CSE 847 (Spring 2022): Machine Learning— Homework 4

Han Meng

1 Logistic Regression: Experiment

I set epsilon as 5e-2 and maxiter as 1000 respectively.

Generally speaking, the Accuracy increases as the number of trainning data increases as the below figure shows. Code can be found at `https://github.com/dearhannah/cse847ss22/tree/`

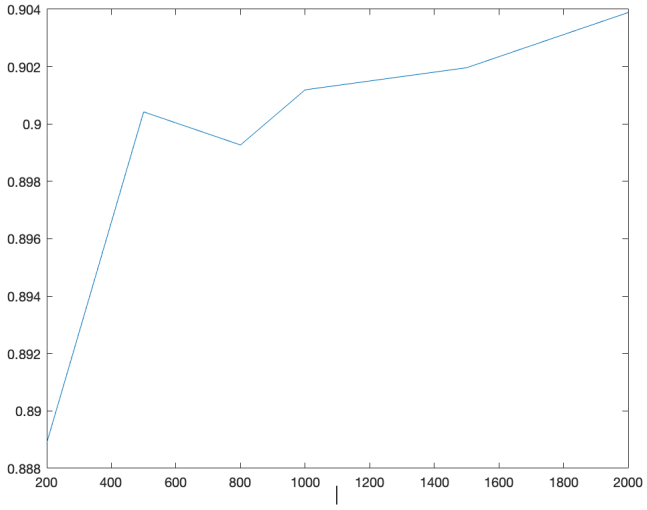

Figure 1: Accuracy vs. Number of Trainning Data

`main/hw4`, main_1.m and logistic_train.m is for question 1. The results produced by my code is:

```
200: 0.88889
500: 0.90042
800: 0.89927
1000: 0.90119
1500: 0.90196
2000: 0.90388
```

2 Sparse Logistic Regression: Experiment

I tried different values of regularization parameter in the suggested list, and here is my results:

```
par: 0, auc: 0.702392, number of features: 116
par: 0.01, auc: 0.629665, number of features: 105
par: 0.1, auc: 0.698565, number of features: 14
par: 0.2, auc: 0.679426, number of features: 5
par: 0.3, auc: 0.644976, number of features: 3
par: 0.4, auc: 0.622967, number of features: 2
par: 0.5, auc: 0.62201, number of features: 1
```

```
par: 0.6, auc: 0.62201, number of features: 1
par: 0.7, auc: 0.62201, number of features: 1
par: 0.8, auc: 0.62201, number of features: 1
par: 0.9, auc: 0.62201, number of features: 1
par: 1, auc: 0.5, number of features: 0
```

Code can be found in the same directory as question 1 and main_2.m is for question 2.