Final Exam
Machine Learning
2024-2025

Duration: 90 minutes
Dr. Abbas Rammal

Problem 1: Evaluation of classifiers

1. Given a data set $D = \{o_1, \dots, o_n\}$ with known class labels $C(o_i) \in C = \{A, B, C\}$ of the objects. In order to evaluate the quality of a classifier K, each object $o_i \in D$ is additionally classified using K, yielding class label $K(o_i)$. The results are given in the table below.

i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$C(o_i)$	A	B	A	C	C	B	A	A	A	B	B	C	C	C	B
$K(o_i)$	A	B	B	C	B	B	A	A	A	C	A	A	C	A	B

   a) Setup the confusion matrix.
   b) Compute the accuracy / classification error.
   c) For each class $i \in C$ compute precision and recall.
   d) Compute the F1-measure for all classes.
   e) Explain how K-Fold Cross-Validation is implemented.

2. To construct a Receiver Operating Characteristic (ROC) curve, we need to calculate the True Positive Rate (TPR) and False Positive Rate (FPR) at different cutoff points. The cutoff point represents the threshold used to classify samples as positive or negative. Construct the ROC curve on cutoff point = 0.3 and 0.7.

Instance	Actual	Probability Yes	Probability No
S1	Yes	0.80	0.20
S2	No	0.25	0.75
S3	Yes	0.40	0.60
S4	No	0.65	0.35
S5	Yes	0.70	0.30
S6	No	0.50	0.50

Problem 2: Clustering

Consider the following 6 data points in a 2-dimensional feature space:

$$x_1 = (0,0);\ x_2 = (0,1);\ x_3 = (-1,2);\ x_4 = (2,0);\ x_5 = (3,0);\ x_6 = (4,-1)$$

Activation Function:

Use the sigmoid function for all hidden and output neurons:

$$\sigma(x) = \frac{1}{1+e^{-x}}, \quad \sigma'(x) = \sigma(x)(1-\sigma(x))$$

Learning Rate:

$\eta = 0.05$

Training Sample (T1):

- Input: [0.8 (humidity), 0.5 (temperature)]
- Target output: [1] (It will rain)

Tasks:

1. **Forward Propagation**
   - Compute the output of the hidden and output layers using the given input.

2. **Loss Calculation**
   - Use the squared error loss:

$$E = \frac{1}{2}(t-o)^2$$

3. **Backpropagation**
   - Compute the error at the output neuron.
   - Propagate the error to the hidden layer.
   - Update all the weights accordingly.
4. **Weight Update**

Show the updated weights after applying one round of training with this sample.

Let's calculate the squared distances between each pair of points using the Euclidean distance formula:

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	0	1	5	4	9	17
x_2	1	0	2	5	10	20
x_3	5	2	0	12	20	34
x_4	4	5	12	0	1	5
x_5	9	10	20	1	0	2
x_6	17	20	34	5	2	0

a) Perform K-means clustering on this dataset. Use the first and last data points as initial centers (K = 2). Given the final parameters, which cluster would $x^* = (1, 1)$ belong to?

b) Perform agglomerative Hierarchical Clustering using single linkage as the cluster distance measure. Draw the associated tree.

c) Perform agglomerative Hierarchical Clustering using complete linkage as the cluster distance measure. Draw the associated tree.

d) Use the Nearest Neighbor clustering algorithm and Euclidean distance to cluster this dataset. Suppose that the threshold t is 4.

Exercise 3: Neural Network for Weather Prediction

You are designing a neural network to predict whether it will rain or not rain tomorrow based on two input features:

- Humidity level
- Temperature

Network Architecture:

- Input layer: 2 neurons (Humidity, Temperature)
- Hidden layer: 2 neurons
- Output layer: 1 neuron (Rain: 1, No Rain: 0)

Initial Weights:

- Input → Hidden:
  - w(1→3) = 0.2, w(1→4) = -0.3
  - w(2→3) = 0.4, w(2→4) = 0.1
- Hidden → Output:
  - w(3→5) = -0.2, w(4→5) = 0.3