Machine Learning
DSECL ZG565
Problems

Dr. Monali Mavani

BITS Pilani

Pilani Campus

Question 1

Consider the hypothesis function $h(\mathbf{w}, \mathbf{x}) = w_0 + w_1 x_1 + w_2 x_2 + w_3 x_1^2 + w_4 x_2^2$; with parameters

w =< w_0 , w_1 , w_2 , w_3, $w_4 > =$ < -20, -2, -4, 1, 1>.

Here x_1 and x_2 are two features.

- Derive the equation of the decision boundary g(x_1, x_2) for logistic regression given by the equation:

$$y = \frac{1}{1+\exp\{-h(w,x)\}}$$

- Draw the decision boundary and predict the class labels [C_0, C_1] for the examples given by A(-2, 2), B(6, 6) and C(-5, 5).

Question 2

Consider the loss function of linear regression given by: $J(\theta_0 ,\ \theta_1)$.
Given $(\theta_0 ,\ \theta_1)$=0, 0.5, Estimate $\partial J/\partial\theta_1$ using the data points below:

x	**2**	**4**	**7.0**	**8.0**	**10.0**
y	1	2	2.5	3.5	5.5

Question 3

Vijay is a certified Data Scientist and he has applied for two companies -Google and Microsoft. He feels that he has a 60% chance of receiving an offer from Google and 50% chance of receiving an offer from Microsoft. If he receives an offer from Microsoft, he has belief that there are 80% chances of receiving an offer Google.

- What is the probability that both the companies will make an offer to him?
- If Vijay receives an offer from Microsoft, what is the probability that he will not receive an offer from Google?
- What are his chances of getting an offer from Microsoft, considering he has an offer from Google?

Question 4

Suppose that the lifetime of *Badger* brand light bulbs is modeled by an exponential distribution with (unknown) parameter λ. We test 5 bulbs and find they have lifetimes of 2, 3, 1, 3, and 4 years, respectively. What is the MLE for λ?

Question 5

The sales of a company (in million dollars) for each year are shown in the table below.

a) Find the least square regression line y = a x + b.
b) Use the least squares regression line as a model to estimate the sales of the company in 2012.

x (year)	2005	2006	2007	2008	2009
y (sales)	12	19	29	37	45

y = a x + b

$$a = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}$$

$$b = \frac{1}{n}(\sum_{i=1}^{n} y_i - a\sum_{i=1}^{n} x_i)$$

Question 6

Suppose we have a sample of real values, called $x_1, x_2, ..., x_n$. Each sampled from p.d.f. $p(x)$ which has the following form:

$$f(x) = \begin{cases} \alpha e^{-\alpha x}, & if \;\; x \geq 0 \\ 0, & otherwise \end{cases}$$

where α is an unknown parameter. Which one of the following expressions is the maximum likelihood estimation of α? (Assume that in our sample, all x_i are large than 1.)

Ans: $\dfrac{n}{\sum_{i=1}^{n} x_i}$

Question 7

- Derive the maximum likelihood estimator (MLE) for the mean μ of a univariate normal distribution. Assume N samples, x_1, ,x_N independently drawn from a normal distribution with known variance σ^2 and unknown mean μ. Show all intermediate steps and assumptions.

Question 8

- Given N independent measurements $x_1, x_2, \ldots, x_N$, determine the optimal parameters of the model, i.e. the parameters that maximize the probability density function (PDF). To model this data, assume Gaussian distribution.

Question 9

- Consider a dataset for binary classification problem with class labels [C_1, C_0]. The features are given by F_1, F_2 and F_3. Each of these features have two values as given in the dataset below. Apply Naïve Bayes classifier by computing the probabilities to classify the new example: <F_1=x_1, F_2=y_2, F_3=z_1>

Sl No	F_1	F_2	F_3	Class
1	x_1	y2	z1	C_1
2	x2	y_1	z2	C_0
3	x_1	y_1	z2	C_1
4	x2	y2	z1	C_0
5	x2	y_1	z1	C1
6	x2	y2	z1	C_0
7	x_1	y_1	z2	C_1
8	x_1	y2	z2	C_0

Thank you