Artificial Intelligence
&
Neural Networks

CSE-351

Lecture-10

Rajorshi Bhattacharya
Lecturer, Dept. of CSE
Bangladesh University of Business & Technology

Confusion Matrix

A confusion matrix is a table that is often used to describe the performance of a classification model (or "classifier") on a set of test data for which the true values are known.

- It allows the visualization of the performance of an algorithm.
- We compare each class with every other class and see how many samples are misclassified.
- We actually come across several key metrics that are very important in the field of machine learning.

Confusion Matrix Classification

Let's consider a binary classification case where the output is either 0 or 1:

1. **True positives:** These are the samples for which we predicted 1 as the output and the ground truth is 1 too.
2. **True negatives:** These are the samples for which we predicted 0 as the output and the ground truth is 0 too.
3. **False positives:** These are the samples for which we predicted 1 as the output but the ground truth is 0. This is also known as a Type I error.
4. **False negatives:** These are the samples for which we predicted 0 as the output but the ground truth is 1. This is also known as a Type II error.

Confusion Matrix Classification

True Positive(TP):

- You projected positive and its turn out to be true. For example, you had predicted that France would win the world cup, and it won.

True Negative(TN):

- When you predicted negative, and it's true. You had predicted that England would not win and it lost.

False Positive(FP):

- Your prediction is positive, and it is false.
- You had predicted that England would win, but it lost.

False Negative(FN):

- Your prediction is negative, and result it is also false.
- You had predicted that France would not win, but it won.

Confusion Matrix Classification

		Predicted Class		
		Positive	**Negative**	
Actual Class	**Positive**	True Positive (TP)	False Negative (FN) **Type II Error**	**Sensitivity** $\frac{TP}{(TP + FN)}$
	Negative	False Positive (FP) **Type I Error**	True Negative (TN)	**Specificity** $\frac{TN}{(TN + FP)}$
		Precision $\frac{TP}{(TP + FP)}$	**Negative Predictive Value** $\frac{TN}{(TN + FN)}$	**Accuracy** $\frac{TP + TN}{(TP + TN + FP + FN)}$

Example Scenario

- Suppose we have a binary classification problem to predict whether an email is **Spam** or **Not Spam**. Here, we have four outcomes in our confusion matrix:
- **True Positive (TP)**: Predicted Spam, and it is actually Spam.
- **True Negative (TN)**: Predicted Not Spam, and it is actually Not Spam.
- **False Positive (FP)**: Predicted Spam, but it is actually Not Spam (also known as a Type I error).
- **False Negative (FN)**: Predicted Not Spam, but it is actually Spam (also known as a Type II error).

Example Scenario

Let's say we test our model with 10 emails, and we have the following results:

Actual	Predicted
Spam	Spam
Not Spam	Spam
Spam	Spam
Not Spam	Not Spam
Not Spam	Spam
Spam	Not Spam
Not Spam	Not Spam
Spam	Spam
Not Spam	Not Spam
Spam	Not Spam

Calculate TP, TN, FP, FN

Based on the results above, we categorize each outcome:

1. **True Positives (TP)**: Predicted Spam and actually Spam → 3
2. **True Negatives (TN)**: Predicted Not Spam and actually Not Spam → 3
3. **False Positives (FP)**: Predicted Spam but actually Not Spam → 2
4. **False Negatives (FN)**: Predicted Not Spam but actually Spam → 2

Example Scenario

Create the Confusion Matrix

Using the counts from above, we can create the confusion matrix:

	Predicted: Spam	Predicted: Not Spam
Actual Spam	TP = 3	FN = 2
Actual Not Spam	FP = 2	TN = 3

- Confusion Matrix:

$$\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}$$

Example Scenario

Calculate Performance Metrics

Using the confusion matrix, we can calculate various metrics to evaluate the model's performance:

1. **Accuracy**: Measures the overall correctness of the model.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} = \frac{3+3}{3+3+2+2} = \frac{6}{10} = 0.6\ (60\%)$$

2. **Precision**: Measures the accuracy of positive predictions.

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{3}{3+2} = \frac{3}{5} = 0.6\ (60\%)$$

3. **Recall** (Sensitivity): Measures the ability to find all positive cases.

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{3}{3+2} = \frac{3}{5} = 0.6\ (60\%)$$

4. **F1 Score**: The harmonic mean of Precision and Recall.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = 2 \times \frac{0.6 \times 0.6}{0.6 + 0.6} = 0.6\ (60\%)$$

Questions ??

Thanks'