Practice Problems Week-05: Classification Metrics

Instructions: Solve the following problems step-by-step. Show all your work.

1. **Basic Confusion Matrix:** Given: TP=40, FN=10, FP=15, TN=35. Create the confusion matrix and calculate Accuracy.

2. **Precision and Recall:** Using the values from question 1, calculate Precision and Recall.

3. **F1-Score:** Calculate the F1-score using the Precision and Recall from question 2.

4. **Specificity: Specificity** (True Negative Rate) measures the proportion of actual negatives that are correctly identified. Calculate the Specificity using the values from question 1.

5. **Finding FN and FP:** A classifier has Accuracy=0.85, Precision=0.75, Recall=0.80, and there are 100 positive instances. Find the number of False Negatives (FN) and False Positives (FP).

6. **Cost of Errors:** In a medical test, a False Negative (missing a disease) is 5 times more costly than a False Positive. If there are 10 FNs and 20 FPs, what is the total cost, assigning a cost of 1 unit to a False Positive?

7. **Prevalence:** Prevalence refers to the proportion of a population that has a particular condition or characteristic. In classification, it represents the frequency of the positive class in the dataset. Prevalence is important because it affects how we interpret metrics like Precision and helps understand the base rate of the positive class. In a dataset of 1000 samples, the positive class has 150 instances. What is the prevalence of the positive class?

8. **Effect of Threshold:** A model predicts the probability of the positive class. The current threshold is 0.5. If we increase the threshold to 0.7, what will happen to Precision and Recall in general?

9. **Balanced Accuracy:** Balanced Accuracy is the arithmetic mean of Recall (Sensitivity) and Specificity. It's particularly useful for imbalanced datasets because it gives equal weight to both classes, unlike standard accuracy which can be misleading when classes are unbalanced. For a model with Recall=0.9 and Specificity=0.7, calculate the Balanced Accuracy and explain why it might be more informative than standard accuracy for an imbalanced dataset.

10. **ROC Point:** A classifier at a certain threshold has a False Positive Rate (FPR) of 0.1 and a True Positive Rate (TPR) of 0.6. Plot this point on a conceptual ROC graph. Is this a good classifier?

11. **AUC Interpretation:** What does an AUC score of 0.5 imply about the model's performance?

12. **Manual ROC Point Calculation:** From the confusion matrix in Q1, calculate the TPR and FPR.

13. **Probability Ranking:** For 5 instances, their true labels and predicted probabilities for the positive class are: (1, 0.9), (0, 0.8), (1, 0.6), (0, 0.4), (1, 0.3). Sort the instances by predicted probability and calculate the TPR and FPR at each threshold to create a small ROC curve.

14. **Weighted Average Precision:** A model is tested on two classes, A and B. For Class A (80 instances), Precision=0.9. For Class B (20 instances), Precision=0.6. Calculate the Weighted Average Precision.

15. **Micro vs. Macro Recall: Macro-average** recall computes recall independently for each class and then takes the average. **Micro-average** recall aggregates the contributions of all classes to calculate the average recall. For a 3-class problem, the recalls for the classes are: Class 1: 0.8, Class 2: 0.5, Class 3: 0.9. Calculate both the Macro-average Recall and the Micro-average Recall (assume equal number of instances per class for simplicity).

16. **Threshold for Specificity:** We want a classifier with at least 95% Specificity. If we have 100 true negative samples, what is the maximum number of false positives we can allow?

17. **Cost-Sensitive Threshold:** The cost of a False Negative is 10 and the cost of a False Positive is 1. The model's probability outputs are calibrated. Should the decision threshold be higher or lower than 0.5 to minimize cost? Explain.

18. **PR Curve vs ROC Curve:** In a scenario with high class imbalance, why might the Precision-Recall (PR) curve be more informative than the ROC curve?

19. **Geometric Mean:** The Geometric Mean (G-Mean) is defined as $\sqrt{\text{Recall} \times \text{Specificity}}$ and is particularly useful for imbalanced datasets. While the F1-score focuses on balancing Precision and Recall (both related to the positive class), the G-Mean balances Recall (True Positive Rate) and Specificity (True Negative Rate), ensuring good performance on both classes. Calculate the G-Mean for the confusion matrix in question 1 and explain how it differs from the F1-score calculated earlier.

20. **Precision-Recall Tradeoff:** A model currently has Precision=0.8 and Recall=0.6. If we want to increase Recall to 0.8, what is likely to happen to Precision, and why?