

I Still Want to Ride My Bicycle, I Still Want to Ride My Bike

Fanaee-T, Hadi and Gamma, J. "Event labeling combining ensemble detectors and background knowledge", Progress in Artificial Intelligence (2013); <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. In particular, while having people occasionally use the bike-sharing system is good, the business model often relies on having a large number of casual users. Therefore it might be useful to better understand the characteristics that are associated with increased or decreased numbers of casual users.

We are interested in modeling bike share usage for the Washington D.C./Arlington, VA/MD area. Your research group has been given a random sample of hourly data with the following variables:

Casual: number of casual bike users

Day: 0 = Sunday, ..., 6 = Saturday

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

You have been asked to analyze whether or not there are group differences in the number of casual users depending on the daily weather and the day of the week.