Korea Aerospace University

2AI4203: Machine Learning

HW01: Foundations of Machine Learning

Jiyoon Shin

KAU College of AI Convergence

Sep. 24 – Sep. 30, 2026

Problem 1. Linear Regression and Gradient Descent

Consider the following training dataset:

$$(x, y) = \{(1, 3), (2, 5), (3, 7)\}.$$

A linear regression model is defined as

$$\hat{y} = wx + b.$$

Suppose the current parameters are

$$w = 1, \qquad b = 1.$$

Problem 1. Linear Regression and Gradient Descent

(a) [5pts] Compute the prediction $\hat{y}$ for each training example.

(b) [5pts] Compute the Mean Squared Error (MSE).

(c) [5pts] Suppose the gradients are

$$\frac{\partial L}{\partial w} = -\frac{28}{3}, \frac{\partial L}{\partial b} = -4$$

Using a learning rate of $\eta = 0.1$, compute the updated values of w and b after one gradient descent step.

(d) [5pts] What may happen if the learning rate is too large? Briefly explain.

Problem 2. Logistic Regression

Consider a binary classification model with the following linear score:

$$z = 2x_1 - x_2 - 1.$$

For a sample with

$$x_1 = 1, \qquad x_2 = 2, \qquad y = 1,$$

answer the following questions.

Problem 2. Logistic Regression

(a) [4pts] Compute the linear score z.

(b) [5pts] Compute the probability of Class 1 using the sigmoid function.

(c) [3pts] Using a threshold of 0.5, determine the predicted class.

(d) [5pts] Compute the Binary Cross-Entropy (BCE) loss.

(e) [3pts] For another sample with $y = 1$, Model A predicts $p = 0.9$, while Model B predicts $p = 0.6$. Which model has the smaller BCE loss? Briefly explain.

Problem 3. Multiclass Classification

A three-class classifier produces the following softmax probabilities:

$$\mathbf{p} = [0.2, 0.7, 0.1].$$

The classes are ordered as

$$[\text{Class } 1, \text{Class } 2, \text{Class } 3].$$

Problem 3. Multiclass Classification

(a) [5pts] What class will the model predict?

(b) [5pts] Suppose the true class is Class 2. Compute the cross-entropy loss.

(c) [5pts] Suppose another model produces

$$p = [0.1, 0.85, 0.05]$$

for the same sample. Which model has the smaller cross-entropy loss? Briefly explain.

Problem 4. Model Evaluation and Performance Metrics

Consider the following confusion matrix for a binary classifier:

	Predicted Positive	Predicted Negative
Actual Positive	36	4
Actual Negative	12	48

Problem 4. Model Evaluation and Performance Metrics

(a) [4pts] Identify TP, TN, FP, and FN.

(b) [4pts] Compute the accuracy.

(c) [4pts] Compute the precision.

(d) [4pts] Compute the recall.

(e) [4pts] Compute the F1 score.

(f) [5pts] Suppose this classifier is used for disease screening, where missing a patient with the disease is particularly costly. Between precision and recall, which metric should receive more attention? Briefly explain.

Problem 5. Generalization and Regularization

Three models produce the following training and validation errors:

Model	Training Error	Validation Error
A	22%	24%
B	6%	8%
C	1%	17%

Problem 5. Generalization and Regularization

(a) [4pts] Which model most likely suffers from underfitting? Explain your answer.

(b) [4pts] Which model most likely suffers from overfitting? Explain your answer.

(c) [3pts] Which model shows the best generalization among the three? Explain your answer.

(d) [3pts] As model complexity increases, what generally happens to bias and variance?

Consider the regularized objective

$$L_{total} = L_{data} + \lambda R(\theta).$$

(e) [2pts] What may happen when λ is too small?

(f) [2pts] What may happen when λ is too large?

(g) [2pts] Briefly describe the difference between L1 and L2 regularization in terms of their effects on model weights.

Korea Aerospace University

Good Luck!

jiyoonshin@kau.ac.kr