**Predicting the Number of rented Bike By hour**

Nada Al-Towerky

**Abstract**

The goal of this project was to use Regression models to predict the number of rented Bike By hour in Seoul to help provide enough bikes during the day.

I worked with data which is in UCI website.

**Design**

This project originated in Data Science Camp as a project that tests how well students are in the skills they learned during camp

I chose this data which is related to the weather because I believe that a clean environment and good weather give a healthy body

That is why we encourage the use of bicycles by providing them in sufficient number on a day with good weather, it is bad to have someone who wants to use a bicycle and does not find one to rent (we lost a cyclist)

**Data**

The dataset contains 8760 points with 14 features for each, 3 of which are categorical. A few feature highlights include Temperature, Hour, Visibility, Wind speed, seasons and is this day holiday or not.

an in-depth analysis of 4 (snowfall - rainfall - Hour - season)

**Algorithms**

1. *1. Change the date into categories (month - day name - day number of the week)*
2. *2. Grouping by date*

*Models*

Regular Linear Regression,Lasso Regression.

*Model Evaluation and Selection*

The entire training dataset of 8760 records was split using sklearn.model\_selection to Split arrays or matrices into random train and test subsets.

The official metric was R² .

R^2: 0.450

Mean Absolute Error: 356.657

Mean Squared Error: 229462.502

Root Mean Squared Error: 479.022

**Holdout**

Alpha: 7.087

R^2: 0.462

Mean Squared Error: 223710.438

**Tools**

* Numpy and Pandas for data manipulation
* Scikit-learn for modeling
* matplotlib.pyplot and Seaborn for plotting
* statsmodels.api for modeling and summry

**Communication**

In addition to the slides and visuals presented, Seoul Bike sharing will be embedded on my page in Githup here.