

# Progress Report of the Big Data Competition

Xishuang Dong

September 28, 2016

**Date: 09/27/16**

- Construct the baseline systems. (Finished by 09/29/16)
  1. Define problems: Forecasting Hourly Female Demand of Rental Bikes in New York City (Done by Xishuang)
  2. Build samples for training and testing model: Selecting features from attributes of data including bike data, taxi data, and weather data. (Done by Safat and Joshua)
  3. Select machine learning models to build the systems: Linear Regression (LR), Artificial Neural Network (ANN) and Support Vector Regression (SVR) (Done by Xishuang, Joshua, and Safat)
  4. Visualize the statistic results on bike data with pie charts: (1) the distributions of gender; (2) the distributions of numbers of hourly rental bikes in different stations; (3) the distributions of numbers of hourly returned bikes in one days. (Done by Travon and DeAhna)
- Next step. (Finished by 09/30/16)
  1. Figure out how to model Female. (Done by Xishuang)
  2. Build samples for constructing the models. (Done by Safat and Joshua)
  3. Select machine learning models to build user models. (Done by Xishuang, Joshua, and Safat)