PS8: Predictive Modelling

Pooja Sadarangani

2022-12-04

Collaborators: Pranali Oza

1 Is COMPAS fair?

1. (2pt) Load the COMPAS data, and perform a basic sanity checks.

```
## [1] 6172    8
```

```
##   age c_charge_degree             race          age_cat  sex priors_count
## 1  69               F            Other Greater than 45 Male            0
## 2  34               F African-American         25 - 45 Male            0
## 3  24               F African-American    Less than 25 Male            4
## 4  44               M            Other         25 - 45 Male            0
## 5  41               F        Caucasian         25 - 45 Male           14
## 6  43               F            Other         25 - 45 Male            3
##   decile_score two_year_recid
## 1            1              0
## 2            3              1
## 3            4              1
## 4            1              0
## 5            6              1
## 6            4              0
```

```
## [1] "age"             "c_charge_degree" "race"            "age_cat"
## [5] "sex"             "priors_count"    "decile_score"    "two_year_recid"
```

```
## [1] FALSE
```

Sanity check is performed, there are no NA values in the data set.

2. (2pt) Filter the data to keep only only Caucasians and African-Americans.

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr   0.3.4
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.3      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
## [1] "African-American" "Caucasian"
```

3. (2pt) Create a new dummy variable based off of COMPAS risk score (decile_score), which indicates if an individual was classified as low risk (score 1-4) or high risk (score 5-10).

```
##   age c_charge_degree             race      age_cat    sex priors_count
## 1  34               F African-American      25 - 45   Male            0
## 2  24               F African-American Less than 25   Male            4
## 3  41               F        Caucasian      25 - 45   Male           14
## 4  39               M        Caucasian      25 - 45 Female            0
## 5  27               F        Caucasian      25 - 45   Male            0
## 6  23               M African-American Less than 25   Male            3
##   decile_score two_year_recid risk_level
## 1            3              1   low risk
## 2            4              1   low risk
## 3            6              1  high risk
## 4            1              0   low risk
## 5            4              0   low risk
## 6            6              1  high risk
```

4. (4pt) Now analyze the offenders across this new risk category:

##(a) What is the recidivism rate for low-risk and high-risk individuals? ##(b) What are the recidivism rates for African-Americans and Caucasians?

Recidivism Rate for low risk individuals

```
## [1] 0.3519797
```

The recidivists rate for low risk individuals is 0.3519797

Recidivism Rate for high risk individuals

```
## [1] 0.6689109
```

The recidivists rate for high risk individuals is 0.6689109

Recidivism Rate for African-American individuals

```
## [1] 0.5625197
```

The recidivists rate for Caucasian individuals is 0.5625197

Recidivism Rate for Caucasian individuals

```
## [1] 0.4146457
```

The recidivists rate for African-American individuals is 0.4146457

5. (5 pt) Now create a confusion matrix comparing COMPAS predictions for recidivism (is/is not low risk) and the actual two-year recidivism and interpret the results. To keep things coherent, let's call recidivists "positive".

```
##    
##     high risk low risk
##   0       923     1872
##   1      1602      881
```

According to the above results, we can make the following interpretations:

No. of true positives = 1602. This means 1602 recidivists were correctly classified as high risk

No. of false positives = 881. This means that 881 recidivists were falsely classified as low risk

No. of true negatives = 1872. This means that 1872 not recidivists were correctly classified as low risk

No. of false negatives = 923. This means that 923 recidivists were falsely classified as high risk

6. (10pt) Note the accuracy of the COMPAS classification, and also how its errors were distributed. Would you feel comfortable having a judge use COMPAS to inform your sentencing guidelines? At what point would the error/misclassification risk be acceptable for you?

```
## [1] 0.6582039
```

```
## The error is 0.3417961
```

```
## [1] 0.6451873
```

```
## [1] 0.6344554
```

The accuracy of COMPAS classification is 65.82039%

The precision of COMPAS classification is 0.6451873

The recall of COMPAS classification is 0.6344554

I would not feel comfortable having a judge use COMPAS to inform sentencing guidelines because as we can see, the accuracy is only 65% for a decision whhich affect someone's enitre life. While humans are prone to making errors as well, jury can provide varied views and eliminate biases. Also, judges have the human intelligence to consider other factors as well, along with those mentioned in question such as inevitable conditions like self-defense or blackmailings.

7. (15pt) Now, you repeat your confusion matrix calculation and analysis from 5. But this time do it separately for African-Americans and for Caucasians:

##(a) How accurate is the COMPAS classification for African-American individuals? For Caucasians? ##(b) What are the false positive rates FPR=FP/N=FP/(FP+TN)? ##(c) The false negative rates FNR=FN/P=FN/(FN+TP)?

Confusion Matrix for race="African-American"

```
##
##                0    1
##   high risk  641 1188
##   low risk   873  473
```

```
## [1] 0.6491339
```

The accuracy of COMPAS for African-American individuals is 64.91339%

```
## [1] 0.4233818
```

```
## [1] 0.7152318
```

```
## [1] 0.2847682
```

```
## [1] 0.5766182
```

Confusion Matrix for race="Caucasian"

```
##
##     high risk low risk
##   0       282      999
##   1       414      408
```

```
## [1] 0.6718973
```

The accuracy of COMPAS for Caucasian individuals is 67.18973%

```
## [1] 0.2201405
```

```
## [1] 0.5036496
```

```
## [1] 0.4963504
```

```
## [1] 0.7798595
```

8. (10pt) If you have done this correctly, you will find that COMPAS's true negative and true positive percentages are similar for African-American and Caucasian individuals, but that false positive rates and false negative rates are different. Look again at the overall recidivism rates in the dataset for Black and White individuals. In your opinion, is the COMPAS algorithm 'fair'? Justify your answer.

Ans. In my opinion, the COMPAS algorithm is not 'fair', because a lot of data has been incorrectly classified. Hundreds of cases have been falsely identified therefore, I don't this this algorithm is fair. And using factors such as 'sex' and 'race' could develop biases in the algorithm, being unfair to certain communities in the long-run.

2 Make your own COMPAS!

1. (4pt) You should not use variables score_text, decile_score and two_year_recid.Explain why!

Ans. These are the variables that we need to predict, thus we cannot use them as a predictor.

2. (6pt) Before we start: what do you think, what is an appropriate model performance measure for this task? A, P, R, For something else? Maybe you want to report multiple measures? Explain!

Ans. In my opinion, all three factors (A,P,R) are crucial. Accuracy would help us determine how many cases ave been correctly identified. However, this would not be an appropriate performance measure if number of true cases and false cases are imbalance. In such cases, it would be better to rely on precision and recall. Precision will tell us how many positive cases have been correcly identified. And recall will tell us how many positive class predictions have been made from positive dataset

3. (6pt) Split your data into training and validation set. Develop a model on the training set and compute its performance on validation set. Tweak your model to get as good performance as you can get. Report the performance.

```
## [1] "age"            "sex"            "race"           "priors_count"
## [5] "two_year_recid"

## [1] 3694

## [1] 3694

## [1] 3694    5

## [1] 1584    5

##
## Call:
## glm(formula = two_year_recid ~ age + priors_count, family = binomial(),
##     data = train)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
```

```
## -2.7886  -1.0535  -0.5931   1.1163   2.2967
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   0.939420   0.113710   8.262   <2e-16 ***
## age          -0.047055   0.003366 -13.980   <2e-16 ***
## priors_count  0.167703   0.009689  17.308   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 5109.5  on 3693  degrees of freedom
## Residual deviance: 4556.9  on 3691  degrees of freedom
## AIC: 4562.9
##
## Number of Fisher Scoring iterations: 4

##
## results1   0   1
##        0 781 518
##        1  64 221

## The accuracy of model 1 is: 0.6433081
```

4. (3pt) Add sex to the model. Does it help to improve the performance? And by "improve" we mean not just a tiny bit.

```
##
## Call:
## glm(formula = two_year_recid ~ age + priors_count + sex, family = binomial(),
##     data = train)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -2.775  -1.034  -0.580   1.107   2.258
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   0.659655   0.135204   4.879 1.07e-06 ***
## age          -0.046635   0.003370 -13.840  < 2e-16 ***
## priors_count  0.163315   0.009716  16.809  < 2e-16 ***
## sexMale       0.346157   0.091032   3.803 0.000143 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 5109.5  on 3693  degrees of freedom
## Residual deviance: 4542.3  on 3690  degrees of freedom
## AIC: 4550.3
##
## Number of Fisher Scoring iterations: 4

##
```

```
## results2   0   1
##        0 777 514
##        1  68 225

## The accuracy of model 2 is: 0.6401515
```

From the above result, we can tell that the accuracy has not improved.

5. (3pt) And finally add race. Does the model improve?

```
##
## Call:
## glm(formula = two_year_recid ~ age + priors_count + sex + race,
##     family = binomial(), data = train)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.7708  -1.0383  -0.5808   1.1075   2.2620
##
## Coefficients:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)    0.671898   0.135882   4.945 7.63e-07 ***
## age           -0.045964   0.003447 -13.335  < 2e-16 ***
## priors_count   0.161673   0.009868  16.383  < 2e-16 ***
## sexMale        0.342607   0.091121   3.760  0.00017 ***
## raceCaucasian -0.068834   0.075466  -0.912  0.36171
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 5109.5  on 3693  degrees of freedom
## Residual deviance: 4541.5  on 3689  degrees of freedom
## AIC: 4551.5
##
## Number of Fisher Scoring iterations: 4

##
## results3   0   1
##        0 778 508
##        1  67 231

## The accuracy of model 1 is: 0.6433081
```

From the above results, we can tell that the accuracy has not changed.

6. (8pt) Discuss the results. Did you manage to get as good results as COMPAS? Did you manage to do even better? Does gender and race help to improve your predictions? What should judges do when having access to such models? Should they use those?

The accuracy of my model was a little lesser than that of COMPAS. Gender and Race did not help me in improving the model prodections. Judges must not use these models as they can be harmful and introduce biases in the system. Although race and gender did not have much effect in model predictions, having a lot of data could lead to these factors having some impact in the output.