2023 Gen 12 Final

▼ Source

Artificial Intelligent and Machine Learning

Time: 1h30

Students are allowed to use their personal calculator

Question 1 (7points)

1a. Explain the idea of gradient descent algorithm, define the cost function and the rule of updating coefficient of linear regression.

1b. Explain why the cost function of logistic regression can be estimated by the following formulation:

$$E(\theta) = -\frac{1}{m}\sum_{i=1}^{m}[y^{(i)} \log h_\theta(x^{(i)}) + (1-y^{(i)}) \log(1-h_\theta(x^{(i)})]$$

1c. Assuming at a certain iteration of batch-gradient descent, we reach the global **maximum** value of the cost function ($\frac{\partial}{\partial_\theta} E(\theta)$ is perfectly 0). What will happen next and why?

- The algorithm stays at the global maximum
- The algorithm updates the coefficient toward the minimum value of cost function

Is the answer remaining the same if we are using mini-batch gradient descent or stochastic-gradient descent?

Question 2 (5 points)

Given the training data for a classification problem, to identify if a person is male/female base on his/her weight and high. Stating at $\theta_0 = \theta_1 = \theta_2 = 0$, and learning rate α=0.0001. Calculate the coefficients after the first iteration with batch-gradient descent.

Training Example	Height (cm)	Weight (kg)	Male/Female
1	172	68	Male
2	163	52	Female
3	158	50	Female
4	180	75	Male

Question 3 (5 points):

Given the training data for a linear regression problem as follow:

Input	Output
0	0
1	2
-1	-2
2	3

After the first iteration, the values of two coefficients are:

Θ_0=2 and Θ_1=1

What are the initial values of Θ_0 and Θ_1? Given the learning rate α equals to 4.

Exercise 4 (3 points)

Compute the Accuracy, Precision, Recall, Specificity and F-measure (F score) of the following confusion matrix (1: positive, 0: negative)

	Predicted: 0	Predicted: 1
Actual:0	900	100
Actual:1	25	975

Question 1

1. Explain the idea of gradient descent algorithm, define the cost function and the rule of updating coefficient of linear regression.
2. Explain why the cost function of logistic regression can be estimated by the following formulation:

$$E(\theta) = -\frac{1}{m}\sum_{i=1}^{m}[y^{(i)} log\ h_\theta(x^{(i)}) + (1-y^i) log(1-h_\theta(x^{(i)}))]$$

3. Assuming at a certain iteration of batch-gradient descent, we reach the global maximum value of the cost function ($\frac{\partial}{\partial_0}E(\theta)$ is perfectly 0). What will happen next and why?
   - The algorithm stays at the global maximum.
   - The algorithm updates the coefficient toward the minimum value of cost function.

   Is the answer remaining the same if we are using mini-batch gradient descent or stochastic-gradient descent?

Question 2

Given the training data for a classification problem, to identify if a person is male/female base on his/her weight and high. Stating at $\theta_0 = \theta_1 = \theta_2 = 0$, and learning rate $\alpha = 0.0001$. Calculate the coefficients after the first iteration with batch-gradient descent.

Training example	Height (cm)	Weight (kg)	Male/Female