Evaluation: From precision, recall, and F-measure to ROC, informedness, markedness and correlation

Krishna Kalyan

About the Author

David M W Powers

Follow

Professor of Computer and Cognitive Science, Director, Knowledge and Interaction Technology Centre

Computational Psycholinguistics, Cognitive Linguistics, Computational Cognitive Science, Sentic Computing, BCI

Verified email at flinders.edu.au - Homepage

Title 1–20	Cited by	Year
Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation DMW Powers International Journal of Machine Learning Technology 2 (1), 37-63	1113	2011
Evaluation: from precision, recall and F-factor to ROC, informedness, markedness and correlation DMW Powers School of Informatics and Engineering Technical Reports (24p ver.of ECAI ...	~~1110~~ *	2007
Measuring semantic similarity in the taxonomy of WordNet D Yang, DMW Powers Proceedings of the Twenty-eighth Australasian conference on Computer Science ...	166	2005
Indexing system and method R Steele, D Powers US Patent 20,030,191,737	116 *	2003
Applications and explanations of Zipf's law DMW Powers Proceedings of the joint conferences on new methods in language processing ...	97	1998

Google Scholar

Get my own profile

Citation indices	All	Since 2012
Citations	3180	2197
h-index	24	16
i10-index	48	29

Co-authors View all…

Trent Wilson Lewis
Martin Luerssen
Richard Leibbrandt
Adham Atyabi
Dongqiang Yang
Sean Patrick Fitzgibbon
Kenneth Pope
Kenneth Treharne

Introduction

- Contingency table

	+R	−R	
+P	tp	fp	pp
−P	fn	tn	pn
	rp	rn	1

	+R	−R	
+P	A	B	A+B
−P	C	D	C+D
	A+C	B+D	N

- Measure to evaluate Machine Learning System

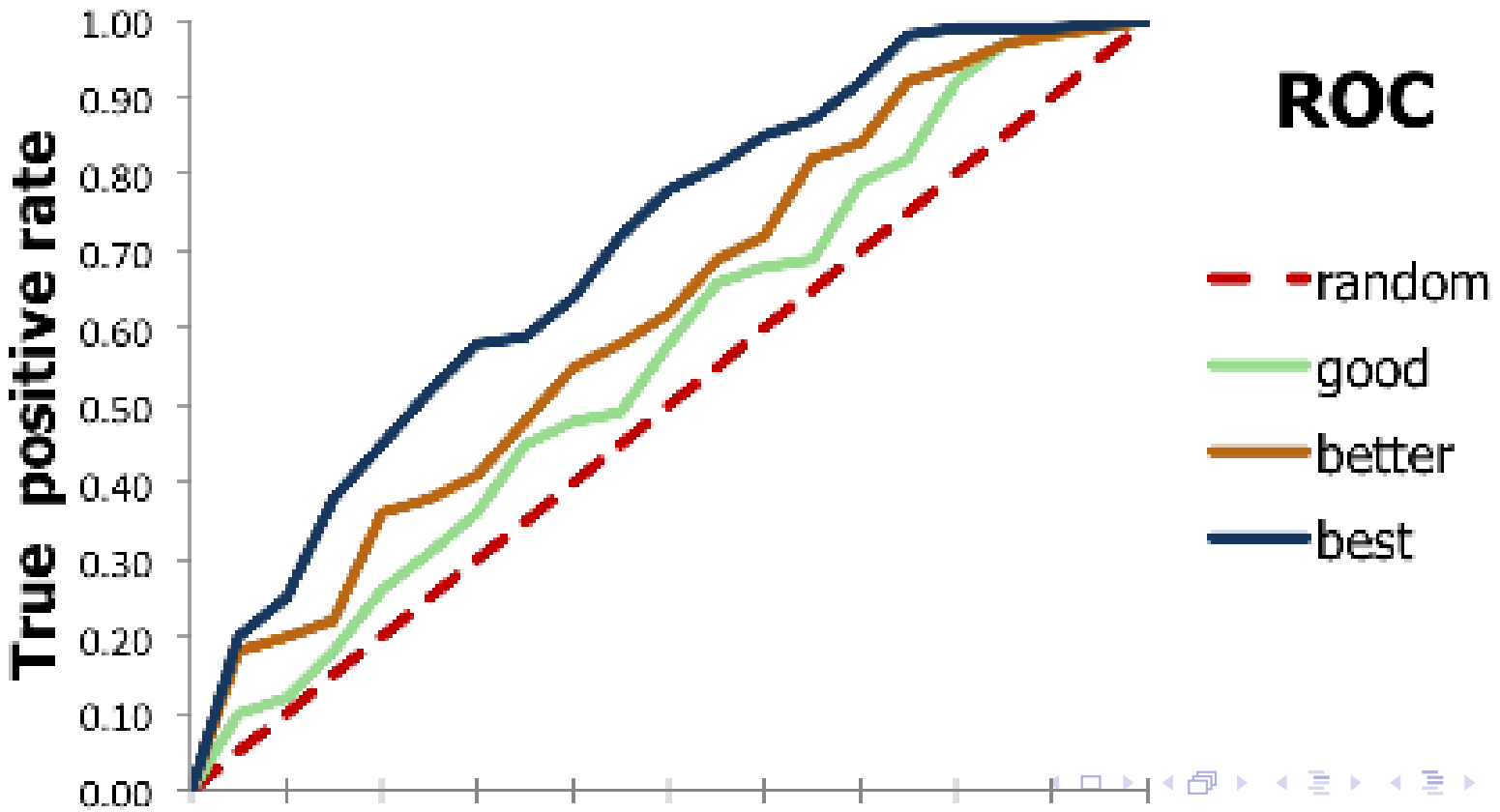
1. Precision
2. Recall
3. F Measure
4. Issues with them

- Alternate Techniques

1. Accuracy
2. Cohen Kappa

ROC Analysis

- ROC analysis give geometric insights into the nature of the measures and their sensitivity to skew
- Compare Classifiers
- Choose parameters based on maximization of AUC



Informedness and Markedness

- Markedness

Markedness quantifies how marked a condition is for the specified predictor, and specifies the probability that a condition is marked by the predictor (versus chance).

Markedness is a deep measure of how consistently the outcome has the Predictor as a Marker by combining surface measures about what proportion of Predictions are correct

Precision + Inverse Precision - 1

- Informedness

Informedness quantifies how informed a predictor is for the specified condition, and specifies the probability that a prediction is informed in relation to the condition (versus chance).

Informedness is a deep measure of how consistently the predictor predicts the outcome by combining surface measures about what proportion of outcomes are correctly predicted.

Regression

- Linear Regression
- Estimating Coefficients

1. Rp
2. Rr
3. Rg

Conclusion

Infromedness usually is a better evaluation measure in binary classification.

Further work to research into the multiclass application of the technique.

Explore the relationship between Infromedness and Markedness

We have a better and a more intuitive understanding of the mearues that we could use for evaluating our ML system.

- Accuracy
- Cohen Kappa
- Precision
- Recall
- F1-Score
- Bookmaker Informedness / Delta P'
- Markedness

Skipped

- Significance
- Montecarlo Simulation
- Evenness (Used in multiclass problems)