Shahid Bahonar
University of Kerman

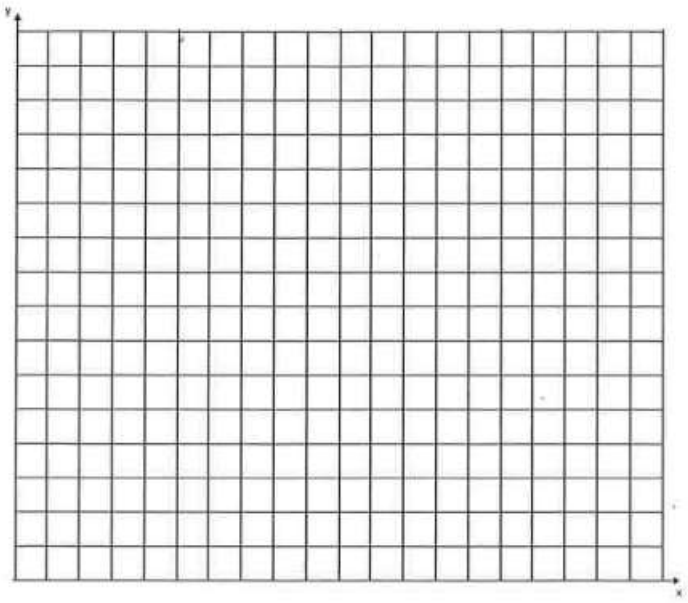
Machine Learning Course Autumn semester

By Dr. Mahdi Eftekhari Homework #3

## Linear Regression

1. A convenience store manager notices that sales of soft drinks are higher on hotter days, so he assembles the data in the table.

(a) Make a scatter plot of the data.



| High Temperature (°F) | Number of cans sold |
|---|---|
| 55 | 340 |
| 58 | 335 |
| 64 | 410 |
| 68 | 460 |
| 70 | 450 |
| 75 | 610 |
| 80 | 735 |
| 84 | 780 |

(b) Find and graph a linear regression equation that models the data

(e) Use the model to predict soft-drink sales if the temperature is 95F

(d) What does the model predict for the temperature if the number of cans sold was only95?

2) A company wants to launch a new product and wants to know whether it will turn out to be a success or failure. We have information on the last 100 products this company launched, including if it was a success/failure, price, weight, color, and several other variables

b)We have information on several Bay Area Tech Companies, including size, industry, revenue, average employee salary, and more. We want to know which features influence the average employee salary.

c)You are given data of 100 individuals and their sequenced DNA and want to know whether these individuals will exhibit a particular disease based off their genomic mutations. We have information on 10,000 individual genomes and whether or not they exhibit the particular disease.

3)Suppose you have been given the following scenario for training and validation error for Linear Regression.

| Hyperparameter setting | Learning Rate | Number of iterations | Training Error | Validation Error |
|---|---|---|---|---|
| 1 | 0.1 | 1000 | 100 | 110 |
| 2 | 0.2 | 600 | 105 | 105 |
| 3 | 0.3 | 400 | 110 | 110 |
| 4 | 0.4 | 300 | 120 | 130 |
| 5 | 0.4 | 250 | 130 | 150 |

Which of the following hyperparameter settings is seemingly the best? **State your reason**

4) . Obtain a linear regression for the data in Table 1 assuming that y is the independent variable.

Table1: Example data for simple linear regression

| x | 1.0 | 2.0 | 3.0 | 4.0 | 5.0 |
|---|---|---|---|---|---|
| y | 1.00 | 2.00 | 1.30 | 3.75 | 2.25 |