Deep Learning
Assignment1
(Linear Regression + Logistic Regression)
Submission date: Sept 15, 2024

Problem 1 (House Price Prediction)

In this task you will design linear regression model to estimate the price of a dependent variable. You can assume that the dataset is in csv format. You have to perform the following tasks. For each task implement a function in python

Load Data - Load the dataset given in CSV format and make matrix X and vector Y where X store the independent variables(features) and Y is the dependent variable (estimated value)

Normalize Data – Normalize the data using zero mean normalization

Split the data – Split the data into train and test sets with 70% data as training data and 30% data as test data (you can use sklearn model selection for this task).

H_function: This function computes the estimated price of the house for all the examples in the train set.

Cost_function: This function must compute the cost of a given hypothesis for all the training examples in the train set. Use mean squared error as cost function.

Gradient_descent: This function computes the gradients of all the learnable parameters and update them according to the learning_rate. The learnable parameters must be updated num_epoch times. Here learning_rate and num_epoch are user defined parameters.

Plot_data: This function must plot the cost after each epoch for both the train and test set. X-axis will have number of epochs and Y axis will have value of cost function.

Use your implementation of linear regression to estimate the price of a house. The required dataset is attached with this problem statement. Check your model with different values of learning_rate and num_epoch and plot the costs of train and test set in a single plot. Report the best value of num_epoch and learning_rate.

Problem 2 (Churn Analysis)

Extend the implementation of linear regression to logistic regression. Implement the sigmoid function and update the cost_function (use cross entropy as cost function) and Gradient_descent functions accordingly.

Use your implementation of logistic regression to perform the bank customer churn analysis. The dataset is attached. This dataset includes two categorical features/columns: country and

gender. Use one hot encoding to preprocess these features and then perform the logistic regression. Check your model with different values of learning_rate and num_epoch and plot the costs of train and test set in a single plot. Report the best value of num_epoch and learning_rate.

What to Submit: Submit the python codes for both the problems and a document that elaborate your experiments for house price prediction and churn Analysis.