

Programming Exercise 2: Logistic Regression

Machine Learning

Introduction

In this exercise, you will implement logistic regression and apply it to two different datasets. Before starting on the programming exercise, we strongly recommend watching the video lectures and completing the review questions for the associated topics.

To get started with the exercise, you will need to download the starter code and unzip its contents to the directory where you wish to complete the exercise. If needed, use the `cd` command in Octave/MATLAB to change to this directory before starting this exercise.

You can also find instructions for installing Octave/MATLAB in the “Environment Setup Instructions” of the course website.

Files included in this exercise

- `ex2.m` - Octave/MATLAB script that steps you through the exercise
- `ex2_reg.m` - Octave/MATLAB script for the later parts of the exercise
- `ex2data1.txt` - Training set for the first half of the exercise
- `ex2data2.txt` - Training set for the second half of the exercise
- `submit.m` - Submission script that sends your solutions to our servers
- `mapFeature.m` - Function to generate polynomial features
- `plotDecisionBoundary.m` - Function to plot classifier’s decision boundary
- [*] `plotData.m` - Function to plot 2D classification data
- [*] `sigmoid.m` - Sigmoid Function
- [*] `costFunction.m` - Logistic Regression Cost Function
- [*] `predict.m` - Logistic Regression Prediction Function
- [*] `costFunctionReg.m` - Regularized Logistic Regression Cost

* indicates files you will need to complete

Throughout the exercise, you will be using the scripts `ex2.m` and `ex2_reg.m`. These scripts set up the dataset for the problems and make calls to functions that you will write. You do not need to modify either of them. You are only required to modify functions in other files, by following the instructions in this assignment.

Where to get help

The exercises in this course use Octave¹ or MATLAB, a high-level programming language well-suited for numerical computations. If you do not have Octave or MATLAB installed, please refer to the installation instructions in the “Environment Setup Instructions” of the course website.

At the Octave/MATLAB command line, typing `help` followed by a function name displays documentation for a built-in function. For example, `help plot` will bring up help information for plotting. Further documentation for Octave functions can be found at the [Octave documentation pages](#). MATLAB documentation can be found at the [MATLAB documentation pages](#).

We also strongly encourage using the online **Discussions** to discuss exercises with other students. However, do not look at any source code written by others or share your source code with others.

1 Logistic Regression

In this part of the exercise, you will build a logistic regression model to predict whether a student gets admitted into a university.

Suppose that you are the administrator of a university department and you want to determine each applicant’s chance of admission based on their results on two exams. You have historical data from previous applicants that you can use as a training set for logistic regression. For each training example, you have the applicant’s scores on two exams and the admissions decision.

Your task is to build a classification model that estimates an applicant’s probability of admission based the scores from those two exams. This outline and the framework code in `ex2.m` will guide you through the exercise.

¹Octave is a free alternative to MATLAB. For the programming exercises, you are free to use either Octave or MATLAB.