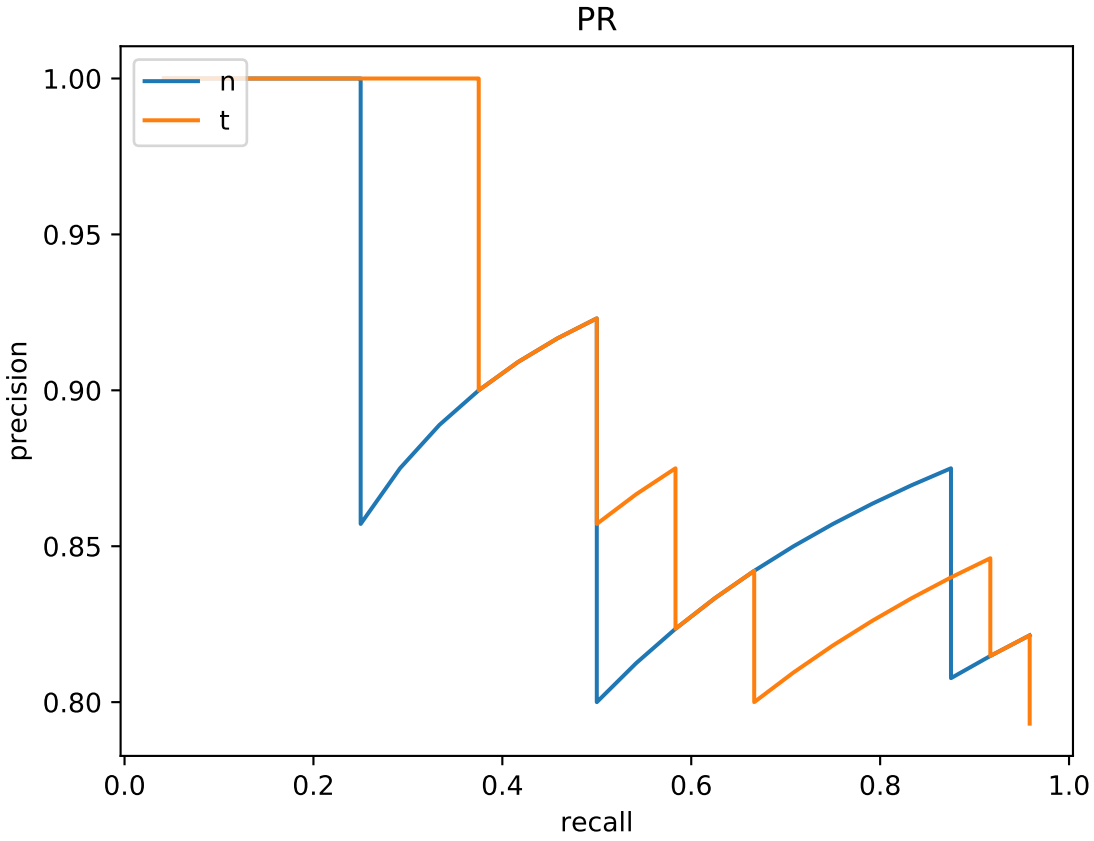
HW03

Jurijs Nazarovs

3/9/2018

Problem 2. PR curve

Image below represents PR curves for both methods: n - Naive Bayesian and t - TAN. Threshold for this plot was defined in every point.



Performance comparison

We notice that a PR curve corresponding to a TAN method is higher than the curve for NB for almost all threshold points. Based on this reason we might think that TAN perfomra better. However, fact that point (1,1) is considered as a best, and curve closes(t to this point represents a better method. In that case we might think that NB performs better, since at the point (0.87, 0.87) n curve is the closest to (1,1).

Thus, it is hard to say which one is better and maybe it is worth to conduct a statistical test (e.g. paired t-test).

Advantages and disadvantages of ROC versus PR curves

The advantave of both curves is that they allow to measure a predictive perfomance of classifier. However, there is an assumption about boinary classification.

ROC

[a] insensitive to changes in class distribution
[a] can identify optimal classification threshold for tasks with differential missclassification costs

PR

[a] well suited for tasks with lots of negative instances
[a] show the fraction of predictions that are dalse positives

Relation to our data

Important to notice that the assumption about binary classification problem is still valid in our data. So, we can apply both of them. We have 81 positive samples in our training set and 61 negatives. Which does not give any priorities to any of these curves. However, in our task it is important to uderstand which which class to assign the data. And we for sure do not want to missclassify metastases. Thus, I would chose PR curve as an assesment curve of the performance in our model.

Problem 3. Accuracy testing

To answer the question whether two methods have same performance, we are using pared t-test, with null hypothesis that test have same performance. Important to notice that paired t-test helps to understand if performance of methods is the same, but not which method perfor better.

$\delta = x_n - x_t$, where x_n - performace of Naive Bayes method and x_t - performacne of TAN. Perfomance in our test is measured by accuracy. Since we consider paired t-test, the formula for t-statistic is following:

$$t = \frac{\bar{\delta}}{\sqrt{\frac{1}{n(n-1)} \sum (\delta_i - \bar{\delta})^2}}$$

From the analysis we have: $n = 10$, $\bar{\delta} = 0.0362$, and $t = 0.0860$.

We consider the t-test with significant level of $\alpha = 0.05$. Since we are working on pairwise t-test, corresponding p-value is with parameters $\alpha/2 = 0.025$, and degrees of freedom $n - 1 = 9$.

P-value corresponing to observed t-statistics $t = 0.0860$ is 0.9333493 (do not forget to multiply by two, since it is 2-sided test). Which is much higher than 0.05, and consequenty we cannot reject null hypothesis that methods have the same performance.

That is, we conclude that method have the same performance.