**Jordan University of Science and Technology**
**Faculty of Computer Science & Information Technology**
**Computer Science Department**
**AI 249- Machine Learning/ Second Exam Spring 2024/2025**

**Name:** **id:**

**CLO1: Understand and explain key machine learning concepts and algorithms.**

**Section 1: Multiple Choice Questions (MCQs) (out of 5)**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | |

**Section 2: True and False Questions (T/F) (out of 5)**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | |

**Section 3: Fill in the Blanks (Mathematics-Based) (out of 5)**

**A.**

| Mean(x) | Mean(y) | B0 | B1 |
|---|---|---|---|
| | | | |

**B.**

| The updated weight is | |
|---|---|

**C.**

| P(Yes\|Sunny) | |
|---|---|

**Section 1: Multiple Choice Questions (MCQ)** Choose the **one** correct answer for each question.

1. **The support vectors in SVM are:**
   a) Points closest to the hyperplane
   b) Randomly selected training points
   c) Points farthest from the hyperplane
   d) Points used only for testing

2. **What is the range of the sigmoid function?**
   a) (-∞, ∞)
   b) (0, ∞)
   c) (0, 1)
   d) (-1, 1)

3. **What is the main purpose of the cost function in machine learning?**
   a) Increase accuracy
   b) Reduce overfitting
   c) Measure model error
   d) Add bias

4. **In Naive Bayes, "naive" refers to:**
   a) Using raw data
   b) Strong assumptions about the data
   c) Assuming feature independence
   d) Ignoring data types

5. **What happens if the learning rate in SGD is too high?**
   a) Converges quickly
   b) Wiggle (wigwag) and may not converge
   c) Gradient becomes zero
   d) Model overfits immediately

6. **Gradient Descent updates parameters in the direction of:**
   a) Higher cost
   b) Zero slope
   c) Gradient increase
   d) Negative gradient

7. What does the B1 coefficient in linear regression represent?
   a) The slope of the line
   b) The intercept
   c) The mean of X
   d) The predicted Y value

8. Which cost function is most suitable for logistic regression?
   a) Mean Absolute Error
   b) Log loss
   c) Euclidean Distance
   d) RMSE

9. Which statement is true for stochastic gradient descent (SGD)?
   a) It always gives the global minimum
   b) It uses all data in one step
   c) It can converge faster than batch gradient descent
   d) It is only used for classification

10. In logistic regression, what does the sigmoid function output represent?
    a) A class label
    b) The cost function
    c) A probability
    d) The slope

**Section 2: True and False Questions (T/F)**

- False – Logistic regression outputs values between 0 and 1 using the sigmoid function.

- True – Stochastic Gradient Descent (SGD) updates weights using one training sample at a time.

- False – A convex function has only one global minimum and no local minima.

- True – In polynomial regression, though input features are raised to powers, the model remains linear in parameters (coefficients).

- True – The learning rate determines the step size taken during each update in gradient descent.

- False – Overfitting happens when a model is too complex, not too simple.

- True – In SVM, a higher C puts more emphasis on reducing classification error, which typically reduces margin width.

- False – Feature scaling is important in gradient descent; otherwise, convergence may be slow or unstable.

- False – Support vectors are critical to determining the decision boundary in SVM.

- True – Linear regression assumes a linear relationship between the dependent and independent variables.

**Section 3: Fill in the Blanks (Mathematics-Based)**

*(Write the correct numerical or conceptual answer in the blank space.)*

## A -

Given the data:
X: [1, 2, 3],
Y: [2, 3, 5]

**Question: Compute mean(x), mean(y)**

**Question: Calculate B1 and B0 using least squares shortcut**

## B-

Given:

- Initial weight W = 2
- Learning rate α = 0.1
- Gradient at W = 3

**Question: Update the weight using one SGD step.**

## C-

Dataset:

- 3 samples:
  - Sunny, Play=Yes
  - Rainy, Play=No
  - Sunny, Play=Yes
- P(Play=Yes) = 2/3, P(Play=No) = 1/3
- P(Sunny|Yes) = 1, P(Sunny|No) = 0

**Question**: If today is sunny, what is P(Yes|Sunny) using Naive Bayes?

**يتوجب على الجميع تعبئة الصفحة الأولى بالحلول حيث لن ينظر لإجاباتك هنا**

Given: X = [1, 2, 3], Y = [2, 3, 5]

- mean(x) = (1+2+3)/3 = 2
- mean(y) = (2+3+5)/3 = 3.33

$$B1 = \frac{(1-2)(2-3.33)+(2-2)(3-3.33)+(3-2)(5-3.33)}{(1-2)^2+(2-2)^2+(3-2)^2} = \frac{1.33+0+1.67}{1+0+1} = \frac{3}{2} = 1.5$$

$$B0 = mean(y) - B1 * mean(x) = 3.33 - 1.5 * 2 = 0.33$$

W = 2, α = 0.1, Gradient = 3

$$W_{new} = W - \alpha * \text{Gradient} = 2 - 0.1 * 3 = 1.7$$

$$P(Play = Yes|Sunny) \propto P(Sunny|Yes) * P(Yes) = 1 * \frac{2}{3} = \frac{2}{3}$$

$$P(Play = No|Sunny) = 0 * \frac{1}{3} = 0$$

Then:

$$P(Yes|Sunny) = \frac{2/3}{2/3+0} = 1.0$$