MiniHW01

April 11, 2020

1 MiniHW01 - Data Mining

1.0.1 *Mohammad Doosti Lakhani - 98722278*

This notebook consists of comparison of different model accuracy measurements:

1. Definitions

    0. Confusion Matrix
    1. Accuracy
    2. Precision
    3. Recall
    4. F1
    5. ROC Graph
    6. AUC Score

2. ROC vs F-measure
3. Relation Between Alpha and Beta in F-measure

1.1 1. Definitions

0. Confusion Matrix
1. Accuracy
2. Precision
3. Recall
4. F1
5. ROC Graph
6. AUC Score

1.1.1 1.0 Confusion Matrix

Becaus all other metircs can be calculated, we first introduce confusion matrix.

Our confusion matrix is like this image:

Where:

- `TN = True Negative`
- `FP = False Positive`
- `FN = False Negative`
- `TP = True Positive`

		Predicted Class	
		No	Yes
Observed Class	No	TN	FP
	Yes	FN	TP

confusion matrix

1.1.2   1.1 Accuracy

Accuracy is how our model can say class 1 is 1 and 0 is 0 regarding all given examples.

Give confusion matrix,

`Accuracy = (TN+TP) / (TN+FP+FN+TP)`

So obviously more is better.

1.1.3   1.2 Precision

Based on confusion matrix,

`precision = TP / (FP+TP)`

Precision is the ratio of correctly predicted positive observations to the total predicted positive observations. High precision relates to the low false positive rate. In our case, it says between all those we labeled as positive, how many of the are actually positive.

1.1.4   1.3 Recall

Based on confusion matrix,

`recall = TP / (TP+FN)`

Recall is the ratio of correctly predicted positive observations to the all observations in actual positive class. In our case, Between all label is positive, how many we said positive.

1.1.5   1.4 F1 Score

Based on available values,

`F1 = 2*(Recall * Precision) / (Recall + Precision)` F1 score is weighted average of precision and recall, so it is very usefull when the cost of error of classes are different from each other (say disease detection).

Here is a graph to alleviate problem:

1.1.6   1.5 ROC-AUC

Let's talk about ROC using AUC value as these are seamlessly interchangable. This graph may help us to understand it better.

AUC is the area under ROC graph. More AUC says our model is more powerful to distinguish between the two classes. AUC = 0.5 says our model almost failed to differentiate. AUC = 1 is perfect model.

For plotting we need two values: 1. `TPR = TP / (TP + FN)` 2. `FPR = FP / (FP + TN)`
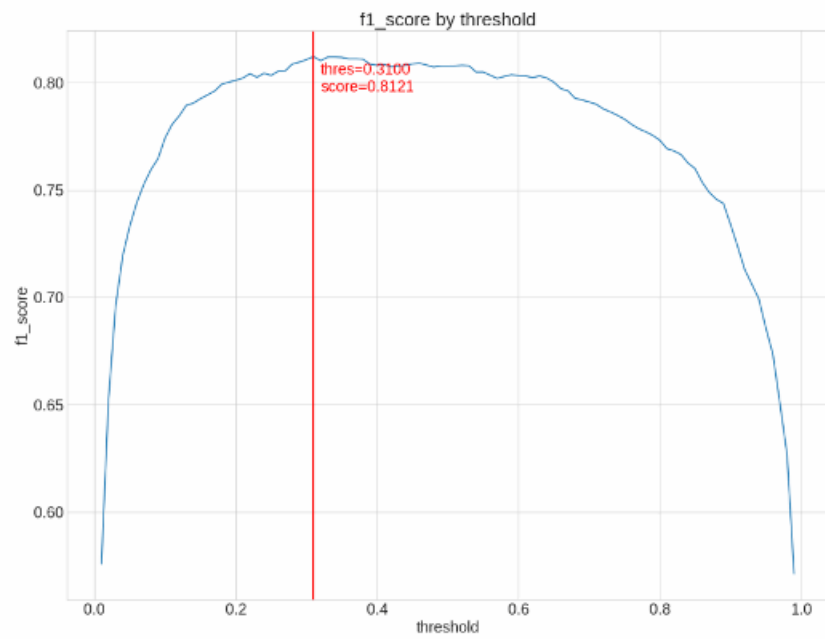
f1 by threshold graph
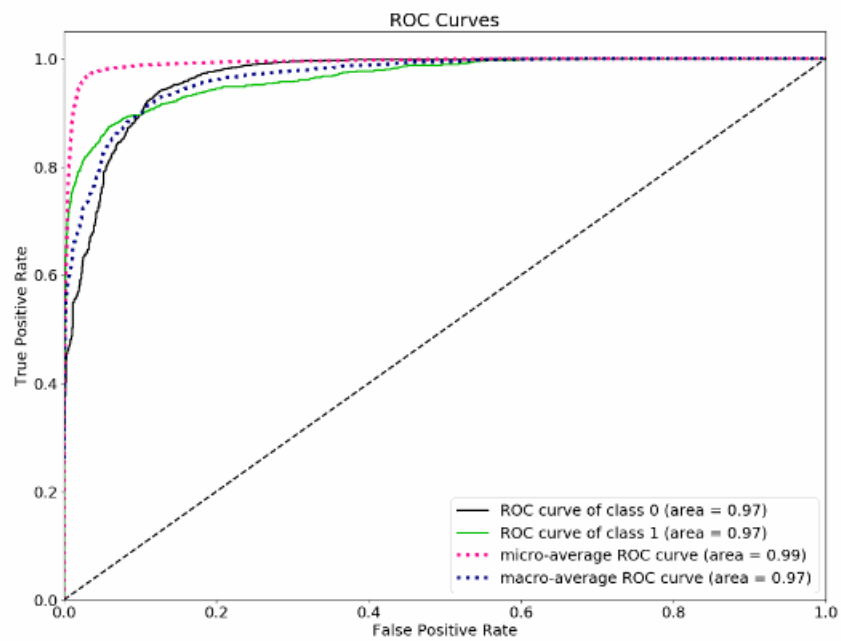
roc auc graph

Now we compare these values regarding different classification thresholds. Increasing threshold label more entries as positive so it increases both `FP` and `TP`. AUC is the aggregated measure of performance among all possible thresholds.

1.1.7   2 F-measure vs. ROC

Simply put, F-measure enable us to give importance to recall or precision based on our needs. It is important to remember that F1 score is calculated from Precision and Recall which, in turn, are calculated on the predicted classes (not prediction scores)

Another point that need to be mentioned is that F-measure focuses on positive class which enable us to have a clear understanding of the model's behavior regarding specific problem as the numbers are easily interpreted.

The challenge in F-measure is to find the best threshold (beta) value to enable approximately best measure over recall and precision. Otherwise the interpretation may be baised toward particular metric. This also is a advantage where we can use it in imbalance datasets to focus on the anomaly (or rare case) rather than averaging over all thresholds which is what ROC does.

It is a chart that visualizes the tradeoff between true positive rate (TPR) and false positive rate (FPR). Basically, for every threshold, we calculate TPR and FPR and plot it on one chart. Of course, the higher TPR and the lower FPR is for each threshold the better and so classifiers that have curves that are more top-left-side are better. The best ROC is the rectangle which has perfect discrimination power for any possible threshold.

One case of imbalance data, ROC is not a good measure as it averages over all thresholds as this will eliminate the effect of anomaly (rare) class. An interpretation for this regarding ROC curve would be false positive rate for highly imbalanced datasets is pulled down due to a large number of true negatives. As ROC is a measure between true negative and true positive, we care about this values more, ROC is the a better measure as only focuses on this rates.

1.2   3 Relation Between Alpha and Beta in F-measure

$$F = \frac{1}{\left(\frac{\alpha}{P} + \frac{1-\alpha}{R}\right)}$$

P: Precision
R: Recall

$$\frac{\partial F}{\partial P} = \frac{\alpha}{\left(\frac{\alpha}{P} + \frac{1-\alpha}{R}\right)^2 P^2} \quad \& \quad \frac{\partial F}{\partial R} = \frac{1-\alpha}{\left(\frac{\alpha}{P} + \frac{1-\alpha}{R}\right)^2 R^2}$$

$$\frac{\partial F}{\partial P} = \frac{\partial F}{\partial R} \quad \text{: giving same importance}$$
to $\underline{R}$ and $\underline{P}$

$$\Rightarrow \frac{\alpha}{P^2} = \frac{1-\alpha}{R^2} \quad \Rightarrow \quad \frac{R}{P} = \sqrt{\frac{1-\alpha}{\alpha}}$$

$$\beta = \sqrt{\frac{1-\alpha}{\alpha}} \quad \Rightarrow \quad \alpha = \frac{1}{\beta^2 + 1} \quad \textcircled{I}$$

$$1 - \alpha = \frac{\beta^2}{\beta^2 + 1} \quad \textcircled{II}$$

$$\textcircled{I} \text{ and } \textcircled{II} \Rightarrow F = \frac{1}{\left(\frac{1}{\beta^2 + 1} \times \frac{1}{P} + \frac{\beta^2}{\beta^2 + 1} \times \frac{1}{R}\right)}$$

alpha beta relation in f-measure