This assignment contains two sections:

1. Experiments and analysis
   1. Logistic regression
   2. Support vector machines with linear kernel
   3. K-fold cross validation implementation and testing result
   4. Support vector machine Gussian kernel
2. Logistic regression and SGD (CSC503 only)

# Pre-processing data

Rescaling

Reassign labels

# Experiments and analysis

# Logistic regression

This part will test the regularization parameter *C* using all (2000) test dataset. The classifier is trained using all 12000 examples.

## SVM with linear kernel

## Implementing K-fold cross validation

## SVM with Gussian kernel

# Logistic regression and SGD

## Implement SGD

The steps of logistic regression with SGD can be summarized as below:

Loop over the following process until some stop condition is met:

1. Calculate the cost.
2. Calculate the gradient.
3. Update the gradient.

The cost is easy to calculated and is defined as:

where:

* is the cross-entropy loss;
* is the sigmoid function;
* is the regulation parameter.

The gradient is calculated and is presented below:

Three different learning rate is tested, they are: constant learn rate, time based learning rate and step-based learning rate, as showing in figure XXXXX

A figure

The implemented code is presented in APENDIX AAAAA

## Testing and analysis