Artificial Intelligence
&
Neural Networks

CSE-407

Lecture-04

Rajorshi Bhattacharya
Lecturer, Dept. of CSE

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** $\frac{TP + TN}{(TP + TN + FP + FN)}$

		Predicted Class		
		Positive	**Negative**	
Actual Class	**Positive**	True Positive (TP) 3	False Negative (FN) **Type II Error** 1	**Sensitivity** $\frac{TP}{(TP+FN)}$
	Negative	False Positive (FP) **Type I Error** 2	True Negative (TN) 4	**Specificity** $\frac{TN}{(TN+FP)}$
		Precision $\frac{TP}{(TP+FP)}$	**Negative Predictive Value** $\frac{TN}{(TN+FN)}$	**Accuracy** $\frac{TP+TN}{(TP+TN+FP+FN)}$

$$\text{Precision} = \frac{3}{3+2} = 0.6 \Rightarrow 60\%$$

$$Recall = \frac{3}{3+1} = 0.75 \Rightarrow 75\%$$

Sensitivity = Recall

$$F1\ Score = \frac{2*Precision*Recall}{Precision+Recall} = \frac{2*0.6*0.75}{0.6+0.75} = 0.6666 = 0.67$$

Questions ??

Thanks'