MENA Exceptionalism and BART

Daniel K Baissa

2023-05-09

In this example I will use BART to predict coups (not the DV) globally and use it to explain my logic for looking at MENA exceptionalism

I will start by loading the libraries and setting up bartMachine

```
library(tidyverse)
library(rJava)

options(java.parameters = "-Xmx5g")
library(bartMachine)
set_bart_machine_num_cores(5)
```

```
## bartMachine now using 5 cores.
```

This is just a subset of data for illustrative perpouses and speed on my laptop.

I will set up the data for BARTmachine

```
df <- read_csv("Data/Vdem_Banks.csv") |>
  select(!c(suvival_time, coup_number, id, ...1))

X <- df |>
  mutate(Region = as.factor(Region)) |>
  na.omit()

y <- as.factor(X$e_pt_coup)

X$e_pt_coup <- NULL

y <- relevel(y, "1")

X <- as.data.frame(X)
```

Here are the variables on the right hand side of this model:

```
# Convert column names to a table for ease of reading
col_table <- matrix(colnames(X), ncol = 2, byrow = TRUE)

# Print the table
col_table
```

```
##       [,1]                    [,2]
##  [1,] "country_name"          "year"
##  [2,] "Polity 5"              "Region"
##  [3,] "GDP Growth"            "GDP Per Capita"
##  [4,] "Inflation"             "Petroleum production per capita"
```

```
##  [5,] "Infant Mortality"          "Clientelism"
##  [6,] "Rule of Law"               "Party Institutionalization"
##  [7,] "Legislative Party Coh"     "National party control"
##  [8,] "Political Polarization"    "Defense Exp Per Cap"
##  [9,] "Boix Democracy"            "Barriers to parties"
## [10,] "Party Ban"                 "Opposition parties autonomy"
## [11,] "Party organizations"       "Party branches"
## [12,] "Party linkages"            "Distinct party platforms"
## [13,] "Candidate selection"       "Party competition across regions"
## [14,] "Subnational party control" "v2exdfpphg"
```

Now let's run a basic BART model. The data are extremely sparce so I will set prob_rule_class = .1 to compensate. This can be validated by looking at the confusion matrix to make sure the model has a good fit.

```
bm <- bartMachine(y = y, X = X, prob_rule_class = .1)
```

```
## bartMachine initializing with 50 trees...
## bartMachine vars checked...
## bartMachine java init...
## bartMachine factors created...
## bartMachine before preprocess...
## bartMachine after preprocess... 134 total features...
## bartMachine training data finalized...
## Now building bartMachine for classification where "1" is considered the target level...Covariate imp
## evaluating in sample data...done
```

```
bm
```

```
## bartMachine v1.3.2 for classification
##
## training data size: n = 1945 and p = 134
## built in 10.9 secs on 5 cores, 50 trees, 250 burn-in and 1000 post. samples
##
## confusion matrix:
##
##            predicted 1 predicted 0 model errors
## actual 1         31.00      13.000        0.295
## actual 0         66.00    1835.000        0.035
## use errors        0.68       0.007        0.041
```

Here we see that the model accurately classifies -2.5% of non-coups and 0.705% of coups with an overall accuracy rate of 0.959%