Applied Machine Learning RS

| | := | |
|---------------------|-------------|-----------|
| Take Home Quiz No 3 | SPRING 2025 | Marks 100 |
| | | |

| NAME: | ROLL NO. |
|-------|----------|
| | |

Q1: You are given transaction data from a tech accessory store. Each row lists the items purchased in a single transaction. Build the **FP-Growth Tree** for Market Basket Analysis and generate association rules. (30 marks)

| TID | Items Bought |
|-----|-------------------------------------|
| T1 | {Laptop, Mouse, Charger, USB, HDMI} |
| T2 | {Laptop, Mouse, Bag, USB, SSD} |
| T3 | {Charger, Mouse, Bag, USB, SSD} |
| T4 | {Mouse, USB, SSD} |
| T5 | {Laptop, Mouse, Charger, HDMI, Bag} |
| T6 | {Laptop, USB, SSD} |
| T7 | {Laptop, Mouse, Charger, SSD} |
| T8 | {Mouse, USB, SSD, HDMI} |

Q2: Short questions: (30 marks)

- Suppose you want to predict tomorrow's temperature in degrees Centigrade, given a lot of historical data. Which classification algorithm would work best for this problem?
- 2. Suppose you want to predict whether or not a certain company will declare bankruptcy within next 7 days. You have training data of similar companies that had previously been at risk of bankruptcy. Would you treat it as a classification or regression problem?
- **3.** Suppose you are given the following confusion matrix. Calculate Precision, Recall, and F1-score.

| | | Actual Values | | | | | | |
|-----------|-----------|---------------|-----------|--|--|--|--|--|
| | | Cancer | No Cancer | | | | | |
| Predicted | Cancer | 45 | 18 | | | | | |
| Values | No Cancer | 12 | 25 | | | | | |

4. Suppose you train a classifier on train split and get the predictions on 20 examples (documents) of test split. Class labels are {0: Sports, 1: Religious, 2: Medical}. You get the following predictions.

| 0 | | 0 | 0 | 0 | 0 | 0 | | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 1 |
|---|--|---|---|---|---|---|--|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
|---|--|---|---|---|---|---|--|---|---|---|---|---|---|---|---|---|---|---|---|---|---|

The following are the actual labels for the test split.

| | | | | | | | | | | | | 1 | | | | | |
|--------|--------|--------|--------|--------|--------|--------|--------|---|---|---|---|---|---|-------|---|---|-------|
| \sim | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| () | | | | | | | | | | | | | | | | | |
| 0 | 0 | 0 | _ | _ | _ | _ | 0 | | | _ | _ | | | _ | _ | _ | |

a) Create a confusion matrix for all three classes above.

Note: You can compute Precision and Recall for each class directly from the confusion matrix and then use them for calculating F1 Score and also Macro Averages.

b) From the confusion matrix above, calculate Macro Precision, Macro Recall, and Macro F1 score.

Note: You can also make a separate confusion matrix for each class and compute Precision and Recall and then use them for calculating F1 Score and also Macro Averages.

Q4: Suppose you are trying to train a Perceptron on single example given below. (20 marks)

| x_1 | x_2 | x_3 | x_4 | x_5 | y |
|-------|-------|-------|-------|-------|---|
| 1 | 0 | 1 | 0 | 1 | 0 |

You initialize the weights randomly. The weights vector, threshold, and learning rate is as follows:

Weights =
$$[0.7, 0.6, 0.5, 0.3, 0.4]$$

 $Threshold = 1.5$
 $\eta = 0.2$

- a) Compute the predicted label.
- b) Update the weights using Perceptron Training Rule $w_i = w_i + \eta(y \hat{y})x_i$
- c) Compute the predicted label using the updated weights.

Q4. Select any famous game and implement Reinforcement Learning Algorithm in Python. Give the complete dry-run of the solution along with the running code. **(20 marks)**