeel*: "Feels Like" Temperature (in degrees Celsius)

*Humidity*: relative humidity in percent

*Windspeed*: windspeed in km/hour

*Casual*: number of casual bike users

*Registered*: number of registered bike users