Confusion Matrix & Performance Measurement Metrics

Dataset Modeling and result evaluations

x	y	z	
a	5	yes	
c	6	no	
e	13	no	
a	2	no	
d	29	yes	
d	11	no	
a	9	yes	
b	99	yes	
c	10	yes	No
a	1	no	No
d	13	yes	Yes
e	71	no	Yes

Confusion Matrix

- A confusion matrix for a two classes (+, -) is shown below.

	C1	C2
C1	True positive	False negative
C2	False positive	True negative

	+	-
+	++	+-
-	-+	--

- There are four quadrants in the confusion matrix, which are symbolized as below.
  - True Positive (TP: f_{++}) : The number of instances that were positive (+) and correctly classified as positive (+).
  - False Negative (FN: f_{+-}): The number of instances that were positive (+) and incorrectly classified as negative (-).
  - False Positive (FP: f_{-+}): The number of instances that were negative (-) and incorrectly classified as (+).
  - True Negative (TN: f--): The number of instances that were negative (-) and correctly classified as (-).

  - **A confusion matrix is a table used to evaluate the performance of a classification model.**

Confusion Matrix

Note:

- N_p = TP (f_{++}) + FN (f_{+-})

  = is the total number of positive instances.

- **Np**: This represents the total number of positive instances. It's calculated by adding the True Positives (TP), which are instances correctly classified as positive, and False Negatives (FN), which are instances incorrectly classified as negative but are actually positive.
- For example, if we have 20 positive instances and the classifier correctly identifies 15 of them as positive (TP = 15), but incorrectly identifies 3 positive instances as negative (FN = 3), then Np = TP + FN = 15 + 3 = 18.

- N_n = FP(f_{-+}) + Tn(f_{--})

  = is the total number of negative instances.

- **Nn**: This represents the total number of negative instances. It's calculated by adding the False Positives (FP), which are instances incorrectly classified as positive but are actually negative, and True Negatives (TN), which are instances correctly classified as negative.
- For example, if we have 30 negative instances and the classifier correctly identifies 25 of them as negative (TN = 25), but incorrectly identifies 2 negative instances as positive (FP = 2), then Nn = FP + TN = 2 + 25 = 27.

Note:

- $N = N_p + N_n$

  = is the total number of instances.

- **N**: This is the total number of instances, calculated by adding Np (total positive instances) and Nn (total negative instances).
- For example, if we have Np = 18 (total positive instances) and Nn = 27 (total negative instances), then N = Np + Nn = 18 + 27 = 45.

- (TP + TN) denotes the number of correct classification
- **(TP + TN)**: This represents the number of correct classifications, which is the sum of True Positives (TP) and True Negatives (TN). It indicates how many instances were correctly classified by the model.

(FP + FN) denotes the number of errors in classification.

(FP + FN): This represents the number of errors in classification, which is the sum of False Positives (FP) and False Negatives (FN). It indicates how many instances were incorrectly classified by the model.

For a perfect classifier, FP = FN = 0

Perfect Classifier: In an ideal scenario, a perfect classifier would have both False Positives (FP) and False Negatives (FN) equal to 0, meaning it makes no errors in classification.

Confusion Matrix Example

- For example,

Class	+	-
+	52 (TP)	18 (FN)
-	21 (FP)	123 (TN)

Calculate the performance evaluation metrics

Accuracy

- It is defined as the fraction of the number of examples that are correctly classified by the classifier to the total number of instances.

$$\text{Accuracy}= \frac{TP+TN}{TP+FP+FN+TN}$$

- This formula means that you divide the total number of correctly classified instances (sum of TP and TN) by the total number of instances (sum of TP, FP, FN, and TN).

- For example, if a classifier correctly identifies 80 out of 100 instances (TP = 60, TN = 20), then the accuracy would be:
- Accuracy= (60+20) / (60 + 10 + 10 + 20) = 80/100 = 0.8
- So, the accuracy of the classifier in this case would be 80%.

Performance Evaluation Metrics

- We now define a number of metrics for the measurement of a classifier.
  - In our discussion, we shall make the assumptions that there are only two classes: + (positive) and – (negative)

- **True Positive Rate (TPR):** It is defined as the fraction of the positive examples predicted correctly by the classifier.

$$TPR = \frac{TP}{P} = \frac{TP}{TP+FN} = \frac{f_{++}}{f_{++}+f_{+-}}$$

  - This metrics is also known as ***Recall***, ***Sensitivity*** or ***Hit rate***.

- **False Positive Rate (FPR):** It is defined as the fraction of negative examples classified as positive class by the classifier.

$$FPR = \frac{FP}{N} = \frac{FP}{FP+TN} = \frac{f_{-+}}{f_{-+}+f_{--}}$$

Performance Evaluation Metrics

- **False Negative Rate (FNR):** It is defined as the fraction of positive examples classified as a negative class by the classifier.

$$FNR = \frac{FN}{P} = \frac{FN}{TP + FN} = \frac{f_{+-}}{f_{++} + f_{+-}}$$

- **True Negative Rate (TNR):** It is defined as the fraction of negative examples classified correctly by the classifier

$$TNR = \frac{TN}{N} = \frac{TN}{TN + FP} = \frac{f_{--}}{f_{--} + f_{-+}}$$

  - This metric is also known as ***Specificity***.

Performance Evaluation Metrics

- Both, Precision and Recall are defined by :
- **Precision: Precision, also known as Positive Predictive Value, measures the accuracy of the positive predictions made by the classifier. It is defined as the fraction of true positive predictions (instances correctly predicted as positive) out of all positive predictions made by the classifier.**

- **Recall: Recall, also known as Sensitivity or True Positive Rate, measures the ability of the classifier to find all the positive instances in the dataset. It is defined as the fraction of true positive predictions (instances correctly predicted as positive) out of all actual positive instances in the dataset.**

- **Recall focuses on the proportion of actual positive instances that are correctly identified by the classifier.**

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

Performance Evaluation Metrics

- **F_1 Score (F_1):** Recall (r) and Precision (p) are two widely used metrics employed in analysis..
  - It is defined in terms of (r or Recall) and (p or Precision) as follows.

$$F_1 Score = \frac{2 * Recall \,.\, Precision}{Recall + Precision} = \frac{2TP}{2TP + FP + FN}$$

Note

- F_1 represents the harmonic mean between recall and precision

- High value of F_1 score ensures that both Precision and Recall are reasonably high.

The F1 Score, also known as the F1 Measure or F1 Score, is a metric that combines both (parallel) precision and recall into a single value. It is defined as the harmonic mean of precision and recall.

Mathematically, the F1 Score is represented as:

$$F_1 Score = \frac{2 * Recall \,.\, Precision}{Recall + Precision}$$

Alternatively, it can also be expressed in terms of True Positives (TP), False Positives (FP), and False Negatives (FN):

$$F_1 Score = \frac{2TP}{2TP + FP + FN}$$

The key points to note about the F1 Score are:

1.Harmonic Mean: The F1 Score calculates the harmonic mean of precision and recall rather than the arithmetic mean. This is important because it gives more weight to lower values. As a result, the F1 Score penalizes extreme values (either high or low) of precision and recall, making it a more balanced metric.

2.Balanced Performance: A high value of the F1 Score indicates that both precision and recall are reasonably high. It ensures a good balance between minimizing false positives (FP) and false negatives (FN).

In summary, the F1 Score provides a single value that summarizes the balance between precision and recall, making it a useful metric for evaluating the overall performance of a classifier.

Analysis with Performance Measurement Metrics

- Based on the various performance metrics, we can characterize a classifier.
- We do it in terms of TPR, FPR, Precision and Recall and Accuracy
- **Case 1: Perfect Classifier**

When every instance is correctly classified, it is called the perfect classifier. In this case, $TP = P$, $TN = N$ and CM is

$$TPR\ ,\ Recall = \mathrm{TP/(TP+FN)} = \frac{P}{P} = 1\ ,\ \mathrm{FN} = 0$$

$$FPR = \frac{0}{N} = 0\ ,\ \mathrm{FPR = FP\ /\ (FP + TN) = 0\ /\ (0 + N) = 0}$$

$$Precision = \frac{P}{P} = 1$$

$$F_1\ Score = \frac{2\times1}{1+1} = 1$$

$$Accuracy = \frac{P+N}{P+N} = 1$$

		Predicted Class	
		+	-
Actual class	+	P	0
	-	0	N

Analysis with Performance Measurement Metrics

- **Case 2: Worst Classifier**

When every instance is wrongly classified, it is called the worst classifier. In this case, $TP = 0$, $TN = 0$ and the CM is

$$TPR = \frac{0}{P} = 0$$

$$FPR = \frac{N}{N} = 1$$

$$Precision = \frac{0}{N} = 0$$

F_1 $Score$ = Not applicable as $Recall + Precision = 0$

$$\text{Accuracy} = \frac{0}{P+N} = 0$$

		Predicted Class	
		+	-
Actual class	+	0	P
	-	N	0

Analysis with Performance Measurement Metrics

- **Case 3: Ultra-Liberal Classifier**

  The classifier always predicts the + class correctly. Here, the False Negative (FN) and True Negative (TN) are zero. The CM is

$$TPR = \frac{P}{P} = 1$$

$$FPR = \frac{N}{N} = 1$$

$$Precision = \frac{P}{P+N}$$

$$F_1\ Score = \frac{2P}{2P+N}$$

$$\text{Accuracy} = \frac{P}{P+N} = 0$$

		Predicted Class	
		+	**-**
Actual class	**+**	**P**	**0**
	-	**N**	**0**

Analysis with Performance Measurement Metrics

- **Case 4: Ultra-Conservative Classifier**

This classifier always predicts the - class correctly. Here, the False Negative (FN) and True Negative (TN) are zero. The CM is

$$TPR = \frac{0}{P} = 0$$

$$FPR = \frac{0}{N} = 0$$

$$Precision = \text{Not applicable}$$

$$(\text{as } TP + FP = 0)$$

$$F_1\ Score = \text{Not applicable}$$

$$\text{Accuracy} = \frac{N}{P+N} = 0$$

		Predicted Class	
		+	-
Actual class	+	0	p
	-	0	N