DATA 300: Statistical Machine Learning

Classification Metrics and the Confusion Matrix

Binary Case.

- **True Positives (TP).**
  It refers to the number of predictions where the classifier correctly predicts the positive class as positive.
- **True Negative (TN).**
  It refers to the number of predictions where the classifier correctly predicts the negative class as negative.
- **False Positive (FP).**
  It refers to the number of predictions where the classifier incorrectly predicts the negative class as positive.
- **False Negative (FN).**
  It refers to the number of predictions where the classifier incorrectly predicts the positive class as negative.

Metrics.

Accuracy: Gives you the fraction of the total examples that were correctly classified by the classifier. It is given by: $\frac{(TP+TN)}{TP+TN+FP+FN}$

Misclassification Rate: Tells you what fraction of the predictions from your model were incorrect. $\frac{(FP+FN)}{TP+TN+FP+FN}$. Or 1-Accuracy.

Precision: It tells you what fraction of predictions as a positive class were actually positive.

$$\frac{(TP)}{TP+FP}$$

Recall: It tells you what fraction of all positive samples were correctly predicted as positive by the classifier. It is also called True Positive Rate, Sensitivity and Probability of Detection:

$$\frac{(TP)}{TP+FN}$$

Specificity: It tells you what fraction of all negative samples are correctly predicted as negative by the classifier. It is also known as True Negative Rate (TNR)

$$\frac{(TN)}{TN+FP}$$

F1 Score: This is the harmonic mean of Precision and recall: $\mathbf{2}\frac{\boldsymbol{PrecsionxRecall}}{\boldsymbol{Precision+Recall}} = \frac{\boldsymbol{2TP}}{\boldsymbol{2TP+FP+FN}}$