

Title of the Course: Deep Learning

Code and Credit Structure: IT549 (3-0-2-4)

Experiment 1:

Study Linear Regression and Logistics Regression

Outcome:

This experiment is divided into three parts

- 1) you will implement linear regression with one variable and get to see it work on data
- 2) you will implement linear regression with multiple variables.
- 3) you will implement logistic regression

PART I:

In this part of experiment, you will implement linear regression with a one variable to prediction of Graduate Admission from an Indian perspective. Suppose you are the student and seeking admission to oversea university for your PG programs. You have data set of GRE score of students (out of 340) and chance of Admit (ranging from 0 to 1).

You would like to use this data set to predict what is your chance to Admit in oversea university based on your GRE score.

The file GREdata.csv contains the dataset for our linear regression problems. The first column is the GRE score of students and second column is chance of Admit.

You have to write a code in Python with following requirements.

- 1) Plotting of Data:
Before starting on any task, it is often useful to understand the data by visualizing it. For this dataset, you can use a scatter plot to visualize the data, since it has only two properties to plot (GRE Score and Chance to Admit).
- 2) Gradient Descent:
In this part, you will fit the linear regression parameters θ to our dataset using gradient descent.
 - A) Initialize the theta and learning rate
 - B) Fix some iteration
 - C) Compute cost function.
- 3) Plot linear regression fit:
Plot linear fit by keeping previous data plot visible.
- 4) Predict chance of Admit for your GRE score.

PART II:

In this part, you will implement linear regression with multiple variables to predict the chance of Admit. Other than GRE score many various parameter depends to admit in oversea university. The dataset contains following parameters which are considered important during the application for PG Programs.

The parameters included are:

- a. GRE Scores (out of 340)
- b. TOEFL Scores (out of 120)
- c. University Rating (out of 5)
- d. Statement of Purpose and Letter of Recommendation Strength (out of 5)
- e. Undergraduate GPA (out of 10)
- f. Research Experience (either 0 or 1)

Select any two parameters to predict the chance of admit.

You have to write a code in Python with following requirements.

- 1) Feature Normalization:
When features differ by orders of magnitude, first performing feature scaling can make gradient descent converge much more quickly
 - Subtract the mean value of each feature from the dataset.
 - After subtracting the mean, additionally scale (divide) the feature values by their respective “standard deviations”.
- 2) Gradient Descent:
In this part, you will fit the linear regression parameters θ to our dataset using gradient descent.
 - D) Initialize the theta and learning rate
 - E) Fix some iteration
 - F) Compute cost function.
- 3) Plot the convergence graph:
- 4) Predict chance of Admit for your GRE score.

PART III:

In this part of the exercise, you will build a logistic regression model to predict whether a patient has chance of heart attack or not.

This dataset has following attributes

Attribute Information

- 1) age
- 2) sex
- 3) chest pain type (4 values)
- 4) resting blood pressure
- 5) serum cholestoral in mg/dl
- 6) fasting blood sugar > 120 mg/dl
- 7) Resting electrocardiographic results (values 0,1,2)
- 8) maximum heart rate achieved

9) exercise induced angina

10) oldpeak = ST depression induced by exercise relative to rest

11) the slope of the peak exercise ST segment

12) number of major vessels (0-3) colored by fluoroscopy

13) thal: 0 = normal; 1 = fixed defect; 2 = reversible defect

14) target: 0 = less chance of heart attack 1 = more chance of heart attack

Your task is to build a classification model that estimates that patient's probability of heart attack based on any two attributes from 1 to 13.

You have to write a code in Python with following requirements.

1) Plotting of Data:

Before starting on any task, it is often useful to understand the data by visualizing it. For this dataset, both axes will be attributes you choose.

2) Cost function and gradient:

Computes the cost and gradient using theta as the parameter for logistic regression and the gradient of the cost w.r.t. to the parameters (use gradient descent).

3) Predict for input data and calculate accuracy on training and testing set.