```r
#  Import Libraries
library(tidyverse)
```

```
## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.0 ──
```

```
## ✓ ggplot2 3.3.1     ✓ purrr   0.3.4
## ✓ tibble  3.0.1     ✓ dplyr   1.0.0
## ✓ tidyr   1.1.2     ✓ stringr 1.4.0
## ✓ readr   1.3.1     ✓ forcats 0.5.0
```

```
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
library(readxl)
library(h2o)
```

```
##
## ----------------------------------------------------------------------
##
## Your next step is to start H2O:
##     > h2o.init()
##
## For H2O package documentation, ask for help:
##     > ??h2o
##
## After starting H2O, you can use the Web UI at http://localhost:54321
## For more information visit https://docs.h2o.ai
##
## ----------------------------------------------------------------------
```

```
##
## Attaching package: 'h2o'
```

```
## The following objects are masked from 'package:stats':
##
##     cor, sd, var
```

```
## The following objects are masked from 'package:base':
##
##     &&, %*%, %in%, ||, apply, as.factor, as.numeric, colnames,
##     colnames<-, ifelse, is.character, is.factor, is.numeric, log,
##     log10, log1p, log2, round, signif, trunc
```

```r
#  Read the Excel Sheets
path   <- "bank_term_deposit_marketing_analysis.xlsx"
sheets <- excel_sheets(path)
```

```r
#  Explore Data in each Sheet
sheets %>%
    map(~ read_excel(path  = path, sheet = .)) %>%
    set_names(sheets)
```

```
## New names:
## * `` -> ...2
## * `` -> ...3
## * `` -> ...4
## * `` -> ...5
## * `` -> ...6
## * ...
```

```
## New names:
## * `` -> ...2
## * `` -> ...4
```

```
## $PROCEDURE
## # A tibble: 14 x 1
##    `BANK MARKETING ANALYSIS PROCEDURE`
##    <chr>
##  1 <NA>
##  2 STEP 1: COLLECT INFORMATION
##  3 1) CLIENT INFORMATION: AGE, JOB, MARITAL STATUS, EDUCATION LEVEL
##  4 2) CLIENT LOAN HISTORY: DEFAULT HISTORY, HOME LOAN, PERSONAL LOAN, CURRENT B…
##  5 3) MARKETING HISTORY: CONTACT TYPE, DAY LAST CONTACT, MONTH LAST CONTACT, LA…
##  6 4) SUBSCRIPTION HISTORY: ENROLLED IN TERM LOAN? (Y/N)
##  7 <NA>
##  8 STEP 2: MERGE INFORMATION
##  9 1) PERFORM VLOOKUP
## 10 <NA>
## 11 STEP 3: MARKETING ANALYSIS
## 12 1) DAILY RANGE: WHAT IS NORMAL HIT RATE?
## 13 2) WHAT FEATURES CONTRIBUTE TO TERM LOAN ENROLLMENT?
## 14 - Job Analysis
##
## $`DATA DESCRIPTION`
## # A tibble: 70 x 1
##    bank_info
##    <chr>
##  1 Citation Request:
##  2 This dataset is public available for research. The details are described in …
##  3 Please include this citation if you plan to use this database:
##  4 <NA>
##  5 [Moro et al., 2011] S. Moro, R. Laureano and P. Cortez. Using Data Mining fo…
##  6 In P. Novais et al. (Eds.), Proceedings of the European Simulation and Model…
##  7 <NA>
##  8 Available at: [pdf] http://hdl.handle.net/1822/14838
##  9 [bib] http://www3.dsi.uminho.pt/pcortez/bib/2011-esm-1.txt
## 10 <NA>
## # … with 60 more rows
##
## $`Step 1 - Collect Information`
## # A tibble: 1 x 2
##    Step Description
##   <dbl> <chr>
## 1     1 Collect Client Information
##
## $CLIENT_INFO
## # A tibble: 45,211 x 5
##    ID      AGE JOB          MARITAL  EDUCATION
##    <chr> <dbl> <chr>        <chr>    <chr>
##  1 2836     58 management   married  tertiary
##  2 2837     44 technician   single   secondary
##  3 2838     33 entrepreneur married  secondary
##  4 2839     47 blue-collar  married  unknown
##  5 2840     33 unknown      single   unknown
##  6 2841     35 management   married  tertiary
##  7 2842     28 management   single   tertiary
##  8 2843     42 entrepreneur divorced tertiary
##  9 2844     58 retired      married  primary
## 10 2845     43 technician   single   secondary
## # … with 45,201 more rows
##
## $LOAN_HISTORY
## # A tibble: 45,211 x 5
##    ID    DEFAULT BALANCE HOUSING LOAN
##    <chr> <chr>     <dbl> <chr>   <chr>
```

```
##  1 2836  no         2143 yes     no
##  2 2837  no           29 yes     no
##  3 2838  no            2 yes     yes
##  4 2839  no         1506 yes     no
##  5 2840  no            1 no      no
##  6 2841  no          231 yes     no
##  7 2842  no          447 yes     yes
##  8 2843  yes           2 yes     no
##  9 2844  no          121 yes     no
## 10 2845  no          593 yes     no
## # … with 45,201 more rows
##
## $`MARKETING HISTORY`
## # A tibble: 45,211 x 9
##    ID    CONTACT   DAY MONTH DURATION CAMPAIGN PDAYS PREVIOUS POUTCOME
##    <chr> <chr>   <dbl> <chr>    <dbl>    <dbl> <dbl>    <dbl> <chr>
##  1 2836  unknown     5 may        261        1    -1        0 unknown
##  2 2837  unknown     5 may        151        1    -1        0 unknown
##  3 2838  unknown     5 may         76        1    -1        0 unknown
##  4 2839  unknown     5 may         92        1    -1        0 unknown
##  5 2840  unknown     5 may        198        1    -1        0 unknown
##  6 2841  unknown     5 may        139        1    -1        0 unknown
##  7 2842  unknown     5 may        217        1    -1        0 unknown
##  8 2843  unknown     5 may        380        1    -1        0 unknown
##  9 2844  unknown     5 may         50        1    -1        0 unknown
## 10 2845  unknown     5 may         55        1    -1        0 unknown
## # … with 45,201 more rows
##
## $`SUBSCRIPTION HISTORY`
## # A tibble: 45,211 x 2
##    ID    TERM_DEPOSIT
##    <chr> <chr>
##  1 2836  no
##  2 2837  no
##  3 2838  no
##  4 2839  no
##  5 2840  no
##  6 2841  no
##  7 2842  no
##  8 2843  no
##  9 2844  no
## 10 2845  no
## # … with 45,201 more rows
##
## $`Step 2 - Merge Information`
## # A tibble: 1 x 2
##    Step Description
##   <dbl> <chr>
## 1     2 Perform Data Merge
##
## $CLIENT_MERGE
## # A tibble: 10,006 x 20
##    `VLOOKUP MERGE … ...2  ...3  ...4  ...5  ...6  ...7  ...8  ...9  ...10 ...11
##    <chr>            <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr>
##  1 1. DIFFICULT TO… <NA>  <NA>  <NA>  <NA>  <NA>  <NA>  <NA>  <NA>  <NA>  <NA>
##  2 2. COMPUTATIONA… <NA>  <NA>  <NA>  <NA>  <NA>  <NA>  <NA>  <NA>  <NA>  <NA>
##  3 3. EVERY CELL C… <NA>  <NA>  <NA>  <NA>  <NA>  <NA>  <NA>  <NA>  <NA>  <NA>
##  4 <NA>             <NA>  <NA>  <NA>  <NA>  <NA>  <NA>  <NA>  <NA>  <NA>  <NA>
##  5 <NA>             CLIE… <NA>  <NA>  <NA>  LOAN… <NA>  <NA>  <NA>  MARK… <NA>
##  6 <NA>             2.0   3.0   4.0   5.0   2.0   3.0   4.0   5.0   2.0   3.0
##  7 ID               AGE   JOB   MARI… EDUC… DEFA… BALA… HOUS… LOAN  CONT… DAY
##  8 2836             58    mana… marr… tert… no    2143  yes   no    unkn… 5
```

```
##  9 2837             44    tech… sing… seco… no    29    yes   no    unkn… 5
## 10 2838             33    entr… marr… seco… no    2     yes   yes   unkn… 5
## # … with 9,996 more rows, and 9 more variables: ...12 <chr>, ...13 <chr>,
## #   ...14 <chr>, ...15 <chr>, ...16 <chr>, ...17 <chr>, ...18 <chr>,
## #   ...19 <chr>, ...20 <chr>
##
## $`Step 3 - Marketing Analysis`
## # A tibble: 1 x 2
##    Step Description
##   <dbl> <chr>
## 1     3 Perform Marketing Analysis
##
## $`DAILY RANGE`
## # A tibble: 28 x 4
##    `HIT RATE` ...2  `DAILY SUMMARY`     ...4
##         <dbl> <lgl> <chr>              <dbl>
##  1     0.0386 NA    MEAN             0.0351
##  2     0.0360 NA    MEDIAN           0.0362
##  3     0.0551 NA    SD               0.0138
##  4     0.0613 NA    LOWER CONF       0.00755
##  5     0.0427 NA    UPPER CONF       0.0627
##  6     0.0391 NA    <NA>            NA
##  7     0.0451 NA    <NA>            NA
##  8     0.0166 NA    <NA>            NA
##  9     0.0222 NA    <NA>            NA
## 10     0.0179 NA    <NA>            NA
## # … with 18 more rows
##
## $`JOB ANALYSIS`
## # A tibble: 0 x 0
##
## $Sheet3
## # A tibble: 0 x 0
```

```
# Join Data by ID Column
data_joined_tbl <- sheets[4:7] %>%
    map(~ read_excel(path = path, sheet = .)) %>%
    reduce(left_join)
```

```
## Joining, by = "ID"
## Joining, by = "ID"
## Joining, by = "ID"
```

```
#  Start H2O Cluster
h2o.init(max_mem_size = "4g")
```

```
##  Connection successful!
##
## R is connected to the H2O cluster:
##     H2O cluster uptime:         15 minutes 41 seconds
##     H2O cluster timezone:       Asia/Kolkata
##     H2O data parsing timezone:  UTC
##     H2O cluster version:        3.30.1.3
##     H2O cluster version age:    24 days
##     H2O cluster name:           H2O_started_from_R_priyarajpurohit_exx530
##     H2O cluster total nodes:    1
##     H2O cluster total memory:   3.85 GB
##     H2O cluster total cores:    8
##     H2O cluster allowed cores:  8
##     H2O cluster healthy:        TRUE
##     H2O Connection ip:          localhost
##     H2O Connection port:        54321
##     H2O Connection proxy:       NA
##     H2O Internal Security:      FALSE
##     H2O API Extensions:         Amazon S3, XGBoost, Algos, AutoML, Core V3, TargetEncoder, C
ore V4
##     R Version:                  R version 3.6.3 (2020-02-29)
```

```
#Data Preparation
data_joined_tbl <- data_joined_tbl %>%
    mutate_if(is.character, as.factor)

train <- as.h2o(data_joined_tbl)
```

```
## Warning in use.package("data.table"): data.table cannot be used without R
## package bit64 version 0.9.7 or higher. Please upgrade to take advangage of
## data.table speedups.
```

```
##
  |
  |                                                                      |   0%
  |
  |======================================================================| 100%
```

```
h2o.describe(train)
```

```
##           Label Type Missing Zeros PosInf NegInf   Min    Max         Mean
## 1            ID enum       0     1      0      0     0  45210           NA
## 2           AGE  int       0     0      0      0    18     95 4.093621e+01
## 3           JOB enum       0  5171      0      0     0     11           NA
## 4       MARITAL enum       0  5207      0      0     0      2           NA
## 5     EDUCATION enum       0  6851      0      0     0      3           NA
## 6       DEFAULT enum       0 44396      0      0     0      1 1.802659e-02
## 7       BALANCE  int       0  3514      0      0 -8019 102127 1.362272e+03
## 8       HOUSING enum       0 20081      0      0     0      1 5.558382e-01
## 9          LOAN enum       0 37967      0      0     0      1 1.602265e-01
## 10      CONTACT enum       0 29285      0      0     0      2           NA
## 11          DAY  int       0     0      0      0     1     31 1.580642e+01
## 12        MONTH enum       0  2932      0      0     0     11           NA
## 13     DURATION  int       0     3      0      0     0   4918 2.581631e+02
## 14     CAMPAIGN  int       0     0      0      0     1     63 2.763841e+00
## 15        PDAYS  int       0     0      0      0    -1    871 4.019783e+01
## 16     PREVIOUS  int       0 36954      0      0     0    275 5.803234e-01
## 17     POUTCOME enum       0  4901      0      0     0      3           NA
## 18 TERM_DEPOSIT enum       0 39922      0      0     0      1 1.169848e-01
##           Sigma Cardinality
## 1            NA       45211
## 2    10.6187620          NA
## 3            NA          12
## 4            NA           3
## 5            NA           4
## 6     0.1330489           2
## 7  3044.7658292          NA
## 8     0.4968778           2
## 9     0.3668200           2
## 10           NA           3
## 11    8.3224762          NA
## 12           NA          12
## 13  257.5278123          NA
## 14    3.0980209          NA
## 15  100.1287460          NA
## 16    2.3034410          NA
## 17           NA           4
## 18    0.3214057           2
```

```
y <- "TERM_DEPOSIT"

x <- setdiff(names(train), c(y, "ID"))
```

```
#H2O AutoML Training

aml <- h2o.automl(
    y = y,
    x = x,
    training_frame = train,
    project_name = "term_deposit",
    max_runtime_secs = 300,
    balance_classes = TRUE,
    #max_models = 10,
    seed = 1)
```

```
##
  |
  |                                                                      |   0%
## 18:25:24.295: New models will be added to existing leaderboard term_deposit@@TERM_DEPOSIT (l
eaderboard frame=null) with already 28 models.
  |
  |                                                                      |   1%
  |
  |=                                                                     |   1%
  |
  |==                                                                    |   2%
  |
  |==                                                                    |   3%
  |
  |===                                                                   |   4%
  |
  |===                                                                   |   5%
  |
  |====                                                                  |   5%
  |
  |====                                                                  |   6%
  |
  |=====                                                                 |   7%
  |
  |=====                                                                 |   8%
  |
  |======                                                                |   8%
  |
  |======                                                                |   9%
  |
  |=======                                                               |  10%
  |
  |========                                                              |  11%
  |
  |========                                                              |  12%
  |
  |=========                                                             |  13%
  |
  |==========                                                            |  14%
  |
  |==========                                                            |  15%
  |
  |===========                                                           |  15%
  |
  |===========                                                           |  16%
  |
  |============                                                          |  17%
  |
  |============                                                          |  18%
  |
  |=============                                                         |  18%
  |
  |=============                                                         |  19%
  |
  |==============                                                        |  19%
  |
  |==============                                                        |  20%
  |
  |==============                                                        |  21%
  |
  |===============                                                       |  21%
```

```
|
|===============                                                       |  22%
|
|================                                                      |  23%
|
|=================                                                     |  24%
|
|=================                                                     |  25%
|
|==================                                                    |  25%
|
|==================                                                    |  26%
|
|===================                                                   |  26%
|
|===================                                                   |  27%
|
|===================                                                   |  28%
|
|====================                                                  |  29%
|
|=====================                                                 |  30%
|
|======================                                                |  31%
|
|=======================                                               |  32%
|
|=======================                                               |  33%
|
|=======================                                               |  34%
|
|========================                                              |  34%
|
|========================                                              |  35%
|
|=========================                                             |  35%
|
|=========================                                             |  36%
|
|==========================                                            |  37%
|
|==========================                                            |  38%
|
|===========================                                           |  39%
|
|============================                                          |  40%
|
|=============================                                         |  41%
|
|==============================                                        |  42%
|
|==============================                                        |  43%
|
|==============================                                        |  44%
|
|===============================                                       |  44%
|
|===============================                                       |  45%
|
|================================                                      |  45%
|
|================================                                      |  46%
```

```
  |
  |=================================                                     |  47%
  |
  |=================================                                     |  48%
  |
  |==================================                                    |  48%
  |
  |==================================                                    |  49%
  |
  |===================================                                   |  49%
  |
  |===================================                                   |  50%
  |
  |===================================                                   |  51%
  |
  |====================================                                  |  51%
  |
  |====================================                                  |  52%
  |
  |=====================================                                 |  52%
  |
  |=====================================                                 |  53%
  |
  |=====================================                                 |  54%
  |
  |======================================                                |  54%
  |
  |=======================================                               |  55%
  |
  |=======================================                               |  56%
  |
  |========================================                              |  57%
  |
  |========================================                              |  58%
  |
  |=========================================                             |  58%
  |
  |=========================================                             |  59%
  |
  |==========================================                            |  59%
  |
  |==========================================                            |  60%
  |
  |==========================================                            |  61%
  |
  |===========================================                           |  61%
  |
  |============================================                          |  62%
  |
  |============================================                          |  63%
  |
  |=============================================                         |  64%
  |
  |=============================================                         |  65%
  |
  |==============================================                        |  65%
  |
  |==============================================                        |  66%
  |
  |===============================================                       |  67%
  |
  |===============================================                       |  68%
```

```
|
|================================================                      |  69%
|
|=================================================                     |  69%
|
|=================================================                     |  70%
|
|=================================================                     |  71%
|
|==================================================                    |  71%
|
|==================================================                    |  72%
|
|===================================================                   |  72%
|
|===================================================                   |  73%
|
|===================================================                   |  74%
|
|====================================================                  |  74%
|
|=====================================================                 |  75%
|
|=====================================================                 |  76%
|
|======================================================                |  77%
|
|======================================================                |  78%
|
|=======================================================               |  78%
|
|=======================================================               |  79%
|
|========================================================              |  80%
|
|========================================================              |  81%
|
|=========================================================             |  81%
|
|=========================================================             |  82%
|
|==========================================================            |  82%
|
|==========================================================            |  83%
|
|===========================================================           |  84%
|
|===========================================================           |  85%
|
|============================================================          |  85%
|
|============================================================          |  86%
|
|=============================================================         |  88%
|
|==============================================================        |  88%
|
|===============================================================       |  89%
|
|===============================================================       |  90%
|
|================================================================      |  91%
```

```
  |
  |===============================================================       |  92%
  |
  |================================================================      |  92%
  |
  |================================================================      |  93%
  |
  |================================================================      |  94%
  |
  |=================================================================     |  94%
  |
  |=================================================================     |  95%
  |
  |==================================================================    |  95%
  |
  |==================================================================    |  96%
  |
  |===================================================================   |  97%
## 18:30:15.196: StackedEnsemble_BestOfFamily_AutoML_20201022_182524 [StackedEnsemble best (bui
lt using top model from each algorithm type)] failed: water.exceptions.H2OIllegalArgumentExcept
ion: Failed to find the xval predictions frame. . .  Looks like keep_cross_validation_predictio
ns wasn't set when building the models, or the frame was deleted.
## 18:30:16.209: StackedEnsemble_AllModels_AutoML_20201022_182524 [StackedEnsemble all (built u
sing all AutoML models)] failed: water.exceptions.H2OIllegalArgumentException: Failed to find t
he xval predictions frame. . .  Looks like keep_cross_validation_predictions wasn't set when bu
ilding the models, or the frame was deleted.
  |
  |======================================================================| 100%
```

```
#View AutoML Leaderboard
lb <- aml@leaderboard

print(lb)
```

```
##                                                model_id       auc   logloss
## 1      XGBoost_grid__1_AutoML_20201022_182524_model_2 0.9356780 0.1967353
## 2      XGBoost_grid__1_AutoML_20201022_180957_model_2 0.9356780 0.1967353
## 3 StackedEnsemble_BestOfFamily_AutoML_20201022_180957 0.9350219 0.2166654
## 4          GBM_grid__1_AutoML_20201022_182524_model_2 0.9332449 0.2048228
## 5          GBM_grid__1_AutoML_20201022_180957_model_2 0.9332449 0.2048228
## 6    StackedEnsemble_AllModels_AutoML_20201022_180957 0.9331147 0.2060752
##       aucpr mean_per_class_error      rmse        mse
## 1 0.6303982            0.1673404 0.2488460 0.06192432
## 2 0.6303982            0.1673404 0.2488460 0.06192432
## 3 0.6310380            0.1621620 0.2543648 0.06470145
## 4 0.6190190            0.1749242 0.2540487 0.06454074
## 5 0.6190190            0.1749242 0.2540487 0.06454074
## 6 0.6320502            0.1646369 0.2509036 0.06295260
##
## [53 rows x 7 columns]
```

```
print(lb, n = nrow(lb))
```

```
##                                                  model_id       auc   logloss
## 1       XGBoost_grid__1_AutoML_20201022_182524_model_2 0.9356780 0.1967353
## 2       XGBoost_grid__1_AutoML_20201022_180957_model_2 0.9356780 0.1967353
## 3  StackedEnsemble_BestOfFamily_AutoML_20201022_180957 0.9350219 0.2166654
## 4           GBM_grid__1_AutoML_20201022_182524_model_2 0.9332449 0.2048228
## 5           GBM_grid__1_AutoML_20201022_180957_model_2 0.9332449 0.2048228
## 6     StackedEnsemble_AllModels_AutoML_20201022_180957 0.9331147 0.2060752
## 7           GBM_grid__1_AutoML_20201022_180957_model_1 0.9329840 0.1998278
## 8           GBM_grid__1_AutoML_20201022_182524_model_1 0.9329840 0.1998278
## 9       XGBoost_grid__1_AutoML_20201022_182524_model_1 0.9325188 0.2012839
## 10      XGBoost_grid__1_AutoML_20201022_180957_model_1 0.9325188 0.2012839
## 11                         GBM_5_AutoML_20201022_180957 0.9315950 0.2340688
## 12                         GBM_5_AutoML_20201022_182524 0.9310546 0.2378464
## 13      XGBoost_grid__1_AutoML_20201022_180957_model_4 0.9304512 0.2051540
## 14                         GBM_2_AutoML_20201022_182524 0.9303961 0.2256216
## 15      XGBoost_grid__1_AutoML_20201022_182524_model_4 0.9303429 0.2048351
## 16                         GBM_2_AutoML_20201022_180957 0.9302493 0.2235782
## 17          GBM_grid__1_AutoML_20201022_180957_model_3 0.9302468 0.2123839
## 18                     XGBoost_1_AutoML_20201022_182524 0.9301496 0.2058485
## 19                         GBM_3_AutoML_20201022_182524 0.9296739 0.2264755
## 20                         GBM_3_AutoML_20201022_180957 0.9294973 0.2296272
## 21      XGBoost_grid__1_AutoML_20201022_182524_model_3 0.9294347 0.2087894
## 22      XGBoost_grid__1_AutoML_20201022_180957_model_3 0.9294347 0.2087894
## 23                     XGBoost_3_AutoML_20201022_182524 0.9293704 0.2053813
## 24                     XGBoost_2_AutoML_20201022_182524 0.9291699 0.2121849
## 25                     XGBoost_3_AutoML_20201022_180957 0.9290737 0.2056809
## 26                         GBM_1_AutoML_20201022_182524 0.9290262 0.2223897
## 27          GBM_grid__1_AutoML_20201022_182524_model_3 0.9289238 0.2180805
## 28                         GBM_4_AutoML_20201022_180957 0.9285313 0.2421198
## 29                     XGBoost_2_AutoML_20201022_180957 0.9279268 0.2194895
## 30                         GBM_4_AutoML_20201022_182524 0.9277236 0.2445372
## 31                         GBM_1_AutoML_20201022_180957 0.9263077 0.2388728
## 32          GBM_grid__1_AutoML_20201022_180957_model_4 0.9251986 0.2801308
## 33                     XGBoost_1_AutoML_20201022_180957 0.9212617 0.2567543
## 34                         DRF_1_AutoML_20201022_182524 0.9129661 0.3840614
## 35                         XRT_1_AutoML_20201022_182524 0.9079611 0.2985438
## 36                         GLM_1_AutoML_20201022_180957 0.9069214 0.2400166
## 37                         GLM_1_AutoML_20201022_182524 0.9067093 0.2397973
## 38                         XRT_1_AutoML_20201022_180957 0.9020821 0.3256563
## 39          GBM_grid__1_AutoML_20201022_182524_model_4 0.9018473 0.3381990
## 40          GBM_grid__1_AutoML_20201022_180957_model_5 0.8950291 0.3385584
## 41 DeepLearning_grid__1_AutoML_20201022_182524_model_1 0.8946814 0.2777782
## 42 DeepLearning_grid__3_AutoML_20201022_182524_model_1 0.8857072 0.2913048
## 43                DeepLearning_1_AutoML_20201022_182524 0.8852609 0.2795636
## 44 DeepLearning_grid__2_AutoML_20201022_180957_model_1 0.8842133 0.2556144
## 45 DeepLearning_grid__2_AutoML_20201022_182524_model_1 0.8780543 0.2740725
## 46 DeepLearning_grid__2_AutoML_20201022_180957_model_2 0.8765819 0.5752285
## 47 DeepLearning_grid__1_AutoML_20201022_180957_model_1 0.8719760 0.2679993
## 48 DeepLearning_grid__3_AutoML_20201022_180957_model_1 0.8692540 0.3134560
## 49 DeepLearning_grid__1_AutoML_20201022_180957_model_2 0.8643338 0.5129829
## 50 DeepLearning_grid__2_AutoML_20201022_182524_model_2 0.8597253 0.5016307
## 51 DeepLearning_grid__1_AutoML_20201022_182524_model_2 0.8540015 0.3872542
## 52                         DRF_1_AutoML_20201022_180957 0.8306920 1.6339417
## 53                DeepLearning_1_AutoML_20201022_180957 0.8300519 0.3654677
##        aucpr mean_per_class_error      rmse        mse
## 1  0.6303982            0.1673404 0.2488460 0.06192432
## 2  0.6303982            0.1673404 0.2488460 0.06192432
## 3  0.6310380            0.1621620 0.2543648 0.06470145
## 4  0.6190190            0.1749242 0.2540487 0.06454074
## 5  0.6190190            0.1749242 0.2540487 0.06454074
## 6  0.6320502            0.1646369 0.2509036 0.06295260
```

```
## 7  0.6147040              0.1774892 0.2508573 0.06292938
## 8  0.6147040              0.1774892 0.2508573 0.06292938
## 9  0.6178210              0.1624412 0.2513834 0.06319361
## 10 0.6178210              0.1624412 0.2513834 0.06319361
## 11 0.6141615              0.1808666 0.2723323 0.07416490
## 12 0.6126850              0.1844423 0.2743283 0.07525604
## 13 0.6066089              0.1688106 0.2540210 0.06452666
## 14 0.6138666              0.1863686 0.2666936 0.07112549
## 15 0.6054021              0.1770424 0.2538020 0.06441547
## 16 0.6138750              0.1829718 0.2654296 0.07045286
## 17 0.6127362              0.1677809 0.2586092 0.06687871
## 18 0.6103736              0.1763310 0.2532572 0.06413921
## 19 0.6117829              0.1738852 0.2675636 0.07159029
## 20 0.6107208              0.1788176 0.2694537 0.07260529
## 21 0.6006786              0.1745010 0.2564203 0.06575138
## 22 0.6006786              0.1745010 0.2564203 0.06575138
## 23 0.6137425              0.1819538 0.2527652 0.06389025
## 24 0.6081160              0.1686919 0.2537222 0.06437494
## 25 0.6131600              0.1703810 0.2529245 0.06397081
## 26 0.6110695              0.1727226 0.2644352 0.06992598
## 27 0.6108366              0.1762797 0.2619030 0.06859316
## 28 0.6087903              0.1698850 0.2766374 0.07652827
## 29 0.6077885              0.1752221 0.2550462 0.06504856
## 30 0.6059858              0.1856767 0.2778675 0.07721037
## 31 0.6042230              0.1889074 0.2735924 0.07485281
## 32 0.5789561              0.1869553 0.2938106 0.08632469
## 33 0.5922735              0.1922122 0.2653683 0.07042034
## 34 0.5646684              0.1847008 0.2838098 0.08054803
## 35 0.5580078              0.1865876 0.2825120 0.07981301
## 36 0.5496125              0.1970992 0.2670000 0.07128898
## 37 0.5506528              0.2074031 0.2667996 0.07118201
## 38 0.5464658              0.1895219 0.2832367 0.08022302
## 39 0.5068837              0.1905388 0.3140848 0.09864925
## 40 0.5146294              0.1915177 0.3142071 0.09872612
## 41 0.5253416              0.2154294 0.2907789 0.08455237
## 42 0.4886886              0.2224349 0.2974873 0.08849869
## 43 0.4933892              0.2194074 0.2768825 0.07666391
## 44 0.4938987              0.2259416 0.2762350 0.07630578
## 45 0.4869933              0.2303299 0.2778627 0.07720770
## 46 0.4767087              0.2491287 0.4220132 0.17809512
## 47 0.5003377              0.2511806 0.2825800 0.07985148
## 48 0.4669421              0.2442798 0.3099441 0.09606537
## 49 0.4669419              0.2591776 0.4013121 0.16105144
## 50 0.4443261              0.2552644 0.4026107 0.16209537
## 51 0.4682556              0.2576376 0.3332676 0.11106727
## 52 0.4493111              0.1980239 0.3036473 0.09220170
## 53 0.3886545              0.2766686 0.3449946 0.11902127
##
## [53 rows x 7 columns]
```

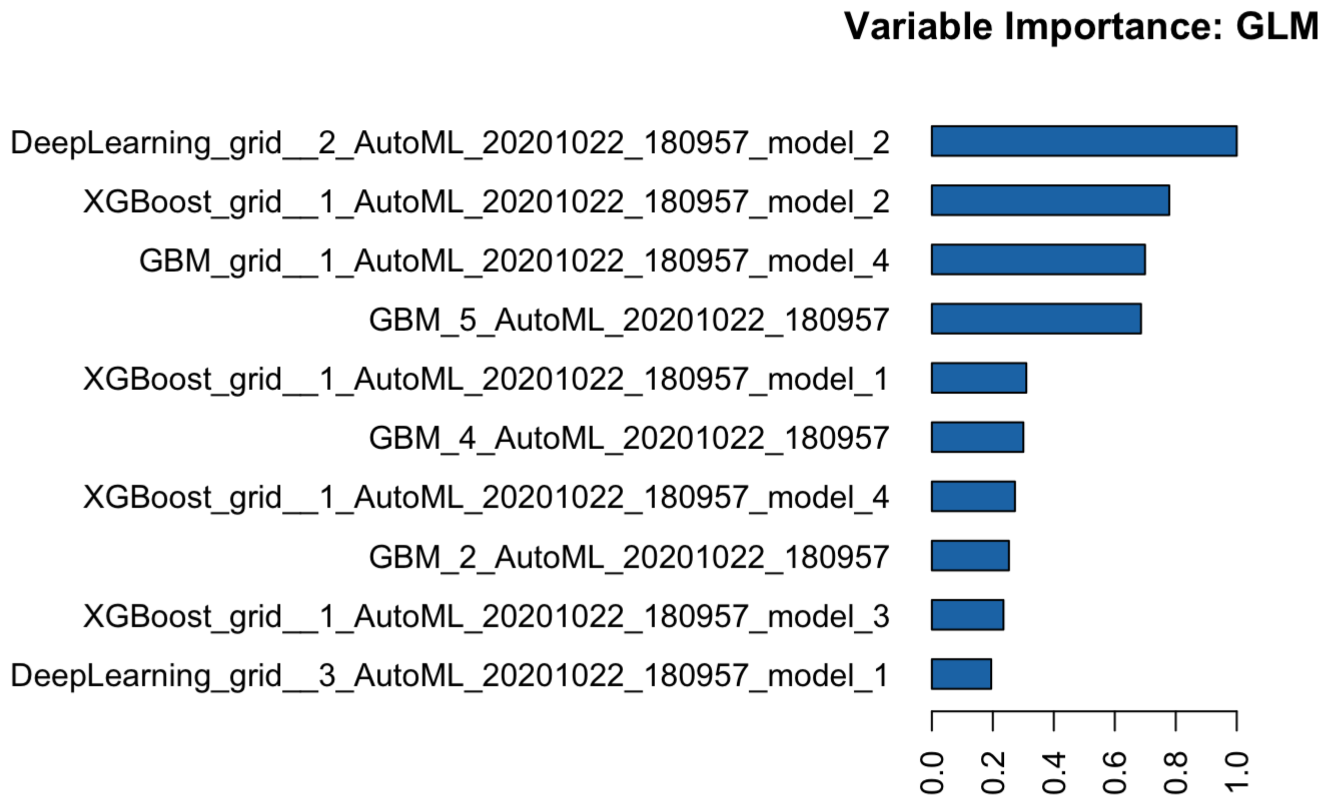
```
# Ensemble Exploration ----

model_ids <- as.data.frame(aml@leaderboard$model_id)[,1]
se <- h2o.getModel(grep("StackedEnsemble_AllModels", model_ids, value = TRUE)[1])
metalearner <- h2o.getModel(se@model$metalearner$name)

h2o.varimp(metalearner)
```

```
##                                                variable relative_importance
## 1  DeepLearning_grid__2_AutoML_20201022_180957_model_2         0.346650843
## 2       XGBoost_grid__1_AutoML_20201022_180957_model_2         0.269919970
## 3           GBM_grid__1_AutoML_20201022_180957_model_4         0.242369833
## 4                         GBM_5_AutoML_20201022_180957         0.237822020
## 5       XGBoost_grid__1_AutoML_20201022_180957_model_1         0.107409688
## 6                         GBM_4_AutoML_20201022_180957         0.104062635
## 7       XGBoost_grid__1_AutoML_20201022_180957_model_4         0.094490941
## 8                         GBM_2_AutoML_20201022_180957         0.087640253
## 9       XGBoost_grid__1_AutoML_20201022_180957_model_3         0.081433076
## 10 DeepLearning_grid__3_AutoML_20201022_180957_model_1         0.067484828
## 11                        GBM_1_AutoML_20201022_180957         0.063001922
## 12                        GBM_3_AutoML_20201022_180957         0.030522733
## 13               DeepLearning_1_AutoML_20201022_180957         0.026747155
## 14          GBM_grid__1_AutoML_20201022_180957_model_5         0.020859414
## 15 DeepLearning_grid__1_AutoML_20201022_180957_model_2         0.006714708
## 16          GBM_grid__1_AutoML_20201022_180957_model_2         0.000000000
## 17          GBM_grid__1_AutoML_20201022_180957_model_1         0.000000000
## 18          GBM_grid__1_AutoML_20201022_180957_model_3         0.000000000
## 19                    XGBoost_3_AutoML_20201022_180957         0.000000000
## 20                    XGBoost_2_AutoML_20201022_180957         0.000000000
## 21                    XGBoost_1_AutoML_20201022_180957         0.000000000
## 22                        GLM_1_AutoML_20201022_180957         0.000000000
## 23                        XRT_1_AutoML_20201022_180957         0.000000000
## 24 DeepLearning_grid__2_AutoML_20201022_180957_model_1         0.000000000
## 25 DeepLearning_grid__1_AutoML_20201022_180957_model_1         0.000000000
## 26                        DRF_1_AutoML_20201022_180957         0.000000000
##    scaled_importance  percentage
## 1         1.00000000 0.193970690
## 2         0.77865084 0.151035441
## 3         0.69917566 0.135619586
## 4         0.68605637 0.133074828
## 5         0.30984978 0.060101776
## 6         0.30019438 0.058228911
## 7         0.27258247 0.052873009
## 8         0.25281996 0.049039663
## 9         0.23491383 0.045566397
## 10        0.19467666 0.037761566
## 11        0.18174461 0.035253127
## 12        0.08805036 0.017079190
## 13        0.07715878 0.014966541
## 14        0.06017413 0.011672018
## 15        0.01937024 0.003757258
## 16        0.00000000 0.000000000
## 17        0.00000000 0.000000000
## 18        0.00000000 0.000000000
## 19        0.00000000 0.000000000
## 20        0.00000000 0.000000000
## 21        0.00000000 0.000000000
## 22        0.00000000 0.000000000
## 23        0.00000000 0.000000000
## 24        0.00000000 0.000000000
## 25        0.00000000 0.000000000
## 26        0.00000000 0.000000000
```

```
h2o.varimp_plot(metalearner)
```



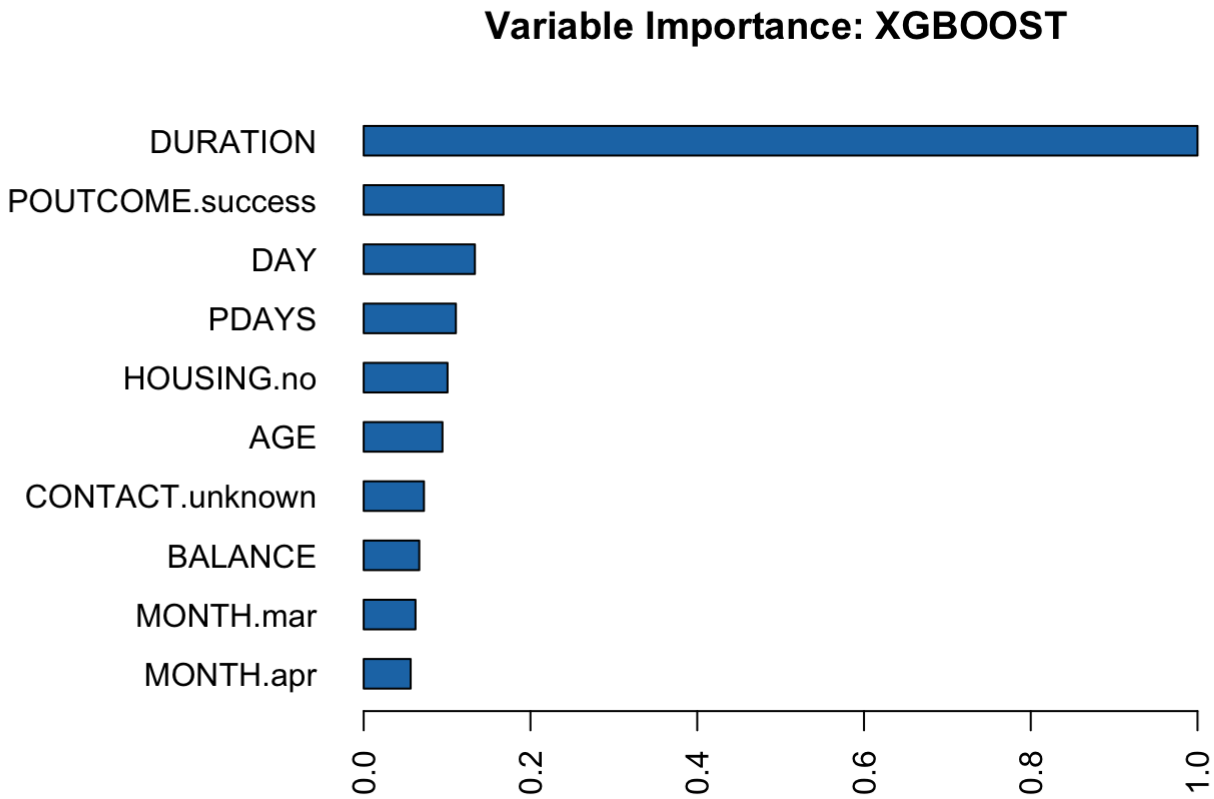

```
#  Baselearner Variable Importance

xgb <- h2o.getModel(grep("XGBoost", model_ids, value = TRUE)[1])

h2o.varimp(xgb)
```

```
## Variable Importances:
##           variable relative_importance scaled_importance percentage
## 1         DURATION        12787.352539          1.000000   0.412424
## 2 POUTCOME.success         2143.389404          0.167618   0.069130
## 3              DAY         1705.303711          0.133359   0.055000
## 4            PDAYS         1413.857666          0.110567   0.045600
## 5       HOUSING.no         1285.888184          0.100559   0.041473
##
## ---
##             variable relative_importance scaled_importance percentage
## 46 EDUCATION.unknown           12.700191          0.000993   0.000410
## 47      JOB.services            8.025841          0.000628   0.000259
## 48       JOB.unknown            6.445179          0.000504   0.000208
## 49       DEFAULT.yes            5.249918          0.000411   0.000169
## 50  JOB.entrepreneur            4.777218          0.000374   0.000154
## 51    JOB.unemployed            1.840618          0.000144   0.000059
```

```
h2o.varimp_plot(xgb)
```


Variable Importance: XGBOOST
DURATION
POUTCOME.success
DAY
PDAYS
HOUSING.no
AGE
CONTACT.unknown
BALANCE
MONTH.mar
MONTH.apr
0.0
0.2
0.4
0.6
0.8
1.0