ROC Curves

FCA Collin, Ph.D.

Thursday, April 15, 2021

Contents

2021-01-11, FCA Collin (update 2021-01-12)

Presentation

The Receiver Operating Characteristic (ROC) is a general representation of a binary classifier; it accounts for:

- the sensitivity (the proportion of real positive case detected),
- the false positive rate (1 - sensitivity),
- the general performance via the area under the ROC curve (AUC).

The classifier for the example identifies tumour tissues (yes/no) in the case of lung cancer (LC) or eventually identifies the Squamous Cell Carcinoma (SCC) histological subtype. The dataset used for the graphic requires at least columns for the sensitivity, the false-positive rate and a decision rule represented by a threshold proportion, in that case varying from 0 to 1. Additionally, information about the classifier itself can include the model identification (LC, SCC), as well as the AUC (+ confidence interval estimation) as a overall evaluation of the models.

Data

Lets dtaplot being as example dataset such as:

Table 1: First row of the dataset behind the ROC curves.

Threshold	Sensitivity	Specificity	FalseAlarm	ntree	auc.ci	auc	Diag
Inf	0.0000000	1	0	10000	0.925-0.984	0.954	2) SCC
Inf	0.0000000	1	0	10000	0.938-0.989	0.964	1) LC
0.96445	0.0097087	1	0	10000	0.925-0.984	0.954	2) SCC
0.95860	0.0194175	1	0	10000	0.925-0.984	0.954	2) SCC
0.95375	0.0291262	1	0	10000	0.925-0.984	0.954	2) SCC
0.95285	0.0388350	1	0	10000	0.925-0.984	0.954	2) SCC

Some decision rules (threshold) were of interest:

- Min. error is the threshold for which the classifier overall error was at its minimum.
- Sens. 90%: has its better not to miss true positive patients, the threshold may be determined so as to catch 90% of true positive patient (sensibility), eventually at a cost in terms of overall error as it then automatically increase the false-alarm rate.

Lets threshold being the supplementary data characterising this two decision rules:

Table 2: 'Threshold, the decision rules.

Threshold	Sensitivity	Specificity	FalseAlarm	ntree	auc.ci	auc	Diag	target	Rational
0.48965	0.9368421	0.9375000	0.0625000	10000	0.938-0.989	0.964	1) LC	0.0103501	Min. error
0.49350	0.8640777	0.9813953	0.0186047	10000	0.925-0.984	0.954	2) SCC	0.0065001	Min. error
0.69835	0.9000000	0.9687500	0.0312500	10000	0.938-0.989	0.964	1) LC	0.0000000	Sens. 90%
0.35290	0.9029126	0.9441860	0.0558140	10000	0.925-0.984	0.954	2) SCC	0.0029126	Sens. 90%

Graphic

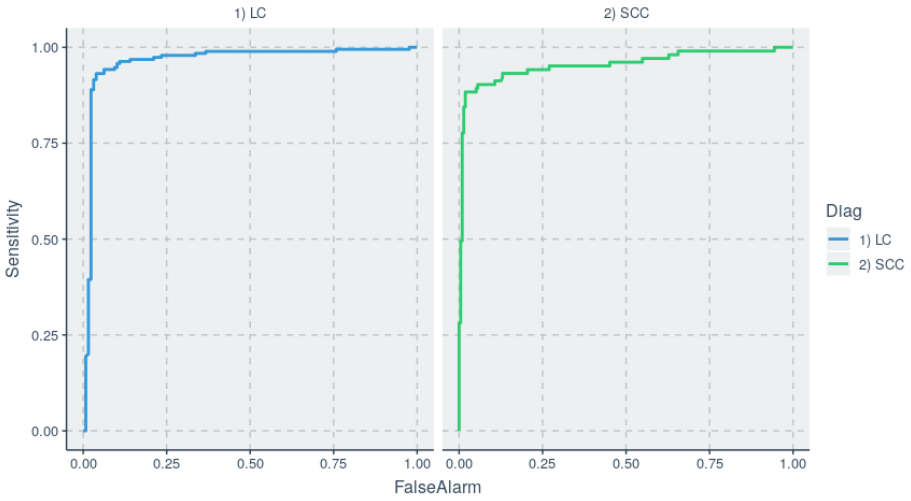
The example is based on the package ggplot2, plus the optional ggthemr which provides graphical themes, for instance the theme flat .

```
library(ggplot2)
if(require(ggthemr)) ggthemr::ggthemr("flat")
```

Basic

The minimal ROC representation is simply a line plot representing the sensitivity as a function of the false alarm rate, for one or the other model.

```
{
  ggplot(
    data = dtaplot,
    mapping = aes(
      x = FalseAlarm,
      y = Sensitivity,
      color = Diag
    )
    ) + geom_line(
    lwd = 1
    ) + facet_grid(
    . ~ Diag
  )
}
```



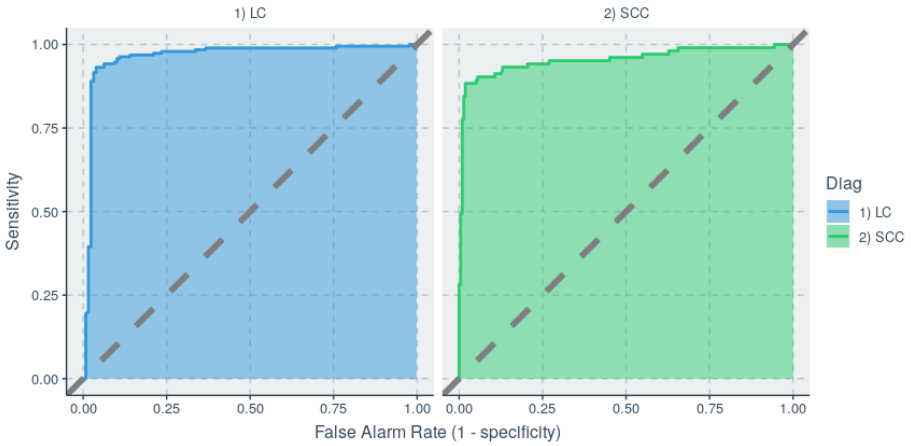
Aesthetic improvement

Aesthetic can help improving the reading:

- the area under the curve is of interest, therefore it can be filled.

- the bisector delimits a model performing as good as a decision made flipping a coin (the reference model).
- sensibility and false-alarm rate are define from 0 to 1, the length of this two axis should equal.

```
{
  ggplot(
    data = dtaplot,
    mapping = aes(
      x = FalseAlarm,
      y = Sensitivity,
      ymax = Sensitivity,
       fill = Diag,
       color = Diag
    )
    ) + geom_ribbon(
    ymin = 0, alpha = .5, color = NA
    ) + geom_line(
    lwd = 1
    ) + geom_abline(
    slope = 1, intercept = 0, col = "gray50", lwd = 2, lty = 2
    ) + xlab(
    "False Alarm Rate (1 - specificity)"
    ) + facet_grid(
    . ~ Diag
    ) + theme(
    asp = 1
  )
}
```

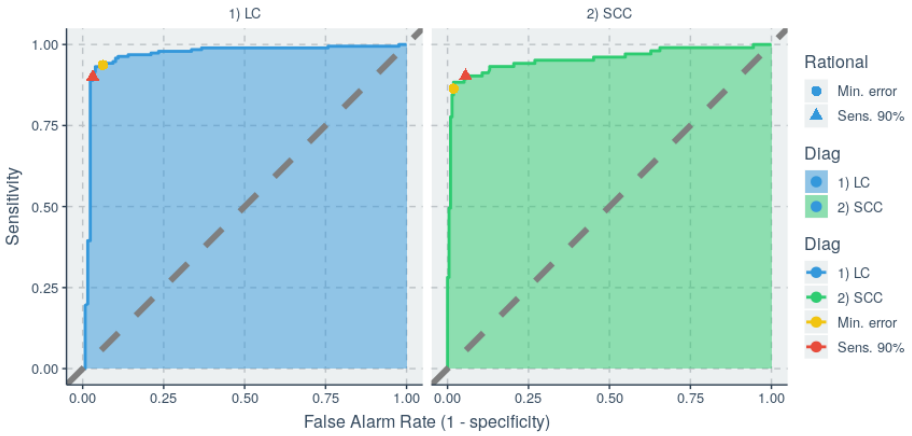


Other data for more annotations

The use of the model for diagnostic demands to define a threshold, various rational can be used, in the example two thresholds were defined: the minimal error, the 90% detection of positive case. Points can identify this threshold and performance on the ROC curve.

```
{
  ggplot(
    data = dtaplot,
    mapping = aes(
      x = FalseAlarm,
      y = Sensitivity,
      ymax = Sensitivity,
      fill = Diag,
      color = Diag
    )
    ) + geom_ribbon(
    ymin = 0, alpha = .5, color = NA
    ) + geom_line(
    lwd = 1
    ) + geom_abline(
    slope = 1, intercept = 0, col = "gray50", lwd = 2, lty = 2
    ) + xlab(
    "False Alarm Rate (1 - specificity)"
    ) + facet_grid(
    . ~ Diag
    ) + theme(
```

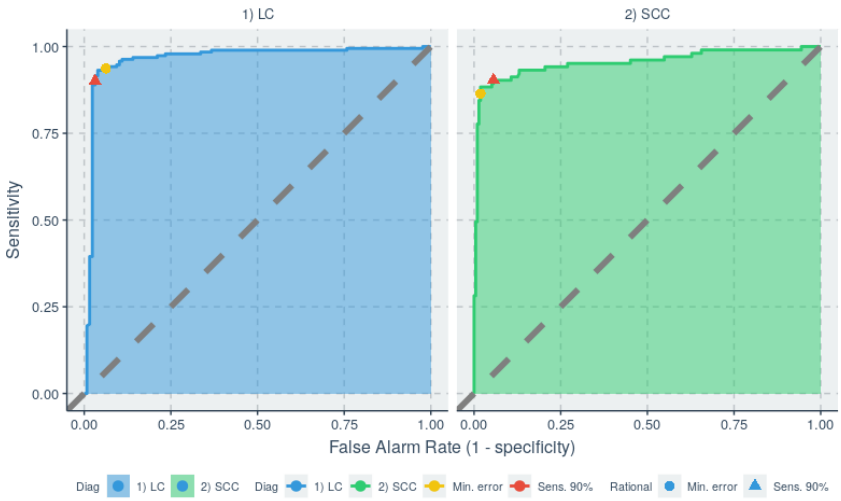
```
    asp = 1
    ) + geom_point(
    data = threshold,
    mapping = aes(shape = Rational, colour = Rational),
    size = 3
  )
}
```



The plot can be further personalised manipulating the theme locally to address for instance the positioning of the legend and other settings.

```
{
  ggplot(
    data = dtaplot,
    mapping = aes(
      x = FalseAlarm,
      y = Sensitivity,
      ymax = Sensitivity,
      fill = Diag,
      color = Diag
    )
    ) + geom_ribbon(
    ymin = 0, alpha = .5, color = NA
    ) + geom_line(
    lwd = 1
    ) + geom_abline(
    slope = 1, intercept = 0, col = "gray50", lwd = 2, lty = 2
    ) + xlab(
```

```
    "False Alarm Rate (1 - specificity)"
    ) + facet_grid(
    . ~ Diag
    ) + geom_point(
    data = threshold,
    mapping = aes(shape = Rational, colour = Rational),
    size = 3
    ) + theme(
    asp = 1, legend.position = "bottom",
    legend.text = element_text(size = 8),
    legend.title = element_text(size = 8),
    legend.background = element_rect(fill = "transparent"),
    plot.background = element_rect(fill = alpha("white", .5), colour = 'white')
  )
}
```

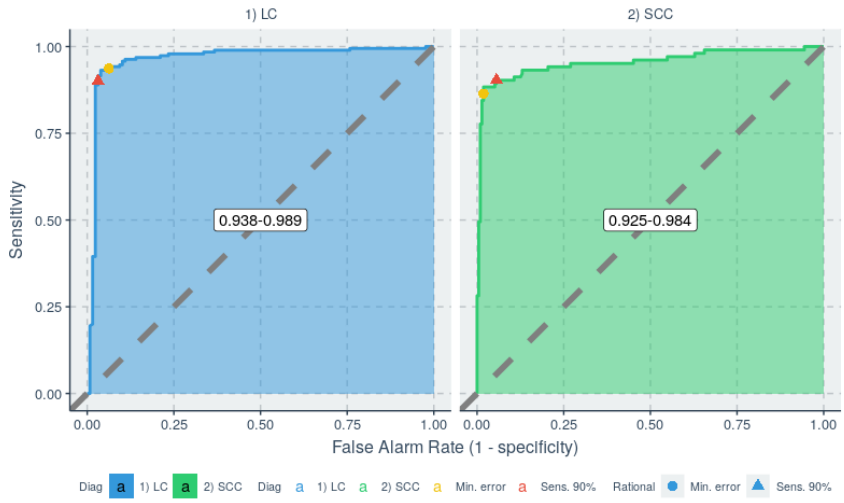


Final

Finally, as a last piece of annotation, the AUC given with its confidence interval may help for further comparison.

```
{
  ggplot(
    data = dtaplot,
    mapping = aes(
      x = FalseAlarm,
      y = Sensitivity,
```

```
      ymax = Sensitivity,
      fill = Diag,
      color = Diag
    )
    ) + geom_ribbon(
    ymin = 0, alpha = .5, color = NA
    ) + geom_line(
    lwd = 1
    ) + geom_abline(
    slope = 1, intercept = 0, col = "gray50", lwd = 2, lty = 2
    ) + xlab(
    "False Alarm Rate (1 - specificity)"
    ) + facet_grid(
    . ~ Diag
    ) + geom_point(
    data = threshold,
    mapping = aes(shape = Rational, colour = Rational),
    size = 3
    ) + theme(
    asp = 1, legend.position = "bottom",
    legend.text = element_text(size = 8),
    legend.title = element_text(size = 8),
    legend.background = element_rect(fill = "transparent"),
    plot.background = element_rect(fill = alpha("white", .5), colour = 'white')
    ) + geom_label(
    data = aggregate(auc.ci ~ Diag + ntree, data = dtaplot, unique),
    mapping = aes(label = auc.ci, ymax = NULL, fill = NULL, color = NULL ,
      x = .5, y = .5
    )
  )
}
```




```
sessionInfo()
```

```
## R version 4.0.4 (2021-02-15)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Debian GNU/Linux 10 (buster)
##
## Matrix products: default
## BLAS:   /usr/lib/x86_64-linux-gnu/openblas/libblas.so.3
## LAPACK: /usr/lib/x86_64-linux-gnu/libopenblasp-r0.3.5.so
##
## locale:
##  [1] LC_CTYPE=en_GB.UTF-8       LC_NUMERIC=C
##  [3] LC_TIME=en_GB.UTF-8        LC_COLLATE=en_GB.UTF-8
##  [5] LC_MONETARY=en_GB.UTF-8    LC_MESSAGES=en_GB.UTF-8
##  [7] LC_PAPER=en_GB.UTF-8       LC_NAME=C
##  [9] LC_ADDRESS=C               LC_TELEPHONE=C
## [11] LC_MEASUREMENT=en_GB.UTF-8 LC_IDENTIFICATION=C
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets
methods   base
##
## other attached packages:
## [1] ggthemr_1.1.0 ggplot2_3.3.3
##
## loaded via a namespace (and not attached):
```

```
##  [1] pillar_1.5.1      compiler_4.0.4    highr_0.8
tools_4.0.4
##  [5] digest_0.6.27     evaluate_0.14     lifecycle_1.0.0
tibble_3.1.0
##  [9] gtable_0.3.0      pkgconfig_2.0.3   rlang_0.4.10
DBI_1.1.1
## [13] yaml_2.2.1        xfun_0.22         withr_2.4.1
stringr_1.4.0
## [17] dplyr_1.0.5       knitr_1.31        generics_0.1.0
vctrs_0.3.7
## [21] grid_4.0.4        tidyselect_1.1.0  glue_1.4.2
R6_2.5.0
## [25] fansi_0.4.2       rmarkdown_2.6     farver_2.1.0
purrr_0.3.4
## [29] magrittr_2.0.1    scales_1.1.1      ellipsis_0.3.1
htmltools_0.5.1.1
## [33] assertthat_0.2.1  colorspace_2.0-0  labeling_0.4.2
utf8_1.2.1
## [37] stringi_1.5.3     munsell_0.5.0     crayon_1.4.1
```