CS550/DSL501: Machine Learning (2023–24–M)
Assignment-2

Full name: Munjam Navadeep **ID: 12241060**

1 which performance parameter you should prefer for model performance?

Given the confusion matrix:

- True Positives (TP) = 80
- False Negatives (FN) = 20
- False Positives (FP) = 100
- True Negatives (TN) = 9,800

We want to determine which performance metric to prefer for model evaluation.

1.1 1. Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{80 + 9800}{80 + 9800 + 100 + 20} = \frac{9880}{10000} = 0.988 = 98.8\%$$

1.2 2. Precision

$$Precision = \frac{TP}{TP + FP} = \frac{80}{80 + 100} = \frac{80}{180} \approx 0.4444 = 44.44\%$$

1.3 3. Recall (Sensitivity)

$$Recall = \frac{TP}{TP + FN} = \frac{80}{80 + 20} = \frac{80}{100} = 0.8 = 80\%$$

1.4 4. Specificity

$$Specificity = \frac{TN}{TN + FP} = \frac{9800}{9800 + 100} = \frac{9800}{9900} \approx 0.9899 = 98.99\%$$

1.5 5. F1 Score

The F1 Score is the harmonic mean of Precision and Recall:

$$F1Score = \frac{2 \times Precision \times Recall}{Precision + Recall} = \frac{2 \times 0.4444 \times 0.8}{0.4444 + 0.8} \approx \frac{0.7111}{1.2444} \approx 0.5715 = 57.15\%$$

1.6 Analysis

- **Accuracy** is high (98.8%), but in the context of imbalanced datasets (where one class is much more prevalent than the other), accuracy can be misleading. A high accuracy might simply reflect the model's ability to predict the majority class well.
- **Precision** is relatively low (44.44%), indicating that when the model predicts a patient has the disease, it's only correct 44.44% of the time. This is important in medical diagnosis, where false positives can lead to unnecessary stress and further medical tests.
- **Recall** is decent (80%), meaning the model correctly identifies 80% of the actual positive cases. However, the remaining 20% of cases are missed (false negatives), which can be critical in medical applications.
- **Specificity** is very high (98.99%), meaning the model is very good at identifying patients who do not have the disease.
- **F1 Score** balances precision and recall and is moderate at 57.15%.

1.7 Conclusion

In a medical setting where the disease is rare, the **Recall (Sensitivity)** is typically prioritized because missing a true positive (a patient who actually has the disease) can be more harmful than a false positive. Therefore, **Recall (80%)** should be the preferred performance parameter in this context. However, improving **Precision (44.44%)** should also be considered to reduce the number of false positives.