**Key points**

* + A very common approach to evaluating accuracy and F1-score is to compare them graphically by plotting both. A widely used plot that does this is the **receiver operating characteristic (ROC) curve**. The ROC curve plots sensitivity (TPR) versus 1 - specificity or the false positive rate (FPR).
  + However, ROC curves have one weakness and it is that neither of the measures plotted depend on prevalence. In cases in which prevalence matters, we may instead make a **precision-recall plot**, which has a similar idea with ROC curve.

**Code**

Note: your results and plots may be slightly different.

p <- 0.9

n <- length(test\_index)

y\_hat <- sample(c("Male", "Female"), n, replace = TRUE, prob=c(p, 1-p)) %>%

factor(levels = levels(test\_set$sex))

mean(y\_hat == test\_set$sex)

# ROC curve

probs <- seq(0, 1, length.out = 10)

guessing <- map\_df(probs, function(p){

y\_hat <-

sample(c("Male", "Female"), n, replace = TRUE, prob=c(p, 1-p)) %>%

factor(levels = c("Female", "Male"))

list(method = "Guessing",

FPR = 1 - specificity(y\_hat, test\_set$sex),

TPR = sensitivity(y\_hat, test\_set$sex))

})

guessing %>% qplot(FPR, TPR, data =., xlab = "1 - Specificity", ylab = "Sensitivity")

cutoffs <- c(50, seq(60, 75), 80)

height\_cutoff <- map\_df(cutoffs, function(x){

y\_hat <- ifelse(test\_set$height > x, "Male", "Female") %>%

factor(levels = c("Female", "Male"))

list(method = "Height cutoff",

FPR = 1-specificity(y\_hat, test\_set$sex),

TPR = sensitivity(y\_hat, test\_set$sex))

})

# plot both curves together

bind\_rows(guessing, height\_cutoff) %>%

ggplot(aes(FPR, TPR, color = method)) +

geom\_line() +

geom\_point() +

xlab("1 - Specificity") +

ylab("Sensitivity")

library(ggrepel)

map\_df(cutoffs, function(x){

y\_hat <- ifelse(test\_set$height > x, "Male", "Female") %>%

factor(levels = c("Female", "Male"))

list(method = "Height cutoff",

cutoff = x,

FPR = 1-specificity(y\_hat, test\_set$sex),

TPR = sensitivity(y\_hat, test\_set$sex))

}) %>%

ggplot(aes(FPR, TPR, label = cutoff)) +

geom\_line() +

geom\_point() +

geom\_text\_repel(nudge\_x = 0.01, nudge\_y = -0.01)

# plot precision against recall

guessing <- map\_df(probs, function(p){

y\_hat <- sample(c("Male", "Female"), length(test\_index),

replace = TRUE, prob=c(p, 1-p)) %>%

factor(levels = c("Female", "Male"))

list(method = "Guess",

recall = sensitivity(y\_hat, test\_set$sex),

precision = precision(y\_hat, test\_set$sex))

})

height\_cutoff <- map\_df(cutoffs, function(x){

y\_hat <- ifelse(test\_set$height > x, "Male", "Female") %>%

factor(levels = c("Female", "Male"))

list(method = "Height cutoff",

recall = sensitivity(y\_hat, test\_set$sex),

precision = precision(y\_hat, test\_set$sex))

})

bind\_rows(guessing, height\_cutoff) %>%

ggplot(aes(recall, precision, color = method)) +

geom\_line() +

geom\_point()

guessing <- map\_df(probs, function(p){

y\_hat <- sample(c("Male", "Female"), length(test\_index), replace = TRUE,

prob=c(p, 1-p)) %>%

factor(levels = c("Male", "Female"))

list(method = "Guess",

recall = sensitivity(y\_hat, relevel(test\_set$sex, "Male", "Female")),

precision = precision(y\_hat, relevel(test\_set$sex, "Male", "Female")))

})

height\_cutoff <- map\_df(cutoffs, function(x){

y\_hat <- ifelse(test\_set$height > x, "Male", "Female") %>%

factor(levels = c("Male", "Female"))

list(method = "Height cutoff",

recall = sensitivity(y\_hat, relevel(test\_set$sex, "Male", "Female")),

precision = precision(y\_hat, relevel(test\_set$sex, "Male", "Female")))

})

bind\_rows(guessing, height\_cutoff) %>%

ggplot(aes(recall, precision, color = method)) +

geom\_line() +

geom\_point()