CMPE 462
Assignment 2

Mahir Efe KAYA - 2016400195

PART1

Description: Implementing a Logistic Regression model from scratch to predict a classification problem with 5-fold cross validation.

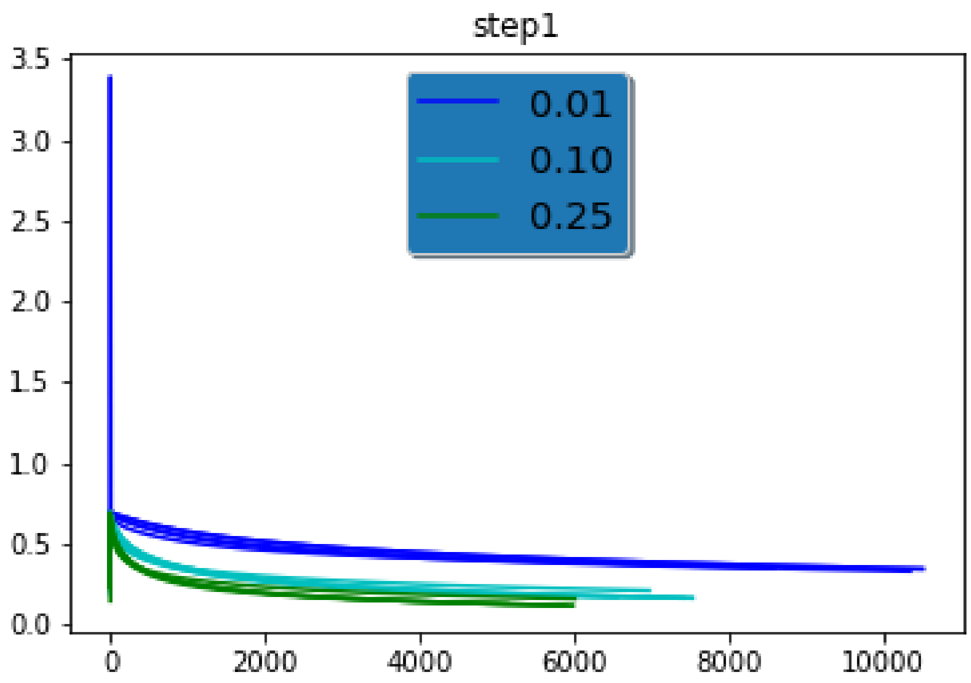
How to Run: pip install -r requirements.txt

python main.py

Then the figures will be saved in the same folder, and the averages will be printed in the console.

Step1

Implementing LR with batch gradient descent.



As the legend explains;

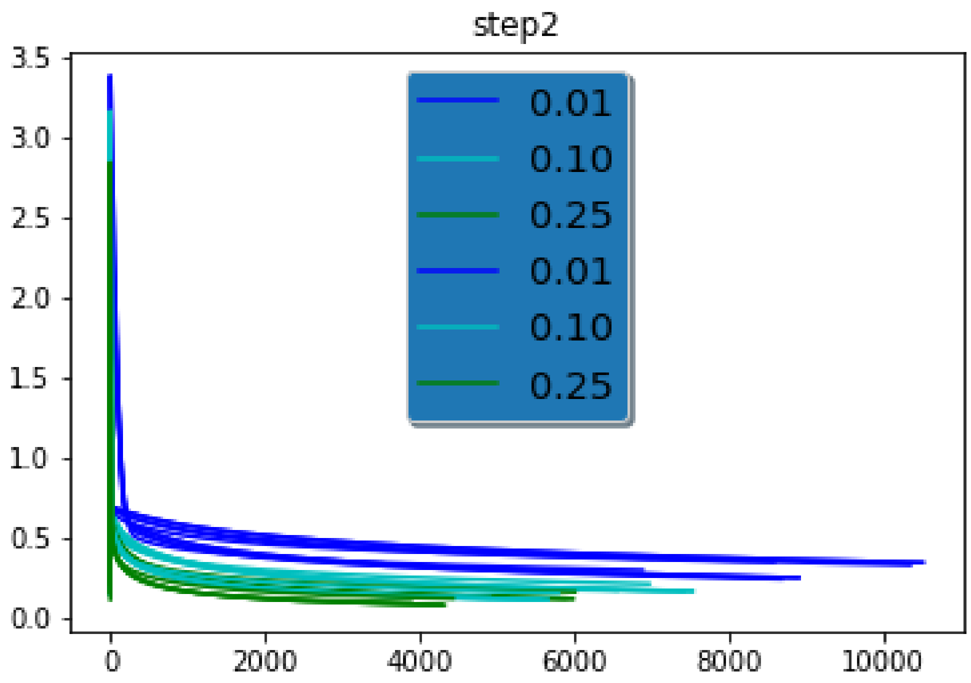
- Blue is for 0.01 as stepsize
- Light blue is for 0.10 as stepsize
- Green is for 0.25 as stepsize

There are multiple lines with the same color. It is done so to show the losses of different folds.

From the plot we can see that the bigger the stepsize is, the faster and the more accurate it gets. Accuracy got even me surprised, but the reason is probably its getting a big step just before the threshold.

Step2

Implementing LR with stochastic descent.



As the legend explains;

- Blue is for 0.01 as stepsize
- Light blue is for 0.10 as stepsize
- Green is for 0.25 as stepsize

For the same reason as before, there are multiple lines with the same color to show the losses of different folds.

From the plot we can see that the bigger the stepsize is, the faster and the more accurate it gets. Accuracy got even me surprised, but the reason is probably its getting a big step just before the threshold.

After several runs, we can safely assume that using stochastic gradient descent for LR is faster and slightly more accurate than batch gradient descent LR.

PART2

P(GiveBirth|mammals) = 6/7
P(GiveBirth|non-mammals) = 1/13

P(CanFly|mammals) = 1/7
P(CanFly|non-mammals) = 3/13

P(LiveInWater = “yes”|mammals) = 2/7
P(LiveInWater = “no”|mammals) = 5/7
P(LiveInWater = “sometimes”|mammals) = 0/7

P(LiveInWater = “yes”|non-mammals) = 3/13
P(LiveInWater = “no”|non-mammals) = 6/13
P(LiveInWater = “sometimes”|non-mammals) = 4/13

P(HaveLegs|mammals) = 5/7
P(HaveLegs|non-mammals) = 9/13

P(mammals) = 7/20

$$P(test = mammals) =$$

$$= P(GiveBirth|mammals) * P(CanFly|mammals)' * P(LiveInWater"yes"|mammals) * P(HaveLegs|mammals)' * P(mammals)$$

$$= \frac{6}{7} * \frac{6}{7} * \frac{2}{7} * \frac{2}{7} * \frac{7}{20}$$

$$= 0.021 = 2.1 * 10^{-2}$$

$$P(test = non - mammals) =$$

$$= P(GiveBirth \mid non - mammals) * P(CanFly \mid non - mammals)' * P(LiveInWater"yes"|non - mammals) * P(HaveLegs|non - mammals)' * P(non - mammals)$$

$$= \frac{1}{13} * \frac{10}{13} * \frac{3}{13} * \frac{4}{13} * \frac{13}{20}$$

$$= 0.0027 = 2.7 * 10^{-3}$$

Since P(test=mammals) > P(test=nonmammals). Test is mammal.