Capstone 1 Project Proposal: Seoul Bike Sharing Demand Prediction

Tzu-Ying Chen

# Problem

Seoul, South Korea is one of the world largest metropolitan area, with around 24 million people living in the city. The transportation within the city has become a public issue to be solved. Rental bikes are then introduced to Seoul for the eco-friendly solution for the commuters and the enhancement of mobility comfort. In order to gain revenue and achieve promised service level, the bike rental companies need to ensure the rental bikes are available and accessible while maximizing the usage and minimizing the idle time.

# Goal and Utility

The goal of this project is to analyze the bike rental activities at different time on different days and in different weather, and predict the appropriate bike count required at each hour for the stable supply of the rental bikes. The analysis and the prediction can ensure a stable supply of the rental bikes for the publics and help the bike sharing companies schedule regular maintenance for the bikes.

# Data

The data set for this project contains 8,760 observations including weather information, such as temperature, humidity, and windspeed, the number of bikes rented per hour, datetime information, such as holidays and date.

The data set is available on UCI Machine Learning Repository. The page contains information about the data set and its source and relevant scientific publications:

<https://archive.ics.uci.edu/ml/datasets/Seoul+Bike+Sharing+Demand#>

# Approach

It is a supervised regression task. The count of the rental bikes will be our target variable, while the datetime, weather information, and holidays will be used as the predictive variables. Linear regression, ridge regression, or Random Forest regression are some of the machine learning models that could be used in this project. Data exploratory analysis will provide deeper understandings of the data and address the problem more clearly, and cross-validation could ensure the outcomes of the model is reliable.

# Deliverables

The deliverables for the project will include the source code, the data set, and a paper addressing the purpose, approach, findings and results of the project.