ORES Preparation IV

Tom Gülenman

17. Dezember 2018

Disclaimer: No guarantee for the correctness of information / explanations / sources is given.

Goals

1. Adjust crucial metrics list to match the damaging model metrics
2. Check out mail attachments
3. Check out new Confluence pages and goals
4. Research
    - Check out FAT Conference Docs
    - In what other cases than confusion matrices are those parameters explained?
    - Are there already visualizations of some of these parameters in any contexts?
    - Are there any applications, where I can filter for these parameters → visualizations or just about anything?

1 Crucial metrics: damaging-model

Metrics simple list:

!f1	
!precision	
!recall	
accuracy	✓
counts	
f1	✓
filter_rate	✓
fpr	✓
match_rate	✓
pr_auc	✓
precision	✓
rates	
recall	✓
roc_auc	✓

The metrics are the same for the damaging and itemquality models, but a few changes will be made to the explanatory parts to better fit the damaging model (TODO: ...right?). Also the structure of explanations will be changed to the following:

For each metric (if possible) there will be:

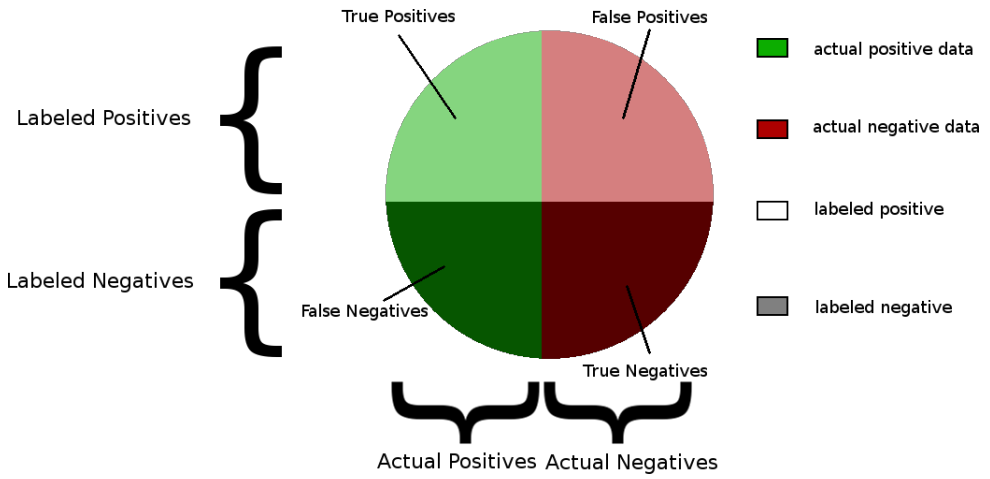
1. An intuitive explanation
2. The formula based on the **confusion matrix**
3. Its meaning based on the "**confusion circle**"
4. Its meaning based on the **loan threshold** representation by Google (Link)

Explanations: References

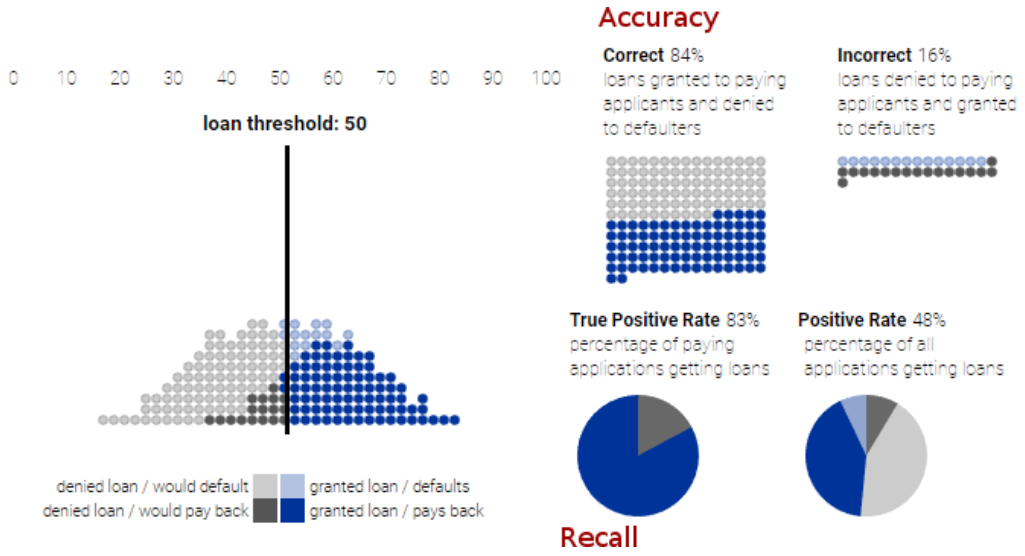
- Confusion Matrix

		Actual	
		Positive	Negative
Predicted	Positive	True Positive	False Positive
	Negative	False Negative	True Negative

- "Confusion Circle"

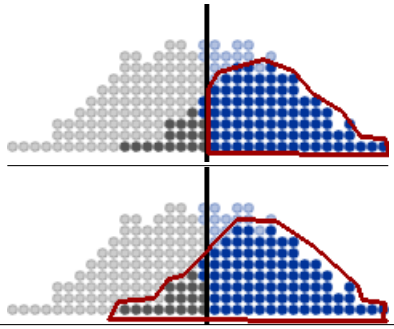


- Loan Threshold



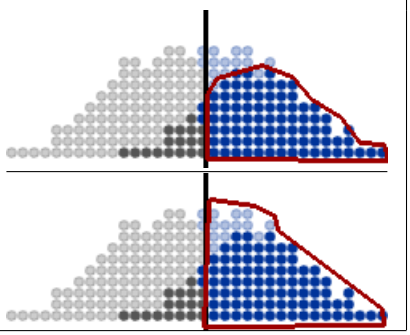
1.1 recall

1. Recall (≡ True Positive Rate) is defined as the ability of a model to find all relevant cases within the dataset.

2.	3.	4.
$= \frac{\texttt{TP}}{\texttt{TP+FN}}$	$=$ 	$=$ 

1.2 precision

1. Ability of the model to find only relevant cases within the dataset

2.	3.	4.
$= \frac{\texttt{TP}}{\texttt{TP+FP}}$	$=$ 	$=$ 

1.3 f1

1. F1-Score, the harmonic mean of recall and precision, a metric from 0 (worst) to 1 (best), used to evaluate the accuracy of a model by taking recall and precision into account

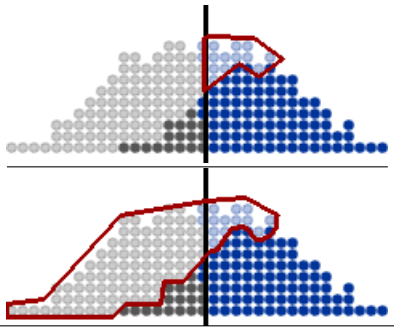
2.	3.	4.
-	-	-

$$= 2 * \frac{\texttt{precision*recall}}{\texttt{precision+recall}}$$

Compared to the simple average (of recall and precision), the harmonic mean punishes extreme values (e.g. precision 1.0 and recall 0.0 $\rightarrow$ average 0.5, but F1 = 0)

1.4 fpr

1. The false positive rate (**FPR**) is the probability of a false alarm

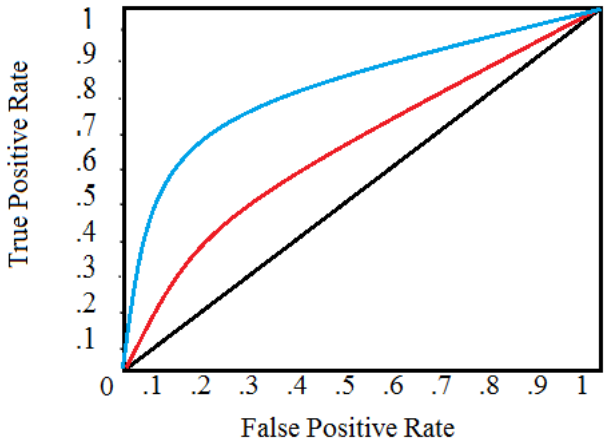
2.	3.	4.
$= \frac{\texttt{FP}}{\texttt{FP+TN}}$	$=$ 	$=$ 

1.5 roc_auc

1. The **area under** the **curve** of the **ROC**-curve, a measure between 0.5 (worthless) and 1.0 (perfect: getting no FPs), rates the ability of a model to achieve a blend of recall and precision

2.	3.	4.
-	-	-

The receiver operating characteristic (ROC) curve plots the TPR versus FPR as a function of the model's threshold for classifying a positive



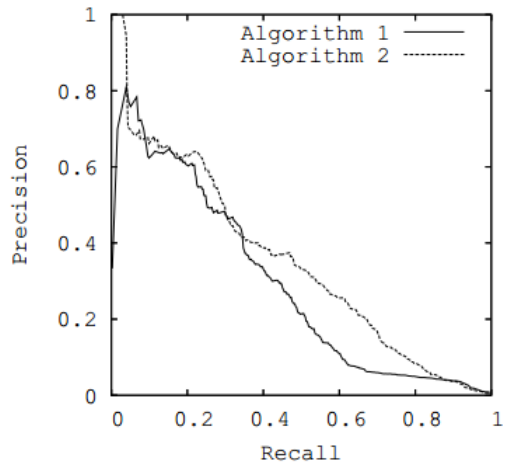
Increasing the threshold → moving up a curve (≡ model) to the top right corner, where all data is predicted as positive (threshold = 1.0) and vice versa

1.6 pr_auc

(see: link 1 and link 2)

1. The **area under** the **curve** of the **PR**-curve, same: similar objective as the **roc_auc**, but PR curves are better than ROC curves if the populations are imbalanced

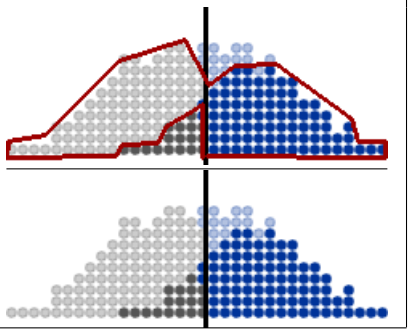
The PR-curve plots the Precision versus the Recall



Instead of the top left corner for the ROC-curve, here, we want to be in the top right corner for our classifier to be perfect

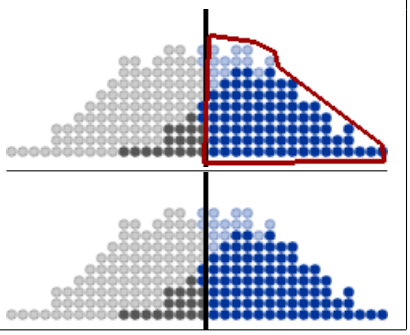
1.7 accuracy

1. Measuring the portion of correctly predicted data

2.	3.	4.
$= \frac{TP+TN}{Total}$	$=$ 	$=$ 

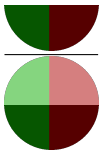
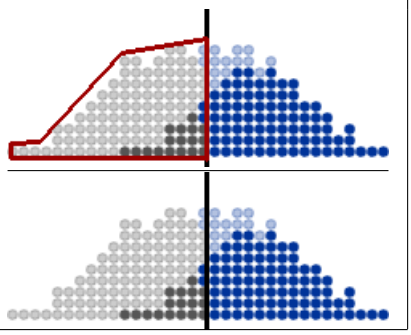
1.8 match_rate

1. The proportion of observations matched/not-matched

2.	3.	4.
$= \frac{\texttt{TP+FP}}{\texttt{Total}}$	$=$ 	$=$ 

1.9 filter_rate

1. The proportion of observations filtered/not-filtered

2.	3.	4.
$= 1 - \texttt{match\_rate} = \frac{\texttt{TN+FN}}{\texttt{Total}}$	$=$ 	$=$ 

Additional Information

- Confluence Link (Bachelor Thesis): `https://intern.hcc.mi.fu-berlin.de/confluence/pages/viewpage.action?pageId=33269659`
- Amir's mail links
  - New Filters for Edit Review Documentation Link
  - ORES API Call Link

    "Basically it asks for threshold from the API when "recall is at its maximum when precision is at least 0.995"'

Questions

- Q: Should I ask Aaron how he would like us to work together? I'm not sure how he meant it.

  A:
- Q: In what situations exactly do we want to optimize the threshold in the context of user centered threshold optimization?

  A:
- Q: VPN recommendation?

  A: