Sheet 7

A. 1. Since $\theta \in R^{(d+1)\times 1}$

, and since $d = 2 \Rightarrow \therefore \theta \in R^{3\times 1}$

Thus, we need 3 parameters to find to solve linear regression.

2. Closed form Solution :

$$\theta = (X^T X)^{-1} X^T y$$

$$X_{10\times 3} = \begin{bmatrix} 1 & 1 & 5 \\ 1 & 2 & 6 \\ 1 & 2 & 10 \\ 1 & 2 & 12 \\ 1 & 3 & 17 \\ 1 & 3 & 12 \\ 1 & 4 & 6 \\ 1 & 4 & 5 \\ 1 & 4 & 7 \\ 1 & 5 & 10 \end{bmatrix} \quad y_{10\times 1} = \begin{bmatrix} 10 \\ 40 \\ 50 \\ 60 \\ 70 \\ 50 \\ 30 \\ 20 \\ 40 \\ 70 \end{bmatrix}$$

$$X^T y_{3\times 1} = \begin{bmatrix} 440 \\ 1380 \\ 4560 \end{bmatrix}$$

$$X^T X_{3\times 3} = \begin{bmatrix} 10 & 30 & 90 \\ 30 & 104 & 270 \\ 90 & 270 & 948 \end{bmatrix}$$

$$\underbrace{\theta}_{3\times 1} = \underbrace{(X^T X)^{-1}}_{3\times 3} \underbrace{X^T y}_{3\times 1} = \begin{bmatrix} -7.987 \\ 4.2857 \\ 4.3478 \end{bmatrix}$$

$$(X^T X)^{-1} = \begin{bmatrix} 1.3298 & -0.214 & -0.065 \\ -0.214 & 0.0714 & 0 \\ -0.065 & 0 & 7.24\times 10^{-3} \end{bmatrix}$$

Model :

$$h(x) = \sum_{j=0}^{d} \theta_j \cdot x_j$$

$$= -7.987 + 4.2857 x_1 + 4.3478 x_2$$

Sheet#7 Ensemble+NeuralNets+ Linear Regression

April 21, 2019

1 Sheet#7 Ensemble+NeuralNets+ Linear Regression

1.1 Question 1 on data given below.

```
In [20]: import numpy as np
         from sklearn.linear_model import LinearRegression
         from sklearn.linear_model import Ridge
```

```
In [3]: X = np.array([[1,1,5],[1,2,6],[1,2,10],[1,2,12],[1,3,17],[1,3,12],[1,4,6],[1,4,5],[1,4
        Y = np.array([10,40,50,60,70,50,30,20,40,70])
```

1.1.1 Use Scikit-learn package for Finding the linear regression solution.

```
In [19]: reg = LinearRegression().fit(X, Y)
         print("Theta 0 is ",reg.intercept_)
         print("Theta 1 is",reg.coef_[1])
         print("Theta 2 is",reg.coef_[2])
```

```
Theta 0 is  -7.987577639751535
Theta 1 is 4.2857142857142865
Theta 2 is 4.34782608695652
```

Comparing to the Closed form solution obtained perviously, it is exactly the same answer we got from the sklearn algorithm.

1.1.2 Add L-2 regularization to the obtained solution. We use Ridge regression from Scikit-learn to do so.

For alpha = 0.1,

```
In [34]: clf = Ridge(alpha=0.1)
         clf.fit(X, Y)
         print("Theta 0 is ",clf.intercept_)
         print("Theta 1 is",clf.coef_[1])
         print("Theta 2 is",clf.coef_[2])
```

```
Theta 0 is  -7.868057374397225
Theta 1 is 4.255319148936171
Theta 2 is 4.344677769732079
```

For alpha = 1,

```
In [39]: clf = Ridge(alpha=1)
         clf.fit(X, Y)
         print("Theta 0 is ",clf.intercept_)
         print("Theta 1 is",clf.coef_[1])
         print("Theta 2 is",clf.coef_[2])
```

```
Theta 0 is  -6.84892086330936
Theta 1 is 3.9999999999999996
Theta 2 is 4.316546762589929
```

For alpha = 10,

```
In [43]: clf = Ridge(alpha=10)
         clf.fit(X, Y)
         print("Theta 0 is ",clf.intercept_)
         print("Theta 1 is",clf.coef_[1])
         print("Theta 2 is",clf.coef_[2])
```

```
Theta 0 is  0.013513513513508713
Theta 1 is 2.5000000000000004
Theta 2 is 4.054054054054054
```

For alpha = 100,

```
In [47]: clf = Ridge(alpha=100)
         clf.fit(X, Y)
         print("Theta 0 is ",clf.intercept_)
         print("Theta 1 is",clf.coef_[1])
         print("Theta 2 is",clf.coef_[2])
```

```
Theta 0 is  19.73197700132685
Theta 1 is 0.5263157894736842
Theta 2 is 2.5210084033613445
```

Regressor 2 (Ridge Regression $\lambda = 20$)

P1: $y = 72.378$

P2: $y = 21.229$

P3: $y = 28.7297$

$y = 0.0135 + 2.5 x_1 + 4.054 x_2$

Regressor 3 (Ridge Regression $\lambda = 100$)

P1: $y = 62.647$

P2: $y = 30.868$

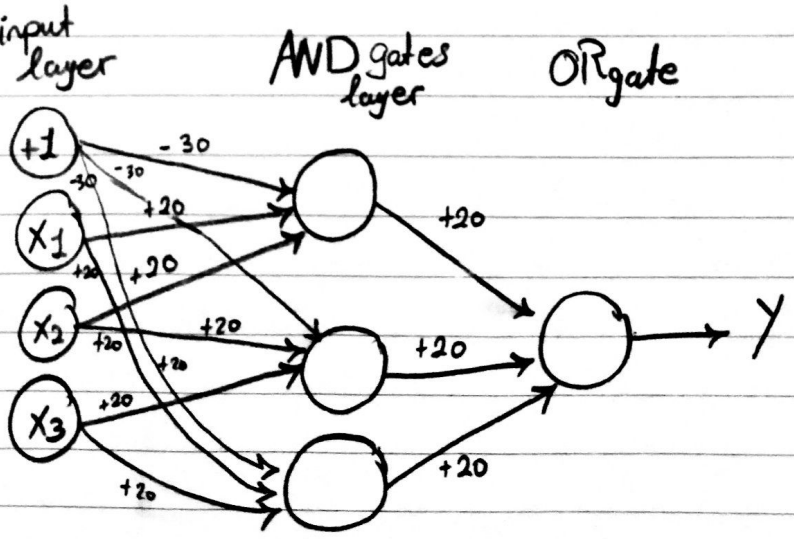
P3: $y = 32.447$

$y = 19.7319 + 0.526 x_1 + 2.522 x_2$

4.

3-input Majority Function

X_1	X_2	X_3	Y
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	0	1



($Y = X_1X_2 + X_2X_3 + X_1X_3$)

3.

Regressor 1: (From linear Regression without Regularization)

$y = -7.9875 + 4.2857 x_1 + 4.3478 x_2$

For P1 (3, 16): $y = 74.4347$

For P2 (2, 4): $y = 17.9751$

For P3 (5, 4): $y = 30.8322$

Regressor 2: (From Ridge Regression with $\lambda = 0.1$)

$y = -7.868 + 4.255 x_1 + 4.344 x_2$

For P1: $y = 74.4127$

For P2: $y = 18.021$

For P3: $y = 30.787$

Regressor 3: (From Ridge Regression with $\lambda = 1$)

$y = -6.8489 + 3.999 x_1 + 4.316 x_2$

P1: $y = 74.216$

P2: $y = 18.417$

P3: $y = 30.417$