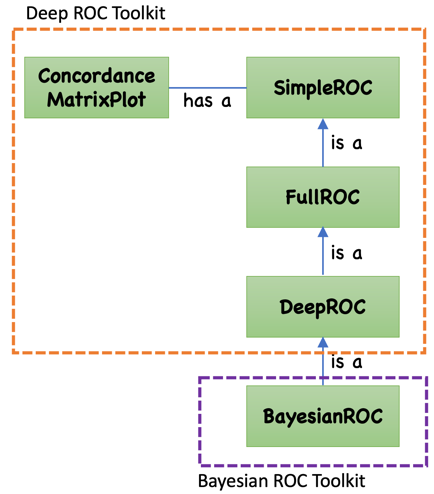
# Bayesian ROC toolkit

André Carrington

The Bayesian ROC toolkit is object-oriented code developed for two papers: “The ROC Diagonal is not Layperson's Chance: a New Baseline Shows the Useful Area” [1] and a manuscript to be submitted. It is object-oriented code written in Python 3.8.

It uses the Deep ROC toolkit [2-5] which includes measures from several sources [6-9]. The BayesianROC class is a subclass of DeepROC (Figure 1) that offers additional functionality (Table 1).

Binary chance is a specific choice of a Bayesian prior, so it does not have its own class for functionality, but it does have a unique file and function for testing.

Figure Class Hierarchy

DeepROC implements pre-test and post-test measures, for the whole ROC plot or for groups, where groups are contiguous risk groups, non-contiguous demographic groups, or any selection of instances. Groups may be overlapping and not cover the whole ROC curve, or they may be mutually exclusive and perfectly cover the ROC curve.

Table 1. Additional functionality in BayesianROC as a subclass

|  |  |  |
| --- | --- | --- |
| **BayesianROC** |  |  |
| Functionality | Short name | Description |
| Plots | Chance ROC plot  for each group | An ROC plot with partial area measures and a binary chance baseline,  for a contiguous group of risk,  in a range of FPR, TPR, threshold, or percentile threshold. |
|  | Chance mean ROC plot  for each group | A mean ROC plot with partial area measures and a binary chance baseline or Bayesian prior baseline,  for acontiguous group of risk,  ina range of FPR or TPR. |
| Pre-Test Measures  (continuous) | Partial area measures  for an ROC plot, or for a mean ROC plot  ( AUCiΩ, AUCinΩ )  AUCiπ, AUCinπ  pAUCπ+, pAUCπ-  pAUCxπ+, pAUCxπ- | Partial area measures with  for a contiguous group of risk  in a range of FPR, TPR, threshold, or percentile threshold.  Relative to a Bayesian prior, π which may be defined as binary chance π=Ω:  Concordant partial AUC, normalized (group AUC).  Partial AUC (vertical) positive, negative  Horizontal partial AUC positive, negative |

## Creating an ROC object

Method 1.

Create an object with ROC data as classification scores (probabilities) and labels.

from bayesianroc.BayesianROC import BayesianROC

roc = BayesianROC(predicted\_scores=scores, labels=labels, poslabel=1, BayesianPrior=(0.5, 0.5), costs=costs, quiet=quiet)

Method 2.

Create an empty object, then set the ROC data with classification scores (probabilities) and labels.

from bayesianroc.BayesianROC import BayesianROC

roc = BayesianROC(predicted\_scores=None, labels=None, poslabel=None, BayesianPrior=None, costs=None, quiet=quiet)

roc.set\_scores\_labels(predicted\_scores=scores, labels=labels, poslabel=1)

roc.set\_costs(costs=costs)

roc.set\_BayesianPrior(BayesianPrior=(0.5, 0.5))

Method 3.

Create an empty object, then set the ROC data with known points in ROC space: (fpr, tpr).

from bayesianroc.BayesianROC import BayesianROC

roc = BayesianROC(predicted\_scores=None, labels=None, poslabel=None, BayesianPrior=None, costs=None, quiet=quiet)

roc.set\_fpr\_tpr(fpr=fpr, tpr=tpr)

roc.set\_costs(costs=costs)

roc.set\_BayesianPrior(BayesianPrior=(0.5, 0.5))

Now use it…

roc.plot(plotTitle, saveFileName=’plot.png’, showPlot=True)

auc = roc.getAUC()

c = roc.getC()

## Groups

With BayesianROC, you can define one group (like a region of interest) or as many groups as you like to measure in an ROC plot. The groups can be overlapping or not. Group measures can be compared to each other and to the whole.

## Example Use of the Classes in Test Functions

TestChanceROC.py, creates a BayesianROC object and performs tests for the Chance ROC paper.

It currently uses the Wisconsin Diagnostic Breast Cancer (wdbc) data in data.csv.

A number of questions are asked interactively – **please generally hit enter for the default, except for a few of them as I explain next.** The questions are old, unchanged and probably need better explanation (and some are not actually used in the code yet). The only settings I changed in testing were:

* different costs (e.g., of FN, FP)
* sometimes setting costs as ‘rates’ instead of the default ‘individuals’
* the region of interest (ROI) is hardcoded on lines 283-284 as groupAxis and groups

The answer to prevalence questions may not fully/properly propagate. There are also 3 parameters in the file that can be changed re the wdbc data:

dropSizeTexture = False

dropSize = False

dropShape = False

## References

1. Carrington AM, Fieguth PW, Mayr F, James N, Holzinger A, Pickering JW, Aviv RI. The ROC Diagonal is not Layperson's Chance: a New Baseline Shows the Useful Area.Submitted to *International IFIP Cross-Domain Conference for Machine Learning and Knowledge Extraction*. Springer. CD-MAKE-25
2. Carrington AM, Manuel DG, Fieguth PW, Ramsay T, Osmani V, Wernly B, Bennett C, Hawken S, Magwood O, Sheikh Y, McInnes M, Holzinger A. Deep ROC analysis and AUC as balanced average accuracy for improved classifier selection, audit and explanation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Early Access, January 25, 2022. doi:10.1109/TPAMI.2022.3145392
3. http://deeproc.org
4. https://pypi.org/project/deeproc/
5. https://github.com/Big-Life-Lab/deepROC
6. Carrington AM, Fieguth PW, Qazi H, Holzinger A, Chen HH, Mayr F and Manuel DG. A new concordant partial AUC and partial C statistic for imbalanced data in the evaluation of machine learning algorithms, *BMC Medical Informatics and Decision Making* 20, 4 (2020) doi:10.1186/s12911-019-1014-6