

...In 2018 the renowned scientific journal *science* broke a story that researchers had re-engineered the commercial criminal risk assessment software COMPAS with a simple logistic regression ([Science: The accuracy, fairness, and limits of predicting recidivism](#)).

According to this article, COMPAS uses 137 features, the authors just used two. In this post, I will up the ante by showing you how to achieve similar results using just *one simple rule* based on only *one feature* which is found *automatically in no-time* by the *OneR* package, so read on!

Algorithms for predicting recidivism are more and more used to assess a criminal defendant's likelihood of committing a crime, especially in the US. These predictions are used in pretrial, parole, and sentencing decisions. The article gives some background:

One widely used criminal risk assessment tool, Correctional Offender Management Profiling for Alternative Sanctions (COMPAS by the company Equivant), has been used to assess more than 1 million offenders since it was developed in 1998. The recidivism prediction component of COMPAS—the recidivism risk scale—has been in use since 2000. This software predicts a defendant's risk of committing a misdemeanor or felony within 2 years of assessment from 137 features about an individual and the individual's past criminal record.

For our analysis, we will use the original dataset which is provided alongside the article: [BROWARD_CLEAN.csv](#) (to get information on the different attributes extensive metadata is provided too: [readme.txt](#)). To decode and aggregate the charge-ids we use an additional file: [CHARGE_ID.csv](#). Please download both files and change the paths in the following code accordingly:

```
broward <- read.csv("data/BROWARD_CLEAN.csv") # change path accordingly
charges <- read.csv("data/CHARGE_ID.csv") # change path accordingly
round(100 * prop.table(table(broward$compas_correct))[2], 1) # COMPAS
performance
##      1
## 65.4

data <- broward[c("race", "sex", "age", "juv_fel_count",
"juv_misd_count", "priors_count", "charge_id",
"charge_degree..misd.fel.", "two_year_recid")]
names(data)[names(data) == "charge_degree..misd.fel."] <-
"charge_degree" # rename column
data$race <- as.factor(data$race)
data$sex <- as.factor(data$sex)
# feature engineering: create aggregate real name version out of
charge_id
data$charge_id <- charges[data$charge_id, 3]
names(data)[names(data) == "charge_id"] <- "charge_name"
data$charge_degree <- as.factor(data$charge_degree)
data$two_year_recid <- as.factor(data$two_year_recid)

str(data)
## 'data.frame':    7214 obs. of  9 variables:
##  $ race          : Factor w/ 6 levels "1","2","3","4",...: 6 2 2 2 6
6 1 6 1 1 ...
```

```
## $ sex : Factor w/ 2 levels "0","1": 1 1 1 1 1 1 1 1 1 2 1
...
## $ age : int 69 34 24 23 43 44 41 43 39 21 ...
## $ juv_fel_count : int 0 0 0 0 0 0 0 0 0 0 ...
## $ juv_misd_count: int 0 0 0 1 0 0 0 0 0 0 ...
## $ priors_count : int 0 0 4 1 2 0 14 3 0 1 ...
## $ charge_name : chr "Assault with a Deadly Weapon" "Battery"
"Possession of Cocaine" "Possession of Cannabis/Marijuana" ...
## $ charge_degree : Factor w/ 2 levels "0","1": 2 2 2 2 2 1 2 2 1 2
...
## $ two_year_recid: Factor w/ 2 levels "0","1": 1 2 2 1 1 1 2 1 1 2
...
```

Ok, so after a little bit of data wrangling and feature engineering we have a dataset with 7214 cases and 8 potential features to predict a defendant's risk of committing a misdemeanor or felony within 2 years (more precisely 1 in `two_year_recid` in the last column means that the person recidivated within two years after the previous charge). We see that COMPAS gets it right about 65.4% of the time.

First, we split our dataset into a training (80%) and a test set (20%) and after that, we fire up the **OneR** package (on CRAN):

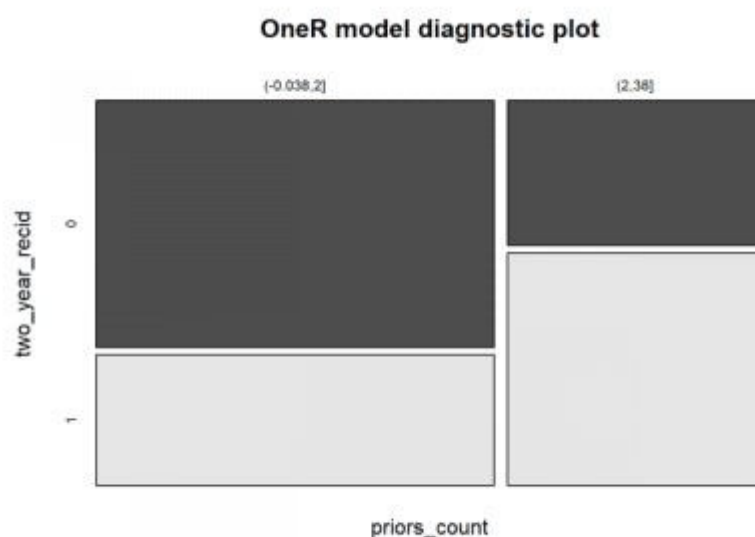
```
set.seed(123)
random <- sample(1:nrow(data), 0.8 * nrow(data))
data_train <- data[random, ]
data_test <- data[-random, ]

library(OneR)
data_bin <- optbin(data_train, method = "infogain")
model <- OneR(data_bin, verbose = TRUE)
##
##      Attribute      Accuracy
## 1 * priors_count    63.94%
## 2  charge_name     58.5%
## 3   age            57.29%
## 4  juv_misd_count  57.11%
## 5  juv_fel_count  56.56%
## 6   race           56.37%
## 7   sex            54.77%
## 7  charge_degree  54.77%
## ---
## Chosen attribute due to accuracy
## and ties method (if applicable): '*'

summary(model)
##
## Call:
## OneR.data.frame(x = data_bin, verbose = TRUE)
##
## Rules:
## If priors_count = (-0.038,2] then two_year_recid = 0
## If priors_count = (2,38]      then two_year_recid = 1
```

```
##
## Accuracy:
## 3690 of 5771 instances classified correctly (63.94%)
##
## Contingency table:
##           priors_count
## two_year_recid (-0.038,2] (2,38]  Sum
##           0          * 2297      864 3161
##           1          1217 * 1393 2610
##           Sum          3514    2257 5771
## ---
## Maximum in each column: '*'
##
## Pearson's Chi-squared test:
## X-squared = 405.93, df = 1, p-value < 2.2e-16

plot(model)
```



```
prediction <- predict(model, data_test)
eval_model(prediction, data_test)
##
## Confusion matrix (absolute):
##           Actual
## Prediction    0    1  Sum
##           0    590  283  873
##           1    212  358  570
##           Sum   802  641 1443
##
## Confusion matrix (relative):
##           Actual
## Prediction    0    1  Sum
##           0   0.41 0.20 0.60
##           1   0.15 0.25 0.40
##           Sum 0.56 0.44 1.00
##
```

```
## Accuracy:
## 0.657 (948/1443)
##
## Error rate:
## 0.343 (495/1443)
##
## Error rate reduction (vs. base rate):
## 0.2278 (p-value = 3.218e-15)
```

There are several takeaways, most notable we get this simple rule:

```
## If priors_count = (-0.038,2] then two_year_recid = 0
## If priors_count = (2,38]      then two_year_recid = 1
```

It very simply states that if an offender has recidivated in the past, he/she will do so again!

With this simple rule, we achieve a better (than COMPAS) out-of-sample accuracy of 65.7%! In the science article, the authors achieved a slightly better accuracy with a logistic regression with two features (priors count and age), OneR only needs one (priors count) for that.

On top of that COMPAS is a black box, logistic regression is somewhat interpretable and only OneR gives a fully interpretable simple rule.

To summarize:

	Used algorithm	No of used features	Accuracy	Interpretability
COMPAS	proprietary	137	65.4%	black box
Linear classifier	logistic regression	2 (priors count, age)	66.6%	somewhat interpretable
OneR	One Rule classifier	1 (priors count)	65.7%	fully interpretable

OneR also gives us a list of the single best predictors in descending order:

```
##      Attribute      Accuracy
## 1 * priors_count  63.94%
## 2  charge_name   58.5%
## 3   age          57.29%
## 4  juv_misd_count 57.11%
## 5  juv_fel_count 56.56%
## 6   race         56.37%
## 7   sex          54.77%
## 7  charge_degree 54.77%
```

For the second one, charge_name, it makes intuitive sense that certain charges lead to a higher recidivism rate. Let us have a look at the ones with a 100% “guarantee”:

```
model_charge_name <- OneR(two_year_recid ~ charge_name, data)
prob_charge_name <- predict(model_charge_name, data.frame(charge_name =
sort(unique(data$charge_name))), type = "prob")[ , 2]
names(prob_charge_name[prob_charge_name == 1])
## [1] "Accessory After the Fact"
## [2] "Aiding Escape"
## [3] "Aiding Prostitution"
## [4] "Bribery"
## [5] "Dealing Heroin"
```

```
## [6] "Discharging Firearm From Vehicle"
## [7] "Exhibition of a Weapon on School Property"
## [8] "Fabricating Physical Evidence"
## [9] "Insurance Fraud"
## [10] "Interference with Custody"
## [11] "Littering"
## [12] "Manslaughter"
## [13] "Obstruction of Officer with Violence"
## [14] "Possession of a Motor Vehicle with Altered Vehicle
Identification Number"
## [15] "Possession of Alcohol Under 21 Years of Age"
## [16] "Possession with Intent to Sell Counterfeits"
## [17] "Principal In The First Degree"
## [18] "Providing a Contradicting Statement"
## [19] "Shooting into a Home"
## [20] "Sound Articles Over 100"
## [21] "Unauthorized Interference with a Railroad"
## [22] "Voyeurism"
```

When you go through the list, you will think to yourself that most of them make sense, e.g. somebody dealing with heroin or a voyeur has a very high probability of continuing to do so in the future. Now let us look at the ones with a zero probability of re-offending:

```
names(prob_charge_name[prob_charge_name == 0])
## [1] "Abuse"
## [2] "Armed Burglary"
## [3] "Carrying an Open Beverage in Public"
## [4] "Causing Public Danger"
## [5] "Compulsory Education Attendance Violation"
## [6] "Computer Pornography"
## [7] "Consuming Alcoholic Beverage in Public"
## [8] "Contributing to the Delinquency Of A Minor"
## [9] "Dealing Ecstasy"
## [10] "Dealing Stolen Property"
## [11] "Disrupting a School Function"
## [12] "Elder Molestation"
## [13] "Exploitation of an Elderly Person of $20,000-$100,000"
## [14] "Failure to Obey Drivers License Restrictions"
## [15] "Failure to Obey Sex Offender Laws"
## [16] "False Information to Law Enforcement Officer During
Investigation"
## [17] "Illegal Gambling"
## [18] "Interference with Traffic Control Railroad Divide"
## [19] "Intoxicated Dangering of Safety of Another"
## [20] "Money Laundering"
## [21] "Murder"
## [22] "Neglect of an Elderly Person"
## [23] "Obtaining Controlled Substance by Fraud"
## [24] "Offense Against Intellectual Property"
## [25] "Operating Motorcycle without a Valid Drivers License"
## [26] "Possession of a Tobacco Product Under 18 Years of Age"
## [27] "Possession of Child Pornography"
```

```
## [28] "Possession of Weapon on School Property"
## [29] "Purchasing a Controlled Substance"
## [30] "Refusing to Supply DNA Sample"
## [31] "Selling Counterfeit Goods"
## [32] "Sex Offender Failing to Comply with Law"
## [33] "Simulation of Legal Process"
## [34] "Unauthorized Loud Noise"
## [35] "Unlawful Disturbance"
## [36] "Unlicensed Telemarketing"
## [37] "Use of 2 Way Device to Facilitate Felony"
## [38] "Using a Computer for Child Exploitation"
## [39] "Video Voyeurism on Child"
```

By inspecting that list you are in for a surprise: “Murder” and “Sex Offender Failing to Comply with Law”! The list is suggesting that there is zero probability of murderers and recalcitrant sex offenders to re-offend! Taking this to the extreme the software is giving us the advice to let all of those felons go free because they won’t try anything bad in the future... what is going on here?

Well, the most probable reason is that those serious offenders don’t have a chance to commit any new crime within a two-year’s timeframe because they luckily stay in prison and don’t have a chance to re-offend!

And there you go again: this is why it is so important to have a system that is fully interpretable and not just a black box! From an algorithmic and data-based point of view, the advice makes total sense (the algorithm doesn’t know anything about crimes, murderers, sex offenders, and recidivism per se, and the data suggests this outcome) but in reality following this advice blindly would end in disaster!

As a last point let us have a look at the feature `age` (which was also included in the logistic regression):

```
OneR(optbin(two_year_recid ~ age, data))
##
## Call:
## OneR.data.frame(x = optbin(two_year_recid ~ age, data))
##
## Rules:
## If age = (17.9,32.3] then two_year_recid = 1
## If age = (32.3,96.1] then two_year_recid = 0
##
## Accuracy:
## 4190 of 7214 instances classified correctly (58.08%)
```

This just corroborates something that has long been known in criminology: if you’re young, you’re risky.

As a final note, ongoing research seems to suggest that there is some accuracy limit of around 70%, so OneR is not too far away from this supposed barrier to predicting recidivism. Even with very sophisticated methods, there remains a huge margin of error. Humans don’t seem to be that predictable after all...