HUDK 4050:
CORE METHODS
IN EDM

11/27/17 5:41 PM

In the news

theguardian

How much ...? The rise of dynamic and personalised pricing

MotherJones

Inside Silicon Valley's Big-Money Push to Remake American Education

The New York Times

Cops, Cellphones and Privacy at the Supreme Court

Forbes

LearnPlatform Aims to Bring Transparency and Savings to Murky, $12 Billion EdTech Purchasing Market

QUARTZ

Ed Tech Helps English Language Learners Flourish

EdTech Focus On K-12

Indian school kids are learning to build robots with DIY kits and online courses

wcet

E.U. Regulations that are Enforceable Against U.S. Higher Education Institutions

AltSchool wants to change how kids learn, but fears have surfaced that it's failing students

TC

npr

Net Neutrality: The Long View

IBM pitched its Watson supercomputer as a revolution in cancer care. It's nowhere close

STAT

HACKERNOON

THE VERGE

New York attorney general says the FCC won't help investigate fake net neutrality comments

More than a Million Pro-Repeal Net Neutrality Comments were Likely Faked

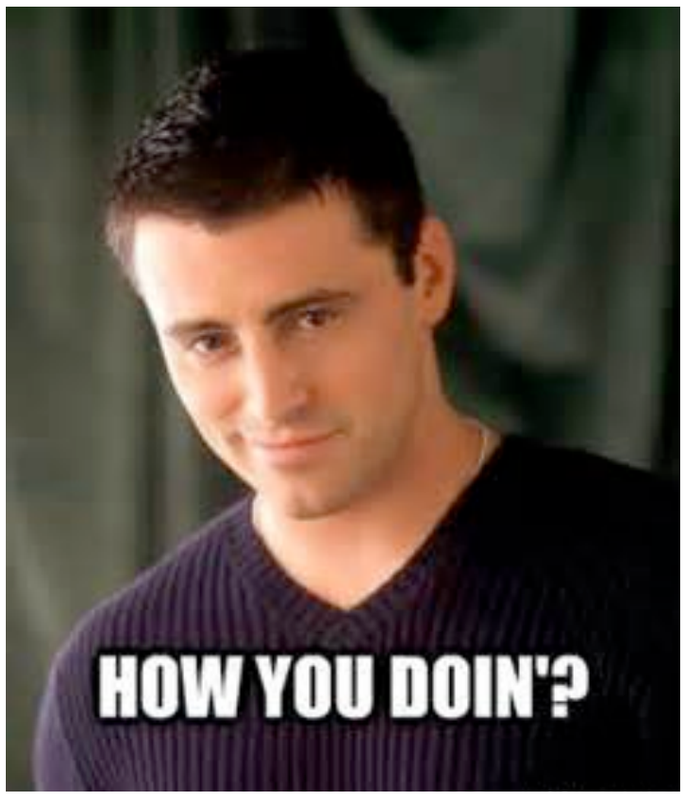
npr

Navy To No Longer Require Pepper-Spraying For Sailors Every 3 Years

Events

Event	Date	Time	Location	URL
Innovative Teaching Co-op Monthly Meetup - December 2017	December 6	4:30pm	Newsela, 475 10th Ave	https://www.eventbrite.com/e/innovative-teaching-co-op-monthly-meetup-december-2017-tickets-39559535636
Global Edtech Landscape	December 12	3:00pm	Online	https://www.eventbrite.com/e/webinar-global-edtech-landscape-tickets-39186587137
MAPLE Landscape Analysis Findings	December 14	4:00pm	Online	https://www.eventbrite.com/e/webinar-findings-a-report-on-personalized-learning-in-massachusetts-tickets-39627055590
iOS Winter Bootcamp	January 1 - 12		UC Berkeley	http://www.bayareamobile.io/
Building Equity in Tech	December 5	7:00pm	BRIC 647 Fulton St	https://www.eventbrite.com/e/the-stoop-series-re-programming-building-equity-in-tech-tickets-36872928923?utm_source=eb_email&utm_medium=email&utm_campaign=order_confirmation_email&utm_term=eventnam
Data Byte: regulating informational infrastructure	December 6	4:00pm	36 West 20th Street	https://datasociety.net/events/databite-no-105-k-sabeel-rahman/
IEEE BigData 2017	December 11-14		Westin Copley Place, Boston located at 10 Huntington Avenue	http://cci.drexel.edu/bigdata/bigdata2017/index.html
TC Innovation Award	December 13	6:00pm	Smith Learning Center	https://listserv.tc.columbia.edu/t/2367677/32524635/11088/8/

Diagnostic Metrics



How to determine how well your model is doing

Diagnostic Metrics

Classification

- Accuracy
- Cohen's Kappa
- ROC/AUC/A'
- Correlation
- RMSE

Regression

- MAE/RMSE
- Pearson's Correlation/R^2
- AIC/BIC

Terms

- **Ground truth**: data that is available, relevant, and most trustworthy to train your model

- **Baseline**: initial measurement

- **Gold standard**: (expensive) comparative measurement

- **Inference**: data that is inferred from logic + data

Diagnostics for Classifiers

Accuracy

- $\frac{\text{correct predictions}}{\text{total predictions}}$

- Gotcha: unequal categories

- EG - Predicting fraudulent credit card transactions

- False positives/negatives (over/ under predict)

Precision & Recall

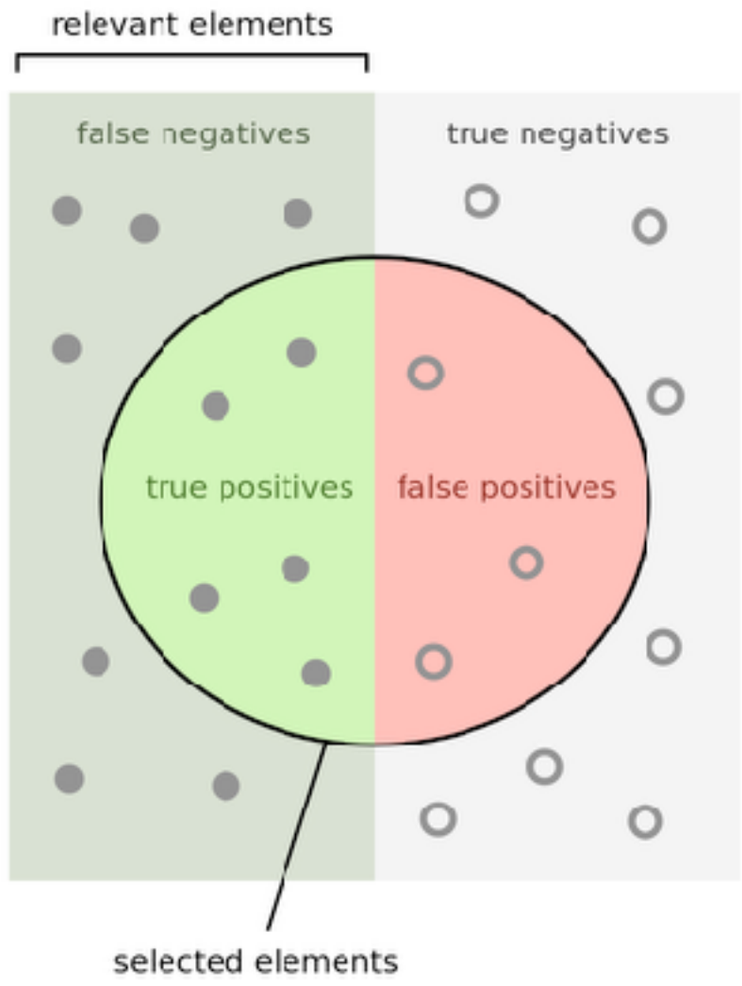
$$\textbf{Precision} = \frac{\text{True Positive}}{\text{True Positive + False Positive}}$$

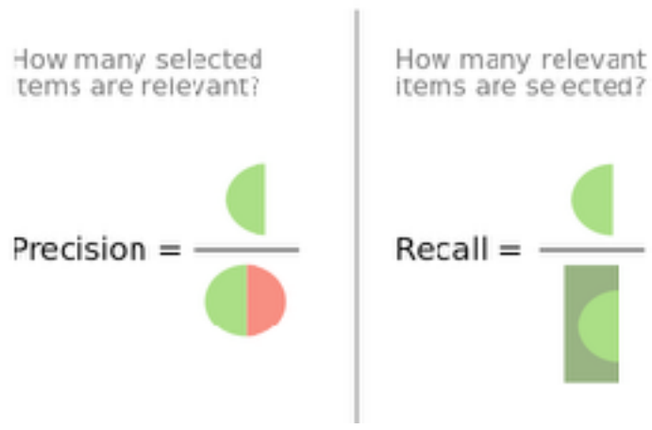
$$\textbf{Recall} = \frac{\text{True Positive}}{\text{True Positive + False Negative}}$$

Precision & Recall

Precision
The fraction (probability) of predictions that are ***relevant***

Recall
The fraction (probability) of relevant instances that are ***predicted***





Walber, 2014

Cohen's Kappa (κ)

- Traditionally used for inter-rater reliability
- We will use it to look at the reliability between the data and our model

$$\kappa = \frac{p_o - p_e}{1 - p_e}$$

Observed Agreement

Expected Agreement

(hypothetical probability of chance agreement)

		Model		
		Yes	No	
Data	Yes	4	2	6
	No	3	3	6
		7	5	12

$p_o = (4 + 3)/12 = 0.58$

$p_e = (7/12) \times (6/12) + (5/12) \times (6/12) = 0.5$

$\kappa = (0.58 - 0.5)/(1 - 0.5) = 0.16$

Is this good? Depends on the context

Gotchas with Kappa

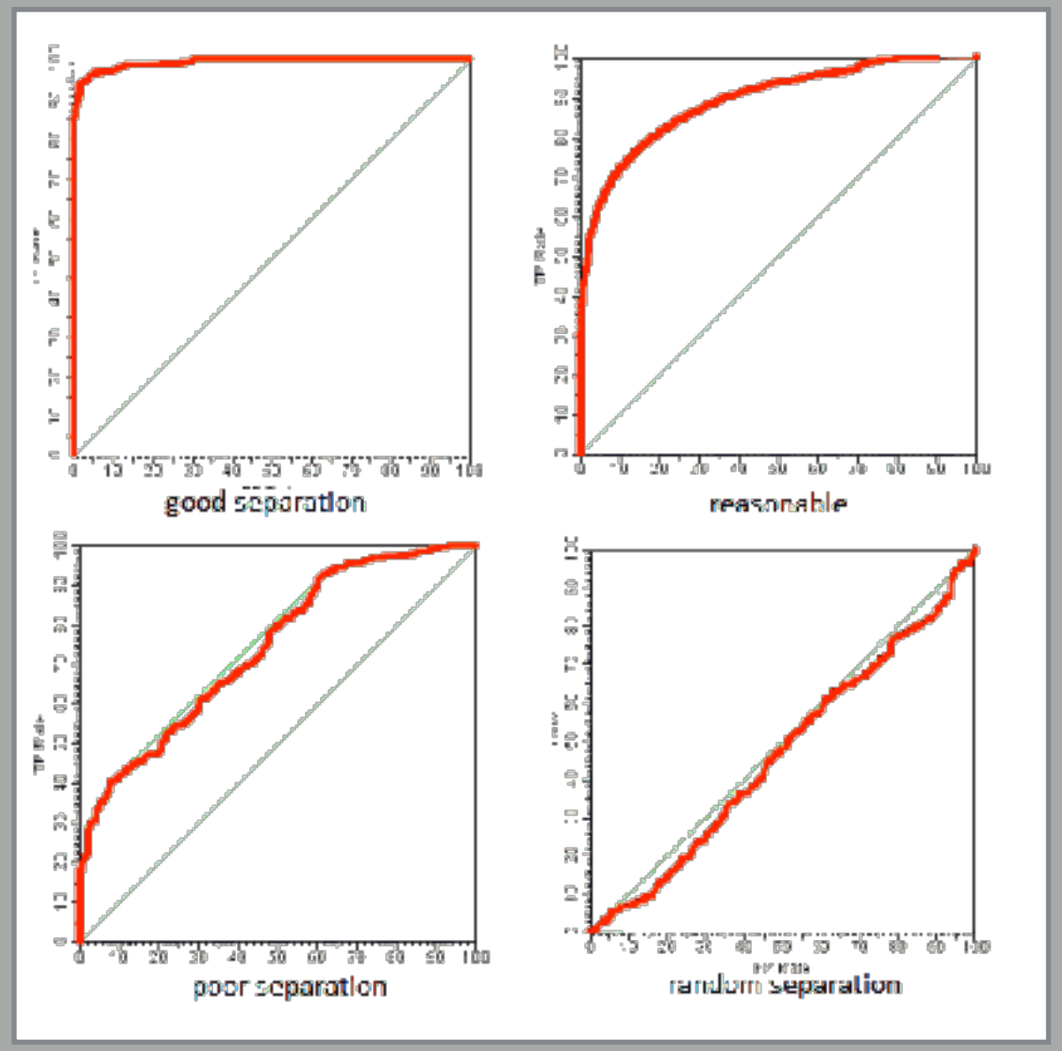
- Again, rare categories pose a problem and will incur a higher penalty than common categories
- Does the marginal probability represent “chance”?

Probabilities

- Model assigns a probability of belonging to a class, rather than a class directly
- Then choose a probability threshold to assign to a class
- Allows us to choose a preference based on the consequences of false positives/negatives
- http://www.navan.name/roc/

Receiver Operating Characteristic (ROC)

- Relationship between the false positive and the true positive rates
- World War II for detecting enemy objects on radar in response to Pearl Harbor
- Demonstrates the sensitivity vs specificity tradeoff



Receiver Operating Characteristic (ROC)

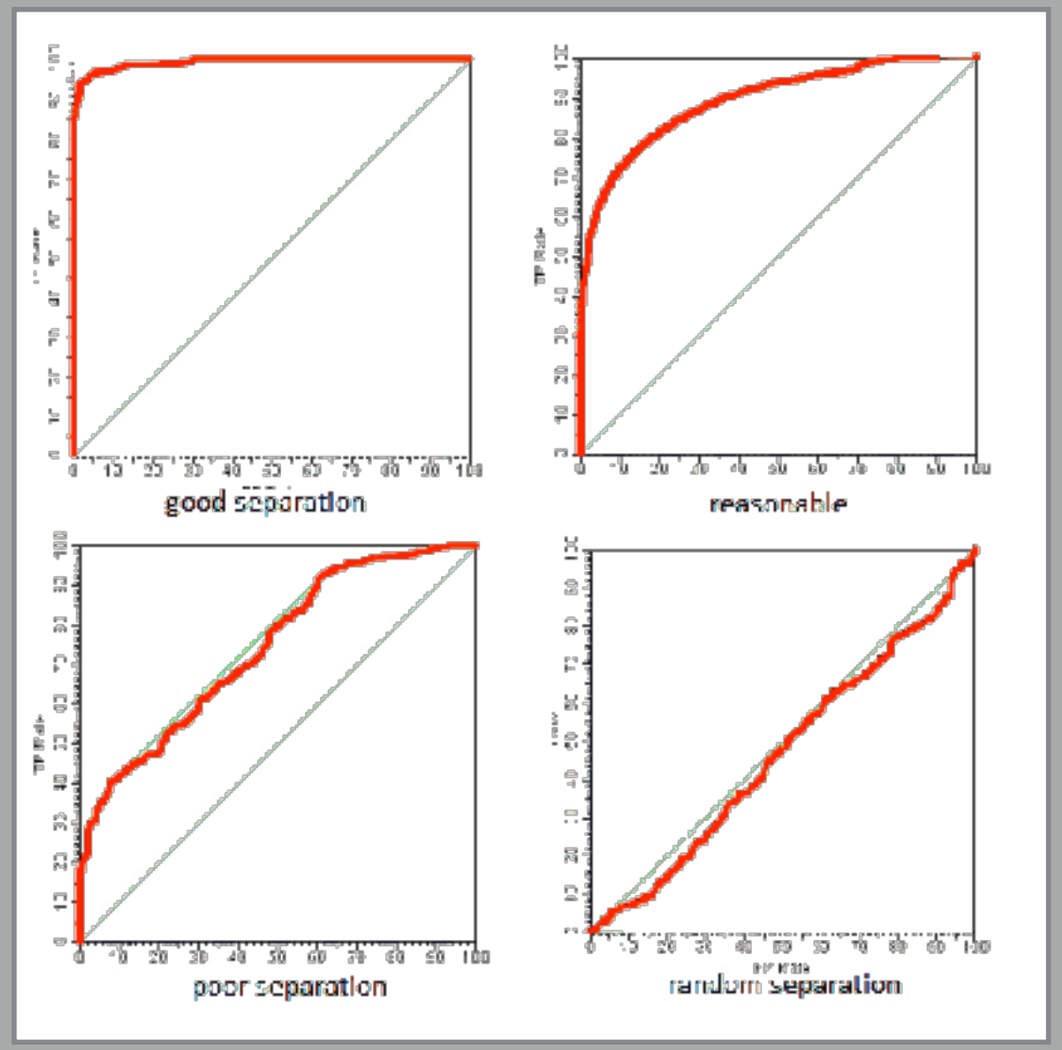
<u>Area Under the ROC Curve:</u>

- AUC is the collapsed metric to compare models

- AUC is the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one

- It is equivalent to the Wilcoxon Sum-Rank Test and can therefore generate a probability test

- Is sometimes called A' (A Prime) depending on how it is calculated



Diagnostics for Regressors

Mean Absolute Error

- Mean of observed values minus predicted values

MAE = $\frac{\sum|x - \overline{x}|}{n}$

Root Mean Squared Error

- Square root of the observed values minus predicted values squared

RMSE = $\sqrt{\frac{\sum_{i=1}^{n}(p_i - a_i)^2}{n}}$

Pearson's Correlation

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

- Measure of the linear dependence between two variables
- Covariance between two variables divided by the product of the standard deviation of those variables
- Development began ~ 1880s by Galton and then Pearson
- Gotcha: must be a linear relationship

Denis Boigelot, 2011

r^2

- The proportion of the variance in the dependent variable that is predicted from the independent variable
- There are several ways to calculate R^2
- If it involves two variables it is the square of the correlation (OLS classes will go more in depth)

Akaike Information Criterion (AIC)

AIC = number of parameters - goodness of fit

- Developed by Akaike in 1971 based on thermodynamics
- Relative estimate of the information lost when a given model is used to represent the process that generates the data
- Model with lowest AIC “wins”
- Represents the trade off between goodness-of-fit with model complexity
- It compares models, cannot give an estimate of model fit in an absolute sense
- Gatcha: Software implementation was not always reliable

Bayesian Information Criterion (BIC)

BIC = number of parameters x sample size - goodness of fit

- Developed by Schwarz in 1978
- Uses Bayes Theorem to penalize the addition of parameters
- Penalty for adding parameters is great than in AIC
- Represents the trade off between goodness-of-fit with model complexity
- Lowest BIC “wins”
- Gotcha: Does poorly when dealing with many parameters

http://bit.ly/cmedma7