5.1

feature of model -> The output is discrete, only binary output,

output -> the class of the input sample, in only true of false (1 or 0)

discriminant function ->

$$g(x) = w^T x = w_0 + w_1 x_1 + w_2 x_2$$

where y_n is the dependent variable of the sample being assessed as congested

x_n is for independent variable of the sample (x_0 as intercept (=1 for all cases), x_1 for Traffic Density, x_2 for Traffic Volume)

w is the parameters (w_0 as intercept, remaining w_n correspond to their specific x_n)

5.2

from the model parameter, it can be deduce that

$$y_n = 1, x_0 = 1, x_1 = 30, x_2 = 1000$$

The equation of the model for the probability is

$$P(y_n = 1|x_n) = \frac{e^{w_0 + w^T x_n}}{1 + e^{w_0 + w^T x_n}} = \frac{e^{w_0 + 30w_1 + 1000w_2}}{1 + e^{w_0 + 30w_1 + 1000w_2}}$$

5.3

Joint probability

$$L(w_0, w_1, w_2) = \frac{e^{w_0 + 50w_1 + 2000w_2}}{1 + e^{w_0 + 50w_1 + 2000w_2}} \cdot \frac{1}{1 + e^{w_0 + 20w_1 + 900w_2}} \cdot \frac{e^{w_0 + 40w_1 + 1500w_2}}{1 + e^{w_0 + 40w_1 + 1500w_2}} \cdot \frac{1}{1 + e^{w_0 + 10w_1 + 500w_2}} \cdot \frac{1}{1 + e^{w_0 + 5w_1 + 300w_2}}$$

5.4

Cross-entropy

$$E(w_0, w_1, w_2) =$$
$$-logL(w_0, w_1, w_2) =$$
$$-[log(\frac{e^{w_0+50w_1+2000w_2}}{1+e^{w_0+50w_1+2000w_2}}) + log(\frac{1}{1+e^{w_0+20w_1+900w_2}}) + log(\frac{e^{w_0+40w_1+1500w_2}}{1+e^{w_0+40w_1+1500w_2}}) + log(\frac{1}{1+e^{w_0+10w_1+500w_2}}) + log(\frac{1}{1+e^{w_0+5w_1+300\iota}}$$

Taking derivitive to the cross-entropy, and using equation (21) from lecture 3

$$\frac{\partial E}{\partial w_j} = \sum_{n=1}^{n=5}(y_n - z_n)x_{nj}$$

Calculate the individual w_n from 0 to 2

$$\frac{\partial E}{\partial w_0} = x_0[(1-z_1) + (-z_2) + (1-z_3) + (-z_4) + (-z_5)]$$
$$= (\frac{1}{1+e^{w_0+50w_1+2000w_2}} - \frac{e^{w_0+20w_1+900w_2}}{1+e^{w_0+20w_1+900w_2}} + \frac{1}{1+e^{w_0+40w_1+1500w_2}} - \frac{e^{w_0+10w_1+500w_2}}{1+e^{w_0+10w_1+500w_2}} - \frac{e^{w_0+5w_1+300w_2}}{1+e^{w_0+5w_1+300w_2}})$$

$$\frac{\partial E}{\partial w_1} = [x_{11}(1-z_1) + x_{21}(-z_2) + x_{31}(1-z_3) + x_{41}(-z_4) + x_{51}(-z_5)]$$
$$= (\frac{50}{1+e^{w_0+50w_1+2000w_2}} - \frac{20e^{w_0+20w_1+900w_2}}{1+e^{w_0+20w_1+900w_2}} + \frac{40}{1+e^{w_0+40w_1+1500w_2}} - \frac{10e^{w_0+10w_1+500w_2}}{1+e^{w_0+10w_1+500w_2}} - \frac{5e^{w_0+5w_1+300w_2}}{1+e^{w_0+5w_1+300w_2}})$$

$$\frac{\partial E}{\partial w_2} = [x_{12}(1-z_1) + x_{22}(-z_2) + x_{32}(1-z_3) + x_{42}(-z_4) + x_{52}(-z_5)]$$
$$= (\frac{2000}{1+e^{w_0+50w_1+2000w_2}} - \frac{900e^{w_0+20w_1+900w_2}}{1+e^{w_0+20w_1+900w_2}} + \frac{1500}{1+e^{w_0+40w_1+1500w_2}} - \frac{500e^{w_0+10w_1+500w_2}}{1+e^{w_0+10w_1+500w_2}} - \frac{300e^{w_0+5w_1+300w_2}}{1+e^{w_0+5w_1+300w_2}})$$

5.5

Initalize

$$w_0 = 1, w_1 = 1, w_2 = 1,$$

discriminant function

$$g(x) = w_0 + w_1x_1 + w_2x_2$$

and

$$z = \frac{1}{1 + e^{-(w_0 + w_1 x_1 + w_2 x_2)}}$$

Setting learning rate

$$\eta = 0.01$$

calculate the cross-entropy loss to w,

$$\Delta w_0 = \eta \sum_{n=1}^{5} (y_n - z_n) = 0.01([1 - \frac{1}{1 + e^{-(1+50+2000)}}] + [0 - \frac{1}{1 + e^{-(1+20+900)}}] + [1 - \frac{1}{1 + e^{-(1+40+1500)}}] + [0 - \frac{1}{1 + e^{-(1+10+500)}}]$$
$$+ [0 - \frac{1}{1 + e^{-(1+5+300)}}])$$
$$= 0.01 \times -3 = -0.03$$

$$\Delta w_1 = \eta \sum_{n=1}^{5} (y_n - z_n) x_{n1} = 0.01(50[1 - \frac{1}{1 + e^{-(1+50+2000)}}] + 20[0 - \frac{1}{1 + e^{-(1+20+900)}}] + 40[1 - \frac{1}{1 + e^{-(1+40+1500)}}] + 10[0 - \frac{}{1 + e^{-(}}$$
$$+ 5[0 - \frac{1}{1 + e^{-(1+5+300)}}])$$
$$= 0.01 \times -35 = -0.35$$

$$\Delta w_2 = \eta \sum_{n=1}^{5} (y_n - z_n) x_{n2} = 0.01(2000[1 - \frac{1}{1 + e^{-(1+50+2000)}}] + 900[0 - \frac{1}{1 + e^{-(1+20+900)}}] + 1500[1 - \frac{1}{1 + e^{-(1+40+1500)}}]$$
$$+ 500[0 - \frac{1}{1 + e^{-(1+10+500)}}] + 300[0 - \frac{1}{1 + e^{-(1+5+300)}}])$$
$$= 0.01 \times -1700 = -17$$

At the 2nd iteration, w_0, w_1, w_2 are 0.97,0.65 and -16 respectively