21. Data Science – Machine Learning – Confusion Matrix

Contents

1. Introduction 2

2. Confusion Matrix 2

3. Name of the each combination 3

4. Table1 4

5. Scenario 5

5.1. True positive 6

5.2. True Negative 6

5.3. False Positive 7

5.4. False Negative 7

6. Type I and Type II errors 8

7. Confusion Matrix 9

8. Table2 9

9. Confusion matrix example 10

10. Performance Metrics 11

11. Accuracy 12

12. Accuracy can be misleading 13

13. Recall or Sensitivity or True Positive Rate 14

14. Specificity or True Negative Rate 15

15. Precision 16

16. F1 Score 18

21. Data Science – Machine Learning – Confusion Matrix

1. Introduction

- ✓ Confusion matrix is a method to explain the results of the classification model.

2. Confusion Matrix

- ✓ In binary classification, the outcomes are,
    - True
    - False

- ✓ To make it more clear let us consider an example of a binary classifier that scans the MRI images and predicts whether a person has cancer or not.
- ✓ The outcomes predicted by classifier and the actual outcomes can only have the following four combinations

Predicted	Actual
Cancer = Yes	Cancer = Yes
Cancer = Yes	Cancer = No
Cancer = No	Cancer = Yes
Cancer = No	Cancer = No

3. Name of the each combination

- ✓ Now let's understand clearly by comparing each other

Name	Predicted	Actual
True Positive	Cancer = Yes	Cancer = Yes
False Positive	Cancer = Yes	Cancer = No
False Negative	Cancer = No	Cancer = Yes
True Negative	Cancer = No	Cancer = No

4. Table1

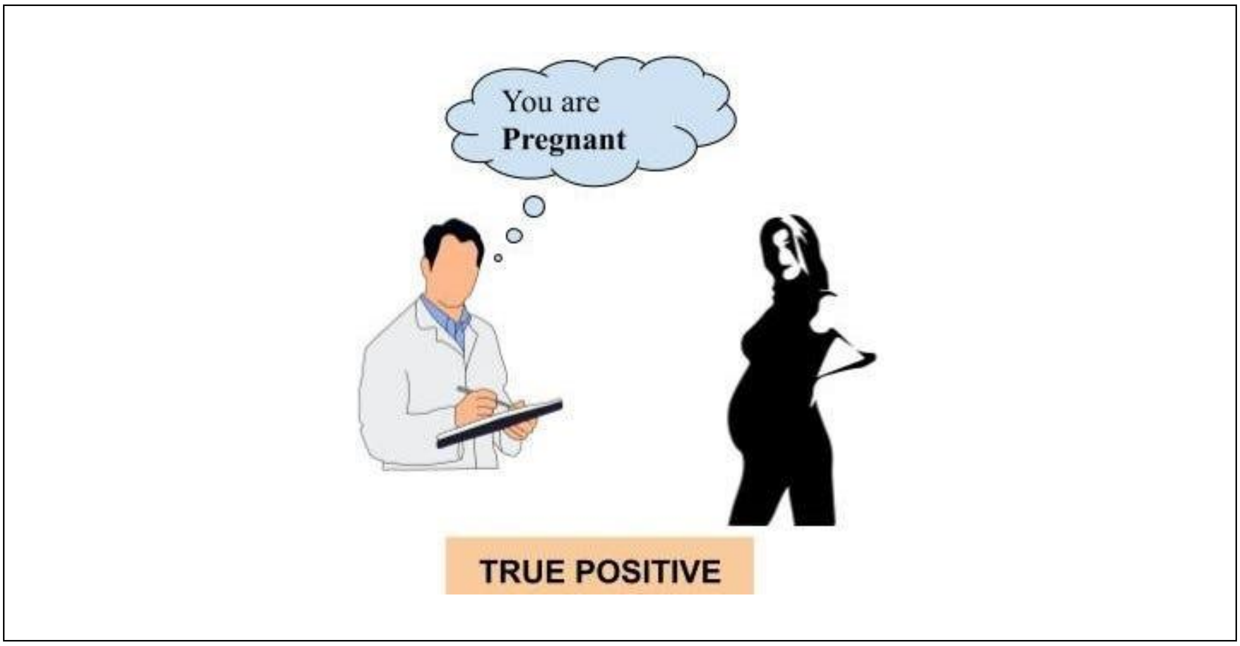
Name	Description
1. True Positive	✓ The model predicts that the patient is having cancer (positive) and indeed he has cancer (true prediction).
2. False Positive	✓ The model predicts that the patient is having cancer (positive) but actually he is not having cancer (false prediction)
3. False Negative	✓ The model predicts that the patient is not having cancer (negative) but he actually has cancer (false prediction).
4. True Negative	✓ The model predicts that the patient is not having cancer (negative) and indeed he does not have cancer (true prediction).

5. Scenario

- ✓ Now consider the above classification (pregnant or not pregnant) carried out by a machine learning algorithm. The output of the machine learning algorithm can be mapped to one of the following categories.

5.1. True positive

- ✓ A person who is actually pregnant (positive) and classified as pregnant (positive). This is called TRUE POSITIVE (TP).



5.2. True Negative

- ✓ A person who is actually not pregnant (negative) and classified as not pregnant (negative). This is called TRUE NEGATIVE (TN).

5.3. False Positive

✓ A person who is actually not pregnant (negative) and classified as pregnant (positive). This is called FALSE POSITIVE (FP).

5.4. False Negative

✓ A person who is actually pregnant (positive) and classified as not pregnant (negative). This is called FALSE NEGATIVE (FN).

6. Type I and Type II errors

Type I error

- ✓ Type I error also called as False Positive.
- ✓ Type I error occurs if you reject the null hypothesis when it was true.

Type II error

- ✓ Type II error also called as False Negative
- ✓ Type II error occurs if you accept the null hypothesis when it was false.

7. Confusion Matrix

- ✓ We can represent above discussion by using matrix also

	Actual Cancer = Yes	Actual Cancer = No
Predicted Cancer = Yes	True Positive (TP)	False Positive (FP)
Predicted Cancer = No	False Negative (FN)	True Negative (TN)

8. Table2

- ✓ True Positive = No. of True Positives from total predictions
- ✓ False Positive = No. of False Positives from total predictions
- ✓ False Negative = No. of False Negatives from total predictions
- ✓ True Negative = No. of True Negatives from total predictions

- ✓ Total number of prediction = TP + FP + FN + TN

9. Confusion matrix example

✓ Let's try to represent this confusion matrix with real numbers.

	Actual Cancer = Yes	Actual Cancer = No
Predicted Cancer = Yes	True Positive 57	False Positive 14
Predicted Cancer = No	False Negative 23	True Negative 171

- 57 = No. of True Positives from total predictions
- 14 = No. of False Positives from total predictions
- 23 = No. of False Negatives from total predictions
- 171 = No. of True Negatives from total predictions

- 57+14+23+171 = 265 = Total no. of predictions

10. Performance Metrics

- ✓ If we have confusion matrix then we can use it to calculate various performance metrics to measure your classification models.

11. Accuracy

- ✓ Accuracy is a measure of the fraction of times the model predicted correctly (both true positive and true negative) out of total no. of predictions.

	Actual Cancer = Yes	Actual Cancer = No
Predicted Cancer = Yes	True Positive 57	False Positive 14
Predicted Cancer = No	False Negative 23	True Negative 171

$$\text{Accuracy} = \frac{TP+TN}{Total Predictions}$$

- ✓ This is the simple and very intuitive metrics.
- ✓ Let us consider the accuracy of the classifier in our example above.

$$\text{Accuracy} = \frac{57+171}{265} = 0.86$$

- ✓ So this means our model can classify cancer and non-cancer cases with 86% accuracy.

12. Accuracy can be misleading

- ✓ We felt like accuracy looks good to measure the performance of the model, but there is a catch!!
- ✓ Assume that your machine model trains badly and only identifies negative scenarios and miss to identify true positive completely. (This happens when there is a class imbalance in training data.)

	Actual Cancer = Yes	Actual Cancer = No
Predicted Cancer = Yes	True Positive 0	False Positive 0
Predicted Cancer = No	False Negative 80	True Negative 185

- ✓ So in our example, say the classifier completely misses to predict any case of cancer and marks all cases as non-cancer, this is how the confusion matrix will look like.
- ✓ So the accuracy of our classifier will be

$$\text{Accuracy} = \frac{185}{265} = 0.70$$

- ✓ So here model could not predict a single cancer case but we got accuracy is 70% which is not so good☺.
- ✓ So, only believing Accuracy is not good way
- ✓ This is why we have several other metrics to measure the performances of the machine learning model.

13. Recall or Sensitivity or True Positive Rate

- ✓ Recall is the fraction of times the model predicts positive cases correctly from the total number of actual positive cases.

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

	Actual Cancer = Yes	Actual Cancer = No
Predicted Cancer = Yes	True Positive 57	False Positive 14
Predicted Cancer = No	False Negative 23	True Negative 171

- ✓ Let's understand, TP + FN is the total number of actual positive cases that the classifier should have identified.
- ✓ In our example, TP = 57 cancer cases classifier predicted correctly.
- ✓ FN = 23 cancer cases our classifier missed.
- ✓ S0, the total cancer cases were TP + FN = 57+23 = 80.

$$\text{Recall} = \frac{57}{57+23} = \frac{57}{80} = 0.71$$

- ✓ So this means our model has 71% recall capability to predict cancer cases from total actual cancer cases.

14. Specificity or True Negative Rate

- ✓ Specificity is the fraction of times the classifier predicts negatives cases correctly from the total number of actual negative cases.
- ✓ This metrics is the opposite of Recall/Sensitivity that we understood above.

	Actual Cancer = Yes	Actual Cancer = No
Predicted Cancer = Yes	True Positive 57	False Positive 14
Predicted Cancer = No	False Negative 23	True Negative 171

$$\text{Specificity} = \frac{TN}{TN+FP}$$

- ✓ Here TN+FP is the total number of actual negative cases.
- ✓ In our example, TN = 171 non-cancer cases were predicted correctly by the classifier and FP = 14 non-cancer cases were incorrectly identified as cancer.
- ✓ So, the total non-cancer cases were TN+FP = 171+14 = 185.

$$\text{Specificity} = \frac{171}{171+14} = \frac{171}{185} = 0.92$$

- ✓ So this means our model has 92% capability to identify negative cases from total number of actual negative cases.

15. Precision

- Precision calculates the fraction of times the model predicts positive cases correctly from the total number of positive cases it predicted.

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

- Let's try to understand TP+FP is the total number of cases predicted by the classifier as positive.
- TP is the total number of cases that are actually positive.

	Actual Cancer = Yes	Actual Cancer = No
Predicted Cancer = Yes	True Positive 57	False Positive 14
Predicted Cancer = No	False Negative 23	True Negative 171

- In our example, the classifier predicted TP+FP = 71 cases of cancer.
- Out of these 71 cases, only TP = 57 cases were correct.

$$\text{Precision} = \frac{57}{57+14} = \frac{57}{71} = 0.80$$

- ✓ So this means our classifier has a precision or correctness of 80% when it predicts cases as cancer.

16. F1 Score

- ✓ So, Recall and precision are the best way to choose.
- ✓ Instead of using many metrics, Can't we chose one metric?
- ✓ Yes we can, here F1 score helps actually
- ✓ The F1 score is a measure of a model's accuracy.

$$\text{F1 Score} = \frac{2 * Precision * Recall}{Precision + Recall}$$

- ✓ Using F1 scores we can now evaluate and compare the performance of the models easily now.