http://poloclub.gatech.edu/cse6242

CSE6242: Data & Visual Analytics

Visualization for Classification ROC, AUC, Confusion Matrix

Duen Horng (Polo) Chau

Associate Professor, College of Computing

Associate Director, MS Analytics

Georgia Tech

Mahdi Roozbahani

Lecturer, Computational Science & Engineering, Georgia Tech

Founder of Filio, a visual asset management platform

Partly based on materials by

Professors Guy Lebanon, Jeffrey Heer, John Stasko, Christos Faloutsos

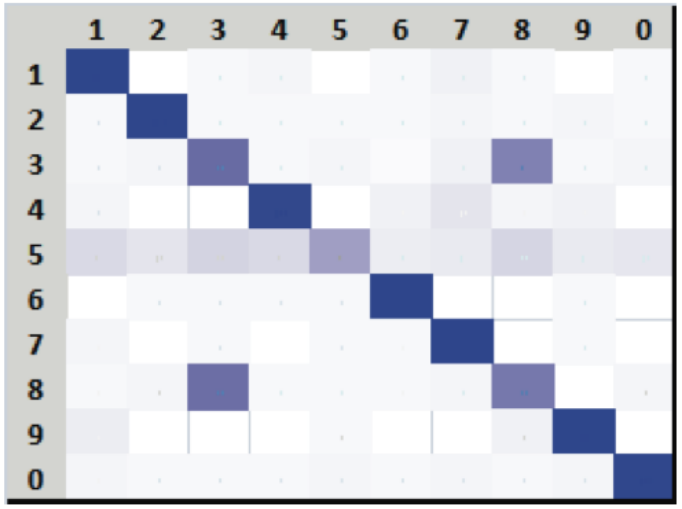
Visualizing Classification Performance

Confusion matrix

		Predicted class		
		Cat	**Dog**	**Rabbit**
Actual class	**Cat**	5	3	0
	Dog	2	3	1
	Rabbit	0	2	11

https://en.wikipedia.org/wiki/Confusion_matrix

	1	2	3	4	5	6	7	8	9	0
1	91	0	1	2	0	1	3	1	0	1
2	1	89	1	1	1	1	2	1	2	1
3	1	2	48	1	2	0	3	40	1	2
4	2	0	0	83	0	3	7	2	3	0
5	10	7	12	10	30	4	5	11	5	6
6	0	1	1	1	1	95	0	0	1	0
7	2	0	1	0	1	1	94	0	1	0
8	1	2	47	1	1	1	2	43	0	2
9	4	0	0	0	1	0	0	3	92	0
0	2	1	1	1	2	1	2	1	2	87



Hard to spot trends and patterns

Much easier!

Figure 2. Representations of confusion matrix for a handwritten digit classification task. (top) standard confusion matrix; (bottom) heat-map confusion matrix. It is much easier to identify underlying patterns in the visual representation; 3 and 8 are often misclassified as each other and 5 is misclassified as many different numbers.

http://research.microsoft.com/en-us/um/redmond/groups/cue/publications/CHI2009-EnsembleMatrix.pdf

Very important:
Find out what "positive" means

		Predicated: Cat	Predicated: Dog
Actual	Cat	5	3
	Dog	2	4

Very important: Find out what “positive” means

Terminology and derivations from a confusion matrix

true positive (TP)

eqv. with hit

true negative (TN)

eqv. with correct rejecti

false positive (FP)

eqv. with false alarm, Type I error

false negative (FN)

eqv. with miss, Type II error

---

sensitivity or true positive rate (TPR)

eqv. with hit rate, recall

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

specificity (SPC) or true negative rate (TNR)

$$SPC = \frac{TN}{N} = \frac{TN}{FP + TN}$$

precision or positive predictive value (PPV)

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

recall (recall)

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

negative predictive value (NPV)

$$NPV = \frac{TN}{TN + FN}$$

fall-out or false positive rate (FPR)

$$FPR = \frac{FP}{N} = \frac{FP}{FP + TN} = 1 - SPC$$

“**False Alarm**” easy to remember in security applications

https://en.wikipedia.org/wiki/Confusion_matrix

Visualizing Classification Performance

using

ROC curve

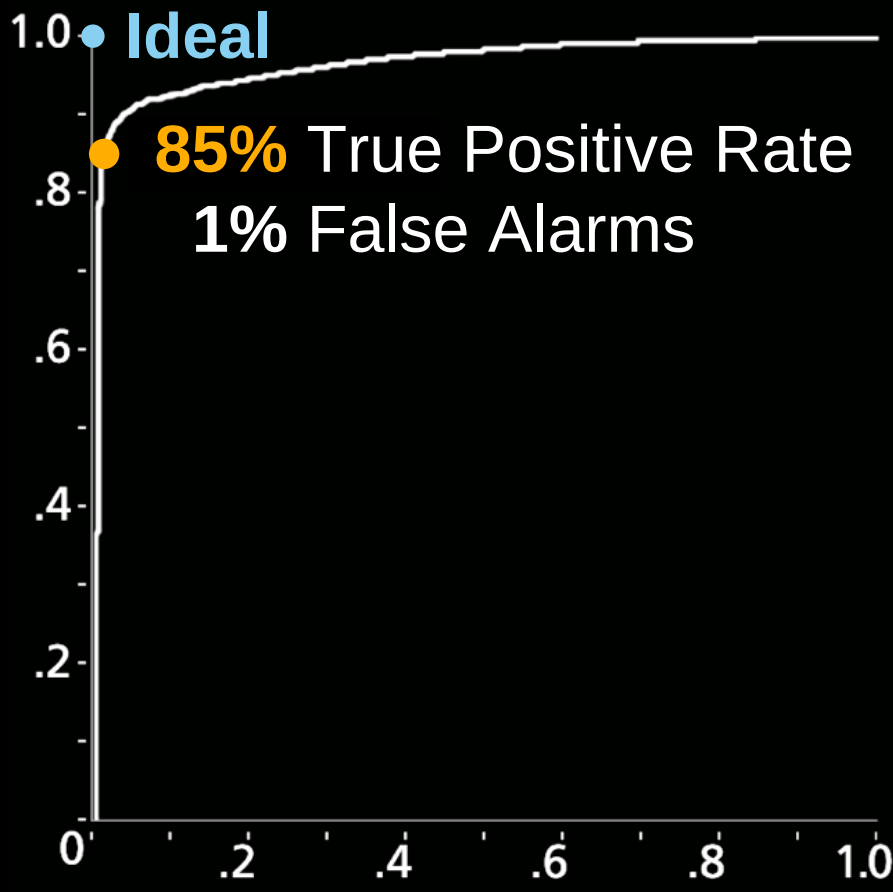
(Receiver Operating Characteristic)

Polonium's ROC Curve

Positive class: malware

Negative class: benign

True Positive Rate
% of **bad** correctly labeled



False Positive Rate (False Alarms)
% of **good** labeled as **bad**

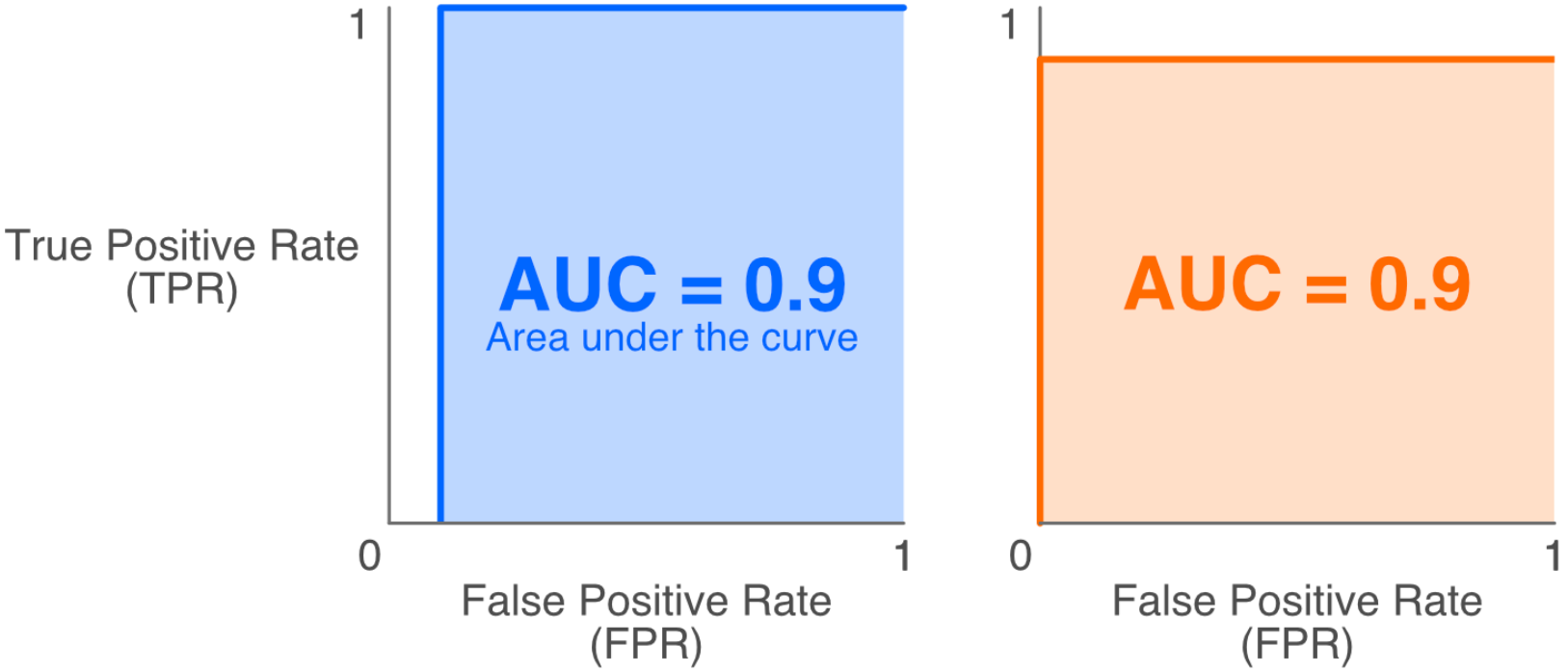
Measuring Classification Performance

using **AUC** (Area under the ROC curve)

If a machine learning algorithm achieves **0.9 AUC** (out of 1.0),

that's a great algorithm, right?

Be Careful with AUC!



Weights in combined models

Bagging / Random forests

- Majority voting

Let people play with the weights?

EnsembleMatrix (adjusting weights)

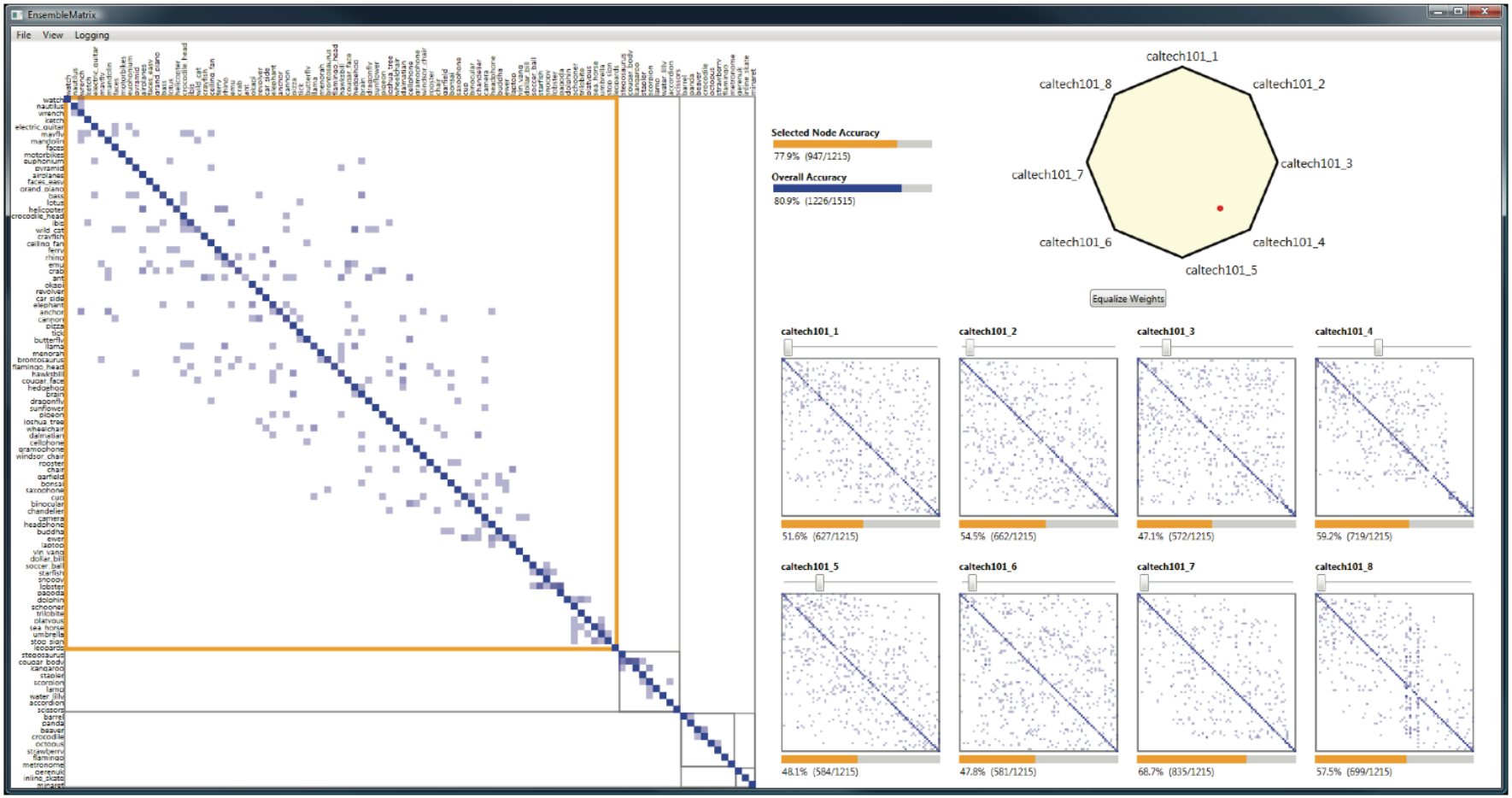


Figure 1. Primary view in EnsembleMatrix. Confusion matrices of component classifiers are shown in thumbnails on the right. The matrix on the left shows the confusion matrix of the current ensemble classifier built by the user.

http://research.microsoft.com/en-us/um/redmond/groups/cue/publications/CHI2009-EnsembleMatrix.pdf

Improving performance

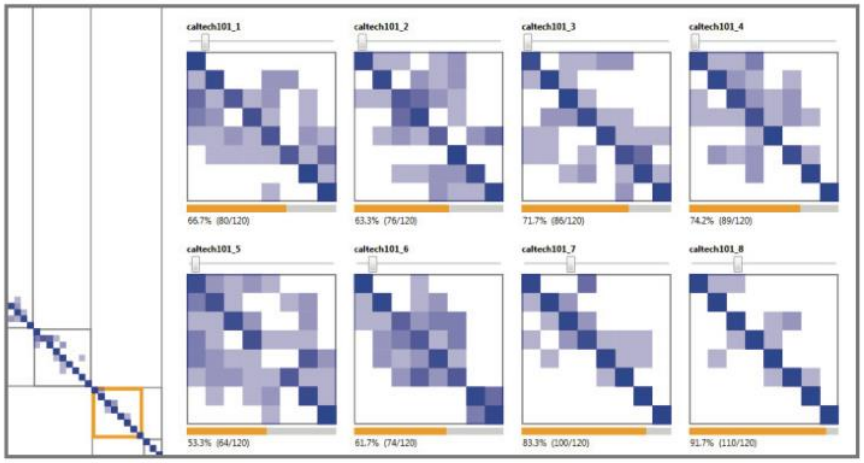


Figure 3. After partitioning the matrix, selecting a partition, outlined in orange, causes the thumbnails to display only the data instances in that partition. The component classifiers demonstrate very different behavior in this partition, including clustering and large differences in accuracy.

- Adjust the weights of the individual classifiers
- Data partition to separate problem areas
  - Adjust weights just for these individual parts
- Caveat: evaluation used one dataset

http://research.microsoft.com/en-us/um/redmond/groups/cue/publications/CHI2009-EnsembleMatrix.pdf