**Project Proposal**

**Project Title:**

Repayment Prediction of Overdue Payments of Online Loan Platform Users

**Team Members:**

Yan Zhou(N19862016)

Yipeng Qu(N12062723)

**Abstract:**

Nowadays users' dishonesty incidents occur frequently with more and more online lenders appear. Therefore it becomes more important to know how to lower the probability of this kind of event. Our project is based on the data set of lending club [1]. We preprocess the data at first, then do feature engineering. Moreover, we will predict the probability of the events of past-due payment with the algorithms of LogisticRegression、SVM、GBDT[2], etc. At last, we will use ROC-AUC to evaluate the effectiveness of the algorithm.

**Milestones:**

The implement of our project could be separated into 4 steps:

1. Collect the raw data.
2. Get featured data by cleaning the raw data, filling the missing values and doing feature engineering.
3. Get the predicted data by using the algorithms of LogisticRegression, SVM, GBDT, etc.
4. Evaluate each predicted result from each model by using ROC-AUC.

**References：**

[1] https://www.lendingclub.com/info/download-data.action

[2] https://en.wikipedia.org/wiki/Gradient\_boosting#Gradient\_tree\_boosting