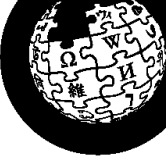




Vandalism Detection on Wikipedia

The class imbalance problem
& new approaches



Contents

Vandalism detection

The class imbalance problem


Content based classifiers



Wikipedia in Numbers

 920 K

 4.7 M

 6 M



Vandalism

“Vandalism is any addition, removal, or change of content, in a *deliberate* attempt to compromise the integrity of Wikipedia.”

en.wikipedia.org/wiki/Wikipedia:Vandalism



Demo

Webis Wikipedia Vandalism Detection Bot

🔄 Tracking and classifying current edits received from en.wikimedia.org IRC channel #en.wikipedia.

Filter classified edits

All Vandalism Regular

Adjust confidence threshold



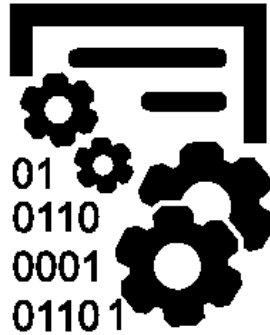
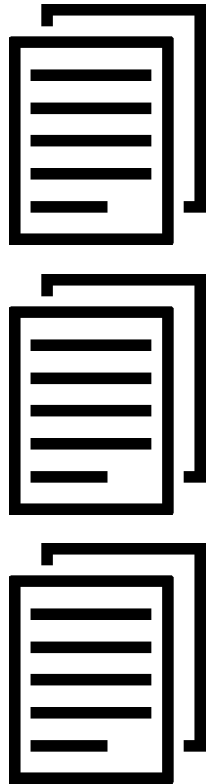
Received 195

Vandalism 6

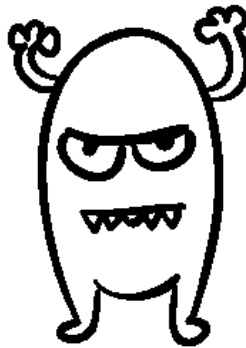
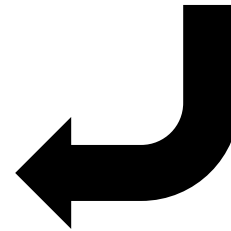
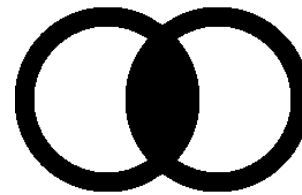
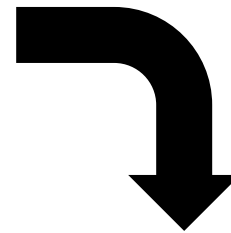
#	Article	Editor	Comment	Edited at	Confidence	
6	Kaun Banega Crorepati	Svpnikhil	/* Other versions */	16:41, 12 Oct 2014	0.5	Show Edit
12	Disney Channel Circle of Stars	80.31.102.189		16:42, 12 Oct 2014	0.7	Show Edit
15	Sean Bell shooting incident	159.92.9.130		16:42, 12 Oct 2014	1	Show Edit
65	Disney Channel Circle of Stars	80.31.102.189	[[WP:AES]]Replaced content with 'gOOD mOORNING'	16:42, 12 Oct 2014	0.8	Show Edit
158	Streptococcus	99.11.160.173	/* Molecular taxonomy and phylogenetics */	16:43, 12 Oct 2014	1	Show Edit
195	Acropolis of Athens	92.75.118.59		16:44, 12 Oct 2014	0.6	Show Edit



Detecting Vandalism

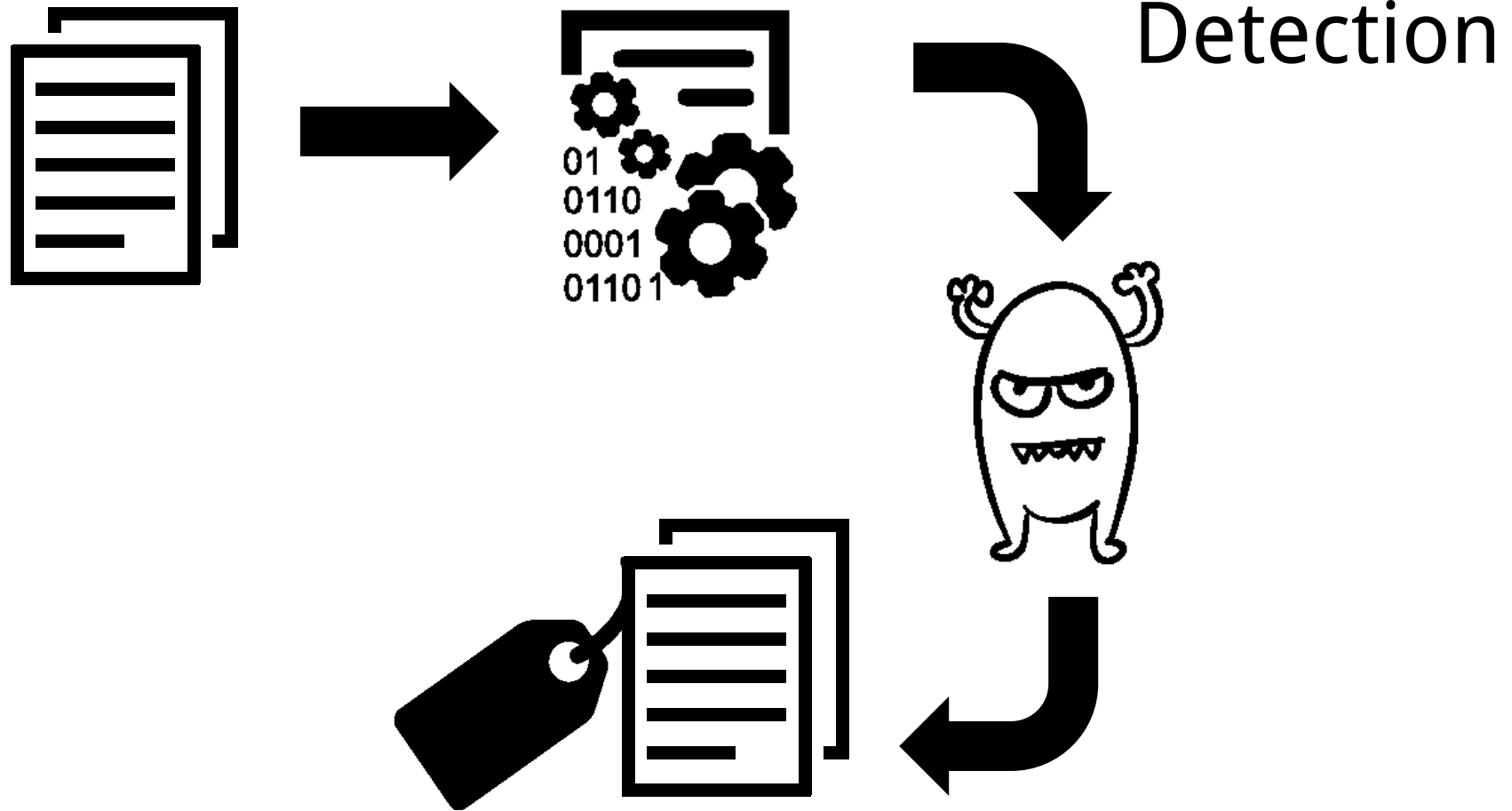


Learning



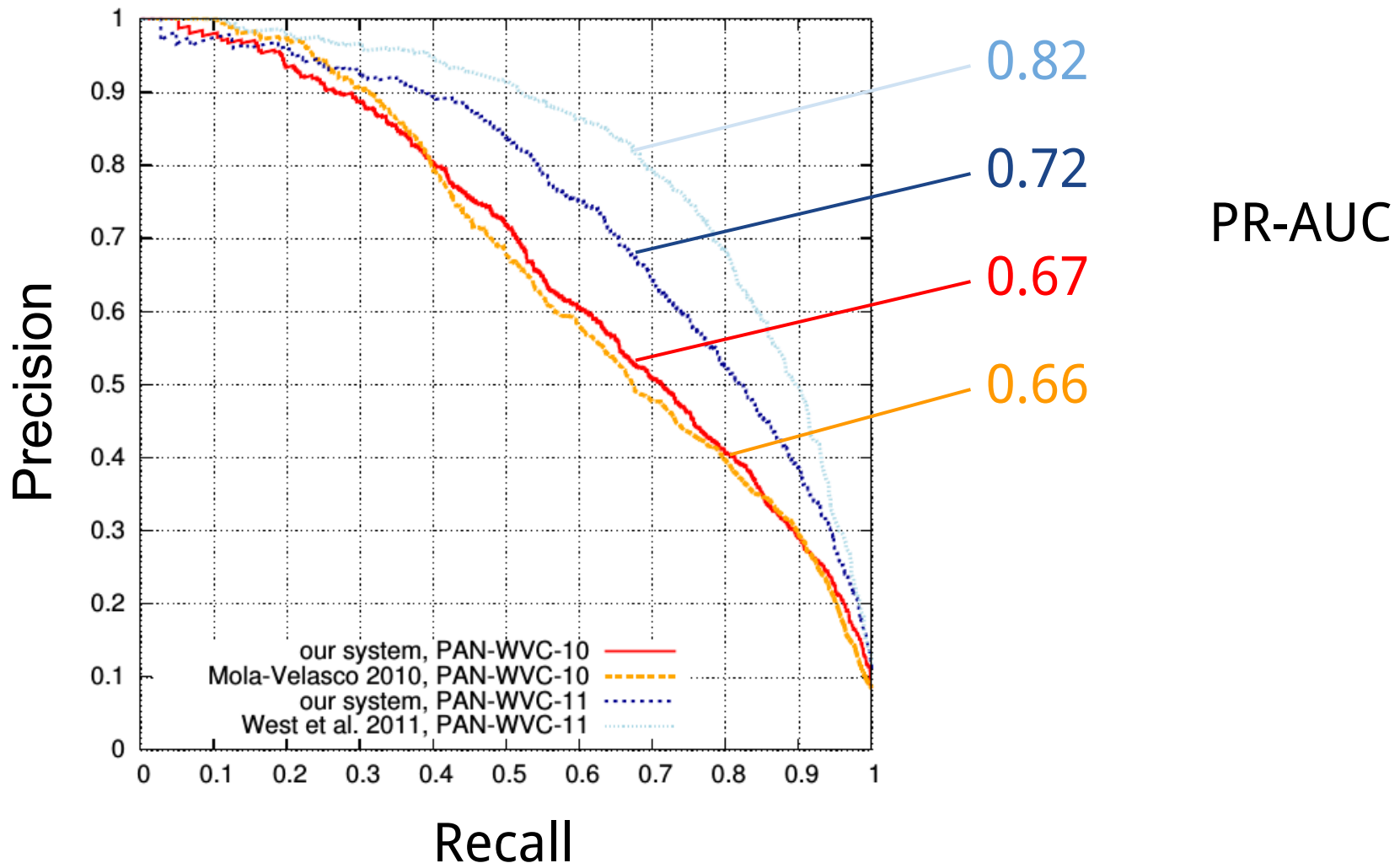


Detecting Vandalism





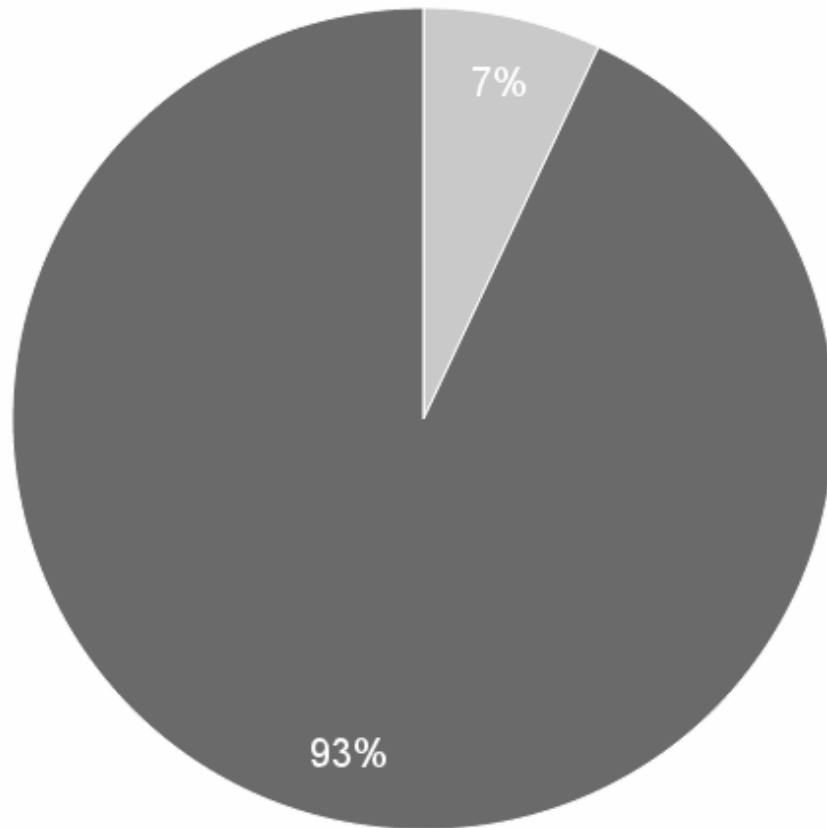
The Detection System





Class Imbalance

Training dataset



■ Vandalism ■ Regular



Class Imbalance Problem

Reasons:

1. minimizing the overall error
2. assuming balanced class distribution
3. assuming equal misclassification cost

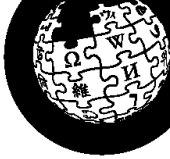


Dataset Resampling

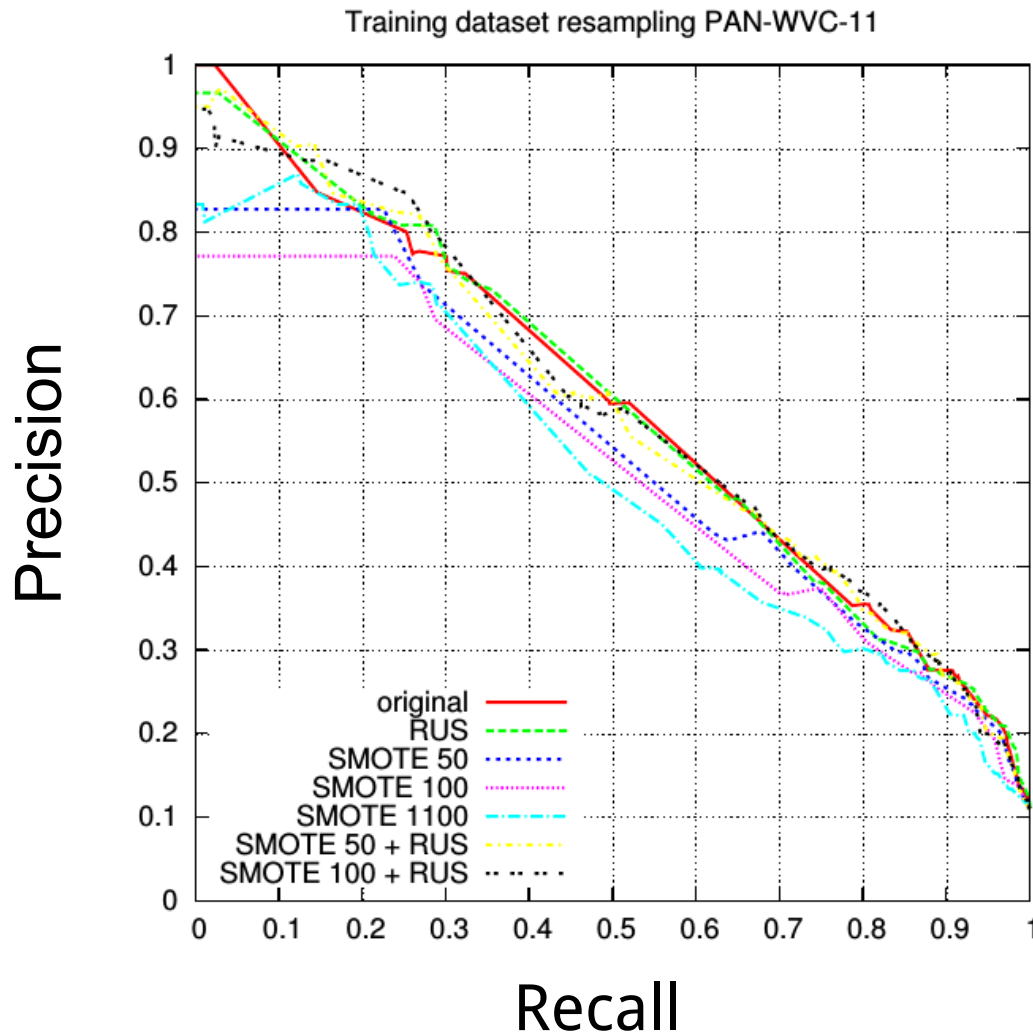
Random Undersampling

SMOTE = Synthetic Minority
Oversampling TEchnique

Chawla, N. V.; Bowyer, K. W.; Hall, L. O. & Kegelmeyer, W. P.: SMOTE: Synthetic Minority Oversampling Technique, *Journal of Artificial Intelligence Research*, AI Access Foundation, **2002**, 16, 321-357



Dataset Resampling

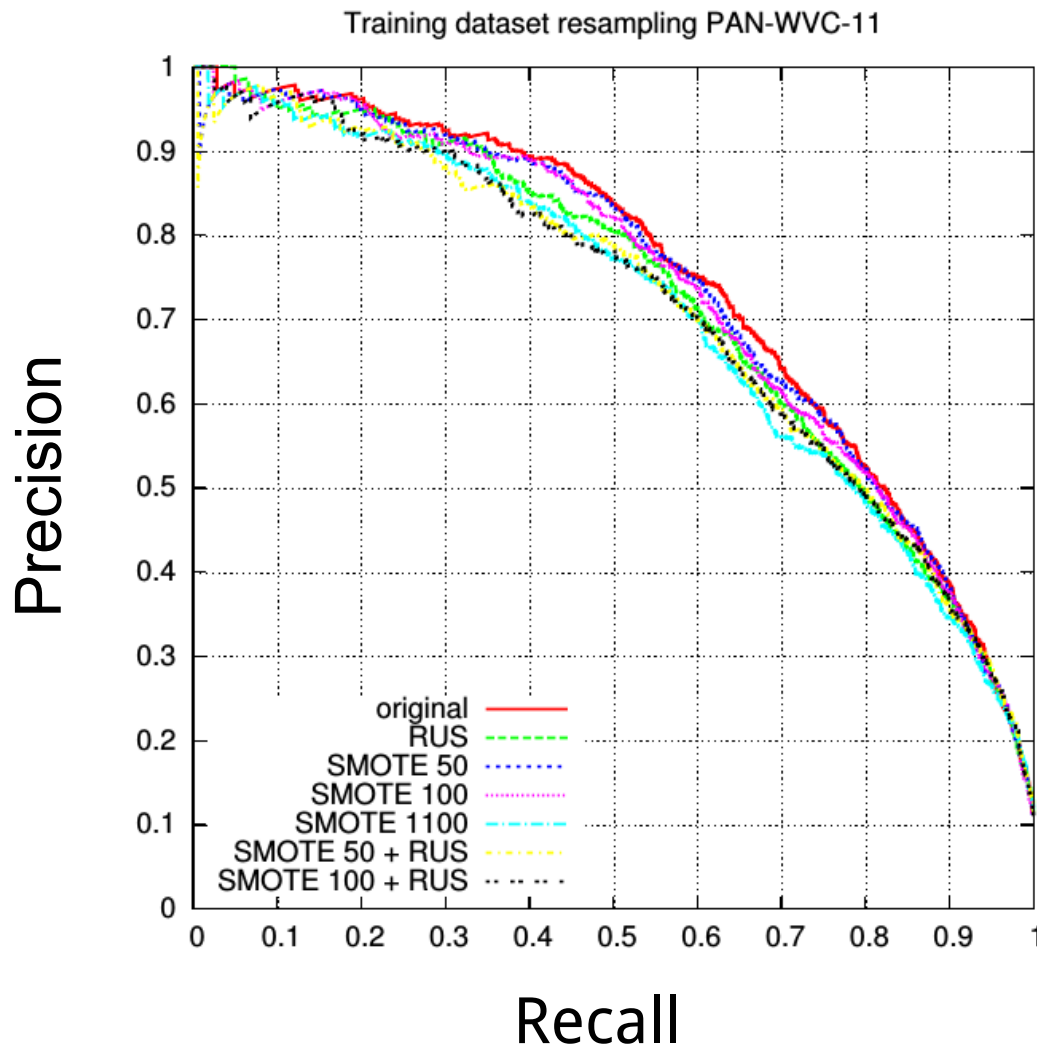


RealAdaBoost

Friedman, J. et al.:
Additive Logistic
Regression: a Statistical
View of Boosting, *The
Annals of Statistics*,
2000, 38



Dataset Resampling



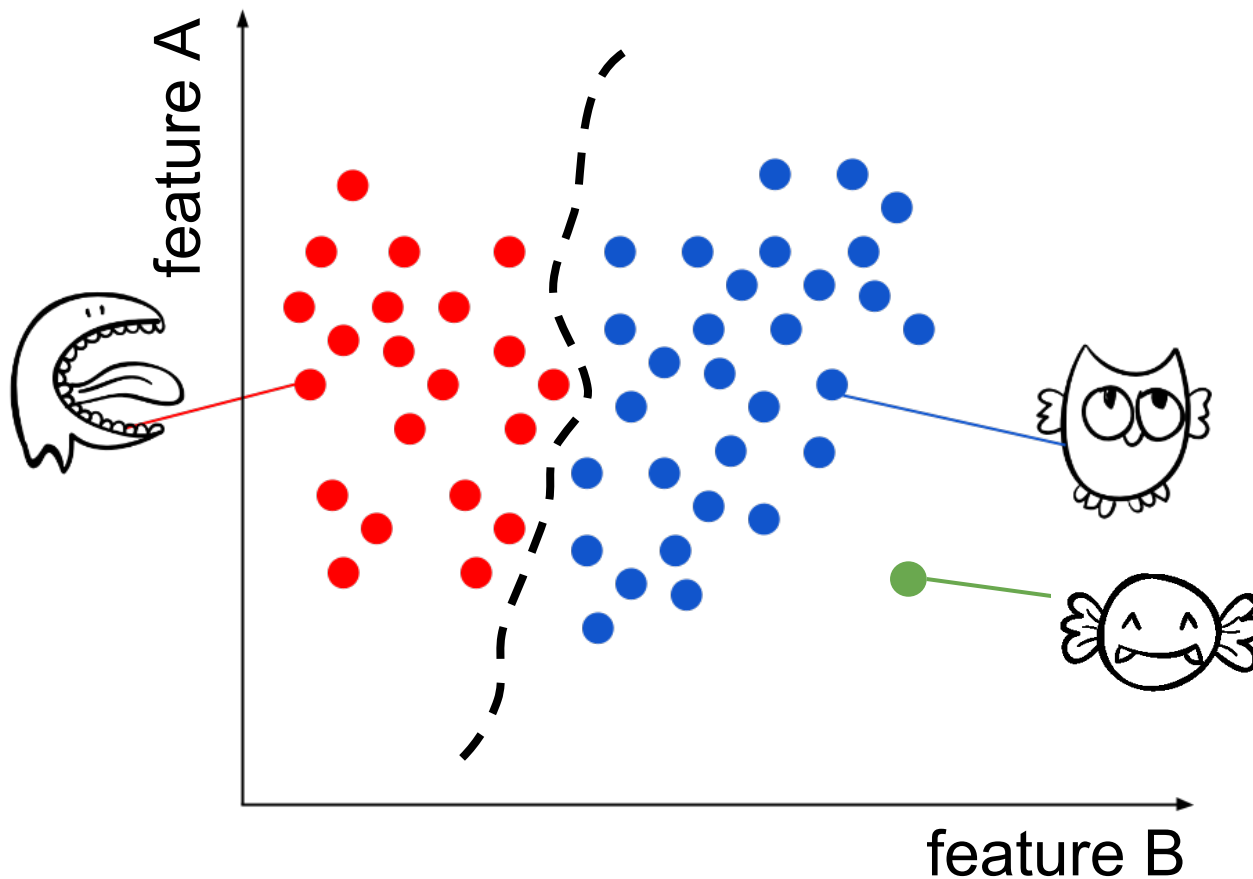
Random Forest

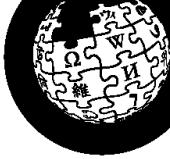
Breiman, L.: Random Forests, *Machine Learning*, Kluwer Academic Publishers, 2001, 45, 5-32



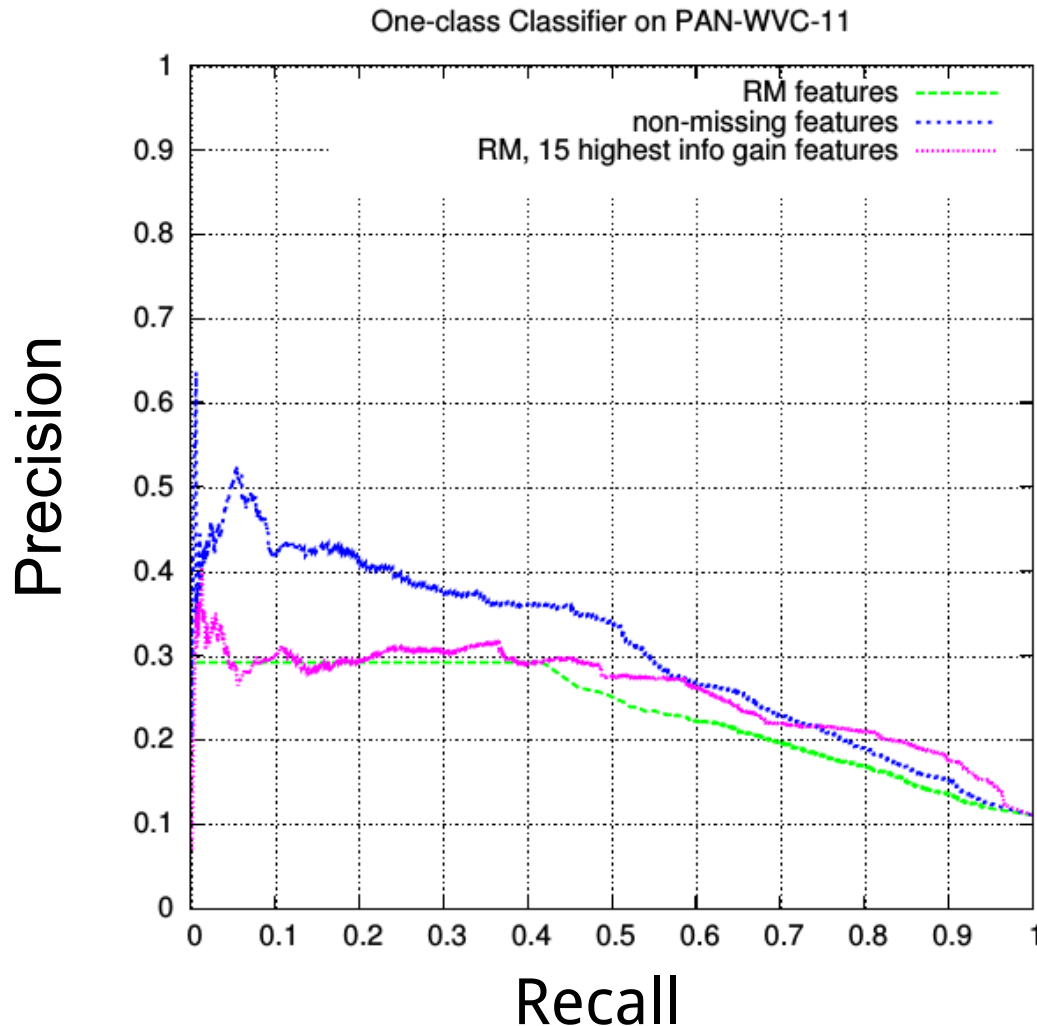
One-class Classification

training solely on vandalism samples





One-class Classification

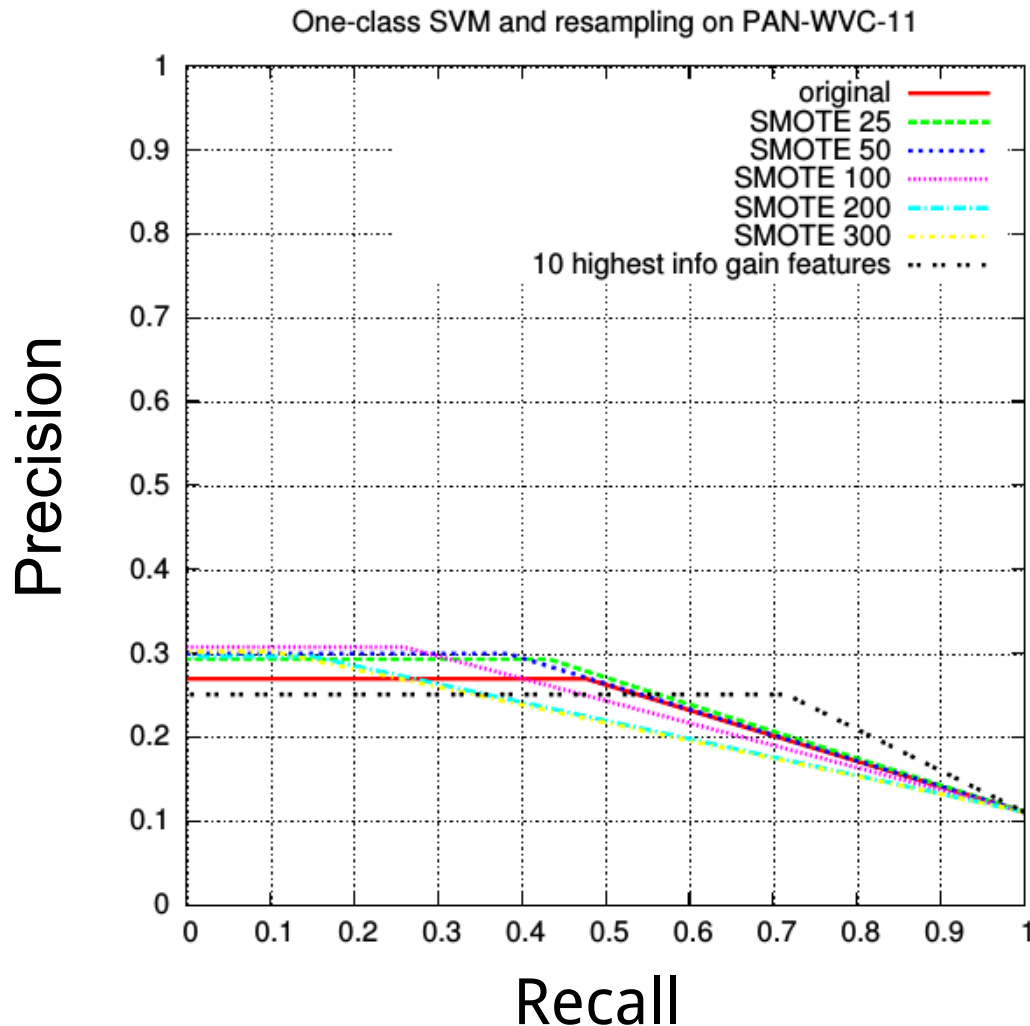


“One-class Classifier”

Hempstalk et al.: One-Class Classification by Combining Density and Class Probability Estimation, *ECML/PKDD* (1), **2008**, 505-519



One-class Classification



One-class SVM

Schölkopf, B. et al.:
Support Vector Method
for Novelty Detection,
*Advances in Neural
Information Processing
Systems 12, 1999, 582-
588*



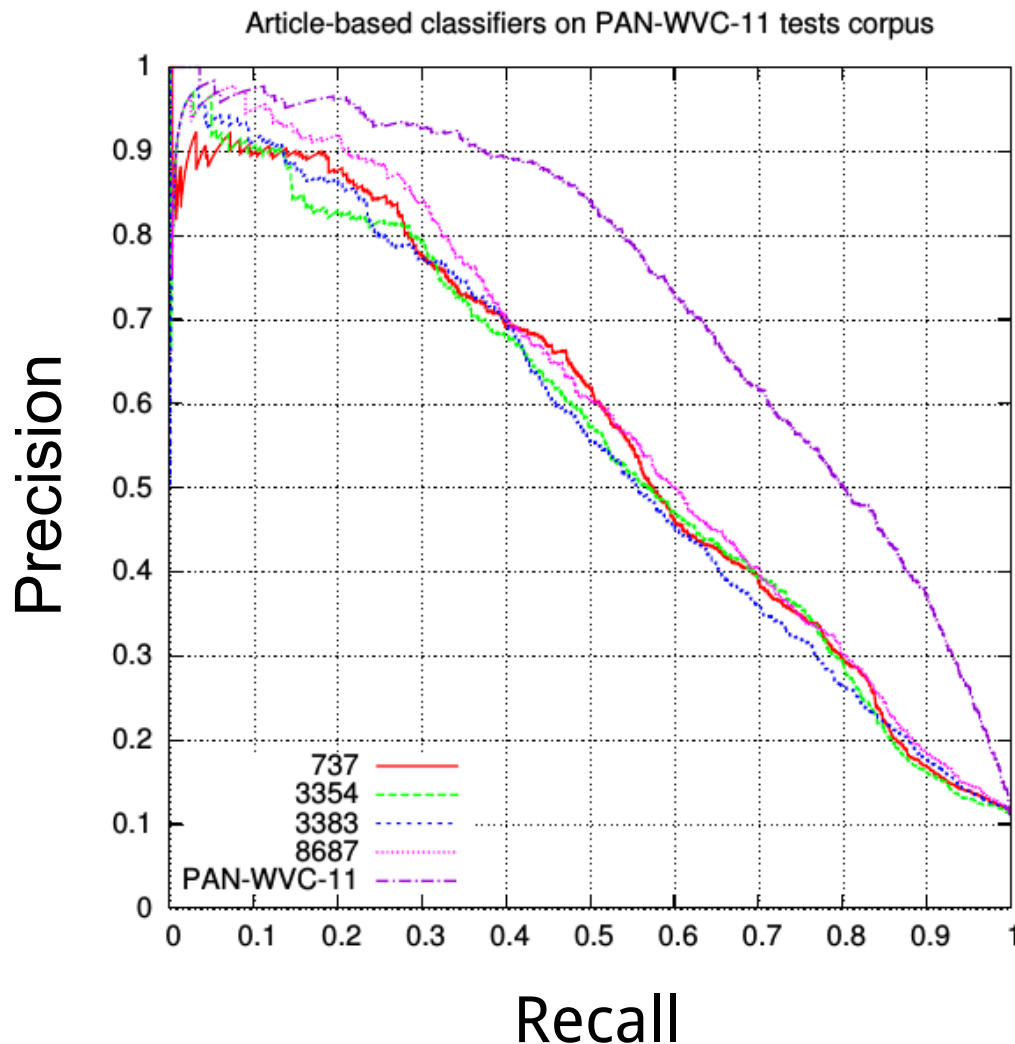
Content-based Classifiers

article-based: automatically compiled simple vandalism edits as training data

category-based: unique vandalism style in each article category



Content-based classifiers



Category:
Geographical places

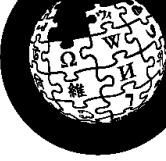


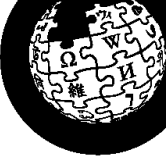
Conclusions

Dataset Resampling: no overall improvement using simple strategies

One-class classification: not suitable with the used settings

Content based classifiers: improved approaches may be promising







Precision & Recall

TP... true positive

FP... false positive

FN ... false negative

$\text{precision} = \text{TP} / (\text{TP} + \text{FP})$

$\text{recall} = \text{TP} / (\text{TP} + \text{FN})$



Detecting Vandalism

Johann Sebastian Bach: Difference between revisions



From Wikipedia, the free encyclopedia

Revision as of 00:19, 8 September 2014 (view source)

Buxtehude (talk | contribs)

← Previous edit

Revision as of 14:54, 8 September 2014 (view source)

Gapi24 (talk | contribs)

Next edit →

Line 9:

[[File:Johann Sebastian Bach.jpg|thumb|250px|<center>Portrait of Bach, aged 61, [[Elias Gottlob Haussmann|Haussmann]], 1748</center>

[[File:Johann Sebastian Bach signature.svg|right|250px|alt=signature written in ink in a flowing script]]]]

{{spaced ndash}}28 July 1750)
was a **German** composer and
musician of the [[Baroque

[[mass in B minor]], [[The Well-Tempered Clavier]], two [[Passions (Bach)|Passions]], keyboard works, and more than 300 [[List of Bach cantatas|cantatas]], of which nearly 100 cantatas have been lost to posterity.<ref name=Stauffer>{{cite web|last=Stauffer|first=George B.|title=Why Bach Moves Us|url=http://www.nybooks.com/articles/archives/2014/feb/20/why-bach-moves-us/work=The New York Review of Books|date=20 February 2014|accessdate=10 April 2014}}</ref> His music is revered for its intellectual depth, technical command, and artistic beauty.

Line 9:

[[File:Johann Sebastian Bach.jpg|thumb|250px|<center>Portrait of Bach, aged 61, [[Elias Gottlob Haussmann|Haussmann]], 1748</center>

[[File:Johann Sebastian Bach signature.svg|right|250px|alt=signature written in ink in a flowing script]]]]

{{spaced ndash}}28 July 1750)
was an **Austrian** composer
and musician of the [[Baroque

[[mass in B minor]], [[The Well-Tempered Clavier]], two [[Passions (Bach)|Passions]], keyboard works, and more than 300 [[List of Bach cantatas|cantatas]], of which nearly 100 cantatas have been lost to posterity.<ref name=Stauffer>{{cite web|last=Stauffer|first=George B.|title=Why Bach Moves Us|url=http://www.nybooks.com/articles/archives/2014/feb/20/why-bach-moves-us/work=The New York Review of Books|date=20 February 2014|accessdate=10 April 2014}}</ref> His music is revered for its intellectual depth, technical command, and artistic beauty.



References

Icons are taken from www.flaticon.com.

Mola Velasco, S. M.: Wikipedia Vandalism Detection Through Machine Learning: Feature Review and New Proposals , Lab Report for PAN at CLEF 2010 *CLEF (Notebook Papers/Labs/Workshops)*, **2010**

West, A. G. & Lee, I.: Multilingual Vandalism Detection using Language, Independent & Ex Post Facto Evidence , Notebook for PAN at CLEF 2011 *CLEF (Notebook Papers/Labs/Workshop)*, **2011**

Chawla, N. V.; Bowyer, K. W.; Hall, L. O. & Kegelmeyer, W. P.: SMOTE: Synthetic Minority Over,sampling Technique, *Journal of Artificial Intelligence Research, AI Access Foundation*, **2002**, 16, 321,357



References (cont.)

Friedman, J. et al.: Additive Logistic Regression: a Statistical View of Boosting, *The Annals of Statistics*, **2000**, 38

Breiman, L.: Random Forests, *Machine Learning*, Kluwer Academic Publishers, **2001**, 45, 5-32

Hempstalk, K.; Frank, E. & Witten, I. H.: One-Class Classification by Combining Density and Class Probability Estimation, *ECML/PKDD (1)*, **2008**, 505, 519

Schölkopf, B.; Williamson, R.; Smola, A.; Shawe-Taylor, J. & Platt, J.: Support Vector Method for Novelty Detection, *Advances in Neural Information Processing Systems 12*, **1999**, 582, 588