

# Ride Hailing Supply and Demand Forecasting using Didi-Tech Dataset

## Project Proposal

Kimberly Williamson, Gopal Menon  
Data Mining  
Spring 2017, University of Utah

### I. BACKGROUND [1]

Very few people in China own cars and they mostly rely on ride hailing services for transportation. Didi Chuxing is the leading ride hailing company in China and processes over 11 million trips, plans over 9 billion routes and collects over 50TB of data per day. Forecasting of ride supply and demand is critical to enabling Didi to maximize utilization of drivers and to ensure that riders can always get a car whenever and wherever they may need a ride. It also helps to predict the volume of drivers and riders at a certain time period in a specific geographic area. For instance, demand tends to surge in residential areas in the mornings and in business districts in the evenings. Supply-demand forecasting allows Didi to predict demand surges and guide drivers to those areas. The aim is to drive towards higher earnings for drivers and no surge pricing for riders.

The Didi Research Institute organized a worldwide algorithm challenge in the year 2016 for forecasting ride supply and demand. They provided a train and test dataset and asked teams to come up with a ride supply and demand forecasting algorithm. Algorithms submitted by the teams were evaluated based on a published metric that measured the average error in the forecast.

### II. PROJECT SCOPE

The 2016 Didi Algorithm challenge is now over and they are planning to organize another competition soon [1]. The project team will use the dataset from the 2016 competition and forecast the ride supply and demand using the regression methods that will be covered in the Data Mining 2017 Spring semester at the School of Computing in the University of Utah. The accuracy of the forecast will be evaluated using the same metric that was used in the competition. In addition to the techniques covered in class, the project team will also select two techniques that are either new or will be built on techniques covered in class with the aim of achieving a higher forecasting accuracy.

### III. PROJECT TEAM

Kimberly Williamson and Gopal Menon will work on the project.

### IV. DATA

The train and test dataset will be downloaded

### V. WHY THIS IS INTERESTING

### VI. WHAT IS NEW

### REFERENCES

- [1] "Algorithm Competition." *Algorithm Competition*. N.p., n.d. Web. 28 Jan. 2017.