CS 480 – Introduction to Artificial Intelligence

Topic: Classifier Evaluation

Mustafa Bilgic

http://www.cs.iit.edu/~mbilgic

https://twitter.com/bilgicm

Types of Errors – Classification

- Assume a target/positive class
  - Spam, HasHeartDisease, etc.
- *False positive*
  - Falsely classifying an object as positive
    - E.g., classifying a legitimate email as spam, diagnosing a healthy patient as having heart disease, and so on
  - Also called *Type I* error
- *False negative*
  - Falsely classifying an object as negative
    - E.g., classifying a spam email as not-spam, claiming that a heart-disease patient is healthy, and so on
  - Also called *Type II* error

Confusion Matrix

		Predicted Class	
		Positive	**Negative**
Actual Class	**Positive**	True Positive	False Negative
	Negative	False Positive	True Negative

Accuracy

		Predicted Class	
		Positive	**Negative**
Actual Class	**Positive**	True Positive	False Negative
	Negative	False Positive	True Negative

$$Accuracy = \frac{Num\ Correct}{Data\ Size} = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision

		Predicted Class	
		Positive	**Negative**
Actual Class	**Positive**	True Positive	False Negative
	Negative	False Positive	True Negative

$$Precision = \frac{True\ Positive}{Predicted\ Positive} = \frac{TP}{TP + FP}$$

True Positive Rate – Recall – Sensitivity

		Predicted Class	
		Positive	**Negative**
Actual Class	**Positive**	True Positive	False Negative
	Negative	False Positive	True Negative

$$TPR = Recall = \frac{True\ Positive}{Actual\ Positive} = \frac{TP}{TP + FN}$$

True Negative Rate – Specificity

		Predicted Class	
		Positive	**Negative**
Actual Class	**Positive**	True Positive	False Negative
	Negative	False Positive	True Negative

$$TNR = Specificity = \frac{True\ Negative}{Actual\ Negative} = \frac{TN}{TN + FP}$$

False Positive Rate – Fall-out

		Predicted Class	
		Positive	**Negative**
Actual Class	**Positive**	True Positive	False Negative
	Negative	False Positive	True Negative

$$FPR = FallOut = \frac{False\ Positive}{Actual\ Negative} = \frac{FP}{TN + FP}$$

False Negative Rate – Miss Rate

		Predicted Class	
		Positive	**Negative**
Actual Class	**Positive**	True Positive	False Negative
	Negative	False Positive	True Negative

$$FNR = Miss\ Rate = \frac{False\ Negative}{Actual\ Positive} = \frac{FN}{TP + FN}$$

F1

		Predicted Class	
		Positive	**Negative**
Actual Class	**Positive**	True Positive	False Negative
	Negative	False Positive	True Negative

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

Harmonic mean of precision & recall

Making a Classification Decision

- Given a probabilistic output for an object, say $\langle p, 1-p \rangle$, how do we decide which class to assign to this object?
- The simplest approach is check whether $p > 0.5$ and make a decision accordingly
- This assumes each mistakes (False Positives and False Negatives) are equally costly

Equal Misclassification Costs?

- Which one is worse for you:
  - Delivering a spam email into your Inbox (False Negative), or
  - Delivering a legitimate email into your Spam folder (False Positive)?
- If one is worse than the other, then, should we use 0.5 as the decision threshold or should we adjust it to your preference?

Cost Matrix

		Predicted Class	
		Positive	**Negative**
Actual Class	**Positive**	0	a
	Negative	b	0

Given a probability distribution of $\langle p, 1-p \rangle$ for $\langle Positive, Negative \rangle$ respectively, and given the above cost matrix, under what conditions (in terms of a, b, and p) would you classify an object as *Positive*?

Exercise

Real Life Measures

- Not as clean as the ones we discussed
- Imagine self-driving cars, medical diagnosis, crime prediction, fraud detection, and so on
- Usually, there is not a single performance measure
- Performance is handled on a case-by-case basis; not on an aggregate level