Model Selection - death_2year

Eduardo Yuki Yada

Global parameters

```r
k <- params$k # Number of folds for cross validation
grid_size <- params$grid_size # Number of parameter combination to tune on each model
repeats <- params$repeats
RUN_ALL_MODELS <- params$RUN_ALL_MODELS
Hmisc::list.tree(params)
```

```
##  params = list 5 (952 bytes)
## .  outcome_column = character 1= death_2year
## .  k = double 1= 10
## .  grid_size = double 1= 30
## .  repeats = double 1= 2
## .  RUN_ALL_MODELS = logical 1= TRUE
```

Minutes to run: 0

Imports

```r
library(tidyverse)
library(yaml)
library(tidymodels)
library(usemodels)
library(vip)
library(bonsai)
library(lightgbm)
library(caret)
library(pROC)

source("aux_functions.R")
```

Minutes to run: 0

Loading data

```r
load('dataset/processed_data.RData')
load('dataset/processed_dictionary.RData')

columns_list <- yaml.load_file("./auxiliar/columns_list.yaml")

outcome_column <- params$outcome_column
features_list <- params$features_list

df <- mutate(df, across(where(is.character), as.factor))
```

Minutes to run: 0.007

```
dir.create(file.path("./auxiliar/model_selection/hyperparameters/"),
           showWarnings = FALSE,
           recursive = TRUE)

dir.create(file.path("./auxiliar/model_selection/performance/"),
           showWarnings = FALSE,
           recursive = TRUE)
```

Minutes to run: 0

Eligible features

```
cat_features_list = readRDS(sprintf(
  "./auxiliar/significant_columns/categorical_%s.rds",
  outcome_column
))

num_features_list = readRDS(sprintf(
  "./auxiliar/significant_columns/numerical_%s.rds",
  outcome_column
))

features_list = c(cat_features_list, num_features_list)
```

Minutes to run: 0

```
eligible_columns = df_names %>%
  filter(momento.aquisicao == 'Admissão t0') %>%
  .$variable.name

exception_columns = c('death_intraop', 'death_intraop_1', 'disch_outcomes_t0')

correlated_columns = c('year_procedure_1', # com year_adm_t0
                       'age_surgery_1', # com age
                       'admission_t0', # com admission_pre_t0_count
                       'atb', # com meds_antimicrobianos
                       'classe_meds_cardio_qtde', # com classe_meds_qtde
                       'suporte_hemod', # com proced_invasivos_qtde,
                       'radiografia', # com exames_imagem_qtde
                       'ecg' # com metodos_graficos_qtde
                       )

eligible_features = eligible_columns %>%
  base::intersect(c(columns_list$categorical_columns, columns_list$numerical_columns)) %>%
  setdiff(c(exception_columns, correlated_columns))

features = base::intersect(eligible_features, features_list)

gluedown::md_order(features, seq = TRUE, pad = TRUE)
```

```
## 01. sex
## 02. age
## 03. education_level
## 04. underlying_heart_disease
## 05. heart_disease
## 06. nyha_basal
## 07. hypertension
## 08. prior_mi
## 09. heart_failure
## 10. af
## 11. cardiac_arrest
## 12. valvopathy
## 13. diabetes
## 14. renal_failure
## 15. hemodialysis
## 16. stroke
## 17. copd
## 18. comorbidities_count
## 19. procedure_type_1
## 20. reop_type_1
## 21. procedure_type_new
## 22. cied_final_1
## 23. cied_final_group_1
## 24. admission_pre_t0_count
## 25. admission_pre_t0_180d
## 26. year_adm_t0
## 27. icu_t0
## 28. dialysis_t0
## 29. admission_t0_emergency
## 30. aco
## 31. antiarritmico
## 32. ieca_bra
## 33. dva
## 34. digoxina
## 35. estatina
## 36. diuretico
## 37. vasodilatador
## 38. insuf_cardiaca
## 39. espironolactona
## 40. antiplaquetario_ev
## 41. insulina
## 42. psicofarmacos
## 43. antifungico
## 44. antiviral
## 45. classe_meds_qtde
## 46. meds_cardiovasc_qtde
## 47. meds_antimicrobianos
## 48. vni
## 49. ventilacao_mecanica
## 50. transplante_cardiaco
## 51. outros_proced_cirurgicos
## 52. icp
## 53. angioplastia
## 54. cateterismo
## 55. cateter_venoso_central
## 56. proced_invasivos_qtde
## 57. transfusao
## 58. interconsulta
## 59. equipe_multiprof
## 60. holter
## 61. teste_esforco
## 62. tilt_teste
## 63. metodos_graficos_qtde
## 64. laboratorio
## 65. cultura
## 66. analises_clinicas_qtde
## 67. citologia
## 68. histopatologia_qtde
## 69. angio_tc
## 70. cintilografia
## 71. ecocardiograma
## 72. endoscopia
## 73. flebografia
## 74. pet_ct
## 75. ultrassom
## 76. tomografia
## 77. ressonancia
## 78. exames_imagem_qtde
## 79. bic
## 80. hospital_stay
```

Minutes to run: 0

Train test split (70%/30%)

```
set.seed(42)

if (outcome_column == 'readmission_30d') {
  df_split <- readRDS("./dataset/split_object.rds")
} else {
  df_split <- initial_split(df, prop = .7, strata = all_of(outcome_column))
}

df_train <- training(df_split) %>% dplyr::select(all_of(c(features, outcome_column)))
df_test <- rsample::testing(df_split) %>% dplyr::select(all_of(c(features, outcome_column)))

df_folds <- vfold_cv(df_train, v = k,
                     strata = all_of(outcome_column))
```

Minutes to run: 0.001

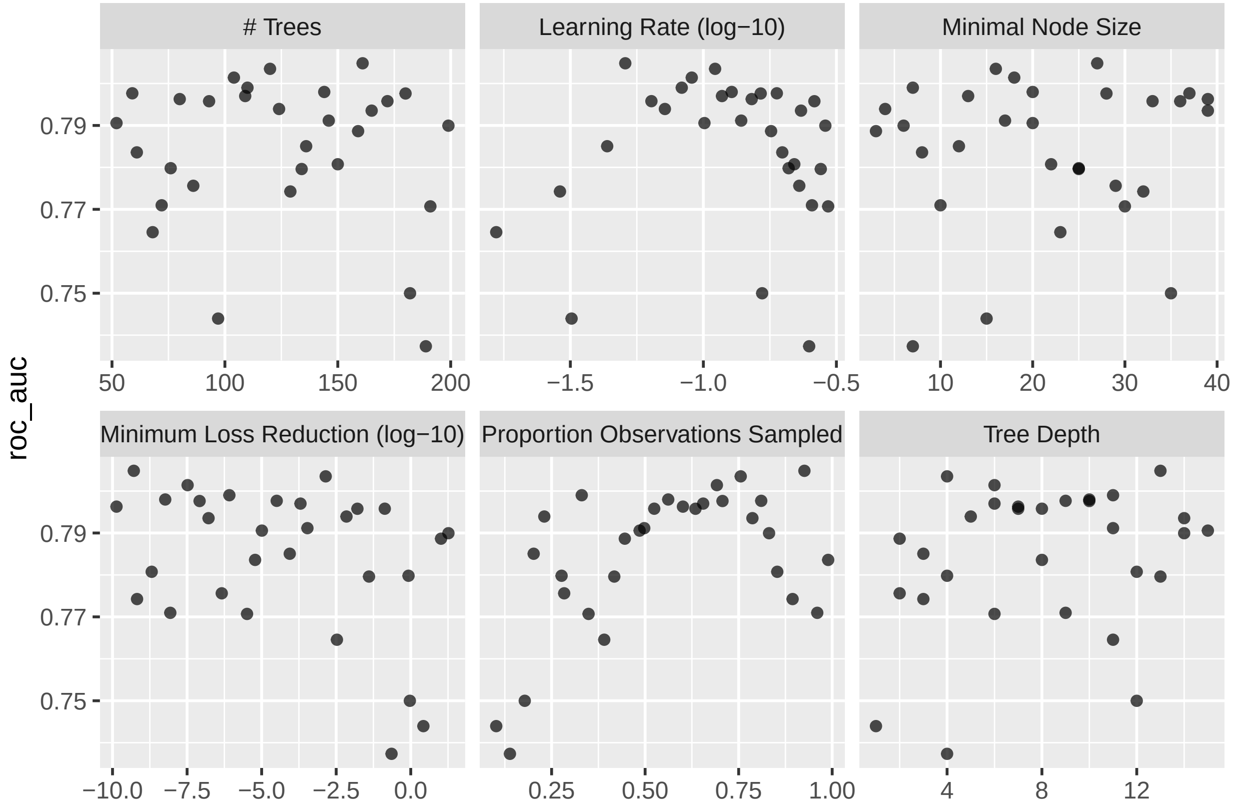
Boosted Tree (XGBoost)

```
xgboost_recipe <-
  recipe(formula = sprintf("%s ~ .", outcome_column) %>% as.formula, data = df_train) %>%
  step_novel(all_nominal_predictors()) %>%
  step_unknown(all_nominal_predictors()) %>%
  step_other(all_nominal_predictors(), threshold = 0.05, other = ".merged") %>%
  step_dummy(all_nominal_predictors())

xgboost_spec <- boost_tree(
  trees = tune(),
  min_n = tune(),
  tree_depth = tune(),
  learn_rate = tune(),
  loss_reduction = tune(),
  sample_size = tune()
) %>%
  set_engine("xgboost",
             nthread = 8) %>%
  set_mode("classification")
```

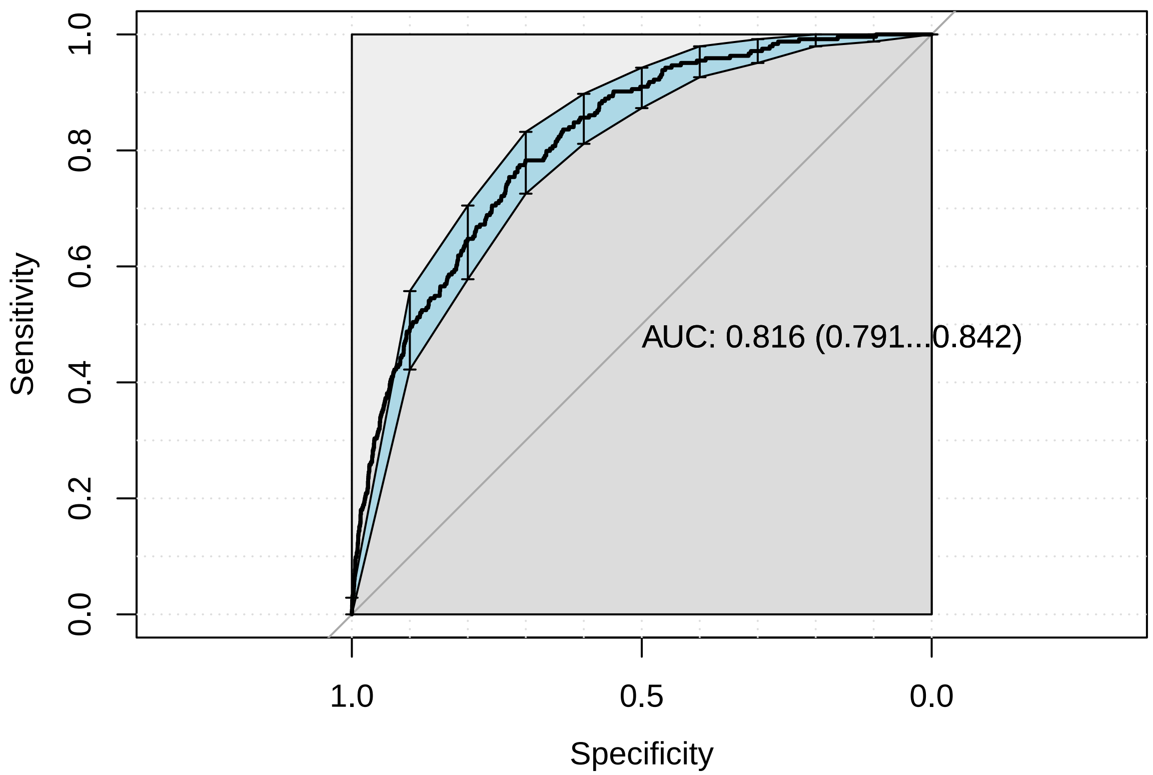
```
xgboost_grid <- grid_latin_hypercube(
  trees(range = c(50L, 200L)),
  min_n(),
  tree_depth(),
  learn_rate(range = c(0.01, 0.3), trans = NULL),
  loss_reduction(),
  sample_prop(range = c(1/10, 1), trans = NULL),
  size = grid_size
)

xgboost_workflow <-
  workflow() %>%
  add_recipe(xgboost_recipe) %>%
  add_model(xgboost_spec)

xgboost_tune <-
  xgboost_workflow %>%
  tune_grid(resamples = df_folds,
            grid = xgboost_grid)

xgboost_tune %>%
  show_best("roc_auc")
```

```
## # A tibble: 5 x 12
##   trees min_n tree_depth learn_rate loss_reduction sample_size .metric .estimator  mean     n std_err .config
##   <int> <int>      <int>      <dbl>          <dbl>       <dbl> <chr>   <chr>      <dbl> <int>   <dbl> <chr>
## 1   161    27         13     0.0509       5.16e-10       0.926 roc_auc binary     0.805    10 0.00962 Preproc
## 2   120    16          4     0.111        1.40e- 3       0.755 roc_auc binary     0.804    10 0.0103  Preproc
## 3   104    18          6     0.0905       3.29e- 8       0.692 roc_auc binary     0.801    10 0.0108  Preproc
## 4   110     7         11     0.0829       8.26e- 7       0.331 roc_auc binary     0.799    10 0.00877 Preproc
## 5   144    20         10     0.128        5.85e- 9       0.562 roc_auc binary     0.798    10 0.0117  Preproc
```

```
best_xgboost <- xgboost_tune %>%
  select_best("roc_auc")

autoplot(xgboost_tune, metric = "roc_auc")
```

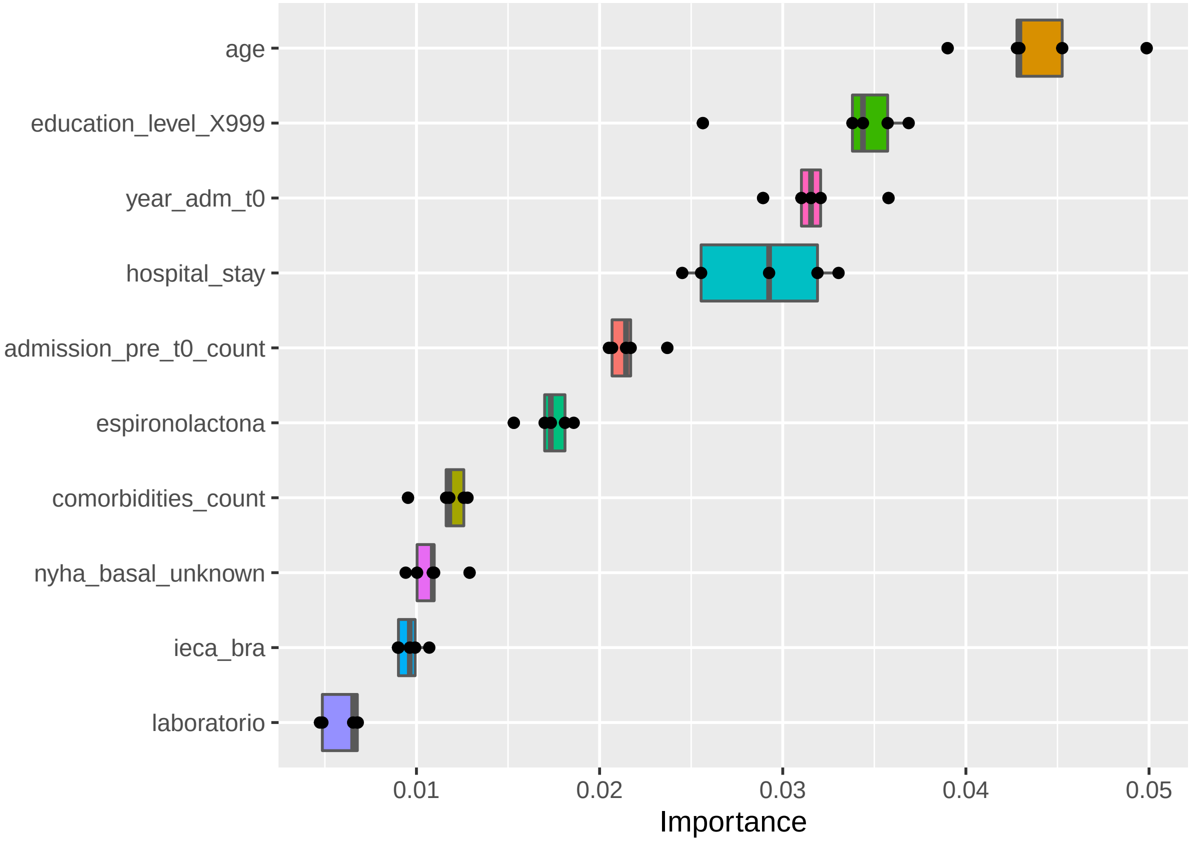


```
final_xgboost_workflow <-
  xgboost_workflow %>%
  finalize_workflow(best_xgboost)

last_xgboost_fit <-
  final_xgboost_workflow %>%
  last_fit(df_split)

final_xgboost_fit <- extract_workflow(last_xgboost_fit)

xgboost_auc <- validation(final_xgboost_fit, df_test)
```

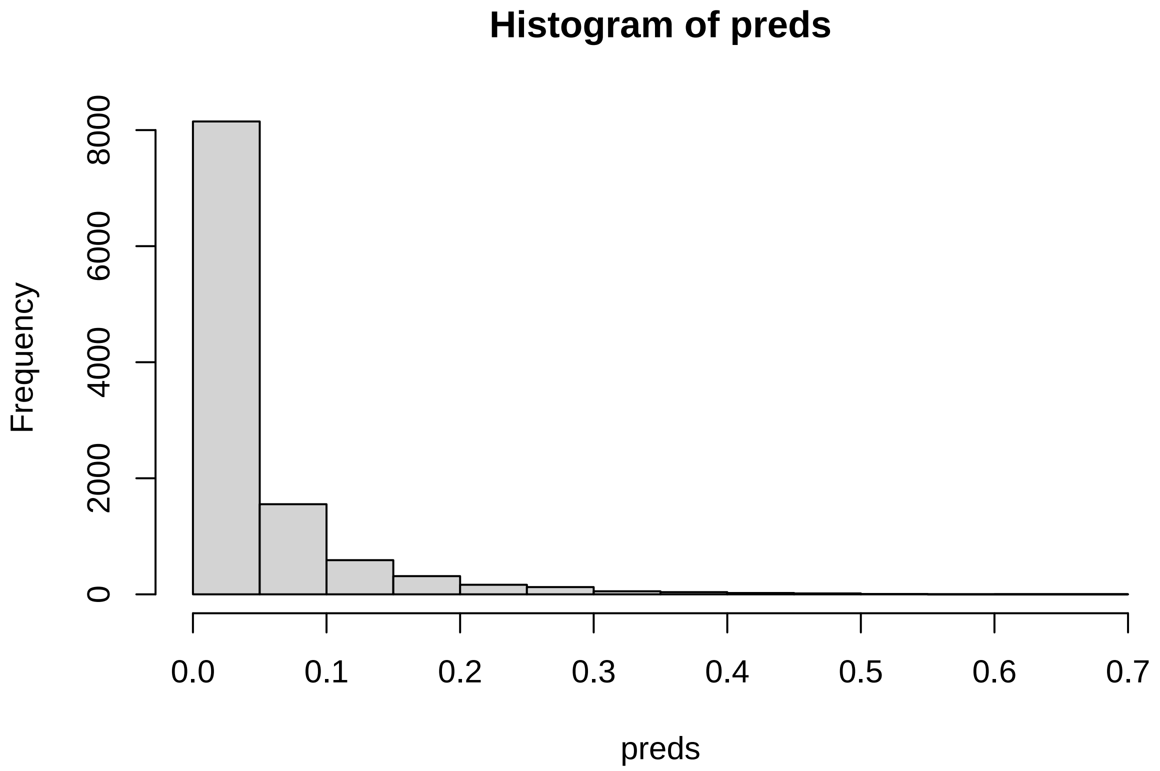


```
## |
```

```
extract_vip(final_xgboost_fit, pred_wrapper = predict,
            reference_class = "0")
```



```
xgboost_parameters <- xgboost_tune %>%
  show_best("roc_auc", n = 1) %>%
  select(trees, min_n, tree_depth, learn_rate, loss_reduction) %>%
  as.list

saveRDS(
  xgboost_parameters,
  file = sprintf(
    "./auxiliar/model_selection/hyperparameters/xgboost_%s.rds",
    outcome_column
  )
)

preds <- predict(final_xgboost_fit, new_data = df_train, type = "prob") %>%
  rename_at(vars(starts_with(".pred_")), ~ str_remove(., ".pred_")) %>%
  .$`1`

hist(preds)
```



Minutes to run: 12.607

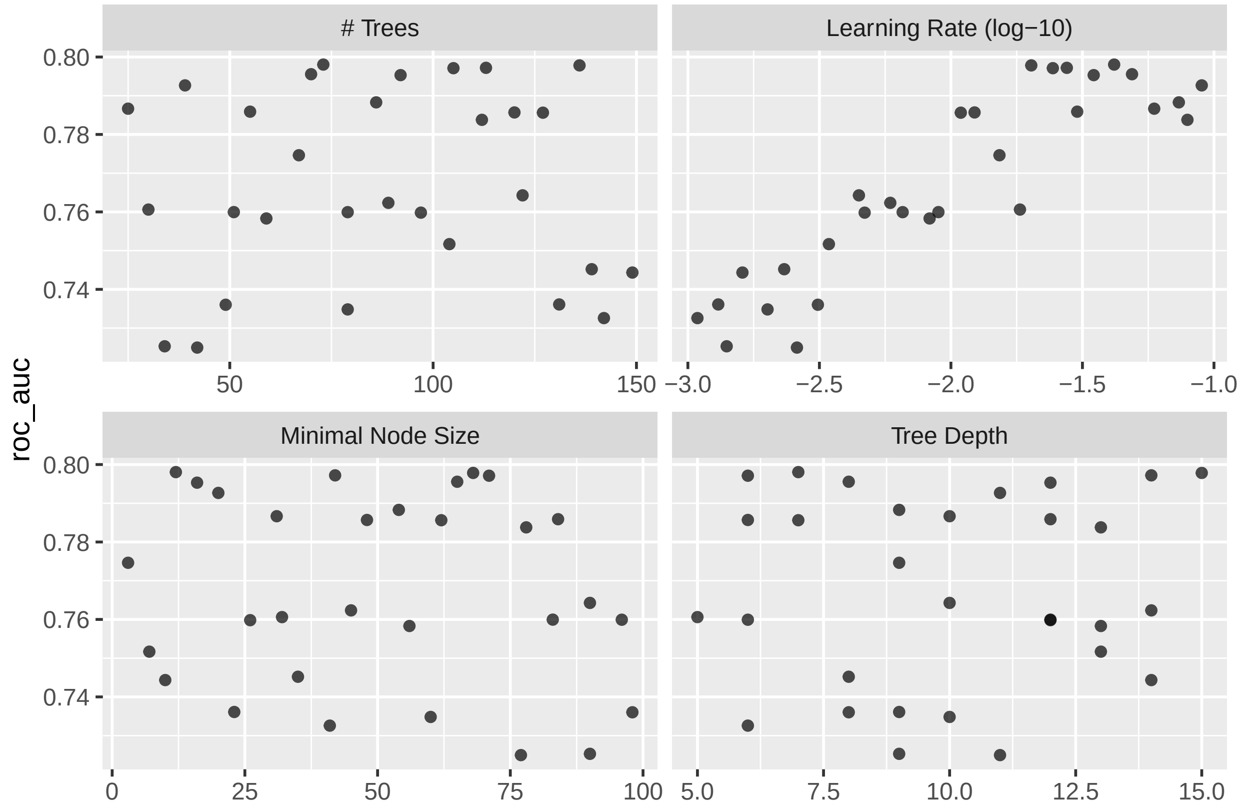
Boosted Tree (LightGBM)

```
lightgbm_recipe <-
  recipe(formula = sprintf("%s ~ .", outcome_column) %>% as.formula, data = df_train) %>%
  step_novel(all_nominal_predictors()) %>%
  step_unknown(all_nominal_predictors()) %>%
  step_other(all_nominal_predictors(), threshold = 0.05, other = ".merged") %>%
  step_dummy(all_nominal_predictors())

lightgbm_spec <- boost_tree(
  trees = tune(),
  min_n = tune(),
  tree_depth = tune(),
  learn_rate = tune(),
  sample_size = 1
) %>%
  set_engine("lightgbm",
             nthread = 8) %>%
  set_mode("classification")

lightgbm_grid <- grid_latin_hypercube(
  trees(range = c(25L, 150L)),
  min_n(range = c(2L, 100L)),
  tree_depth(range = c(5L, 15L)),
  learn_rate(range = c(-3, -1), trans = log10_trans()),
  size = grid_size
)

lightgbm_workflow <-
```

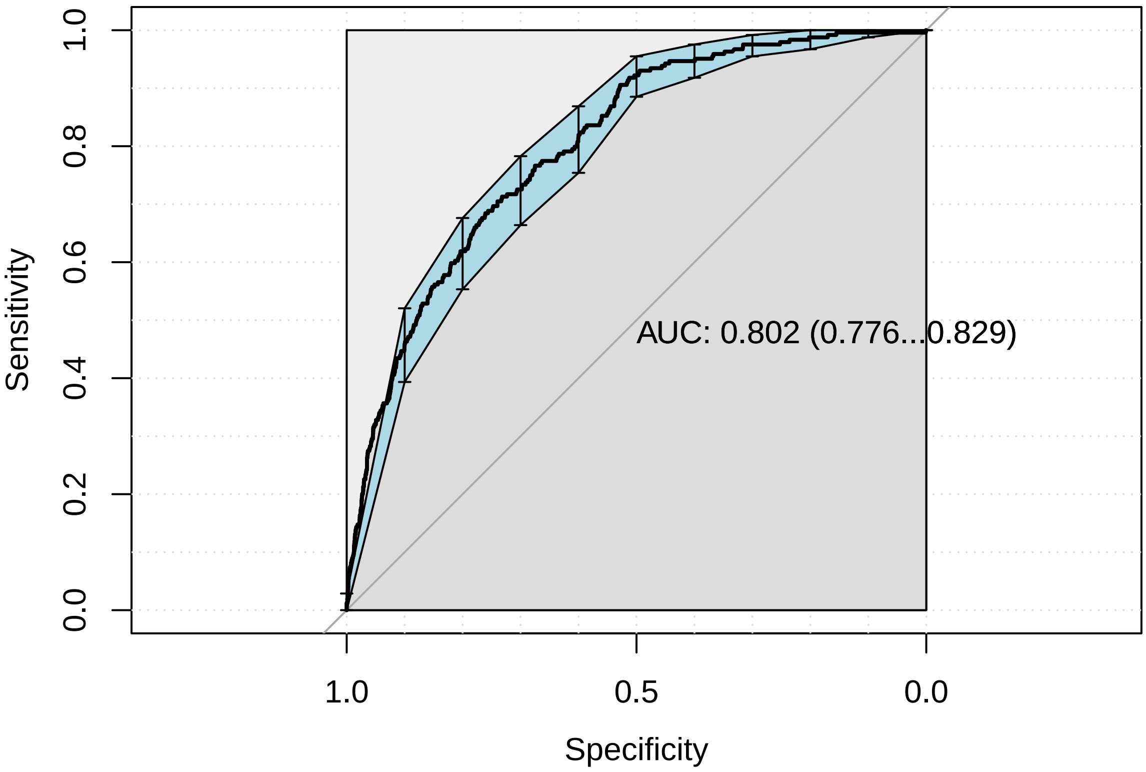
```
  workflow() %>%
  add_recipe(lightgbm_recipe) %>%
  add_model(lightgbm_spec)

lightgbm_tune <-
  lightgbm_workflow %>%
  tune_grid(resamples = df_folds,
            grid = lightgbm_grid)

lightgbm_tune %>%
  show_best("roc_auc")
```

```
## # A tibble: 5 x 10
##   trees min_n tree_depth learn_rate .metric .estimator  mean     n std_err .config
##   <int> <int>      <int>      <dbl> <chr>   <chr>      <dbl> <int>   <dbl> <chr>
## 1    73    12          7     0.0417 roc_auc binary     0.798    10 0.00793 Preprocessor1_Model04
## 2   136    68         15     0.0202 roc_auc binary     0.798    10 0.00914 Preprocessor1_Model21
## 3   113    42         14     0.0276 roc_auc binary     0.797    10 0.00857 Preprocessor1_Model13
## 4   105    71          6     0.0244 roc_auc binary     0.797    10 0.00757 Preprocessor1_Model22
## 5    70    65          8     0.0488 roc_auc binary     0.796    10 0.00875 Preprocessor1_Model20
```

```
best_lightgbm <- lightgbm_tune %>%
  select_best("roc_auc")

autoplot(lightgbm_tune, metric = "roc_auc")
```



```
final_lightgbm_workflow <-
  lightgbm_workflow %>%
  finalize_workflow(best_lightgbm)

last_lightgbm_fit <-
  final_lightgbm_workflow %>%
  last_fit(df_split)

final_lightgbm_fit <- extract_workflow(last_lightgbm_fit)

lightgbm_auc <- validation(final_lightgbm_fit, df_test)
```



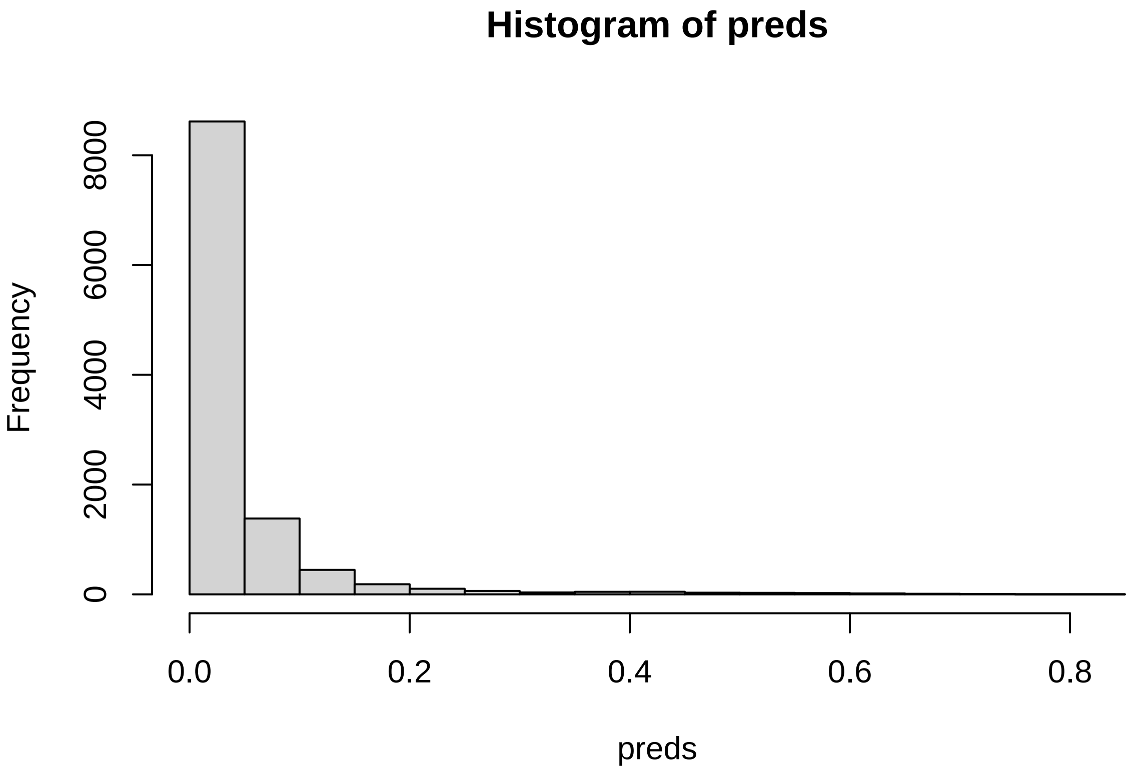
```
##   |
##
##        'Positive' Class : 0
##
```

```
lightgbm_parameters <- lightgbm_tune %>%
  show_best("roc_auc", n = 1) %>%
  select(-.metric, -.estimator, -.config, -mean, -n, -std_err) %>%
  as.list

Hmisc::list.tree(lightgbm_parameters)
```

```
##  lightgbm_parameters = list 4 (736 bytes)
## .  trees = integer 1= 73
## .  min_n = integer 1= 12
## .  tree_depth = integer 1= 7
## .  learn_rate = double 1= 0.041746
```

```
saveRDS(
  lightgbm_parameters,
  file = sprintf(
    "./auxiliar/model_selection/hyperparameters/lightgbm_%s.rds",
    outcome_column
  )
)
```

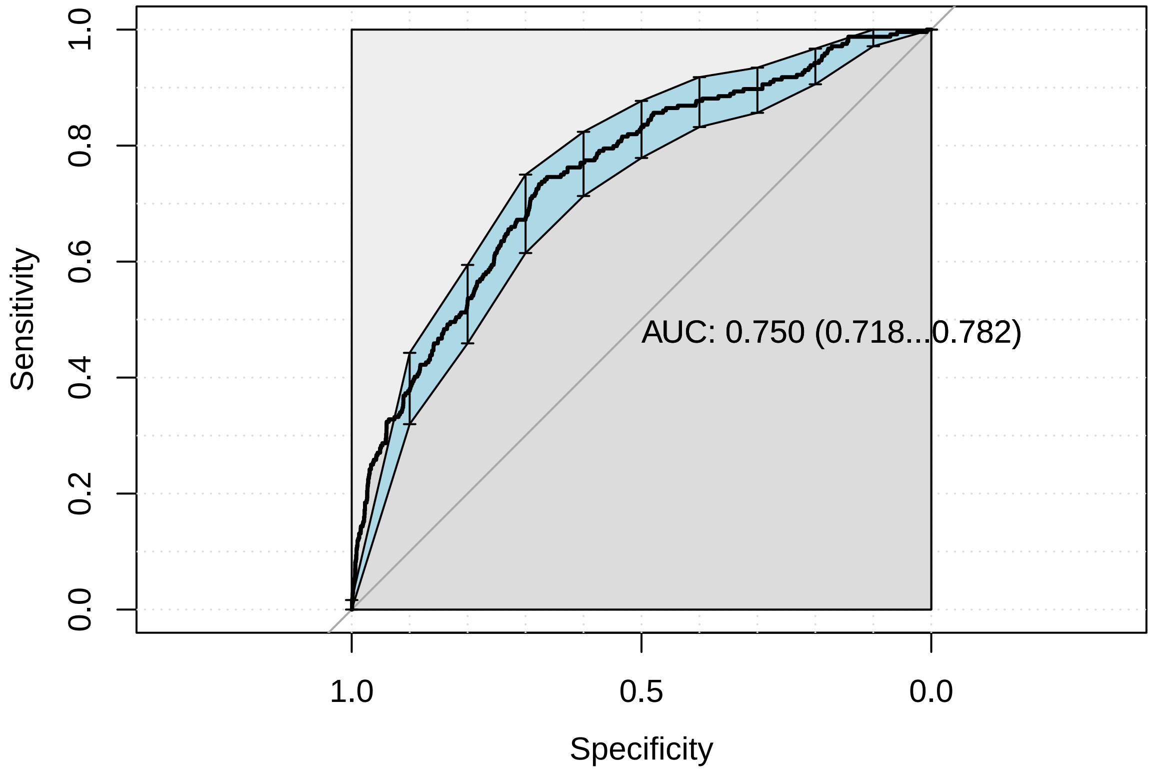
Minutes to run: 3.638



Minutes to run: 0.005

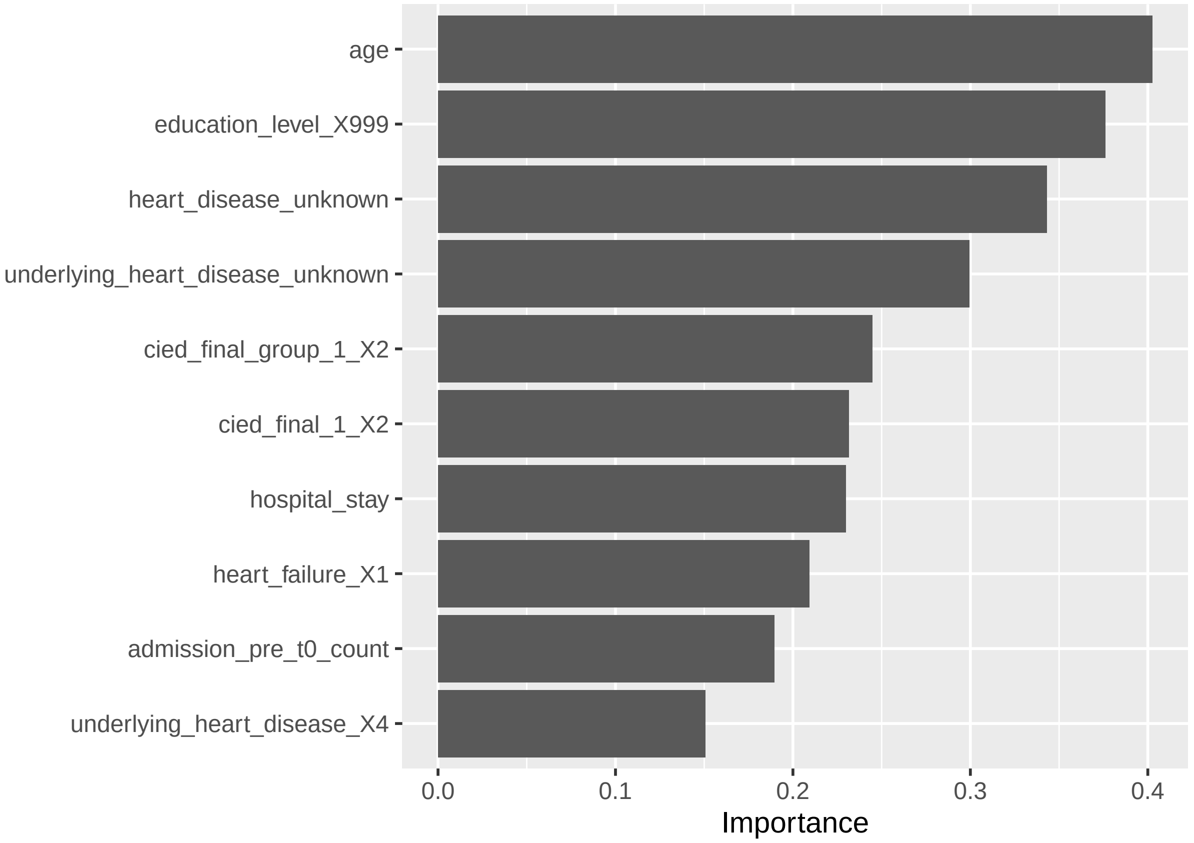
GLM

```
glmnet_recipe <-
  recipe(formula = sprintf("%s ~ .", outcome_column) %>% as.formula, data = df_train) %>%
  step_novel(all_nominal_predictors()) %>%
  step_unknown(all_nominal_predictors()) %>%
  step_other(all_nominal_predictors(), threshold = 0.05, other = ".merged") %>%
  step_dummy(all_nominal_predictors()) %>%
  step_zv(all_predictors()) %>%
  step_normalize(all_numeric_predictors())

glmnet_spec <-
  logistic_reg(penalty = 0) %>%
  set_mode("classification") %>%
  set_engine("glmnet")

glmnet_workflow <-
  workflow() %>%
  add_recipe(glmnet_recipe) %>%
  add_model(glmnet_spec)

glm_fit <- glmnet_workflow %>%
  fit(df_train)

glmnet_auc <- validation(glm_fit, df_test)
```

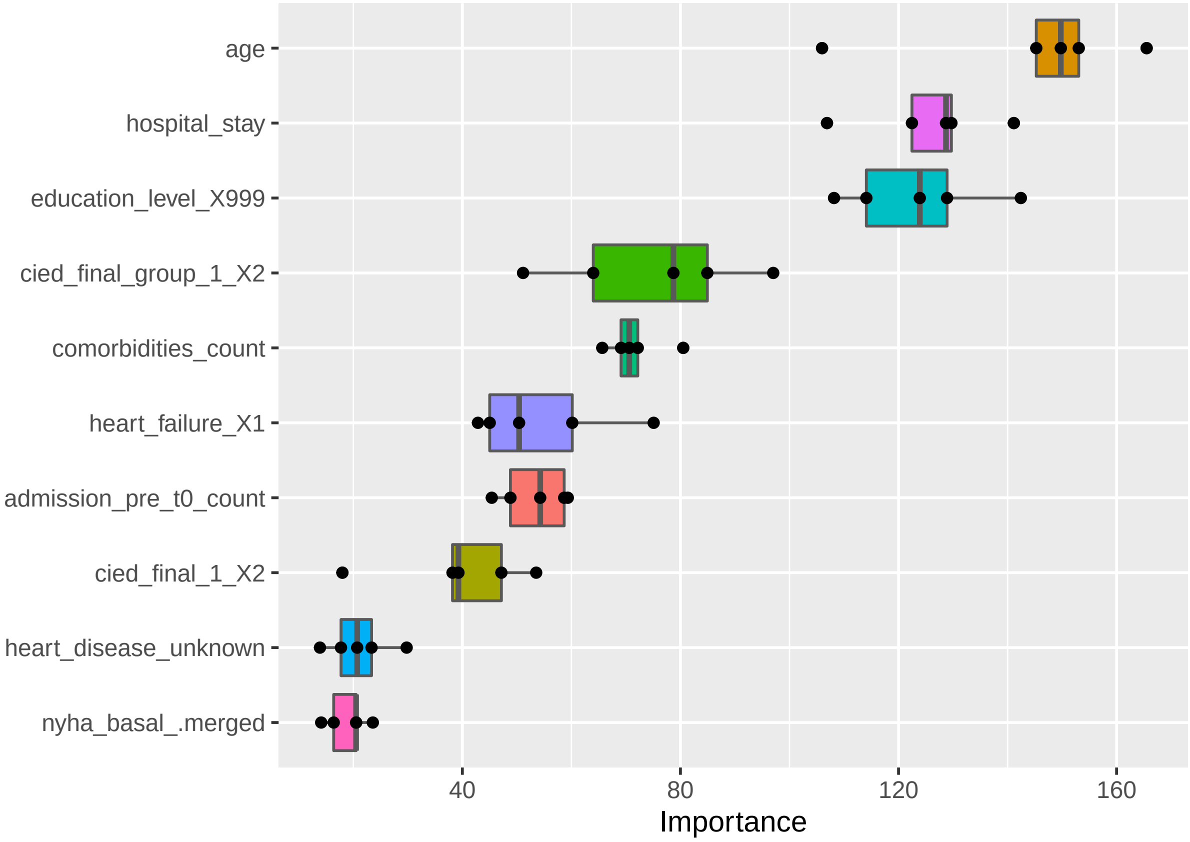


```
##    |
##
##                Accuracy : 0.6797
##                  95% CI : (0.6662, 0.693)
##     No Information Rate : 0.9484
##     P-Value [Acc > NIR] : 1
##
##                   Kappa : 0.1114
##
##  Mcnemar’s Test P-Value : <2e-16
##
##             Sensitivity : 0.6768
##             Specificity : 0.7336
##          Pos Pred Value : 0.9790
##          Neg Pred Value : 0.1099
##              Prevalence : 0.9484
##          Detection Rate : 0.6419
##    Detection Prevalence : 0.6556
##       Balanced Accuracy : 0.7052
##
##        ’Positive’ Class : 0
##
```

```r
pfun_glmnet <- function(object, newdata) predict(object, newx = newdata)

extract_vip(glm_fit, pred_wrapper = pfun_glmnet,
            reference_class = "1", method = 'model')
```



```r
extract_vip(glm_fit, pred_wrapper = pfun_glmnet,
            reference_class = "1", method = 'permute')
```



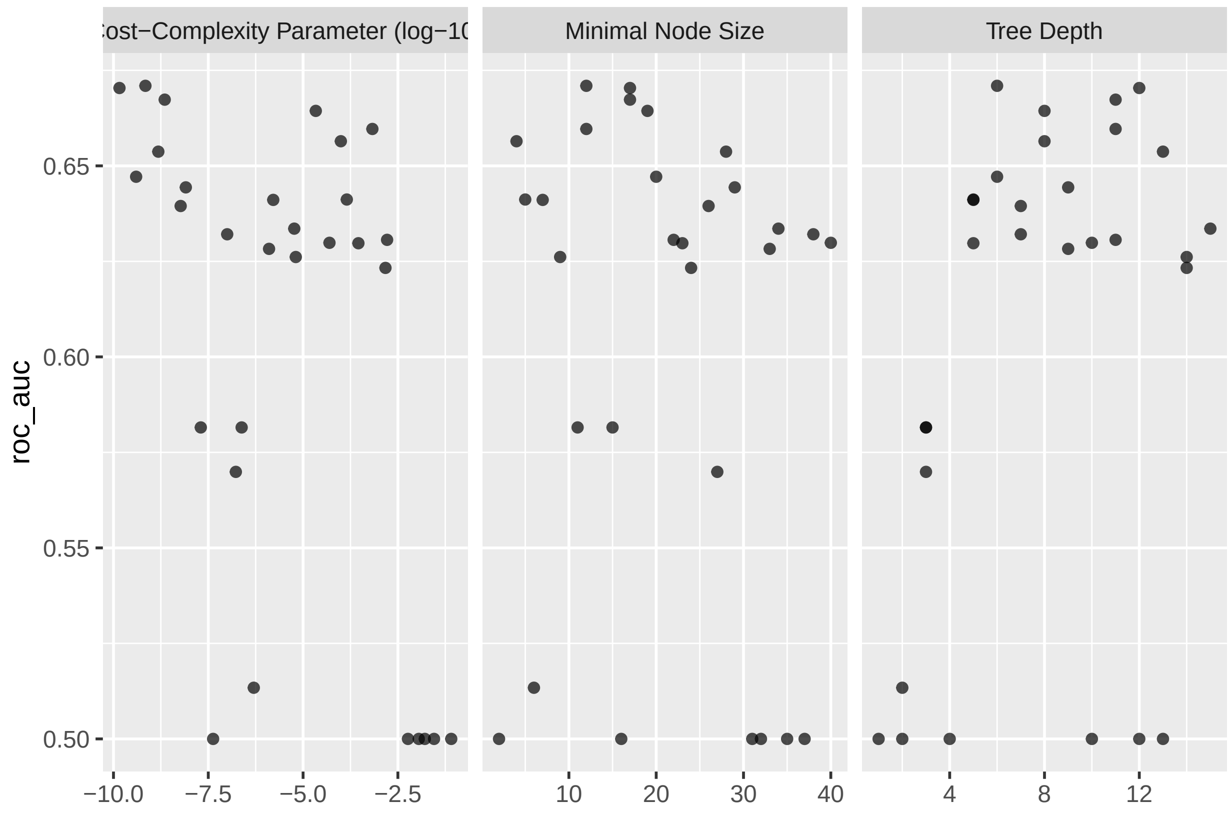
Minutes to run: 2.897

Decision Tree

```
tree_recipe <-
  recipe(formula = sprintf("%s ~ .", outcome_column) %>% as.formula, data = df_train) %>%
  step_novel(all_nominal_predictors()) %>%
  step_unknown(all_nominal_predictors()) %>%
  step_other(all_nominal_predictors(), threshold = 0.05, other = ".merged") %>%
  step_dummy(all_nominal_predictors()) %>%
  step_zv(all_predictors())

tree_spec <-
  decision_tree(cost_complexity = tune(),
                tree_depth = tune(),
                min_n = tune()) %>%
  set_mode("classification") %>%
  set_engine("rpart")

tree_grid <- grid_latin_hypercube(cost_complexity(),
                                  tree_depth(),
                                  min_n(),
                                  size = grid_size)

tree_workflow <-
  workflow() %>%
  add_recipe(tree_recipe) %>%
  add_model(tree_spec)

tree_tune <-
  tree_workflow %>%
  tune_grid(resamples = df_folds,
            grid = tree_grid)

tree_tune %>%
  collect_metrics()
```

```
## # A tibble: 60 x 9
##    cost_complexity tree_depth min_n .metric  .estimator  mean     n std_err .config
##              <dbl>      <int> <int> <chr>    <chr>      <dbl> <int>   <dbl> <chr>
##  1        1.49e- 3         14    24 accuracy binary     0.950    10 0.00154 Preprocessor1_Model01
##  2        1.49e- 3         14    24 roc_auc  binary     0.623    10 0.00941 Preprocessor1_Model01
##  3        6.73e- 4         11    12 accuracy binary     0.940    10 0.00260 Preprocessor1_Model02
##  4        6.73e- 4         11    12 roc_auc  binary     0.660    10 0.0142  Preprocessor1_Model02
##  5        3.96e-10          6    20 accuracy binary     0.950    10 0.00142 Preprocessor1_Model03
##  6        3.96e-10          6    20 roc_auc  binary     0.647    10 0.0120  Preprocessor1_Model03
##  7        1.12e- 2         10    37 accuracy binary     0.954    10 0.00132 Preprocessor1_Model04
##  8        1.12e- 2         10    37 roc_auc  binary     0.5      10 0       Preprocessor1_Model04
##  9        9.98e- 8          7    38 accuracy binary     0.952    10 0.00147 Preprocessor1_Model05
## 10        9.98e- 8          7    38 roc_auc  binary     0.632    10 0.00737 Preprocessor1_Model05
## # ... with 50 more rows
```

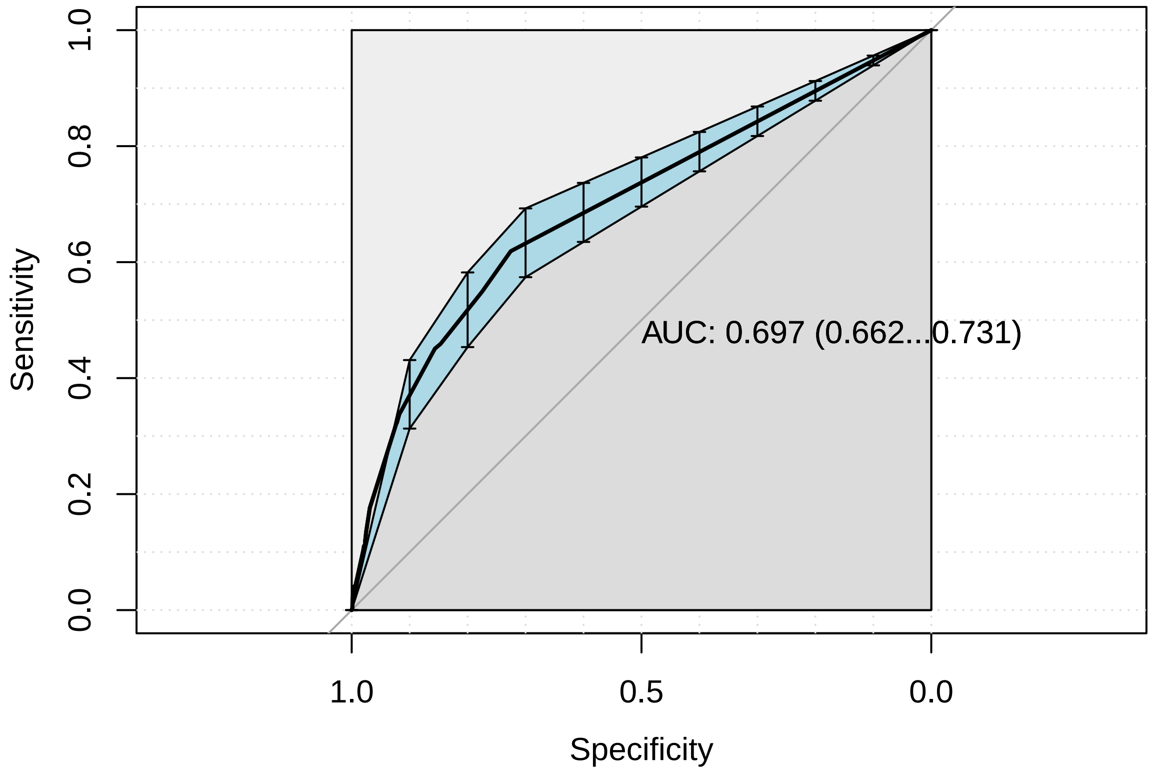
```
autoplot(tree_tune, metric = "roc_auc")
```



```
tree_tune %>%
  show_best("roc_auc")
```

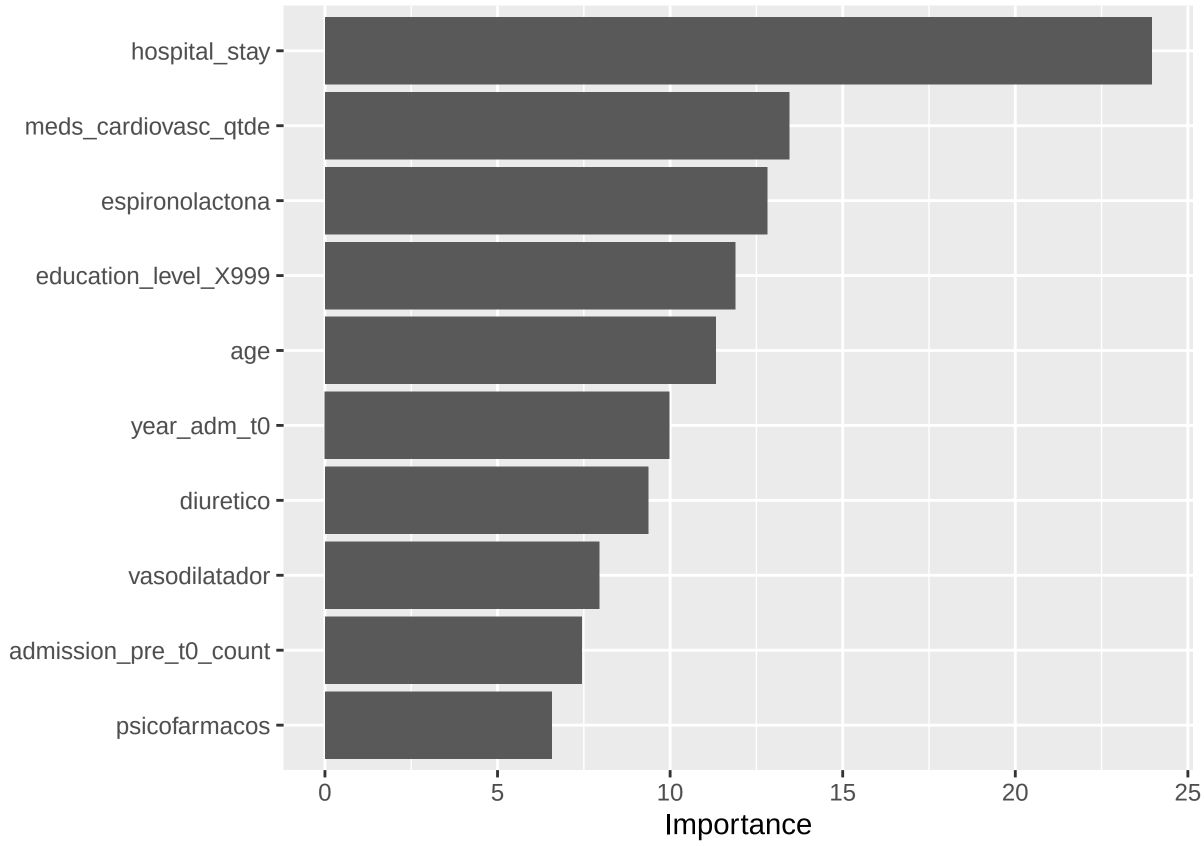
```
## # A tibble: 5 x 9
##   cost_complexity tree_depth min_n .metric .estimator  mean     n std_err .config
##             <dbl>      <int> <int> <chr>   <chr>      <dbl> <int>   <dbl> <chr>
## 1        6.95e-10          6    12 roc_auc binary     0.671    10  0.0116 Preprocessor1_Model20
## 2        1.44e-10         12    17 roc_auc binary     0.670    10  0.0122 Preprocessor1_Model14
## 3        2.25e- 9         11    17 roc_auc binary     0.667    10  0.0145 Preprocessor1_Model23
## 4        2.16e- 5          8    19 roc_auc binary     0.664    10  0.0147 Preprocessor1_Model06
## 5        6.73e- 4         11    12 roc_auc binary     0.660    10  0.0142 Preprocessor1_Model02
```

```
best_tree <- tree_tune %>%
  select_best("roc_auc")

final_tree_workflow <-
  tree_workflow %>%
  finalize_workflow(best_tree)

last_tree_fit <-
  final_tree_workflow %>%
  last_fit(df_split)

final_tree_fit <- extract_workflow(last_tree_fit)

tree_auc <- validation(final_tree_fit, df_test)
```



```
##   |
##                   Kappa : 0.1075
##
##  Mcnemar's Test P-Value : <2e-16
##
##             Sensitivity : 0.7256
##             Specificity : 0.6189
##          Pos Pred Value : 0.9722
##          Neg Pred Value : 0.1093
##              Prevalence : 0.9484
##          Detection Rate : 0.6882
##    Detection Prevalence : 0.7078
##       Balanced Accuracy : 0.6722
##
##        'Positive' Class : 0
##
```

```
extract_vip(final_tree_fit, pred_wrapper = predict,
            reference_class = "0", use_matrix = FALSE,
            method = 'model')
```



```
# extract_vip(final_tree_fit, pred_wrapper = predict,
#             reference_class = "1", use_matrix = FALSE,
#             method = 'permute')
```

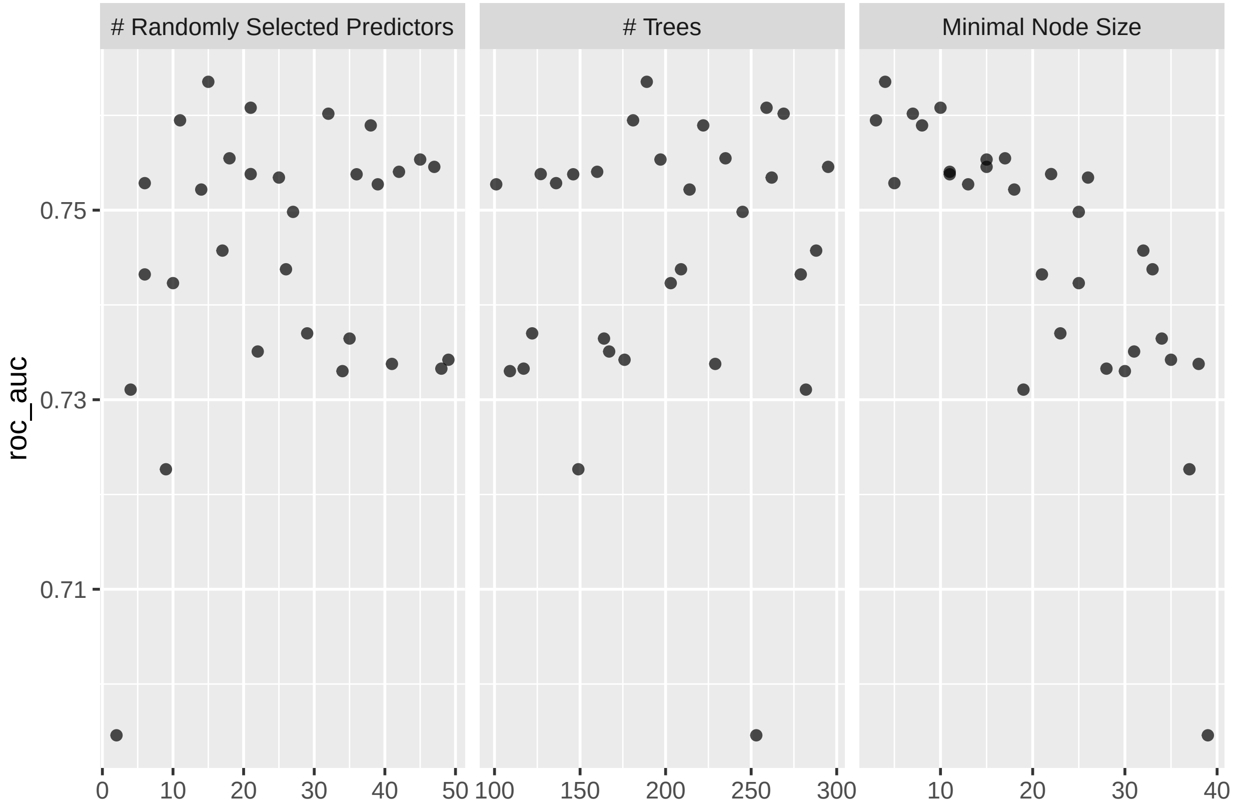
Minutes to run: 9.749

Random Forest

```
rf_recipe <-
  recipe(formula = sprintf("%s ~ .", outcome_column) %>% as.formula,
         data = df_train) %>%
  step_novel(all_nominal_predictors()) %>%
  step_unknown(all_nominal_predictors()) %>%
  step_other(all_nominal_predictors(), threshold = 0.05, other = ".merged") %>%
  step_dummy(all_nominal_predictors()) %>%
  step_zv(all_predictors()) %>%
  step_impute_mean(all_numeric_predictors())

rf_spec <-
  rand_forest(mtry = tune(),
              trees = tune(),
              min_n = tune()) %>%
  set_mode("classification") %>%
  set_engine("randomForest",
             probability = TRUE,
             nthread = 8)

rf_grid <- grid_latin_hypercube(mtry(range = c(1L, 50L)),
                                trees(range = c(100L, 300L)),
                                min_n(),
                                size = grid_size)

rf_workflow <-
  workflow() %>%
  add_recipe(rf_recipe) %>%
  add_model(rf_spec)

rf_tune <-
  rf_workflow %>%
  tune_grid(resamples = df_folds,
            grid = rf_grid)

rf_tune %>%
  collect_metrics()
```

```
## # A tibble: 60 x 9
##     mtry trees min_n .metric  .estimator  mean     n std_err .config
##    <int> <int> <int> <chr>    <chr>      <dbl> <int>   <dbl> <chr>
##  1    25   262    26 accuracy binary     0.954    10 0.00134 Preprocessor1_Model01
##  2    25   262    26 roc_auc  binary     0.753    10 0.00754 Preprocessor1_Model01
##  3    42   160    11 accuracy binary     0.953    10 0.00111 Preprocessor1_Model02
##  4    42   160    11 roc_auc  binary     0.754    10 0.0103  Preprocessor1_Model02
##  5    15   189     4 accuracy binary     0.954    10 0.00133 Preprocessor1_Model03
##  6    15   189     4 roc_auc  binary     0.764    10 0.0104  Preprocessor1_Model03
##  7    39   101    13 accuracy binary     0.953    10 0.00131 Preprocessor1_Model04
##  8    39   101    13 roc_auc  binary     0.753    10 0.00852 Preprocessor1_Model04
##  9    11   181     3 accuracy binary     0.954    10 0.00135 Preprocessor1_Model05
## 10    11   181     3 roc_auc  binary     0.759    10 0.0114  Preprocessor1_Model05
## # ... with 50 more rows
```

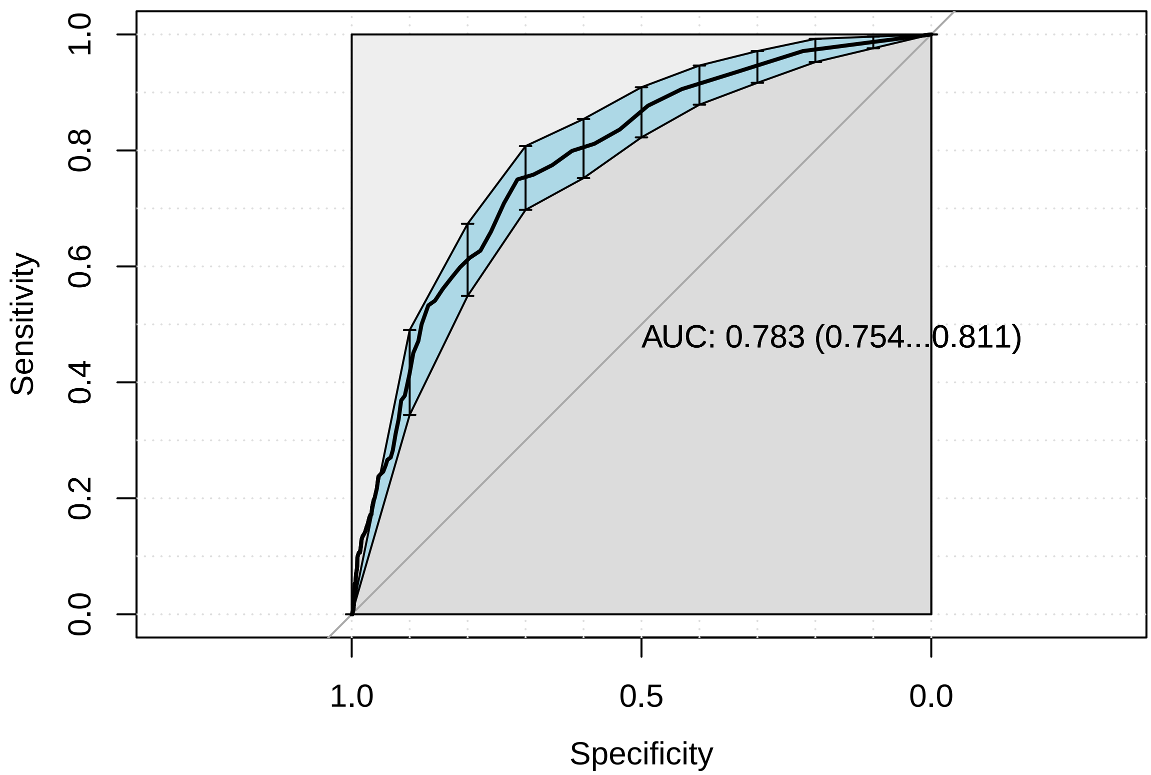
```
autoplot(rf_tune, metric = "roc_auc")
```



```
rf_tune %>%
  show_best("roc_auc")
```

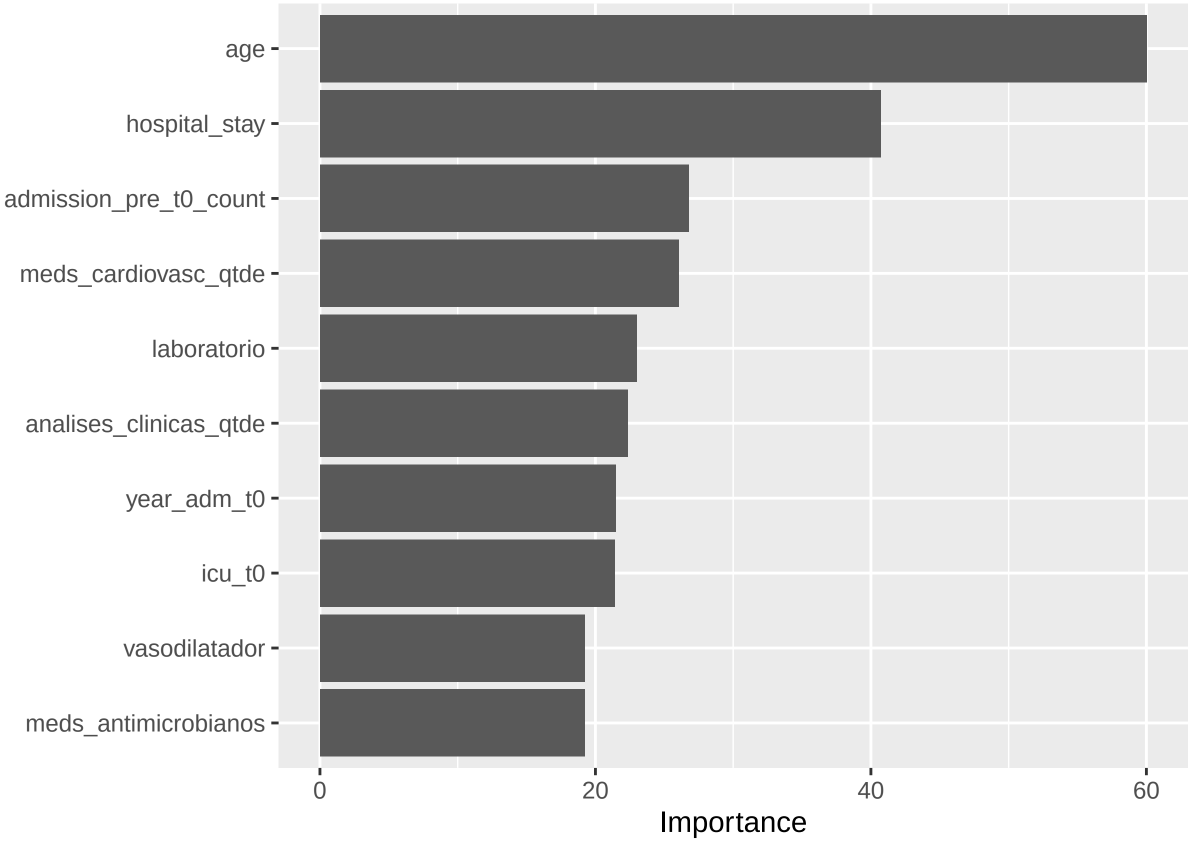
```
## # A tibble: 5 x 9
##    mtry trees min_n .metric .estimator  mean     n std_err .config
##   <int> <int> <int> <chr>   <chr>      <dbl> <int>   <dbl> <chr>
## 1    15   189     4 roc_auc binary     0.764    10 0.0104  Preprocessor1_Model03
## 2    21   259    10 roc_auc binary     0.761    10 0.00878 Preprocessor1_Model14
## 3    32   269     7 roc_auc binary     0.760    10 0.00932 Preprocessor1_Model13
## 4    11   181     3 roc_auc binary     0.759    10 0.0114  Preprocessor1_Model05
## 5    38   222     8 roc_auc binary     0.759    10 0.00982 Preprocessor1_Model10
```

```
best_rf <- rf_tune %>%
  select_best("roc_auc")

final_rf_workflow <-
  rf_workflow %>%
  finalize_workflow(best_rf)

last_rf_fit <-
  final_rf_workflow %>%
  last_fit(df_split)

final_rf_fit <- extract_workflow(last_rf_fit)

rf_auc <- validation(final_rf_fit, df_test)
```



```
##   |
```

```
pfun_rf <- function(object, newdata) predict(object, data = newdata)

extract_vip(final_rf_fit, pred_wrapper = predict,
            reference_class = "1", use_matrix = FALSE,
            method = 'model')
```



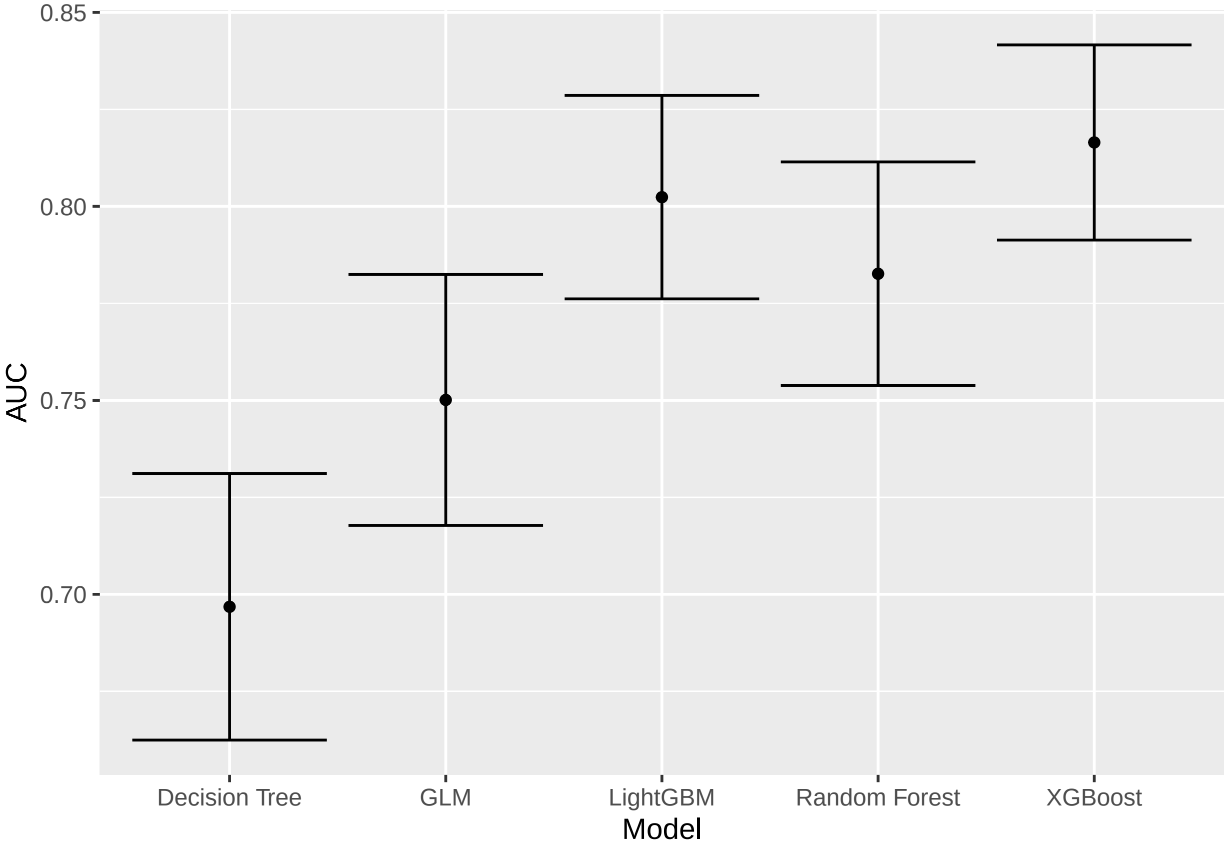
```
# extract_vip(final_rf_fit, pred_wrapper = predict,
#             reference_class = "1", use_matrix = FALSE,
#             method = 'permute')
```

Minutes to run: 129.867

Models Comparison

```
if (RUN_ALL_MODELS) {
  df_auc <- tibble::tribble(
    ~Model, ~`AUC`, ~`Lower Limit`, ~`Upper Limit`,
    'XGBoost', as.numeric(xgboost_auc$auc), xgboost_auc$ci[1], xgboost_auc$ci[3],
    'LightGBM', as.numeric(lightgbm_auc$auc), lightgbm_auc$ci[1], lightgbm_auc$ci[3],
    'GLM', as.numeric(glmnet_auc$auc), glmnet_auc$ci[1], glmnet_auc$ci[3],
    'Decision Tree', as.numeric(tree_auc$auc), tree_auc$ci[1], tree_auc$ci[3],
    'Random Forest', as.numeric(rf_auc$auc), rf_auc$ci[1], rf_auc$ci[3]
  ) %>%
    mutate(Target = outcome_column)
} else {
  df_auc <- tibble::tribble(
    ~Model, ~`AUC`, ~`Lower Limit`, ~`Upper Limit`,
    'LightGBM', as.numeric(lightgbm_auc$auc), lightgbm_auc$ci[1], lightgbm_auc$ci[3]
  ) %>%
    mutate(Target = outcome_column)
}
```

```r
df_auc %>%
  ggplot(aes(x = Model, y = AUC, ymin = `Lower Limit`, ymax = `Upper Limit`)) +
    geom_point() +
    geom_errorbar()
```



```r
saveRDS(df_auc, sprintf("./auxiliar/model_selection/performance/%s.RData", outcome_column))
```

Minutes to run: 0.002