

# Choose Your Own Project: Wine Quality

HarvardX PH125.9x Data Science: Capstone

Shunsuke Kobayashi

final: 2021-12-08

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Data Preprocessing and Explanatory Data Analysis</b>	<b>3</b>
2.1	Initial Data Preprocessing . . . . .	3
2.2	Explanatory Data Analysis (EDA) . . . . .	4
2.3	Prepare the train / test data set . . . . .	7
<b>3</b>	<b>Methods and Analysis</b>	<b>10</b>
3.1	Logistic Regression . . . . .	10
3.2	CART (Classification and Regression Tree) . . . . .	10
3.3	Random Forest . . . . .	10
3.4	SVM (Support-vector machine) . . . . .	11
3.5	Neural Net . . . . .	11
3.6	Gradient Boosting . . . . .	11
3.7	Deep Learning . . . . .	11
<b>4</b>	<b>Results</b>	<b>11</b>
4.1	Each model . . . . .	11
4.1.1	logit . . . . .	11
4.1.2	rpart . . . . .	14
4.1.3	rf . . . . .	16
4.1.4	svm . . . . .	19
4.1.5	nnet . . . . .	21
4.1.6	xgbLinear . . . . .	25
4.1.7	xgbDART . . . . .	30
4.1.8	xgbTree . . . . .	48

4.1.9	torch . . . . .	59
4.1.10	Tabnet . . . . .	62
4.2	Compare the models . . . . .	65
4.3	Additional: Google Colaboratory . . . . .	66
<b>5</b>	<b>conclusion</b>	<b>66</b>
	<b>Reference</b>	<b>66</b>

# 1 Introduction

The goal of this project is to use wine quality data sets for classification.

In the previous MovieLens project, we did machine learning using 10 million data, and we see this as an application problem in BtoC business such like Netflix, Amazon etc. In this CYO project, we set it up as a way to apply it to BtoB business, especially with small data sets in the field of product manufacturing and technology development. The red wine quality data set is a set of about 1,600 tabular tables, consisting of information such as alcohol and acid content, pH, and density, in addition to the target quality (This is similar to the structure of information in the chemical industry to which I belong).

In addition, the goal was to learn and compare not only learned methods but also newly algorithms developed and apply new technologies. In addition to the gradient boosting method, we also worked on the deep learning library **torch**, whose R version will be implemented in September 2020, and the classification in **Tabnet** that utilizes it, and created 10 models. In the new technology, we also worked on using the **GPU** with Google Colaboratory and succeeded in reducing the computation time from 53 minutes when using the CPU to 35 minutes when using the GPU.

## 2 Data Preprocessing and Explanatory Data Analysis

In the UCI repository, there are two datasets related to the Portuguese wine Vinho Verde, red and white. Due to privacy and logistical issues, only physicochemical (input) and sensory (output) variables are available.

The red wine quality data set is a set of about 1,600 tabular tables, consisting of information such as alcohol and acid content, pH, and density, in addition to the target quality (This is similar to the structure of information in the chemical industry to which I belong).

A similar dataset is the Water portability dataset from Kaggle, but we chose this one because it requires sign-in for automatic download and the red wine dataset has less data.

### 2.1 Initial Data Preprocessing

It was confirmed that the 1599 rows of data consisted of only numerical data and that there was no missing data.

```
# Wine Quality Data Set
# https://archive.ics.uci.edu/ml/datasets/Wine+Quality
# https://archive.ics.uci.edu/ml/machine-learning-databases/wine-quality/winequality-red.csv

url <- c("https://archive.ics.uci.edu/ml/machine-learning-databases/wine-quality/winequality-red.csv")
wine <- read_delim(url, delim = ";")

# modify the column name

colnames(wine) <- c("fixed_acidity", "volatile_acidity", "citric_acid",
                    "residual_sugar", "chlorides", "free_sulfur_dioxide",
                    "total_sulfur_dioxide", "density", "pH",
                    "sulphates", "alcohol", "quality")

### Check the data -----

# whole the data

head(wine)
```

```
## # A tibble: 6 x 12
##   fixed_acidity volatile_acidity citric_acid residual_sugar chlorides
##   <dbl>         <dbl>         <dbl>         <dbl>         <dbl>
## 1         7.4         0.7           0             1.9         0.076
## 2         7.8         0.88          0             2.6         0.098
## 3         7.8         0.76          0.04          2.3         0.092
## 4        11.2         0.28          0.56          1.9         0.075
## 5         7.4         0.7           0             1.9         0.076
## 6         7.4         0.66          0             1.8         0.075
## # ... with 7 more variables: free_sulfur_dioxide <dbl>,
## #   total_sulfur_dioxide <dbl>, density <dbl>, pH <dbl>, sulphates <dbl>,
## #   alcohol <dbl>, quality <dbl>
```

```
wine %>% summary()
```

```
##   fixed_acidity   volatile_acidity   citric_acid   residual_sugar
##   Min.      : 4.60   Min.      :0.1200   Min.      :0.000   Min.      : 0.900
##   1st Qu.: 7.10   1st Qu.:0.3900   1st Qu.:0.090   1st Qu.: 1.900
##   Median : 7.90   Median :0.5200   Median :0.260   Median : 2.200
##   Mean    : 8.32   Mean    :0.5278   Mean    :0.271   Mean    : 2.539
##   3rd Qu.: 9.20   3rd Qu.:0.6400   3rd Qu.:0.420   3rd Qu.: 2.600
##   Max.    :15.90   Max.    :1.5800   Max.    :1.000   Max.    :15.500
##   chlorides      free_sulfur_dioxide total_sulfur_dioxide   density
##   Min.      :0.01200   Min.      : 1.00     Min.      : 6.00     Min.      :0.9901
##   1st Qu.:0.07000   1st Qu.: 7.00     1st Qu.: 22.00     1st Qu.:0.9956
##   Median :0.07900   Median :14.00     Median : 38.00     Median :0.9968
##   Mean    :0.08747   Mean    :15.87     Mean    : 46.47     Mean    :0.9967
##   3rd Qu.:0.09000   3rd Qu.:21.00     3rd Qu.: 62.00     3rd Qu.:0.9978
##   Max.    :0.61100   Max.    :72.00     Max.    :289.00     Max.    :1.0037
##   pH            sulphates          alcohol          quality
##   Min.      :2.740   Min.      :0.3300   Min.      : 8.40   Min.      :3.000
##   1st Qu.:3.210   1st Qu.:0.5500   1st Qu.: 9.50   1st Qu.:5.000
##   Median :3.310   Median :0.6200   Median :10.20   Median :6.000
##   Mean    :3.311   Mean    :0.6581   Mean    :10.42   Mean    :5.636
##   3rd Qu.:3.400   3rd Qu.:0.7300   3rd Qu.:11.10   3rd Qu.:6.000
##   Max.    :4.010   Max.    :2.0000   Max.    :14.90   Max.    :8.000
```

```
sum(is.na(wine))
```

```
## [1] 0
```

## 2.2 Explanatory Data Analysis (EDA)

The relationship between the quality data and each Attribution, which is the target for classification, was confirmed. The relationship with quality and the correlation between each were confirmed. It can be seen that while alcohol and volatile acidity are likely to have an effect on quality, residual sugar has a small relationship not only with quality but also with other Attributions.

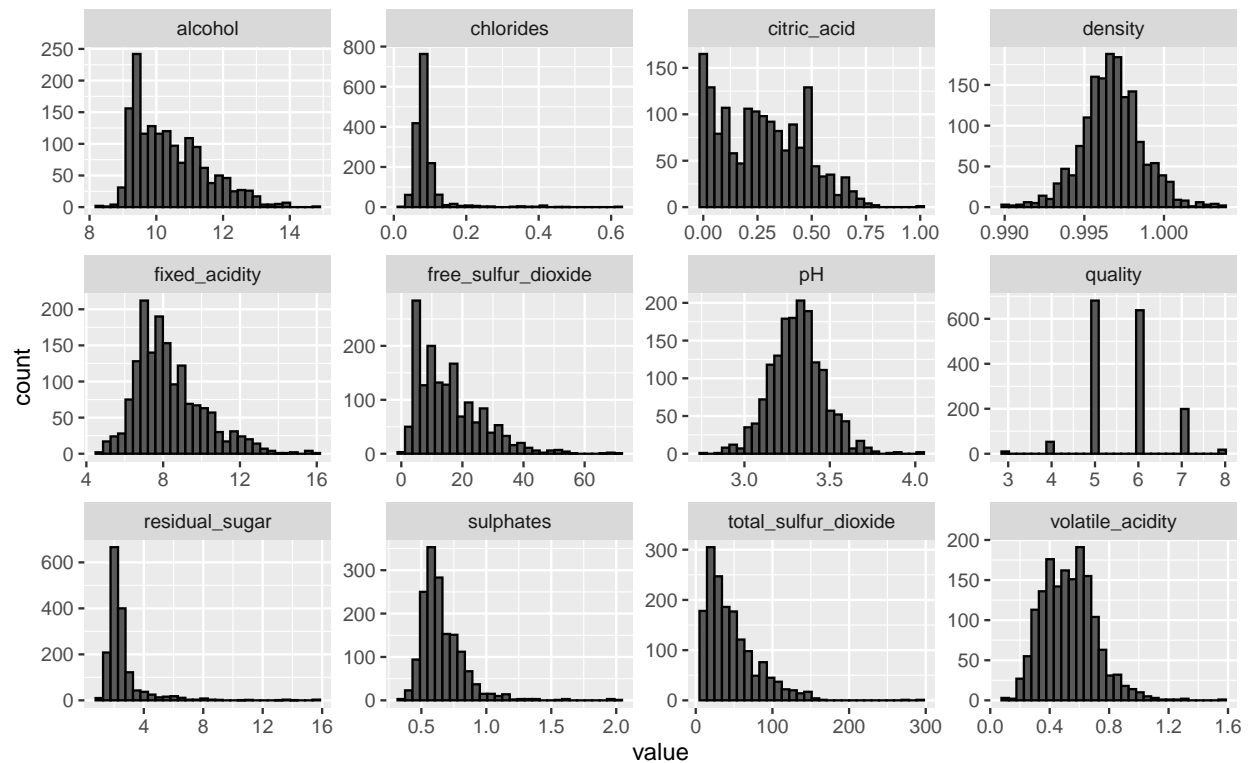
```
# Target: quality
```

```
table(wine$quality)
```

```
##
## 3 4 5 6 7 8
## 10 53 681 638 199 18
```

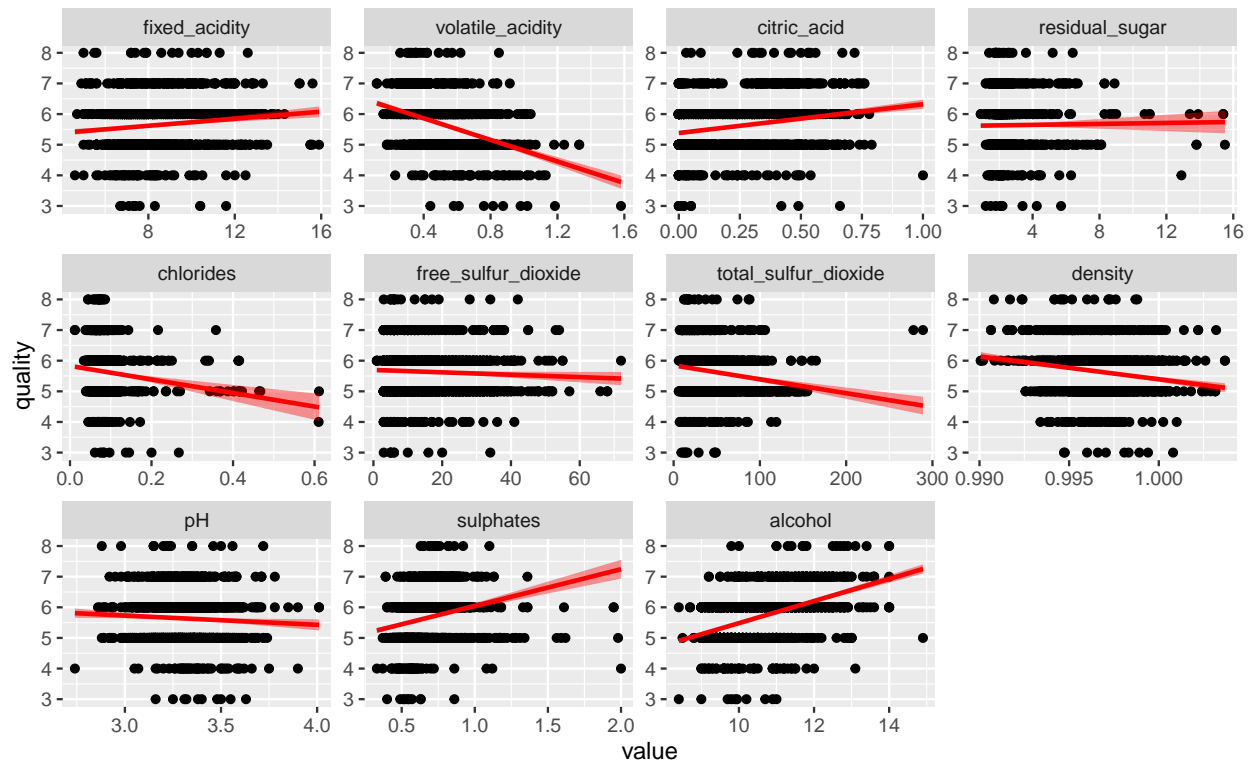
```
# Distribution of each value
```

```
wine %>% gather() %>%
  ggplot(aes(value)) +
  geom_histogram(col="black") +
  facet_wrap(~key, scales = "free")
```



```
# Check the relations between each attribute and quality (target)
```

```
wine %>% reshape2::melt(., "quality") %>%
  ggplot(aes(value, quality)) +
  geom_point() +
  geom_smooth(method = lm, col = "red", fill = "red") +
  facet_wrap(~variable, scales = "free")
```

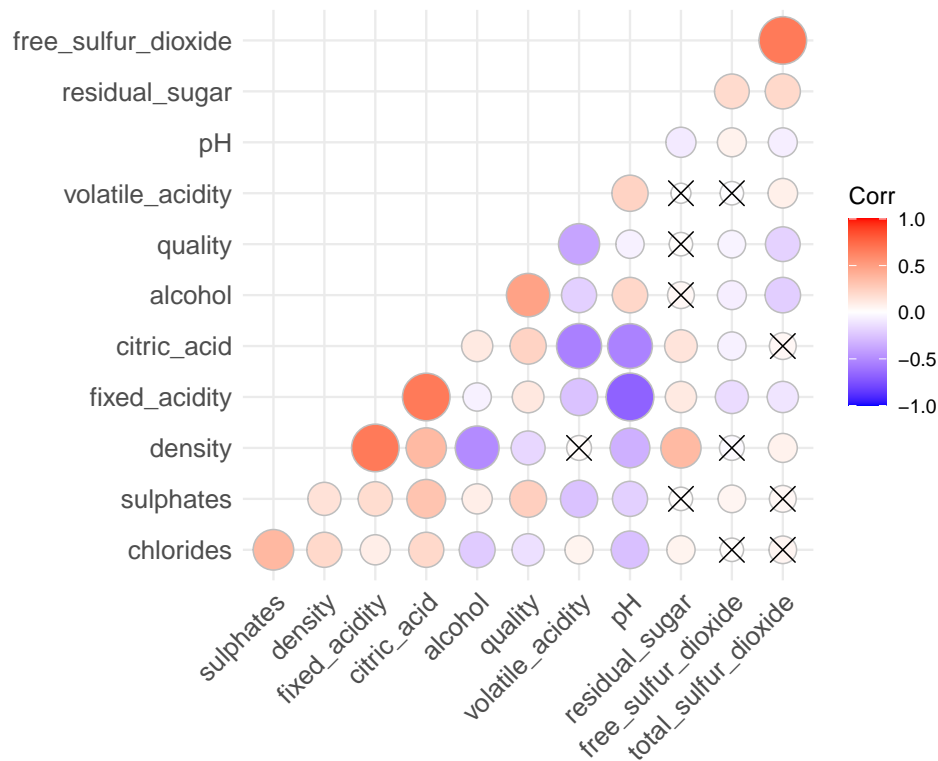


```
# Check the correlation
## Compute a matrix of correlation p-values

p.mat <- cor_pmat(wine)

# X means no significant coefficient

round(cor(wine),2) %>%
ggcorrplot(., method = "circle", hc.order = TRUE, type = "lower", p.mat = p.mat)
```



## 2.3 Prepare the train / test data set

In order to build a model as a classification problem for quality, we divided the data. Since the number of data is small (1600), we decided to split the data 80:20. In order to make it a classification problem, a binomial classification was made for quality, with 1 being above 6 and 0 being below 6. In order to check the magnitude of the effect of each Attribution, the numbers were centered and standardized. The **recipe** library was used for the conversion. By using this library, it is easy to complete missing values and to make dummy variables for categorical variables (although we did not use this library, we commented it out in the code).

```
### Data preparation: Split the data set ----
# Preparation test and train data set for model selection
# Due to the small amount of data, 20% was used as test data.

set.seed(1, sample.kind="Rounding")
test_index <- createDataPartition(y = wine$quality, times = 1, p = 0.2, list = FALSE)
train_set <- wine[-test_index,]
test_set <- wine[test_index,]

# To build each model, centralization and standardization are done with recipe.

rec_pre <- recipe(train_set, quality~.) %>%
  step_center(all_numeric(),-quality) %>% # Centralization
  step_scale(all_numeric(), -quality) %>% # Standardization
  # step_knnimpute(all_predictors(),K=5) # There is no NA
  # step_dummy(all_predictors(), -all_numeric()) # There is no categorical data
  prep()
```

```
# apply the recipe
```

```
train <- bake(rec_pre, new_data = train_set)
test  <- bake(rec_pre, new_data = test_set)
```

```
# check the data
```

```
head(train)
```

```
## # A tibble: 6 x 12
##   fixed_acidity volatile_acidity citric_acid residual_sugar chlorides
##   <dbl>          <dbl>          <dbl>          <dbl>    <dbl>
## 1    -0.521        0.934        -1.39        -0.445   -0.246
## 2    -0.283        1.93         -1.39         0.0476   0.210
## 3    -0.283        1.26         -1.18        -0.163   0.0859
## 4     1.74        -1.38         1.50        -0.445   -0.266
## 5    -0.521        0.934        -1.39        -0.445   -0.246
## 6    -0.521        0.713        -1.39        -0.515   -0.266
## # ... with 7 more variables: free_sulfur_dioxide <dbl>,
## #   total_sulfur_dioxide <dbl>, density <dbl>, pH <dbl>, sulphates <dbl>,
## #   alcohol <dbl>, quality <dbl>
```

```
head(test)
```

```
## # A tibble: 6 x 12
##   fixed_acidity volatile_acidity citric_acid residual_sugar chlorides
##   <dbl>          <dbl>          <dbl>          <dbl>    <dbl>
## 1    -0.461        -0.168         0.471         2.51   -0.349
## 2    -1.59         0.465         -1.39        -0.656   0.0238
## 3    -0.223        -0.554        -0.303        -0.656   0.376
## 4    -0.699        0.989         -1.39        -0.445   -0.163
## 5    -0.342        0.493         -1.18         0.892   -0.0798
## 6    -0.640        1.07          -1.13         1.49   -0.0384
## # ... with 7 more variables: free_sulfur_dioxide <dbl>,
## #   total_sulfur_dioxide <dbl>, density <dbl>, pH <dbl>, sulphates <dbl>,
## #   alcohol <dbl>, quality <dbl>
```

```
# Transform "quality" into binary
```

```
train$quality[train$quality < 6] <- 0
test$quality[test$quality < 6] <- 0
train$quality[train$quality > 6] <- 1
test$quality[test$quality > 6] <- 1
```

```
train$quality <- as.factor(train$quality)
test$quality <- as.factor(test$quality)
```

```
# summary
```

```
summary(train)
```



```
## fixed_acidity    volatile_acidity    citric_acid    residual_sugar
## Min.      :-2.1872    Min.      :-2.2623    Min.      :-1.38758    Min.      :-1.14804
## 1st Qu.   :-0.6993    1st Qu.   :-0.7194    1st Qu.   :-0.92290    1st Qu.   :-0.44474
## Median    :-0.2232    Median    :-0.0581    Median    :-0.07098    Median    :-0.23375
## Mean      : 0.0000    Mean      : 0.0000    Mean      : 0.00000    Mean      : 0.00000
## 3rd Qu.   : 0.5357    3rd Qu.   : 0.6032    3rd Qu.   : 0.78094    3rd Qu.   : 0.04757
## Max.      : 4.5382    Max.      : 5.7830    Max.      : 3.77556    Max.      : 9.12018
## chlorides      free_sulfur_dioxide    total_sulfur_dioxide
## Min.      :-1.57164    Min.      :-1.4219    Min.      :-1.2264
## 1st Qu.   :-0.36992    1st Qu.   :-0.8527    1st Qu.   :-0.7441
## Median    :-0.18344    Median    :-0.1885    Median    :-0.2618
## Mean      : 0.00000    Mean      : 0.0000    Mean      : 0.0000
## 3rd Qu.   : 0.04447    3rd Qu.   : 0.4757    3rd Qu.   : 0.4919
## Max.      :10.83929    Max.      : 5.3145    Max.      : 7.3048
## density        pH        sulphates        alcohol
## Min.      :-3.605932    Min.      :-3.71497    Min.      :-1.9078    Min.      :-1.8854
## 1st Qu.   :-0.603029    1st Qu.   :-0.66652    1st Qu.   :-0.6301    1st Qu.   :-0.8651
## Median    : 0.004086    Median    :-0.01792    Median    :-0.2236    Median    :-0.2158
## Mean      : 0.000000    Mean      : 0.00000    Mean      : 0.0000    Mean      : 0.0000
## 3rd Qu.   : 0.603033    3rd Qu.   : 0.56583    3rd Qu.   : 0.4152    3rd Qu.   : 0.6190
## Max.      : 3.810123    Max.      : 4.52232    Max.      : 7.7907    Max.      : 4.1438
## quality
## 0:595
## 1:683
##
##
##
##
```

```
summary(test)
```

```
## fixed_acidity    volatile_acidity    citric_acid    residual_sugar
## Min.      :-1.9492    Min.      :-1.93166    Min.      :-1.38758    Min.      :-1.14804
## 1st Qu.   :-0.6993    1st Qu.   :-0.82956    1st Qu.   :-0.87127    1st Qu.   :-0.44474
## Median    :-0.1636    Median    :-0.05810    Median    :-0.04517    Median    :-0.23375
## Mean      : 0.1325    Mean      :-0.07475    Mean      : 0.05729    Mean      : 0.02260
## 3rd Qu.   : 0.7886    3rd Qu.   : 0.46539    3rd Qu.   : 0.93583    3rd Qu.   : 0.04757
## Max.      : 4.3596    Max.      : 2.80734    Max.      : 2.53640    Max.      : 9.04985
## chlorides      free_sulfur_dioxide    total_sulfur_dioxide    density
## Min.      :-1.57164    Min.      :-1.42193    Min.      :-1.22645    Min.      :-3.6059
## 1st Qu.   :-0.34920    1st Qu.   :-0.85266    1st Qu.   :-0.71397    1st Qu.   :-0.5023
## Median    :-0.18344    Median    :-0.28338    Median    :-0.26178    Median    : 0.1239
## Mean      :-0.03996    Mean      :-0.05283    Mean      :-0.03245    Mean      : 0.1470
## 3rd Qu.   : 0.04447    3rd Qu.   : 0.47566    3rd Qu.   : 0.37128    3rd Qu.   : 0.8045
## Max.      : 7.85569    Max.      : 3.70157    Max.      : 3.02411    Max.      : 3.8101
## pH        sulphates        alcohol        quality
## Min.      :-2.74206    Min.      :-1.67548    Min.      :-1.60715    0:149
## 1st Qu.   :-0.66652    1st Qu.   :-0.63014    1st Qu.   :-0.86509    1:172
## Median    :-0.01792    Median    :-0.22362    Median    :-0.30854
## Mean      :-0.05328    Mean      :-0.01032    Mean      :-0.04457
## 3rd Qu.   : 0.56583    3rd Qu.   : 0.47328    3rd Qu.   : 0.52628
## Max.      : 3.80886    Max.      : 4.07390    Max.      : 3.30902
```

### 3 Methods and Analysis

We made the following models for the classification problem. The evaluation was done by using Accuracy. Accuracy was calculated by comparing the predictions made by each model using the test data set with the actual quality.

The importance of each item was extracted and graphed if it could be extracted. As mentioned in the introduction, the model is intended to be used in actual business, and it is important to understand how the target changes by changing the value of each item.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

where  $TP = \text{True positive}$ ;  $FP = \text{False positive}$ ;  $TN = \text{True negative}$ ;  $FN = \text{False negative}$

In the analysis of each model using the library **caret**, we decided to use cross validation and made 10 divisions. Automatic parameter tuning was used for tuning.

```
#####  
# MODEL Analysis  
#####  
  
# set the Cross Validation for caret package  
  
trControl = trainControl(method = "cv", number = 10)
```

#### 3.1 Logistic Regression

Logistic regression analysis is a type of “multivariate analysis” in which analysis is performed from multiple variables to predict qualitative probabilities.

$$p = \frac{1}{1 + e^{-(a_1x_1 + a_2x_2 + \dots + a_nx_n + b)}}$$

#### 3.2 CART (Classification and Regression Tree)

CART is a machine learning method that makes predictions based on “Yes or No” conditions for feature values. We selected it as one of the models because it is characterized by easy understanding and interpretation of learning results.

#### 3.3 Random Forest

Since the learning results are easy to understand and interpret, we choose Random Forest as one of the models and train it using multiple decision trees. Random forests are trained using multiple decision trees, which are generated by randomly selecting data from the original training data. Then, when actually evaluating the unknown data, the conclusions of the individual decision trees are combined into an overall conclusion by majority voting.

Decision trees are a very straightforward algorithm, but they are known to fail to generate the desired tree structure depending on the data, and are prone to overlearning. Random forests, on the other hand, have the advantage that the effect of overlearning is much smaller than that of decision trees because the correlation between individual decision trees is low.

### 3.4 SVM (Support-vector machine)

SVM is a machine learning model that can be applied to problems such as classification and regression. It is a method of determining which hyperplane (or straight line in the case of 2D) separating two classes of data is the farthest from each data. Compared to other methods, it has advantages such as being able to obtain a highly accurate model even with a small amount of data, easily maintaining discrimination accuracy even when the number of dimensions (number of features) increases, and making it easy to adjust parameters. On the other hand, the amount of computation increases rapidly when the amount of training data increases, and the principle is two-class classification, so it is difficult to apply to multi-class classification.

### 3.5 Neural Net

A neural network is a model of how the human brain works and is applied to computers. Deep learning is a method of applying neural networks, also known as deep neural networks. The basic model structure of a neural network consists of an **input layer**, a **hidden layer**, and an **output layer**.

### 3.6 Gradient Boosting

Gradient boosting is a machine learning method for tasks such as regression and classification that generates a predictive model in the form of an ensemble of weak prediction models (usually decision trees).

In addition to XGBoost's Linear and Tree, which were also used in the MovieLens project, modeling is also done using a method called DART. In gradient boosting, there was a problem that the gradient was generally applied to fit the data in the more extreme locations towards the end of the step. In order to prevent overlearning, MART (Multiple Additive Regression Trees) is improved by introducing the concept of Drop Out, which is called DART (Dropouts meet Multiple Additive Regression Trees).

### 3.7 Deep Learning

Deep learning, or deep learning, is a machine learning method that uses a neural network that reproduces the mechanism of human neurons, and is characterized by the use of a multilayer neural network. Deep learning is a machine learning method that uses neural networks that replicate the mechanism of human neurons, and is characterized by the use of multi-layered neural networks. It is currently producing significant results in various fields such as image recognition, speech recognition, and translation.

It has been possible to use the library **keras** to build deep learning models in R for some time. However, it requires a **python** environment for installation and is not suitable for automatic installation of projects. The library **torch** has been implemented in R version in September 2020. In this project, we used torch to build a model for deep learning.

We also worked on building a model with tabnet, which is available by installing torch; it is a deep learning for data tables announced by Google Cloud in 2020. Unlike deep learning, which works like a black box, in the case of TabNet the model can interpret the features it chooses.

## 4 Results

### 4.1 Each model

#### 4.1.1 logit

```
# logistic -----
```

```
set.seed(1, sample.kind="Rounding")
```

```
# train the model
```

```
wine_logit <- train(  
  quality ~ .,  
  data = train,  
  method = "glm",  
  family = binomial(),  
  trControl = trControl  
)
```

```
# Check the model
```

```
summary(wine_logit)
```

```
##
```

```
## Call:
```

```
## NULL
```

```
##
```

```
## Deviance Residuals:
```

```
##      Min       1Q   Median       3Q      Max  
## -3.5361  -0.8221   0.2879   0.7995   2.3193
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error z value Pr(>|z|)  
## (Intercept)      0.25808    0.07050   3.660 0.000252 ***  
## fixed_acidity      0.22858    0.18891   1.210 0.226269  
## volatile_acidity  -0.58151    0.09980  -5.827 5.65e-09 ***  
## citric_acid       -0.23073    0.12360  -1.867 0.061938 .  
## residual_sugar     0.11745    0.08566   1.371 0.170373  
## chlorides         -0.20340    0.08437  -2.411 0.015918 *  
## free_sulfur_dioxide 0.19234    0.09640   1.995 0.046024 *  
## total_sulfur_dioxide -0.49020    0.10438  -4.696 2.65e-06 ***  
## density           -0.10475    0.16926  -0.619 0.536022  
## pH                -0.08697    0.12414  -0.701 0.483587  
## sulphates         0.49160    0.08742   5.624 1.87e-08 ***  
## alcohol           0.98504    0.12682   7.767 8.02e-15 ***
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
```

```
## (Dispersion parameter for binomial family taken to be 1)
```

```
##
```

```
##      Null deviance: 1765.6  on 1277  degrees of freedom
```

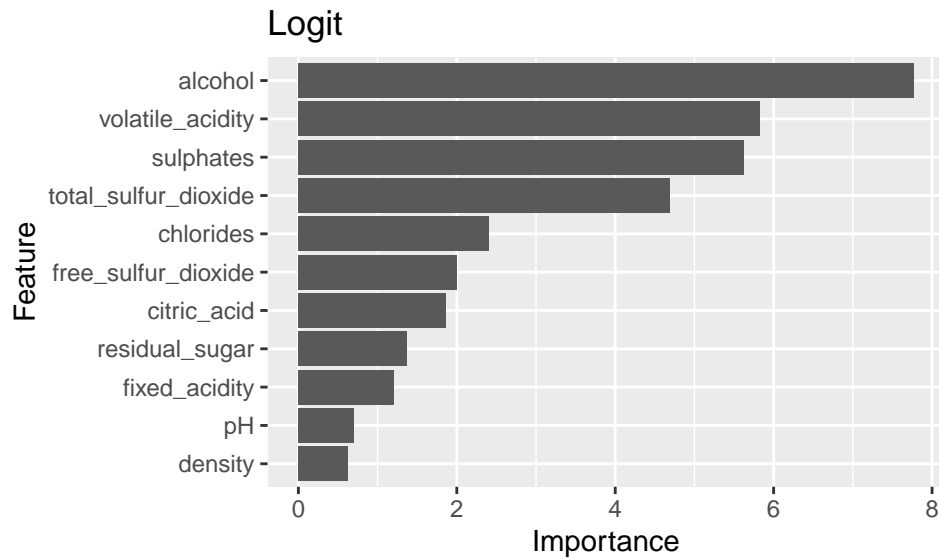
```
## Residual deviance: 1303.5  on 1266  degrees of freedom
```

```
## AIC: 1327.5
```

```
##
```

```
## Number of Fisher Scoring iterations: 4
```

```
p_logit <- varImp(wine_logit, scale = F) %>%
  ggplot() + ggtitle("Logit")
p_logit
```



```
# Evaluate
```

```
confusionMatrix(predict(wine_logit, test), test$quality)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 106  45
##           1  43 127
##
##           Accuracy : 0.7259
##           95% CI : (0.6736, 0.7739)
##       No Information Rate : 0.5358
##       P-Value [Acc > NIR] : 2.246e-12
##
##           Kappa : 0.4494
##
##  McNemar's Test P-Value : 0.9151
##
##           Sensitivity : 0.7114
##           Specificity : 0.7384
##       Pos Pred Value : 0.7020
##       Neg Pred Value : 0.7471
##           Prevalence : 0.4642
##       Detection Rate : 0.3302
##   Detection Prevalence : 0.4704
##       Balanced Accuracy : 0.7249
##
```

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

```
logit_acc <- confusionMatrix(predict(wine_logit, test), test$quality)$overall["Accuracy"]
logit_acc
```

```
## Accuracy
## 0.7258567
```

#### 4.1.2 rpart

```
### Rpart -----
```

```
set.seed(1, sample.kind="Rounding")
```

```
# train the model
```

```
wine_rpart <- train(
  quality ~ .,
  data = train,
  method = "rpart",
  trControl = trControl,
  tuneLength = 10
)
```

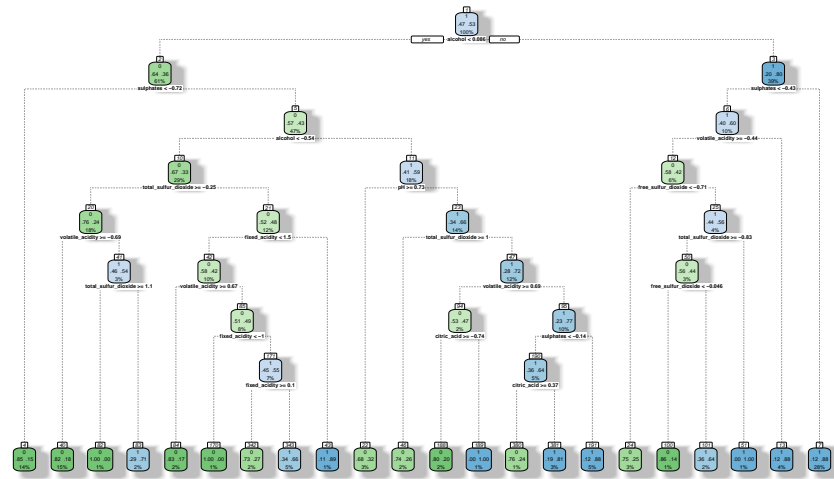
```
# Check the model
```

```
wine_rpart
```

```
## CART
##
## 1278 samples
## 11 predictor
## 2 classes: '0', '1'
##
## No pre-processing
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 1151, 1150, 1150, 1149, 1150, 1151, ...
## Resampling results across tuning parameters:
##
##      cp          Accuracy   Kappa
## 0.006722689 0.7316243 0.4608685
## 0.007563025 0.7221574 0.4398870
## 0.008403361 0.7237323 0.4434475
## 0.010084034 0.7198688 0.4363828
## 0.010924370 0.7198688 0.4363828
## 0.011764706 0.7191303 0.4340019
## 0.018487395 0.7121112 0.4204372
## 0.026890756 0.7081805 0.4142043
## 0.034453782 0.6894418 0.3808617
## 0.357983193 0.6579027 0.3040603
```

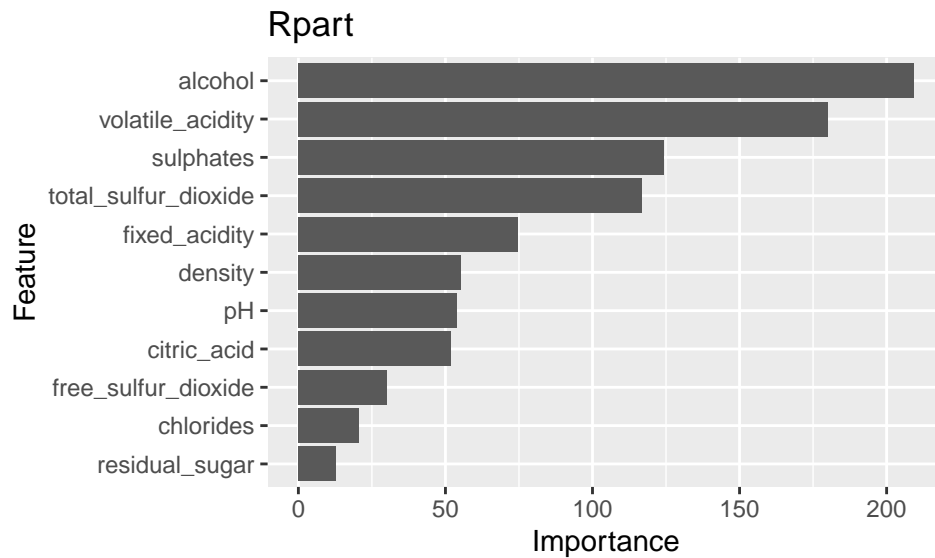
```
##
## Accuracy was used to select the optimal model using the largest value.
## The final value used for the model was cp = 0.006722689.
```

```
fancyRpartPlot(wine_rpart$finalModel)
```



Rattle 2021-Dec-08 22:55:24 Kobayashi

```
p_rpart <- varImp(wine_rpart, scale = F) %>%
  ggplot() + ggtitle("Rpart")
p_rpart
```



```
# Evaluate
```

```
confusionMatrix(predict(wine_rpart, test), test$quality)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 106  38
##           1  43 134
##
##           Accuracy : 0.7477
##           95% CI : (0.6964, 0.7943)
##       No Information Rate : 0.5358
##       P-Value [Acc > NIR] : 4.333e-15
##
##           Kappa : 0.4916
##
##  Mcnemar's Test P-Value : 0.6567
##
##           Sensitivity : 0.7114
##           Specificity : 0.7791
##       Pos Pred Value : 0.7361
##       Neg Pred Value : 0.7571
##           Prevalence : 0.4642
##       Detection Rate : 0.3302
##       Detection Prevalence : 0.4486
##       Balanced Accuracy : 0.7452
##
##       'Positive' Class : 0
##
```

```
rpart_acc <- confusionMatrix(predict(wine_rpart, test), test$quality)$overall["Accuracy"]
rpart_acc
```

```
## Accuracy
## 0.7476636
```

#### 4.1.3 rf

```
### Random forest -----

set.seed(1, sample.kind="Rounding")

# train the model

wine_rf <- train(
  quality ~ .,
  data = train,
  method = "rf",
  trControl = trControl,
  tuneLength = 10
)

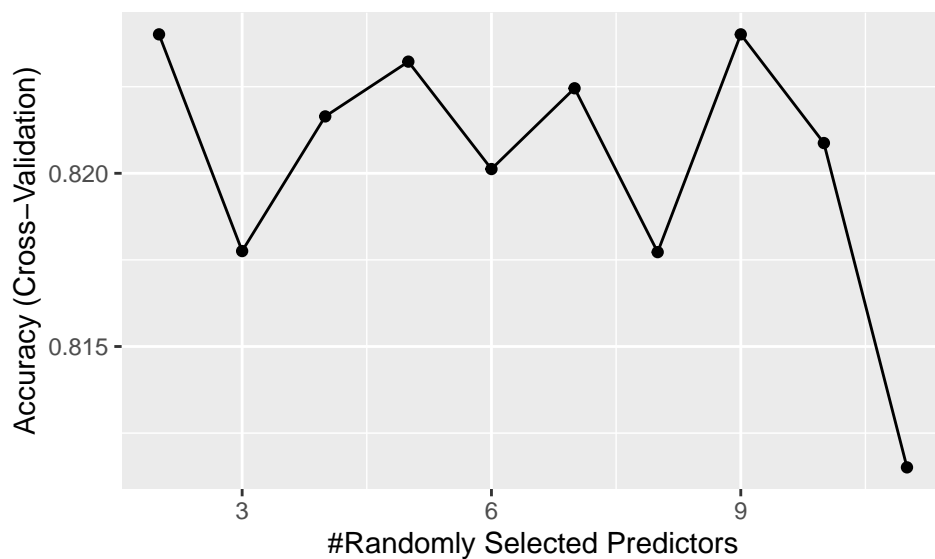
# Check the model
```



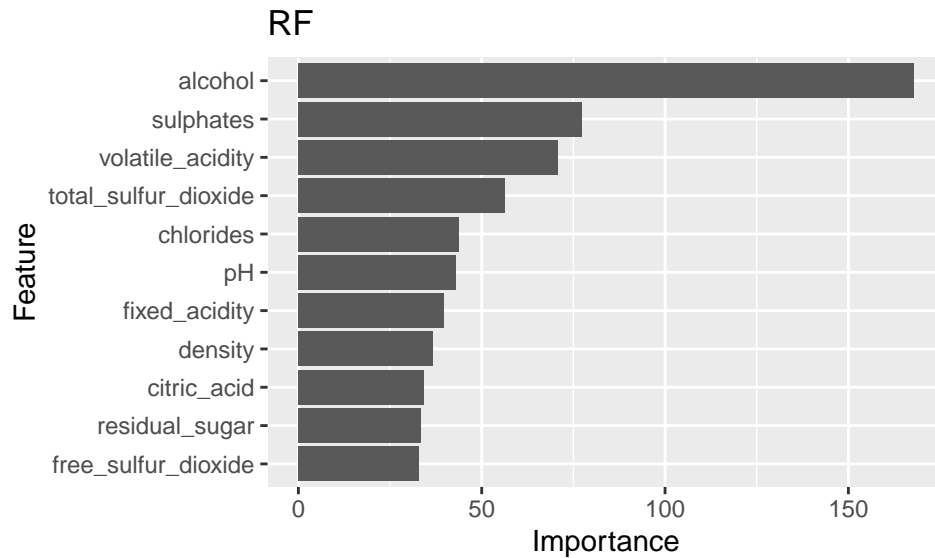
```
wine_rf
```

```
## Random Forest
##
## 1278 samples
## 11 predictor
## 2 classes: '0', '1'
##
## No pre-processing
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 1151, 1150, 1150, 1149, 1150, 1151, ...
## Resampling results across tuning parameters:
##
## mtry Accuracy Kappa
## 2 0.8240111 0.6462582
## 3 0.8177548 0.6333618
## 4 0.8216428 0.6417497
## 5 0.8232236 0.6449335
## 6 0.8201232 0.6385645
## 7 0.8224546 0.6432484
## 8 0.8177244 0.6336485
## 9 0.8240112 0.6464435
## 10 0.8208738 0.6400483
## 11 0.8115108 0.6213133
##
## Accuracy was used to select the optimal model using the largest value.
## The final value used for the model was mtry = 9.
```

```
ggplot(wine_rf)
```



```
p_rf <- varImp(wine_rf, scale = F) %>%
  ggplot() + ggtitle("RF")
p_rf
```



```
# Evaluate
```

```
confusionMatrix(predict(wine_rf, test), test$quality)
```

```
## Confusion Matrix and Statistics
```

```
##
```

```
##           Reference
```

```
## Prediction  0    1
```

```
##           0 115  33
```

```
##           1  34 139
```

```
##
```

```
##           Accuracy : 0.7913
```

```
##           95% CI : (0.7427, 0.8344)
```

```
## No Information Rate : 0.5358
```

```
## P-Value [Acc > NIR] : <2e-16
```

```
##
```

```
##           Kappa : 0.5802
```

```
##
```

```
## McNemar's Test P-Value : 1
```

```
##
```

```
##           Sensitivity : 0.7718
```

```
##           Specificity : 0.8081
```

```
## Pos Pred Value : 0.7770
```

```
## Neg Pred Value : 0.8035
```

```
## Prevalence : 0.4642
```

```
## Detection Rate : 0.3583
```

```
## Detection Prevalence : 0.4611
```

```
## Balanced Accuracy : 0.7900
```

```
##
```

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

```
##
```

```
rf_acc <- confusionMatrix(predict(wine_rf, test), test$quality)$overall["Accuracy"]
rf_acc
```

```
## Accuracy
## 0.7912773
```

#### 4.1.4 svm

```
### SVM -----
```

```
set.seed(1, sample.kind="Rounding")
```

```
# train the model
```

```
wine_svm <- train(
  quality ~ .,
  data = train,
  method = "svmRadial",
  trControl = trControl,
  tuneLength = 10
)
```

```
# Check the model
```

```
wine_svm
```

```
## Support Vector Machines with Radial Basis Function Kernel
```

```
##
```

```
## 1278 samples
```

```
## 11 predictor
```

```
## 2 classes: '0', '1'
```

```
##
```

```
## No pre-processing
```

```
## Resampling: Cross-Validated (10 fold)
```

```
## Summary of sample sizes: 1151, 1150, 1150, 1149, 1150, 1151, ...
```

```
## Resampling results across tuning parameters:
```

```
##
```

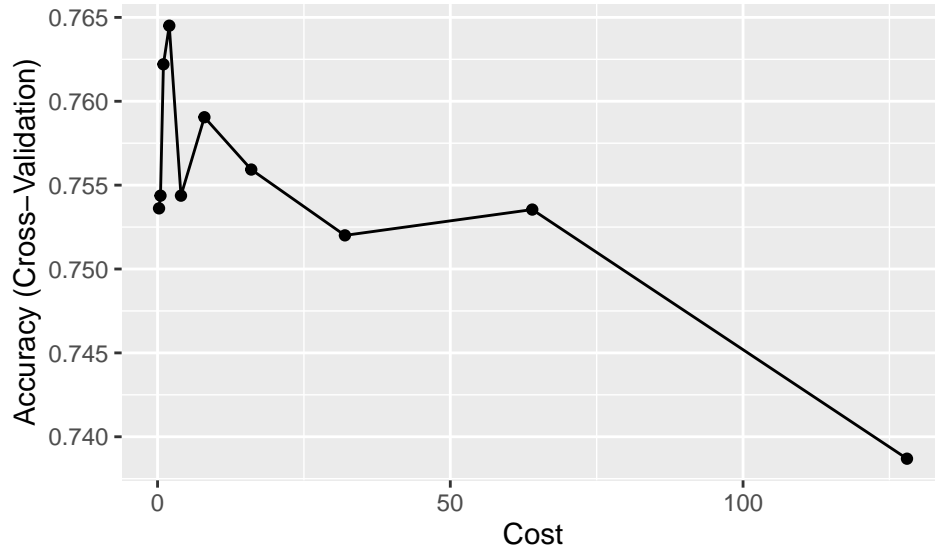
##	C	Accuracy	Kappa
##	0.25	0.7536164	0.5074271
##	0.50	0.7543732	0.5084859
##	1.00	0.7622040	0.5235872
##	2.00	0.7645052	0.5281345
##	4.00	0.7543731	0.5078602
##	8.00	0.7590486	0.5171078
##	16.00	0.7559298	0.5109009
##	32.00	0.7520054	0.5029998
##	64.00	0.7535435	0.5063640
##	128.00	0.7386931	0.4753174

```
##
```

```
## Tuning parameter 'sigma' was held constant at a value of 0.09316274
```

```
## Accuracy was used to select the optimal model using the largest value.
## The final values used for the model were sigma = 0.09316274 and C = 2.
```

```
ggplot(wine_svm)
```



```
# Evaluate
```

```
confusionMatrix(predict(wine_svm, test), test$quality)
```

```
## Confusion Matrix and Statistics
```

```
##
```

```
##           Reference
```

```
## Prediction  0    1
```

```
##           0 108  40
```

```
##           1  41 132
```

```
##
```

```
##           Accuracy : 0.7477
```

```
##           95% CI : (0.6964, 0.7943)
```

```
## No Information Rate : 0.5358
```

```
## P-Value [Acc > NIR] : 4.333e-15
```

```
##
```

```
##           Kappa : 0.4925
```

```
##
```

```
## McNemar's Test P-Value : 1
```

```
##
```

```
##           Sensitivity : 0.7248
```

```
##           Specificity : 0.7674
```

```
## Pos Pred Value : 0.7297
```

```
## Neg Pred Value : 0.7630
```

```
## Prevalence : 0.4642
```

```
## Detection Rate : 0.3364
```

```
## Detection Prevalence : 0.4611
```

```
## Balanced Accuracy : 0.7461
```

```
##
```

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

```
svm_acc <- confusionMatrix(predict(wine_svm, test), test$quality)$overall["Accuracy"]
svm_acc
```

```
## Accuracy
## 0.7476636
```

#### 4.1.5 nnet

```
### Neural Net -----
```

```
set.seed(1, sample.kind="Rounding")
```

```
# train the model
```

```
wine_nnet <- train(
  quality ~ .,
  data = train,
  method = "nnet",
  trControl = trControl,
  tuneLength = 10,
  trace = FALSE
)
```

```
# Check the model
```

```
wine_nnet
```

```
## Neural Network
```

```
##
```

```
## 1278 samples
```

```
## 11 predictor
```

```
## 2 classes: '0', '1'
```

```
##
```

```
## No pre-processing
```

```
## Resampling: Cross-Validated (10 fold)
```

```
## Summary of sample sizes: 1151, 1150, 1150, 1149, 1150, 1151, ...
```

```
## Resampling results across tuning parameters:
```

```
##
```

##	size	decay	Accuracy	Kappa
##	1	0.0000000000	0.7411282	0.4823255
##	1	0.0001000000	0.7411282	0.4823255
##	1	0.0002371374	0.7380032	0.4753936
##	1	0.0005623413	0.7403349	0.4811305
##	1	0.0013335214	0.7411282	0.4823255
##	1	0.0031622777	0.7411282	0.4823255
##	1	0.0074989421	0.7395596	0.4792031
##	1	0.0177827941	0.7387722	0.4776884
##	1	0.0421696503	0.7387722	0.4776884

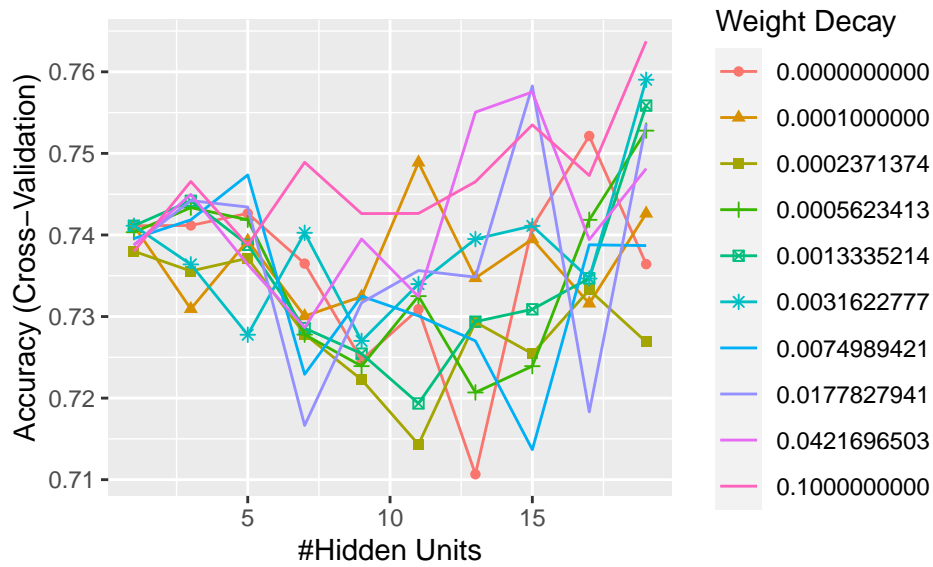
##	1	0.1000000000	0.7379848	0.4760318
##	3	0.0000000000	0.7411651	0.4784584
##	3	0.0001000000	0.7309347	0.4617602
##	3	0.0002371374	0.7355797	0.4707178
##	3	0.0005623413	0.7433558	0.4853736
##	3	0.0013335214	0.7441918	0.4878423
##	3	0.0031622777	0.7363916	0.4725871
##	3	0.0074989421	0.7418546	0.4829034
##	3	0.0177827941	0.7442165	0.4870011
##	3	0.0421696503	0.7449734	0.4871806
##	3	0.1000000000	0.7465662	0.4906912
##	5	0.0000000000	0.7426478	0.4850420
##	5	0.0001000000	0.7392907	0.4752348
##	5	0.0002371374	0.7371484	0.4721912
##	5	0.0005623413	0.7418607	0.4833875
##	5	0.0013335214	0.7387909	0.4763028
##	5	0.0031622777	0.7277672	0.4550144
##	5	0.0074989421	0.7473543	0.4940767
##	5	0.0177827941	0.7434230	0.4856350
##	5	0.0421696503	0.7363855	0.4717189
##	5	0.1000000000	0.7387597	0.4753341
##	7	0.0000000000	0.7364894	0.4707394
##	7	0.0001000000	0.7300495	0.4584486
##	7	0.0002371374	0.7278283	0.4545386
##	7	0.0005623413	0.7277854	0.4528632
##	7	0.0013335214	0.7285785	0.4564756
##	7	0.0031622777	0.7402670	0.4806715
##	7	0.0074989421	0.7229201	0.4418487
##	7	0.0177827941	0.7166457	0.4287281
##	7	0.0421696503	0.7286517	0.4554615
##	7	0.1000000000	0.7489099	0.4970318
##	9	0.0000000000	0.7246781	0.4487092
##	9	0.0001000000	0.7324483	0.4633832
##	9	0.0002371374	0.7222611	0.4414049
##	9	0.0005623413	0.7239212	0.4458059
##	9	0.0013335214	0.7254474	0.4500294
##	9	0.0031622777	0.7270097	0.4504285
##	9	0.0074989421	0.7324978	0.4641461
##	9	0.0177827941	0.7316917	0.4606521
##	9	0.0421696503	0.7395038	0.4774235
##	9	0.1000000000	0.7426174	0.4839751
##	11	0.0000000000	0.7309041	0.4597021
##	11	0.0001000000	0.7488858	0.4957492
##	11	0.0002371374	0.7143204	0.4255914
##	11	0.0005623413	0.7325031	0.4634055
##	11	0.0013335214	0.7193439	0.4373631
##	11	0.0031622777	0.7339806	0.4661220
##	11	0.0074989421	0.7300807	0.4575253
##	11	0.0177827941	0.7356343	0.4693615
##	11	0.0421696503	0.7324427	0.4627939
##	11	0.1000000000	0.7426422	0.4828828
##	13	0.0000000000	0.7106647	0.4184417
##	13	0.0001000000	0.7347250	0.4673072
##	13	0.0002371374	0.7293114	0.4565664

```

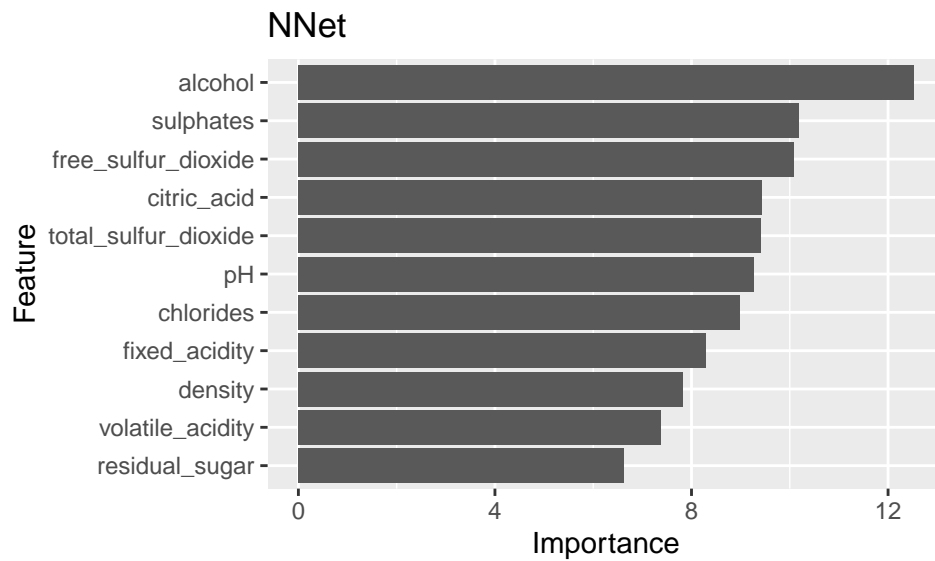
## 13 0.0005623413 0.7206922 0.4388827
## 13 0.0013335214 0.7293720 0.4555830
## 13 0.0031622777 0.7394984 0.4767887
## 13 0.0074989421 0.7270160 0.4514281
## 13 0.0177827941 0.7348292 0.4671441
## 13 0.0421696503 0.7550565 0.5086654
## 13 0.1000000000 0.7465237 0.4915275
## 15 0.0000000000 0.7409996 0.4780454
## 15 0.0001000000 0.7394805 0.4756444
## 15 0.0002371374 0.7254419 0.4472550
## 15 0.0005623413 0.7239164 0.4450928
## 15 0.0013335214 0.7308858 0.4599140
## 15 0.0031622777 0.7411100 0.4810926
## 15 0.0074989421 0.7136859 0.4237930
## 15 0.0177827941 0.7582729 0.5135487
## 15 0.0421696503 0.7575229 0.5117582
## 15 0.1000000000 0.7535003 0.5045235
## 17 0.0000000000 0.7521331 0.5006737
## 17 0.0001000000 0.7315996 0.4599450
## 17 0.0002371374 0.7332608 0.4634188
## 17 0.0005623413 0.7418539 0.4819537
## 17 0.0013335214 0.7346884 0.4671634
## 17 0.0031622777 0.7346575 0.4657347
## 17 0.0074989421 0.7387969 0.4755462
## 17 0.0177827941 0.7183128 0.4350593
## 17 0.0421696503 0.7394313 0.4770899
## 17 0.1000000000 0.7472869 0.4928088
## 19 0.0000000000 0.7364280 0.4702842
## 19 0.0001000000 0.7426296 0.4836545
## 19 0.0002371374 0.7268818 0.4511408
## 19 0.0005623413 0.7527925 0.5029529
## 19 0.0013335214 0.7558563 0.5086475
## 19 0.0031622777 0.7590302 0.5151016
## 19 0.0074989421 0.7386987 0.4738227
## 19 0.0177827941 0.7536167 0.5051497
## 19 0.0421696503 0.7481231 0.4940005
## 19 0.1000000000 0.7637485 0.5260543
##
## Accuracy was used to select the optimal model using the largest value.
## The final values used for the model were size = 19 and decay = 0.1.

```

```
ggplot(wine_nnet)
```



```
p_nnet <- varImp(wine_nnet, scale = F) %>%
  ggplot() + ggtitle("NNet")
p_nnet
```



```
# Evaluate
```

```
confusionMatrix(predict(wine_nnet, test), test$quality)
```

```
## Confusion Matrix and Statistics
```

```
##
```

```
##           Reference
```

```
## Prediction    0    1
```

```
##           0 116 37
```

```
##           1  33 135
```



```
##
##           Accuracy : 0.7819
##           95% CI : (0.7327, 0.8259)
##    No Information Rate : 0.5358
##    P-Value [Acc > NIR] : <2e-16
##
##           Kappa : 0.5624
##
##    Mcnemar's Test P-Value : 0.7199
##
##           Sensitivity : 0.7785
##           Specificity : 0.7849
##           Pos Pred Value : 0.7582
##           Neg Pred Value : 0.8036
##           Prevalence : 0.4642
##           Detection Rate : 0.3614
##    Detection Prevalence : 0.4766
##           Balanced Accuracy : 0.7817
##
##           'Positive' Class : 0
##
```

```
nnet_acc <- confusionMatrix(predict(wine_nnet, test), test$quality)$overall["Accuracy"]
nnet_acc
```

```
## Accuracy
## 0.7819315
```

#### 4.1.6 xgbLinear

```
### XGBoost: xgbLinear----

set.seed(1, sample.kind="Rounding")

# train the model

wine_xgbl <- train(
  quality ~ .,
  data = train,
  method = "xgbLinear",
  trControl = trControl,
  tuneLength = 5
)
```

```
# Check the model

wine_xgbl
```

```
## eXtreme Gradient Boosting
##
## 1278 samples
```

```

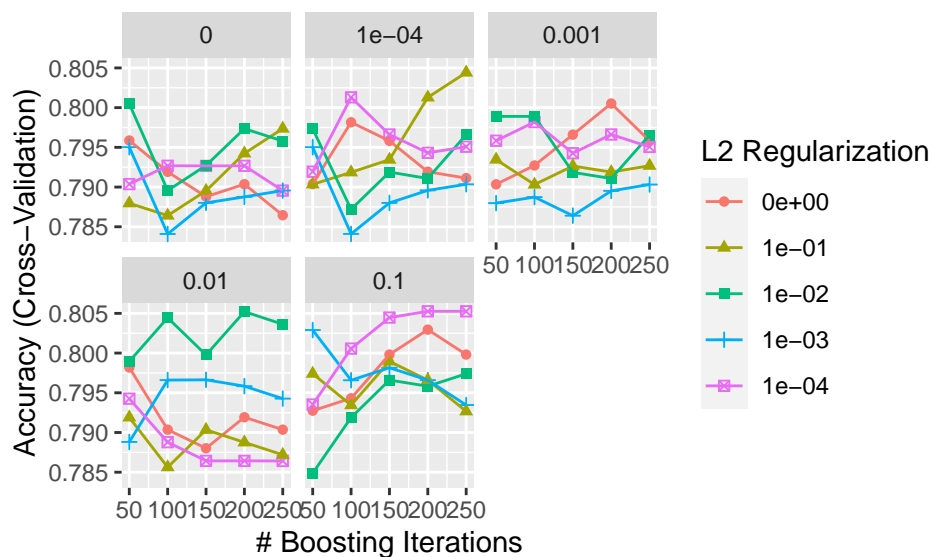
## 11 predictor
## 2 classes: '0', '1'
##
## No pre-processing
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 1151, 1150, 1150, 1149, 1150, 1151, ...
## Resampling results across tuning parameters:
##
##  lambda  alpha  nrounds  Accuracy  Kappa
##  0e+00   0e+00    50      0.7958608  0.5901153
##  0e+00   0e+00   100      0.7919239  0.5819415
##  0e+00   0e+00   150      0.7887863  0.5758016
##  0e+00   0e+00   200      0.7903611  0.5789057
##  0e+00   0e+00   250      0.7864547  0.5711173
##  0e+00   1e-04    50      0.7903736  0.5793526
##  0e+00   1e-04   100      0.7981557  0.5949843
##  0e+00   1e-04   150      0.7958178  0.5899729
##  0e+00   1e-04   200      0.7919299  0.5823010
##  0e+00   1e-04   250      0.7911241  0.5807874
##  0e+00   1e-03    50      0.7903431  0.5793143
##  0e+00   1e-03   100      0.7927174  0.5841713
##  0e+00   1e-03   150      0.7965993  0.5920801
##  0e+00   1e-03   200      0.8005179  0.6001858
##  0e+00   1e-03   250      0.7958119  0.5905590
##  0e+00   1e-02    50      0.7981923  0.5945793
##  0e+00   1e-02   100      0.7903611  0.5789104
##  0e+00   1e-02   150      0.7880173  0.5742684
##  0e+00   1e-02   200      0.7919235  0.5819371
##  0e+00   1e-02   250      0.7903609  0.5790912
##  0e+00   1e-01    50      0.7927477  0.5838754
##  0e+00   1e-01   100      0.7943101  0.5870522
##  0e+00   1e-01   150      0.7998158  0.5983935
##  0e+00   1e-01   200      0.8029531  0.6042649
##  0e+00   1e-01   250      0.7998158  0.5980556
##  1e-04   0e+00    50      0.7903736  0.5791694
##  1e-04   0e+00   100      0.7926808  0.5837737
##  1e-04   0e+00   150      0.7926623  0.5835631
##  1e-04   0e+00   200      0.7926745  0.5836261
##  1e-04   0e+00   250      0.7895555  0.5775625
##  1e-04   1e-04    50      0.7919361  0.5823667
##  1e-04   1e-04   100      0.8012930  0.6013778
##  1e-04   1e-04   150      0.7966175  0.5916675
##  1e-04   1e-04   200      0.7943104  0.5871763
##  1e-04   1e-04   250      0.7950794  0.5887282
##  1e-04   1e-03    50      0.7958365  0.5905345
##  1e-04   1e-03   100      0.7981864  0.5949388
##  1e-04   1e-03   150      0.7942616  0.5870887
##  1e-04   1e-03   200      0.7966175  0.5920263
##  1e-04   1e-03   250      0.7950549  0.5888721
##  1e-04   1e-02    50      0.7942553  0.5867674
##  1e-04   1e-02   100      0.7887863  0.5759726
##  1e-04   1e-02   150      0.7864302  0.5712484
##  1e-04   1e-02   200      0.7864302  0.5710535
##  1e-04   1e-02   250      0.7864179  0.5712193

```

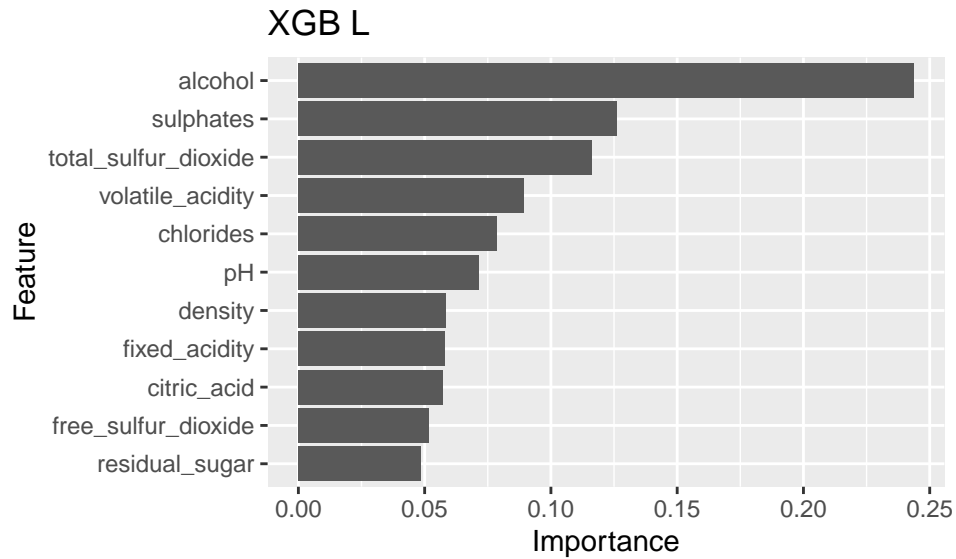
##	1e-04	1e-01	50	0.7935229	0.5855653
##	1e-04	1e-01	100	0.8005604	0.5995975
##	1e-04	1e-01	150	0.8044668	0.6073136
##	1e-04	1e-01	200	0.8052602	0.6087793
##	1e-04	1e-01	250	0.8052603	0.6086133
##	1e-03	0e+00	50	0.7950245	0.5883815
##	1e-03	0e+00	100	0.7840928	0.5662288
##	1e-03	0e+00	150	0.7880114	0.5739719
##	1e-03	0e+00	200	0.7887620	0.5751140
##	1e-03	0e+00	250	0.7895618	0.5770784
##	1e-03	1e-04	50	0.7950245	0.5885734
##	1e-03	1e-04	100	0.7840990	0.5666732
##	1e-03	1e-04	150	0.7880053	0.5744831
##	1e-03	1e-04	200	0.7895862	0.5776089
##	1e-03	1e-04	250	0.7903492	0.5790568
##	1e-03	1e-03	50	0.7879808	0.5742439
##	1e-03	1e-03	100	0.7887377	0.5756211
##	1e-03	1e-03	150	0.7863938	0.5711292
##	1e-03	1e-03	200	0.7895189	0.5774744
##	1e-03	1e-03	250	0.7903186	0.5791545
##	1e-03	1e-02	50	0.7888109	0.5757457
##	1e-03	1e-02	100	0.7966052	0.5914658
##	1e-03	1e-02	150	0.7966357	0.5918528
##	1e-03	1e-02	200	0.7958361	0.5900723
##	1e-03	1e-02	250	0.7942674	0.5867986
##	1e-03	1e-01	50	0.8029043	0.6046522
##	1e-03	1e-01	100	0.7965806	0.5914601
##	1e-03	1e-01	150	0.7981555	0.5949643
##	1e-03	1e-01	200	0.7965929	0.5917418
##	1e-03	1e-01	250	0.7934618	0.5853906
##	1e-02	0e+00	50	0.8004750	0.5995516
##	1e-02	0e+00	100	0.7895615	0.5778349
##	1e-02	0e+00	150	0.7927231	0.5840040
##	1e-02	0e+00	200	0.7973497	0.5934654
##	1e-02	0e+00	250	0.7957931	0.5901931
##	1e-02	1e-04	50	0.7973254	0.5930882
##	1e-02	1e-04	100	0.7871993	0.5732167
##	1e-02	1e-04	150	0.7918869	0.5822658
##	1e-02	1e-04	200	0.7910874	0.5806455
##	1e-02	1e-04	250	0.7965991	0.5920779
##	1e-02	1e-03	50	0.7989002	0.5962895
##	1e-02	1e-03	100	0.7989002	0.5962534
##	1e-02	1e-03	150	0.7918505	0.5816562
##	1e-02	1e-03	200	0.7910752	0.5803013
##	1e-02	1e-03	250	0.7965441	0.5912807
##	1e-02	1e-02	50	0.7989493	0.5959013
##	1e-02	1e-02	100	0.8044427	0.6065880
##	1e-02	1e-02	150	0.7997671	0.5976010
##	1e-02	1e-02	200	0.8052175	0.6085573
##	1e-02	1e-02	250	0.8036305	0.6054588
##	1e-02	1e-01	50	0.7848859	0.5680917
##	1e-02	1e-01	100	0.7918806	0.5818803
##	1e-02	1e-01	150	0.7966049	0.5915145
##	1e-02	1e-01	200	0.7958299	0.5899191

```
## 1e-02 1e-01 250 0.7973924 0.5932490
## 1e-01 0e+00 50 0.7879562 0.5743817
## 1e-01 0e+00 100 0.7864119 0.5713114
## 1e-01 0e+00 150 0.7895309 0.5776237
## 1e-01 0e+00 200 0.7942368 0.5871749
## 1e-01 0e+00 250 0.7973619 0.5933139
## 1e-01 1e-04 50 0.7903427 0.5792903
## 1e-01 1e-04 100 0.7918564 0.5817421
## 1e-01 1e-04 150 0.7934494 0.5853833
## 1e-01 1e-04 200 0.8012867 0.6007985
## 1e-01 1e-04 250 0.8044178 0.6074000
## 1e-01 1e-03 50 0.7934620 0.5854748
## 1e-01 1e-03 100 0.7903123 0.5791110
## 1e-01 1e-03 150 0.7926501 0.5835578
## 1e-01 1e-03 200 0.7918871 0.5822555
## 1e-01 1e-03 250 0.7926868 0.5839328
## 1e-01 1e-02 50 0.7918993 0.5824580
## 1e-01 1e-02 100 0.7856307 0.5693689
## 1e-01 1e-02 150 0.7903428 0.5787425
## 1e-01 1e-02 200 0.7887621 0.5757647
## 1e-01 1e-02 250 0.7871933 0.5725978
## 1e-01 1e-01 50 0.7974049 0.5938090
## 1e-01 1e-01 100 0.7934435 0.5857668
## 1e-01 1e-01 150 0.7989614 0.5964540
## 1e-01 1e-01 200 0.7965930 0.5917967
## 1e-01 1e-01 250 0.7926683 0.5838032
##
## Tuning parameter 'eta' was held constant at a value of 0.3
## Accuracy was used to select the optimal model using the largest value.
## The final values used for the model were nrounds = 250, lambda = 1e-04, alpha
## = 0.1 and eta = 0.3.
```

```
ggplot(wine_xgb1)
```



```
p_xgb1 <- varImp(wine_xgb1, scale = F) %>%
  ggplot() + ggtitle("XGB L")
p_xgb1
```



```
# Evaluate
```

```
confusionMatrix(predict(wine_xgb1, test), test$quality)
```

```
## Confusion Matrix and Statistics
```

```
##
```

```
##           Reference
```

```
## Prediction  0    1
```

```
##           0 117  29
```

```
##           1  32 143
```

```
##
```

```
##           Accuracy : 0.81
```

```
##           95% CI : (0.7627, 0.8514)
```

```
## No Information Rate : 0.5358
```

```
## P-Value [Acc > NIR] : <2e-16
```

```
##
```

```
##           Kappa : 0.6175
```

```
##
```

```
## McNemar's Test P-Value : 0.7979
```

```
##
```

```
##           Sensitivity : 0.7852
```

```
##           Specificity : 0.8314
```

```
## Pos Pred Value : 0.8014
```

```
## Neg Pred Value : 0.8171
```

```
## Prevalence : 0.4642
```

```
## Detection Rate : 0.3645
```

```
## Detection Prevalence : 0.4548
```

```
## Balanced Accuracy : 0.8083
```

```
##
```

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

```
xgbl_acc <- confusionMatrix(predict(wine_xgbl, test), test$quality)$overall["Accuracy"]
xgbl_acc
```

```
## Accuracy
## 0.8099688
```

#### 4.1.7 xgbDART

```
### XGBoost: xgbDART----
```

```
set.seed(1, sample.kind="Rounding")
```

```
# train the model
```

```
wine_xgbd <- train(
  quality ~ .,
  data = train,
  method = "xgbDART",
  trControl = trControl,
  tuneLength = 3
)
```

```
# Check the model
```

```
wine_xgbd
```

```
## eXtreme Gradient Boosting
##
## 1278 samples
## 11 predictor
## 2 classes: '0', '1'
##
```

```
## No pre-processing
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 1151, 1150, 1150, 1149, 1150, 1151, ...
## Resampling results across tuning parameters:
##
```

##	max_depth	eta	rate_drop	skip_drop	subsample	colsample_bytree	nrounds
##	1	0.3	0.01	0.05	0.50	0.6	50
##	1	0.3	0.01	0.05	0.50	0.6	100
##	1	0.3	0.01	0.05	0.50	0.6	150
##	1	0.3	0.01	0.05	0.50	0.8	50
##	1	0.3	0.01	0.05	0.50	0.8	100
##	1	0.3	0.01	0.05	0.50	0.8	150
##	1	0.3	0.01	0.05	0.75	0.6	50
##	1	0.3	0.01	0.05	0.75	0.6	100
##	1	0.3	0.01	0.05	0.75	0.6	150
##	1	0.3	0.01	0.05	0.75	0.8	50

##	1	0.3	0.01	0.05	0.75	0.8	100
##	1	0.3	0.01	0.05	0.75	0.8	150
##	1	0.3	0.01	0.05	1.00	0.6	50
##	1	0.3	0.01	0.05	1.00	0.6	100
##	1	0.3	0.01	0.05	1.00	0.6	150
##	1	0.3	0.01	0.05	1.00	0.8	50
##	1	0.3	0.01	0.05	1.00	0.8	100
##	1	0.3	0.01	0.05	1.00	0.8	150
##	1	0.3	0.01	0.95	0.50	0.6	50
##	1	0.3	0.01	0.95	0.50	0.6	100
##	1	0.3	0.01	0.95	0.50	0.6	150
##	1	0.3	0.01	0.95	0.50	0.8	50
##	1	0.3	0.01	0.95	0.50	0.8	100
##	1	0.3	0.01	0.95	0.50	0.8	150
##	1	0.3	0.01	0.95	0.75	0.6	50
##	1	0.3	0.01	0.95	0.75	0.6	100
##	1	0.3	0.01	0.95	0.75	0.6	150
##	1	0.3	0.01	0.95	0.75	0.8	50
##	1	0.3	0.01	0.95	0.75	0.8	100
##	1	0.3	0.01	0.95	0.75	0.8	150
##	1	0.3	0.01	0.95	1.00	0.6	50
##	1	0.3	0.01	0.95	1.00	0.6	100
##	1	0.3	0.01	0.95	1.00	0.6	150
##	1	0.3	0.01	0.95	1.00	0.8	50
##	1	0.3	0.01	0.95	1.00	0.8	100
##	1	0.3	0.01	0.95	1.00	0.8	150
##	1	0.3	0.50	0.05	0.50	0.6	50
##	1	0.3	0.50	0.05	0.50	0.6	100
##	1	0.3	0.50	0.05	0.50	0.6	150
##	1	0.3	0.50	0.05	0.50	0.8	50
##	1	0.3	0.50	0.05	0.50	0.8	100
##	1	0.3	0.50	0.05	0.50	0.8	150
##	1	0.3	0.50	0.05	0.75	0.6	50
##	1	0.3	0.50	0.05	0.75	0.6	100
##	1	0.3	0.50	0.05	0.75	0.6	150
##	1	0.3	0.50	0.05	0.75	0.8	50
##	1	0.3	0.50	0.05	0.75	0.8	100
##	1	0.3	0.50	0.05	0.75	0.8	150
##	1	0.3	0.50	0.05	1.00	0.6	50
##	1	0.3	0.50	0.05	1.00	0.6	100
##	1	0.3	0.50	0.05	1.00	0.6	150
##	1	0.3	0.50	0.05	1.00	0.8	50
##	1	0.3	0.50	0.05	1.00	0.8	100
##	1	0.3	0.50	0.05	1.00	0.8	150
##	1	0.3	0.50	0.95	0.50	0.6	50
##	1	0.3	0.50	0.95	0.50	0.6	100
##	1	0.3	0.50	0.95	0.50	0.6	150
##	1	0.3	0.50	0.95	0.50	0.8	50
##	1	0.3	0.50	0.95	0.50	0.8	100
##	1	0.3	0.50	0.95	0.50	0.8	150
##	1	0.3	0.50	0.95	0.75	0.6	50
##	1	0.3	0.50	0.95	0.75	0.6	100
##	1	0.3	0.50	0.95	0.75	0.6	150
##	1	0.3	0.50	0.95	0.75	0.8	50

##	1	0.3	0.50	0.95	0.75	0.8	100
##	1	0.3	0.50	0.95	0.75	0.8	150
##	1	0.3	0.50	0.95	1.00	0.6	50
##	1	0.3	0.50	0.95	1.00	0.6	100
##	1	0.3	0.50	0.95	1.00	0.6	150
##	1	0.3	0.50	0.95	1.00	0.8	50
##	1	0.3	0.50	0.95	1.00	0.8	100
##	1	0.3	0.50	0.95	1.00	0.8	150
##	1	0.4	0.01	0.05	0.50	0.6	50
##	1	0.4	0.01	0.05	0.50	0.6	100
##	1	0.4	0.01	0.05	0.50	0.6	150
##	1	0.4	0.01	0.05	0.50	0.8	50
##	1	0.4	0.01	0.05	0.50	0.8	100
##	1	0.4	0.01	0.05	0.50	0.8	150
##	1	0.4	0.01	0.05	0.75	0.6	50
##	1	0.4	0.01	0.05	0.75	0.6	100
##	1	0.4	0.01	0.05	0.75	0.6	150
##	1	0.4	0.01	0.05	0.75	0.8	50
##	1	0.4	0.01	0.05	0.75	0.8	100
##	1	0.4	0.01	0.05	0.75	0.8	150
##	1	0.4	0.01	0.05	1.00	0.6	50
##	1	0.4	0.01	0.05	1.00	0.6	100
##	1	0.4	0.01	0.05	1.00	0.6	150
##	1	0.4	0.01	0.05	1.00	0.8	50
##	1	0.4	0.01	0.05	1.00	0.8	100
##	1	0.4	0.01	0.05	1.00	0.8	150
##	1	0.4	0.01	0.95	0.50	0.6	50
##	1	0.4	0.01	0.95	0.50	0.6	100
##	1	0.4	0.01	0.95	0.50	0.6	150
##	1	0.4	0.01	0.95	0.50	0.8	50
##	1	0.4	0.01	0.95	0.50	0.8	100
##	1	0.4	0.01	0.95	0.50	0.8	150
##	1	0.4	0.01	0.95	0.75	0.6	50
##	1	0.4	0.01	0.95	0.75	0.6	100
##	1	0.4	0.01	0.95	0.75	0.6	150
##	1	0.4	0.01	0.95	0.75	0.8	50
##	1	0.4	0.01	0.95	0.75	0.8	100
##	1	0.4	0.01	0.95	0.75	0.8	150
##	1	0.4	0.01	0.95	1.00	0.6	50
##	1	0.4	0.01	0.95	1.00	0.6	100
##	1	0.4	0.01	0.95	1.00	0.6	150
##	1	0.4	0.01	0.95	1.00	0.8	50
##	1	0.4	0.01	0.95	1.00	0.8	100
##	1	0.4	0.01	0.95	1.00	0.8	150
##	1	0.4	0.50	0.05	0.50	0.6	50
##	1	0.4	0.50	0.05	0.50	0.6	100
##	1	0.4	0.50	0.05	0.50	0.6	150
##	1	0.4	0.50	0.05	0.50	0.8	50
##	1	0.4	0.50	0.05	0.50	0.8	100
##	1	0.4	0.50	0.05	0.50	0.8	150
##	1	0.4	0.50	0.05	0.75	0.6	50
##	1	0.4	0.50	0.05	0.75	0.6	100
##	1	0.4	0.50	0.05	0.75	0.6	150
##	1	0.4	0.50	0.05	0.75	0.8	50



##	1	0.4	0.50	0.05	0.75	0.8	100
##	1	0.4	0.50	0.05	0.75	0.8	150
##	1	0.4	0.50	0.05	1.00	0.6	50
##	1	0.4	0.50	0.05	1.00	0.6	100
##	1	0.4	0.50	0.05	1.00	0.6	150
##	1	0.4	0.50	0.05	1.00	0.8	50
##	1	0.4	0.50	0.05	1.00	0.8	100
##	1	0.4	0.50	0.05	1.00	0.8	150
##	1	0.4	0.50	0.95	0.50	0.6	50
##	1	0.4	0.50	0.95	0.50	0.6	100
##	1	0.4	0.50	0.95	0.50	0.6	150
##	1	0.4	0.50	0.95	0.50	0.8	50
##	1	0.4	0.50	0.95	0.50	0.8	100
##	1	0.4	0.50	0.95	0.50	0.8	150
##	1	0.4	0.50	0.95	0.75	0.6	50
##	1	0.4	0.50	0.95	0.75	0.6	100
##	1	0.4	0.50	0.95	0.75	0.6	150
##	1	0.4	0.50	0.95	0.75	0.8	50
##	1	0.4	0.50	0.95	0.75	0.8	100
##	1	0.4	0.50	0.95	0.75	0.8	150
##	1	0.4	0.50	0.95	1.00	0.6	50
##	1	0.4	0.50	0.95	1.00	0.6	100
##	1	0.4	0.50	0.95	1.00	0.6	150
##	1	0.4	0.50	0.95	1.00	0.8	50
##	1	0.4	0.50	0.95	1.00	0.8	100
##	1	0.4	0.50	0.95	1.00	0.8	150
##	2	0.3	0.01	0.05	0.50	0.6	50
##	2	0.3	0.01	0.05	0.50	0.6	100
##	2	0.3	0.01	0.05	0.50	0.6	150
##	2	0.3	0.01	0.05	0.50	0.8	50
##	2	0.3	0.01	0.05	0.50	0.8	100
##	2	0.3	0.01	0.05	0.50	0.8	150
##	2	0.3	0.01	0.05	0.75	0.6	50
##	2	0.3	0.01	0.05	0.75	0.6	100
##	2	0.3	0.01	0.05	0.75	0.6	150
##	2	0.3	0.01	0.05	0.75	0.8	50
##	2	0.3	0.01	0.05	0.75	0.8	100
##	2	0.3	0.01	0.05	0.75	0.8	150
##	2	0.3	0.01	0.05	1.00	0.6	50
##	2	0.3	0.01	0.05	1.00	0.6	100
##	2	0.3	0.01	0.05	1.00	0.6	150
##	2	0.3	0.01	0.05	1.00	0.8	50
##	2	0.3	0.01	0.05	1.00	0.8	100
##	2	0.3	0.01	0.05	1.00	0.8	150
##	2	0.3	0.01	0.95	0.50	0.6	50
##	2	0.3	0.01	0.95	0.50	0.6	100
##	2	0.3	0.01	0.95	0.50	0.6	150
##	2	0.3	0.01	0.95	0.50	0.8	50
##	2	0.3	0.01	0.95	0.50	0.8	100
##	2	0.3	0.01	0.95	0.50	0.8	150
##	2	0.3	0.01	0.95	0.75	0.6	50
##	2	0.3	0.01	0.95	0.75	0.6	100
##	2	0.3	0.01	0.95	0.75	0.6	150
##	2	0.3	0.01	0.95	0.75	0.8	50

##	2	0.3	0.01	0.95	0.75	0.8	100
##	2	0.3	0.01	0.95	0.75	0.8	150
##	2	0.3	0.01	0.95	1.00	0.6	50
##	2	0.3	0.01	0.95	1.00	0.6	100
##	2	0.3	0.01	0.95	1.00	0.6	150
##	2	0.3	0.01	0.95	1.00	0.8	50
##	2	0.3	0.01	0.95	1.00	0.8	100
##	2	0.3	0.01	0.95	1.00	0.8	150
##	2	0.3	0.50	0.05	0.50	0.6	50
##	2	0.3	0.50	0.05	0.50	0.6	100
##	2	0.3	0.50	0.05	0.50	0.6	150
##	2	0.3	0.50	0.05	0.50	0.8	50
##	2	0.3	0.50	0.05	0.50	0.8	100
##	2	0.3	0.50	0.05	0.50	0.8	150
##	2	0.3	0.50	0.05	0.75	0.6	50
##	2	0.3	0.50	0.05	0.75	0.6	100
##	2	0.3	0.50	0.05	0.75	0.6	150
##	2	0.3	0.50	0.05	0.75	0.8	50
##	2	0.3	0.50	0.05	0.75	0.8	100
##	2	0.3	0.50	0.05	0.75	0.8	150
##	2	0.3	0.50	0.05	1.00	0.6	50
##	2	0.3	0.50	0.05	1.00	0.6	100
##	2	0.3	0.50	0.05	1.00	0.6	150
##	2	0.3	0.50	0.05	1.00	0.8	50
##	2	0.3	0.50	0.05	1.00	0.8	100
##	2	0.3	0.50	0.05	1.00	0.8	150
##	2	0.3	0.50	0.95	0.50	0.6	50
##	2	0.3	0.50	0.95	0.50	0.6	100
##	2	0.3	0.50	0.95	0.50	0.6	150
##	2	0.3	0.50	0.95	0.50	0.8	50
##	2	0.3	0.50	0.95	0.50	0.8	100
##	2	0.3	0.50	0.95	0.50	0.8	150
##	2	0.3	0.50	0.95	0.75	0.6	50
##	2	0.3	0.50	0.95	0.75	0.6	100
##	2	0.3	0.50	0.95	0.75	0.6	150
##	2	0.3	0.50	0.95	0.75	0.8	50
##	2	0.3	0.50	0.95	0.75	0.8	100
##	2	0.3	0.50	0.95	0.75	0.8	150
##	2	0.3	0.50	0.95	1.00	0.6	50
##	2	0.3	0.50	0.95	1.00	0.6	100
##	2	0.3	0.50	0.95	1.00	0.6	150
##	2	0.3	0.50	0.95	1.00	0.8	50
##	2	0.3	0.50	0.95	1.00	0.8	100
##	2	0.3	0.50	0.95	1.00	0.8	150
##	2	0.4	0.01	0.05	0.50	0.6	50
##	2	0.4	0.01	0.05	0.50	0.6	100
##	2	0.4	0.01	0.05	0.50	0.6	150
##	2	0.4	0.01	0.05	0.50	0.8	50
##	2	0.4	0.01	0.05	0.50	0.8	100
##	2	0.4	0.01	0.05	0.50	0.8	150
##	2	0.4	0.01	0.05	0.75	0.6	50
##	2	0.4	0.01	0.05	0.75	0.6	100
##	2	0.4	0.01	0.05	0.75	0.6	150
##	2	0.4	0.01	0.05	0.75	0.8	50

##	2	0.4	0.01	0.05	0.75	0.8	100
##	2	0.4	0.01	0.05	0.75	0.8	150
##	2	0.4	0.01	0.05	1.00	0.6	50
##	2	0.4	0.01	0.05	1.00	0.6	100
##	2	0.4	0.01	0.05	1.00	0.6	150
##	2	0.4	0.01	0.05	1.00	0.8	50
##	2	0.4	0.01	0.05	1.00	0.8	100
##	2	0.4	0.01	0.05	1.00	0.8	150
##	2	0.4	0.01	0.95	0.50	0.6	50
##	2	0.4	0.01	0.95	0.50	0.6	100
##	2	0.4	0.01	0.95	0.50	0.6	150
##	2	0.4	0.01	0.95	0.50	0.8	50
##	2	0.4	0.01	0.95	0.50	0.8	100
##	2	0.4	0.01	0.95	0.50	0.8	150
##	2	0.4	0.01	0.95	0.75	0.6	50
##	2	0.4	0.01	0.95	0.75	0.6	100
##	2	0.4	0.01	0.95	0.75	0.6	150
##	2	0.4	0.01	0.95	0.75	0.8	50
##	2	0.4	0.01	0.95	0.75	0.8	100
##	2	0.4	0.01	0.95	0.75	0.8	150
##	2	0.4	0.01	0.95	1.00	0.6	50
##	2	0.4	0.01	0.95	1.00	0.6	100
##	2	0.4	0.01	0.95	1.00	0.6	150
##	2	0.4	0.01	0.95	1.00	0.8	50
##	2	0.4	0.01	0.95	1.00	0.8	100
##	2	0.4	0.01	0.95	1.00	0.8	150
##	2	0.4	0.50	0.05	0.50	0.6	50
##	2	0.4	0.50	0.05	0.50	0.6	100
##	2	0.4	0.50	0.05	0.50	0.6	150
##	2	0.4	0.50	0.05	0.50	0.8	50
##	2	0.4	0.50	0.05	0.50	0.8	100
##	2	0.4	0.50	0.05	0.50	0.8	150
##	2	0.4	0.50	0.05	0.75	0.6	50
##	2	0.4	0.50	0.05	0.75	0.6	100
##	2	0.4	0.50	0.05	0.75	0.6	150
##	2	0.4	0.50	0.05	0.75	0.8	50
##	2	0.4	0.50	0.05	0.75	0.8	100
##	2	0.4	0.50	0.05	0.75	0.8	150
##	2	0.4	0.50	0.05	1.00	0.6	50
##	2	0.4	0.50	0.05	1.00	0.6	100
##	2	0.4	0.50	0.05	1.00	0.6	150
##	2	0.4	0.50	0.05	1.00	0.8	50
##	2	0.4	0.50	0.05	1.00	0.8	100
##	2	0.4	0.50	0.05	1.00	0.8	150
##	2	0.4	0.50	0.95	0.50	0.6	50
##	2	0.4	0.50	0.95	0.50	0.6	100
##	2	0.4	0.50	0.95	0.50	0.6	150
##	2	0.4	0.50	0.95	0.50	0.8	50
##	2	0.4	0.50	0.95	0.50	0.8	100
##	2	0.4	0.50	0.95	0.50	0.8	150
##	2	0.4	0.50	0.95	0.75	0.6	50
##	2	0.4	0.50	0.95	0.75	0.6	100
##	2	0.4	0.50	0.95	0.75	0.6	150
##	2	0.4	0.50	0.95	0.75	0.8	50

##	2	0.4	0.50	0.95	0.75	0.8	100
##	2	0.4	0.50	0.95	0.75	0.8	150
##	2	0.4	0.50	0.95	1.00	0.6	50
##	2	0.4	0.50	0.95	1.00	0.6	100
##	2	0.4	0.50	0.95	1.00	0.6	150
##	2	0.4	0.50	0.95	1.00	0.8	50
##	2	0.4	0.50	0.95	1.00	0.8	100
##	2	0.4	0.50	0.95	1.00	0.8	150
##	3	0.3	0.01	0.05	0.50	0.6	50
##	3	0.3	0.01	0.05	0.50	0.6	100
##	3	0.3	0.01	0.05	0.50	0.6	150
##	3	0.3	0.01	0.05	0.50	0.8	50
##	3	0.3	0.01	0.05	0.50	0.8	100
##	3	0.3	0.01	0.05	0.50	0.8	150
##	3	0.3	0.01	0.05	0.75	0.6	50
##	3	0.3	0.01	0.05	0.75	0.6	100
##	3	0.3	0.01	0.05	0.75	0.6	150
##	3	0.3	0.01	0.05	0.75	0.8	50
##	3	0.3	0.01	0.05	0.75	0.8	100
##	3	0.3	0.01	0.05	0.75	0.8	150
##	3	0.3	0.01	0.05	1.00	0.6	50
##	3	0.3	0.01	0.05	1.00	0.6	100
##	3	0.3	0.01	0.05	1.00	0.6	150
##	3	0.3	0.01	0.05	1.00	0.8	50
##	3	0.3	0.01	0.05	1.00	0.8	100
##	3	0.3	0.01	0.05	1.00	0.8	150
##	3	0.3	0.01	0.95	0.50	0.6	50
##	3	0.3	0.01	0.95	0.50	0.6	100
##	3	0.3	0.01	0.95	0.50	0.6	150
##	3	0.3	0.01	0.95	0.50	0.8	50
##	3	0.3	0.01	0.95	0.50	0.8	100
##	3	0.3	0.01	0.95	0.50	0.8	150
##	3	0.3	0.01	0.95	0.75	0.6	50
##	3	0.3	0.01	0.95	0.75	0.6	100
##	3	0.3	0.01	0.95	0.75	0.6	150
##	3	0.3	0.01	0.95	0.75	0.8	50
##	3	0.3	0.01	0.95	0.75	0.8	100
##	3	0.3	0.01	0.95	0.75	0.8	150
##	3	0.3	0.01	0.95	1.00	0.6	50
##	3	0.3	0.01	0.95	1.00	0.6	100
##	3	0.3	0.01	0.95	1.00	0.6	150
##	3	0.3	0.01	0.95	1.00	0.8	50
##	3	0.3	0.01	0.95	1.00	0.8	100
##	3	0.3	0.01	0.95	1.00	0.8	150
##	3	0.3	0.50	0.05	0.50	0.6	50
##	3	0.3	0.50	0.05	0.50	0.6	100
##	3	0.3	0.50	0.05	0.50	0.6	150
##	3	0.3	0.50	0.05	0.50	0.8	50
##	3	0.3	0.50	0.05	0.50	0.8	100
##	3	0.3	0.50	0.05	0.50	0.8	150
##	3	0.3	0.50	0.05	0.75	0.6	50
##	3	0.3	0.50	0.05	0.75	0.6	100
##	3	0.3	0.50	0.05	0.75	0.6	150
##	3	0.3	0.50	0.05	0.75	0.8	50

##	3	0.3	0.50	0.05	0.75	0.8	100
##	3	0.3	0.50	0.05	0.75	0.8	150
##	3	0.3	0.50	0.05	1.00	0.6	50
##	3	0.3	0.50	0.05	1.00	0.6	100
##	3	0.3	0.50	0.05	1.00	0.6	150
##	3	0.3	0.50	0.05	1.00	0.8	50
##	3	0.3	0.50	0.05	1.00	0.8	100
##	3	0.3	0.50	0.05	1.00	0.8	150
##	3	0.3	0.50	0.95	0.50	0.6	50
##	3	0.3	0.50	0.95	0.50	0.6	100
##	3	0.3	0.50	0.95	0.50	0.6	150
##	3	0.3	0.50	0.95	0.50	0.8	50
##	3	0.3	0.50	0.95	0.50	0.8	100
##	3	0.3	0.50	0.95	0.50	0.8	150
##	3	0.3	0.50	0.95	0.75	0.6	50
##	3	0.3	0.50	0.95	0.75	0.6	100
##	3	0.3	0.50	0.95	0.75	0.6	150
##	3	0.3	0.50	0.95	0.75	0.8	50
##	3	0.3	0.50	0.95	0.75	0.8	100
##	3	0.3	0.50	0.95	0.75	0.8	150
##	3	0.3	0.50	0.95	1.00	0.6	50
##	3	0.3	0.50	0.95	1.00	0.6	100
##	3	0.3	0.50	0.95	1.00	0.6	150
##	3	0.3	0.50	0.95	1.00	0.8	50
##	3	0.3	0.50	0.95	1.00	0.8	100
##	3	0.3	0.50	0.95	1.00	0.8	150
##	3	0.4	0.01	0.05	0.50	0.6	50
##	3	0.4	0.01	0.05	0.50	0.6	100
##	3	0.4	0.01	0.05	0.50	0.6	150
##	3	0.4	0.01	0.05	0.50	0.8	50
##	3	0.4	0.01	0.05	0.50	0.8	100
##	3	0.4	0.01	0.05	0.50	0.8	150
##	3	0.4	0.01	0.05	0.75	0.6	50
##	3	0.4	0.01	0.05	0.75	0.6	100
##	3	0.4	0.01	0.05	0.75	0.6	150
##	3	0.4	0.01	0.05	0.75	0.8	50
##	3	0.4	0.01	0.05	0.75	0.8	100
##	3	0.4	0.01	0.05	0.75	0.8	150
##	3	0.4	0.01	0.05	1.00	0.6	50
##	3	0.4	0.01	0.05	1.00	0.6	100
##	3	0.4	0.01	0.05	1.00	0.6	150
##	3	0.4	0.01	0.05	1.00	0.8	50
##	3	0.4	0.01	0.05	1.00	0.8	100
##	3	0.4	0.01	0.05	1.00	0.8	150
##	3	0.4	0.01	0.95	0.50	0.6	50
##	3	0.4	0.01	0.95	0.50	0.6	100
##	3	0.4	0.01	0.95	0.50	0.6	150
##	3	0.4	0.01	0.95	0.50	0.8	50
##	3	0.4	0.01	0.95	0.50	0.8	100
##	3	0.4	0.01	0.95	0.50	0.8	150
##	3	0.4	0.01	0.95	0.75	0.6	50
##	3	0.4	0.01	0.95	0.75	0.6	100
##	3	0.4	0.01	0.95	0.75	0.6	150
##	3	0.4	0.01	0.95	0.75	0.8	50

##	3	0.4	0.01	0.95	0.75	0.8	100
##	3	0.4	0.01	0.95	0.75	0.8	150
##	3	0.4	0.01	0.95	1.00	0.6	50
##	3	0.4	0.01	0.95	1.00	0.6	100
##	3	0.4	0.01	0.95	1.00	0.6	150
##	3	0.4	0.01	0.95	1.00	0.8	50
##	3	0.4	0.01	0.95	1.00	0.8	100
##	3	0.4	0.01	0.95	1.00	0.8	150
##	3	0.4	0.50	0.05	0.50	0.6	50
##	3	0.4	0.50	0.05	0.50	0.6	100
##	3	0.4	0.50	0.05	0.50	0.6	150
##	3	0.4	0.50	0.05	0.50	0.8	50
##	3	0.4	0.50	0.05	0.50	0.8	100
##	3	0.4	0.50	0.05	0.50	0.8	150
##	3	0.4	0.50	0.05	0.75	0.6	50
##	3	0.4	0.50	0.05	0.75	0.6	100
##	3	0.4	0.50	0.05	0.75	0.6	150
##	3	0.4	0.50	0.05	0.75	0.8	50
##	3	0.4	0.50	0.05	0.75	0.8	100
##	3	0.4	0.50	0.05	0.75	0.8	150
##	3	0.4	0.50	0.05	1.00	0.6	50
##	3	0.4	0.50	0.05	1.00	0.6	100
##	3	0.4	0.50	0.05	1.00	0.6	150
##	3	0.4	0.50	0.05	1.00	0.8	50
##	3	0.4	0.50	0.05	1.00	0.8	100
##	3	0.4	0.50	0.05	1.00	0.8	150
##	3	0.4	0.50	0.95	0.50	0.6	50
##	3	0.4	0.50	0.95	0.50	0.6	100
##	3	0.4	0.50	0.95	0.50	0.6	150
##	3	0.4	0.50	0.95	0.50	0.8	50
##	3	0.4	0.50	0.95	0.50	0.8	100
##	3	0.4	0.50	0.95	0.50	0.8	150
##	3	0.4	0.50	0.95	0.75	0.6	50
##	3	0.4	0.50	0.95	0.75	0.6	100
##	3	0.4	0.50	0.95	0.75	0.6	150
##	3	0.4	0.50	0.95	0.75	0.8	50
##	3	0.4	0.50	0.95	0.75	0.8	100
##	3	0.4	0.50	0.95	0.75	0.8	150
##	3	0.4	0.50	0.95	1.00	0.6	50
##	3	0.4	0.50	0.95	1.00	0.6	100
##	3	0.4	0.50	0.95	1.00	0.6	150
##	3	0.4	0.50	0.95	1.00	0.8	50
##	3	0.4	0.50	0.95	1.00	0.8	100
##	3	0.4	0.50	0.95	1.00	0.8	150
##	Accuracy	Kappa					
##	0.7497159	0.4983671					
##	0.7442107	0.4873371					
##	0.7575594	0.5135290					
##	0.7473234	0.4918025					
##	0.7426478	0.4828382					
##	0.7403652	0.4784462					
##	0.7528165	0.5045089					
##	0.7520841	0.5027772					
##	0.7614778	0.5217758					

##	0.7457914	0.4898114
##	0.7551665	0.5086517
##	0.7481474	0.4939660
##	0.7442107	0.4869664
##	0.7442103	0.4862419
##	0.7450101	0.4883825
##	0.7458036	0.4898711
##	0.7465544	0.4915522
##	0.7442287	0.4867366
##	0.7614290	0.5205448
##	0.7599154	0.5179690
##	0.7544768	0.5078229
##	0.7536041	0.5056455
##	0.7481719	0.4946523
##	0.7598421	0.5177496
##	0.7465606	0.4918453
##	0.7583346	0.5151970
##	0.7504667	0.4984730
##	0.7411099	0.4806803
##	0.7370932	0.4722076
##	0.7449183	0.4868178
##	0.7520477	0.5026315
##	0.7450344	0.4881204
##	0.7450282	0.4879140
##	0.7520294	0.5026778
##	0.7504972	0.4992803
##	0.7465847	0.4913918
##	0.7340476	0.4661962
##	0.7340358	0.4662813
##	0.7363732	0.4705253
##	0.7184223	0.4356873
##	0.7277853	0.4536676
##	0.7316366	0.4616480
##	0.7340847	0.4667690
##	0.7332665	0.4651501
##	0.7340660	0.4665014
##	0.7184284	0.4370510
##	0.7278036	0.4549562
##	0.7261801	0.4515153
##	0.7230795	0.4447425
##	0.7301235	0.4587427
##	0.7269922	0.4531429
##	0.7129472	0.4256110
##	0.7145221	0.4283863
##	0.7207112	0.4402903
##	0.7465605	0.4922284
##	0.7473296	0.4925656
##	0.7559480	0.5101636
##	0.7504851	0.4998167
##	0.7512845	0.5008646
##	0.7481167	0.4939979
##	0.7418851	0.4821970
##	0.7450100	0.4881239
##	0.7512481	0.5001955

##	0.7426782	0.4835908
##	0.7465849	0.4908374
##	0.7481410	0.4942322
##	0.7496793	0.4982617
##	0.7527800	0.5040115
##	0.7457486	0.4898433
##	0.7489105	0.4960815
##	0.7489469	0.4963322
##	0.7418727	0.4817270
##	0.7465544	0.4913622
##	0.7402858	0.4790882
##	0.7527983	0.5036157
##	0.7505093	0.4986776
##	0.7488859	0.4957171
##	0.7481048	0.4946595
##	0.7528533	0.5042743
##	0.7575347	0.5135199
##	0.7575777	0.5141150
##	0.7473539	0.4928442
##	0.7418729	0.4817203
##	0.7434721	0.4858680
##	0.7457241	0.4891337
##	0.7465053	0.4910096
##	0.7473171	0.4930409
##	0.7395108	0.4774539
##	0.7348350	0.4683394
##	0.7371972	0.4730611
##	0.7575716	0.5137214
##	0.7558870	0.5103797
##	0.7488798	0.4956140
##	0.7512296	0.5002625
##	0.7528350	0.5034979
##	0.7396021	0.4784003
##	0.7528045	0.5037266
##	0.7496732	0.4978552
##	0.7551847	0.5091313
##	0.7434169	0.4847886
##	0.7458035	0.4897076
##	0.7591340	0.5170156
##	0.7418484	0.4823093
##	0.7387536	0.4757101
##	0.7442043	0.4867678
##	0.7528228	0.5043801
##	0.7434474	0.4855301
##	0.7403101	0.4788835
##	0.7372155	0.4722609
##	0.7426114	0.4831121
##	0.7418790	0.4815039
##	0.7254842	0.4511654
##	0.7379481	0.4745015
##	0.7465665	0.4919386
##	0.7293598	0.4582041
##	0.7347737	0.4683771
##	0.7379050	0.4749417



##	0.7333216	0.4658892
##	0.7285848	0.4557254
##	0.7324913	0.4647320
##	0.7301596	0.4590836
##	0.7309346	0.4606884
##	0.7324912	0.4630974
##	0.7199664	0.4393506
##	0.7176103	0.4340939
##	0.7301354	0.4583006
##	0.7481289	0.4940910
##	0.7544465	0.5069671
##	0.7489411	0.4959667
##	0.7441555	0.4858690
##	0.7464932	0.4911927
##	0.7481415	0.4943790
##	0.7481291	0.4948740
##	0.7449916	0.4869869
##	0.7543608	0.5063482
##	0.7559787	0.5106186
##	0.7536347	0.5057075
##	0.7442348	0.4861235
##	0.7457790	0.4906827
##	0.7434717	0.4854868
##	0.7465603	0.4914688
##	0.7457730	0.4906041
##	0.7450467	0.4886374
##	0.7411034	0.4804889
##	0.7552338	0.5082950
##	0.7685036	0.5357125
##	0.7606846	0.5204918
##	0.7434108	0.4857931
##	0.7685278	0.5354183
##	0.7646215	0.5278198
##	0.7520356	0.5019495
##	0.7559112	0.5096608
##	0.7622103	0.5232596
##	0.7575716	0.5141638
##	0.7512540	0.5017475
##	0.7645847	0.5282335
##	0.7512789	0.5008768
##	0.7544038	0.5075689
##	0.7567231	0.5120508
##	0.7536103	0.5051129
##	0.7535980	0.5049734
##	0.7590791	0.5163924
##	0.7598422	0.5182140
##	0.7677161	0.5341326
##	0.7629733	0.5242819
##	0.7504971	0.4988317
##	0.7567658	0.5113737
##	0.7614962	0.5215375
##	0.7598851	0.5186842
##	0.7606969	0.5196739
##	0.7739360	0.5463016

##	0.7700844	0.5391434
##	0.7739539	0.5469298
##	0.7770667	0.5530386
##	0.7559172	0.5098241
##	0.7614535	0.5213812
##	0.7778664	0.5544467
##	0.7481781	0.4955339
##	0.7590423	0.5166632
##	0.7707740	0.5399312
##	0.7457912	0.4898446
##	0.7527801	0.5041687
##	0.7512173	0.5014254
##	0.7434539	0.4852334
##	0.7457914	0.4898988
##	0.7465786	0.4916078
##	0.7379236	0.4742573
##	0.7395107	0.4771375
##	0.7481599	0.4946190
##	0.7418971	0.4820953
##	0.7466215	0.4915612
##	0.7474274	0.4935907
##	0.7465911	0.4924985
##	0.7481535	0.4958670
##	0.7512542	0.5023151
##	0.7340480	0.4669078
##	0.7387357	0.4764342
##	0.7356043	0.4701281
##	0.7512545	0.5006750
##	0.7434357	0.4847269
##	0.7552032	0.5087446
##	0.7536716	0.5063743
##	0.7622288	0.5227966
##	0.7606603	0.5193616
##	0.7512481	0.5006576
##	0.7520659	0.5018525
##	0.7669164	0.5315431
##	0.7575411	0.5141819
##	0.7708349	0.5405970
##	0.7700657	0.5385503
##	0.7520600	0.5024529
##	0.7583103	0.5155132
##	0.7622104	0.5228007
##	0.7489349	0.4959772
##	0.7700966	0.5394220
##	0.7747843	0.5483670
##	0.7449857	0.4872838
##	0.7497591	0.4976229
##	0.7646153	0.5270106
##	0.7637912	0.5258442
##	0.7653967	0.5292132
##	0.7700599	0.5382093
##	0.7630218	0.5240992
##	0.7786785	0.5560278
##	0.7794227	0.5577162

##	0.7536716	0.5056531
##	0.7559543	0.5097284
##	0.7661659	0.5303260
##	0.7450651	0.4885207
##	0.7513155	0.5013403
##	0.7661108	0.5308765
##	0.7489530	0.4959894
##	0.7535797	0.5055404
##	0.7567478	0.5118571
##	0.7536467	0.5060087
##	0.7637547	0.5267672
##	0.7629735	0.5240598
##	0.7333400	0.4653736
##	0.7590183	0.5160234
##	0.7551244	0.5079716
##	0.7513031	0.5019998
##	0.7661596	0.5309009
##	0.7622104	0.5227611
##	0.7543731	0.5070088
##	0.7684181	0.5356609
##	0.7731726	0.5448798
##	0.7551549	0.5082559
##	0.7669164	0.5319052
##	0.7739297	0.5456870
##	0.7567231	0.5120034
##	0.7716101	0.5415155
##	0.7739419	0.5459200
##	0.7512849	0.5012514
##	0.7488982	0.4960361
##	0.7504911	0.4996551
##	0.7457670	0.4899705
##	0.7481292	0.4943643
##	0.7434418	0.4844325
##	0.7356592	0.4698310
##	0.7411100	0.4814780
##	0.7488982	0.4963759
##	0.7434413	0.4852637
##	0.7442352	0.4870470
##	0.7348598	0.4682965
##	0.7418911	0.4814912
##	0.7410975	0.4804275
##	0.7457913	0.4905047
##	0.7356102	0.4702231
##	0.7364222	0.4720085
##	0.7426726	0.4840577
##	0.7551361	0.5090029
##	0.7606967	0.5199586
##	0.7661718	0.5312050
##	0.7504669	0.4988326
##	0.7496793	0.4977747
##	0.7637482	0.5248853
##	0.7497285	0.4980348
##	0.7528719	0.5051305
##	0.7630098	0.5245215

##	0.7505096	0.4993403
##	0.7599275	0.5184302
##	0.7606723	0.5196970
##	0.7512482	0.5003440
##	0.7661109	0.5305865
##	0.7629796	0.5247499
##	0.7590855	0.5163661
##	0.7707618	0.5399110
##	0.7778177	0.5542516
##	0.7582736	0.5143255
##	0.7661292	0.5301032
##	0.7739786	0.5458381
##	0.7661844	0.5314738
##	0.7662331	0.5304651
##	0.7708962	0.5403974
##	0.7763221	0.5508636
##	0.7817972	0.5626044
##	0.7848857	0.5686208
##	0.7770548	0.5529817
##	0.7802165	0.5587359
##	0.7911669	0.5807157
##	0.7708166	0.5404564
##	0.7801984	0.5589429
##	0.7856672	0.5693612
##	0.7677593	0.5344805
##	0.7755171	0.5494091
##	0.7746869	0.5477575
##	0.7653779	0.5291431
##	0.7755410	0.5488784
##	0.7849408	0.5677202
##	0.7685159	0.5365232
##	0.7802107	0.5591523
##	0.7848924	0.5686645
##	0.7669959	0.5331609
##	0.7723671	0.5431638
##	0.7731853	0.5443346
##	0.7684794	0.5354209
##	0.7731547	0.5437403
##	0.7872603	0.5723558
##	0.7692422	0.5374817
##	0.7832808	0.5645761
##	0.7966296	0.5913254
##	0.7693217	0.5378312
##	0.7865095	0.5719374
##	0.8013662	0.6013898
##	0.7441612	0.4871144
##	0.7536042	0.5057047
##	0.7622351	0.5233060
##	0.7535984	0.5062166
##	0.7598851	0.5189953
##	0.7543674	0.5073862
##	0.7552155	0.5093982
##	0.7567722	0.5126022
##	0.7528964	0.5044147

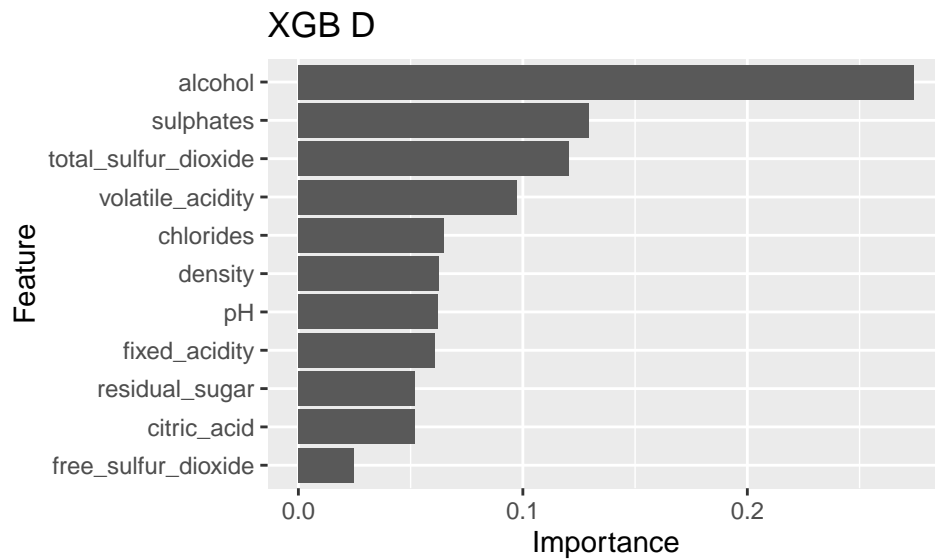
##	0.7450285	0.4888145
##	0.7528473	0.5041834
##	0.7544038	0.5072496
##	0.7489594	0.4966055
##	0.7528718	0.5050760
##	0.7489716	0.4969371
##	0.7465910	0.4922821
##	0.7489531	0.4971388
##	0.7473419	0.4937725
##	0.7654273	0.5289115
##	0.7896283	0.5777093
##	0.7801861	0.5579459
##	0.7559298	0.5094848
##	0.7629917	0.5237699
##	0.7700600	0.5374206
##	0.7708472	0.5407540
##	0.7716591	0.5423870
##	0.7841351	0.5664391
##	0.7544037	0.5078720
##	0.7802290	0.5593268
##	0.7895619	0.5767329
##	0.7637122	0.5257017
##	0.7793620	0.5573121
##	0.7848307	0.5679483
##	0.7739970	0.5474229
##	0.7935594	0.5862204
##	0.7934554	0.5854653
##	0.7606783	0.5203228
##	0.7567597	0.5125926
##	0.7708349	0.5401015
##	0.7685461	0.5356244
##	0.7692727	0.5370790
##	0.7777934	0.5540147
##	0.7676978	0.5339294
##	0.7872540	0.5731519
##	0.7770856	0.5522683
##	0.7614292	0.5216178
##	0.7802473	0.5590781
##	0.7935105	0.5855624
##	0.7769996	0.5527750
##	0.7840920	0.5668851
##	0.7918682	0.5819707
##	0.7653846	0.5291483
##	0.7849470	0.5688832
##	0.7856429	0.5699349
##	0.7700538	0.5380623
##	0.7693217	0.5366251
##	0.7779340	0.5546612
##	0.7645601	0.5278307
##	0.7825357	0.5640843
##	0.7887434	0.5756092
##	0.7748698	0.5480559
##	0.7888595	0.5754108
##	0.7880298	0.5732457

```

## 0.7801492 0.5590974
## 0.7950550 0.5878944
## 0.7871874 0.5720343
## 0.7786359 0.5556366
## 0.7919115 0.5822809
## 0.7918932 0.5823356
## 0.7661964 0.5322750
## 0.7825297 0.5640681
## 0.7778054 0.5541320
## 0.7520538 0.5022460
## 0.7583652 0.5154517
## 0.7598908 0.5186096
## 0.7512725 0.5018899
## 0.7504911 0.4986901
## 0.7528167 0.5029555
## 0.7426417 0.4838367
## 0.7551851 0.5091382
## 0.7575533 0.5141521
## 0.7465910 0.4915078
## 0.7481901 0.4951543
## 0.7481654 0.4944052
## 0.7340419 0.4666193
## 0.7473542 0.4939962
## 0.7528596 0.5046646
## 0.7489472 0.4966711
## 0.7520970 0.5034028
## 0.7544226 0.5082088
## 0.7558382 0.5097792
## 0.7613681 0.5207012
## 0.7731056 0.5443237
## 0.7512911 0.5009146
## 0.7739232 0.5464936
## 0.7856916 0.5699228
## 0.7754740 0.5494951
## 0.7872359 0.5722705
## 0.7840865 0.5662124
## 0.7661593 0.5314979
## 0.7809976 0.5604040
## 0.7833724 0.5650625
## 0.7770670 0.5533568
## 0.7864365 0.5718220
## 0.7817306 0.5619040
## 0.7802349 0.5591755
## 0.7841294 0.5665451
## 0.7856981 0.5697512
##
## Tuning parameter 'gamma' was held constant at a value of 0
## Tuning
## parameter 'min_child_weight' was held constant at a value of 1
## Accuracy was used to select the optimal model using the largest value.
## The final values used for the model were nrounds = 150, max_depth = 3, eta
## = 0.3, gamma = 0, subsample = 1, colsample_bytree = 0.8, rate_drop =
## 0.01, skip_drop = 0.95 and min_child_weight = 1.

```

```
p_xgbd <- varImp(wine_xgbd, scale = F) %>%
  ggplot() + ggtitle("XGB D")
p_xgbd
```



```
# Evaluate
```

```
confusionMatrix(predict(wine_xgbd, test), test$quality)
```

```
## [23:49:02] WARNING: amalgamation/./src/c_api/c_api.cc:718: 'ntree_limit' is deprecated, use 'iterat
```

```
## Confusion Matrix and Statistics
```

```
##
```

```
##           Reference
```

```
## Prediction  0    1
```

```
##           0 118  27
```

```
##           1  31 145
```

```
##
```

```
##           Accuracy : 0.8193
```

```
##           95% CI : (0.7728, 0.8598)
```

```
##           No Information Rate : 0.5358
```

```
##           P-Value [Acc > NIR] : <2e-16
```

```
##
```

```
##           Kappa : 0.6361
```

```
##
```

```
##           McNemar's Test P-Value : 0.6936
```

```
##
```

```
##           Sensitivity : 0.7919
```

```
##           Specificity : 0.8430
```

```
##           Pos Pred Value : 0.8138
```

```
##           Neg Pred Value : 0.8239
```

```
##           Prevalence : 0.4642
```

```
##           Detection Rate : 0.3676
```

```
##           Detection Prevalence : 0.4517
```

```
##      Balanced Accuracy : 0.8175
##
##      'Positive' Class : 0
##
```

```
xgbd_acc <- confusionMatrix(predict(wine_xgbd, test), test$quality)$overall["Accuracy"]
```

```
## [23:49:02] WARNING: amalgamation/./src/c_api/c_api.cc:718: 'ntree_limit' is deprecated, use 'iterat
```

```
xgbd_acc
```

```
## Accuracy
## 0.8193146
```

#### 4.1.8 xgbTree

```
### XGBoost: xgbTree----
```

```
set.seed(1, sample.kind="Rounding")
```

```
# train the model
```

```
wine_xgbt <- train(
  quality ~ .,
  data = train,
  method = "xgbTree",
  trControl = trControl,
  tuneLength = 5
)
```

```
# Check the model
```

```
wine_xgbt
```

```
## eXtreme Gradient Boosting
```

```
##
```

```
## 1278 samples
```

```
## 11 predictor
```

```
## 2 classes: '0', '1'
```

```
##
```

```
## No pre-processing
```

```
## Resampling: Cross-Validated (10 fold)
```

```
## Summary of sample sizes: 1151, 1150, 1150, 1149, 1150, 1151, ...
```

```
## Resampling results across tuning parameters:
```

```
##
```

##	eta	max_depth	colsample_bytree	subsample	nrounds	Accuracy	Kappa
##	0.3	1	0.6	0.500	50	0.7496487	0.4979813
##	0.3	1	0.6	0.500	100	0.7449613	0.4873684
##	0.3	1	0.6	0.500	150	0.7465784	0.4910973
##	0.3	1	0.6	0.500	200	0.7434353	0.4853094



##	0.3	1	0.6	0.500	250	0.7497222	0.4974293
##	0.3	1	0.6	0.625	50	0.7505154	0.4991680
##	0.3	1	0.6	0.625	100	0.7497162	0.4976127
##	0.3	1	0.6	0.625	150	0.7583225	0.5151200
##	0.3	1	0.6	0.625	200	0.7528229	0.5034805
##	0.3	1	0.6	0.625	250	0.7489410	0.4966740
##	0.3	1	0.6	0.750	50	0.7449430	0.4889398
##	0.3	1	0.6	0.750	100	0.7496672	0.4979979
##	0.3	1	0.6	0.750	150	0.7441860	0.4865818
##	0.3	1	0.6	0.750	200	0.7489043	0.4965776
##	0.3	1	0.6	0.750	250	0.7481413	0.4943523
##	0.3	1	0.6	0.875	50	0.7379540	0.4736665
##	0.3	1	0.6	0.875	100	0.7426722	0.4830625
##	0.3	1	0.6	0.875	150	0.7466215	0.4909662
##	0.3	1	0.6	0.875	200	0.7481595	0.4939573
##	0.3	1	0.6	0.875	250	0.7442349	0.4861466
##	0.3	1	0.6	1.000	50	0.7520355	0.5029666
##	0.3	1	0.6	1.000	100	0.7497281	0.4978286
##	0.3	1	0.6	1.000	150	0.7465848	0.4917048
##	0.3	1	0.6	1.000	200	0.7481351	0.4945275
##	0.3	1	0.6	1.000	250	0.7473417	0.4930898
##	0.3	1	0.8	0.500	50	0.7520966	0.5029548
##	0.3	1	0.8	0.500	100	0.7496979	0.4979277
##	0.3	1	0.8	0.500	150	0.7458158	0.4899657
##	0.3	1	0.8	0.500	200	0.7513032	0.5012891
##	0.3	1	0.8	0.500	250	0.7473725	0.4932435
##	0.3	1	0.8	0.625	50	0.7425990	0.4829670
##	0.3	1	0.8	0.625	100	0.7488492	0.4951492
##	0.3	1	0.8	0.625	150	0.7504604	0.4986381
##	0.3	1	0.8	0.625	200	0.7403529	0.4791129
##	0.3	1	0.8	0.625	250	0.7481901	0.4948563
##	0.3	1	0.8	0.750	50	0.7449856	0.4885751
##	0.3	1	0.8	0.750	100	0.7465605	0.4912466
##	0.3	1	0.8	0.750	150	0.7513091	0.5012014
##	0.3	1	0.8	0.750	200	0.7482025	0.4953127
##	0.3	1	0.8	0.750	250	0.7536409	0.5061851
##	0.3	1	0.8	0.875	50	0.7520477	0.5024606
##	0.3	1	0.8	0.875	100	0.7488919	0.4961297
##	0.3	1	0.8	0.875	150	0.7543486	0.5066420
##	0.3	1	0.8	0.875	200	0.7559296	0.5101490
##	0.3	1	0.8	0.875	250	0.7551726	0.5085133
##	0.3	1	0.8	1.000	50	0.7512664	0.5011293
##	0.3	1	0.8	1.000	100	0.7465724	0.4914618
##	0.3	1	0.8	1.000	150	0.7442408	0.4865798
##	0.3	1	0.8	1.000	200	0.7434473	0.4849342
##	0.3	1	0.8	1.000	250	0.7465726	0.4913530
##	0.3	2	0.6	0.500	50	0.7622592	0.5234359
##	0.3	2	0.6	0.500	100	0.7699925	0.5392508
##	0.3	2	0.6	0.500	150	0.7442104	0.4872543
##	0.3	2	0.6	0.500	200	0.7488980	0.4964396
##	0.3	2	0.6	0.500	250	0.7550994	0.5086932
##	0.3	2	0.6	0.625	50	0.7629729	0.5240823
##	0.3	2	0.6	0.625	100	0.7699984	0.5380904
##	0.3	2	0.6	0.625	150	0.7621798	0.5219712

##	0.3	2	0.6	0.625	200	0.7786599	0.5554678
##	0.3	2	0.6	0.625	250	0.7778603	0.5535694
##	0.3	2	0.6	0.750	50	0.7410791	0.4806174
##	0.3	2	0.6	0.750	100	0.7629854	0.5246135
##	0.3	2	0.6	0.750	150	0.7684727	0.5353029
##	0.3	2	0.6	0.750	200	0.7653720	0.5295794
##	0.3	2	0.6	0.750	250	0.7754860	0.5497023
##	0.3	2	0.6	0.875	50	0.7645847	0.5271147
##	0.3	2	0.6	0.875	100	0.7638098	0.5259728
##	0.3	2	0.6	0.875	150	0.7802410	0.5588354
##	0.3	2	0.6	0.875	200	0.7818585	0.5622250
##	0.3	2	0.6	0.875	250	0.7896165	0.5771385
##	0.3	2	0.6	1.000	50	0.7497223	0.4978322
##	0.3	2	0.6	1.000	100	0.7629916	0.5246452
##	0.3	2	0.6	1.000	150	0.7716040	0.5419000
##	0.3	2	0.6	1.000	200	0.7770610	0.5523641
##	0.3	2	0.6	1.000	250	0.7738871	0.5458254
##	0.3	2	0.8	0.500	50	0.7575290	0.5135304
##	0.3	2	0.8	0.500	100	0.7574617	0.5130588
##	0.3	2	0.8	0.500	150	0.7621982	0.5229346
##	0.3	2	0.8	0.500	200	0.7637605	0.5257170
##	0.3	2	0.8	0.500	250	0.7614355	0.5212132
##	0.3	2	0.8	0.625	50	0.7536407	0.5050883
##	0.3	2	0.8	0.625	100	0.7614105	0.5208827
##	0.3	2	0.8	0.625	150	0.7676244	0.5334112
##	0.3	2	0.8	0.625	200	0.7700845	0.5387586
##	0.3	2	0.8	0.625	250	0.7856856	0.5691240
##	0.3	2	0.8	0.750	50	0.7598358	0.5184010
##	0.3	2	0.8	0.750	100	0.7575288	0.5141114
##	0.3	2	0.8	0.750	150	0.7660557	0.5302995
##	0.3	2	0.8	0.750	200	0.7582738	0.5147619
##	0.3	2	0.8	0.750	250	0.7613988	0.5209545
##	0.3	2	0.8	0.875	50	0.7669104	0.5324161
##	0.3	2	0.8	0.875	100	0.7739174	0.5467007
##	0.3	2	0.8	0.875	150	0.7723363	0.5428081
##	0.3	2	0.8	0.875	200	0.7708042	0.5397119
##	0.3	2	0.8	0.875	250	0.7746924	0.5473674
##	0.3	2	0.8	1.000	50	0.7505098	0.4997416
##	0.3	2	0.8	1.000	100	0.7622166	0.5233791
##	0.3	2	0.8	1.000	150	0.7691993	0.5366535
##	0.3	2	0.8	1.000	200	0.7723060	0.5426161
##	0.3	2	0.8	1.000	250	0.7723122	0.5425294
##	0.3	3	0.6	0.500	50	0.7622348	0.5227416
##	0.3	3	0.6	0.500	100	0.7692298	0.5362816
##	0.3	3	0.6	0.500	150	0.7770304	0.5523454
##	0.3	3	0.6	0.500	200	0.7801680	0.5580750
##	0.3	3	0.6	0.500	250	0.7833053	0.5648145
##	0.3	3	0.6	0.625	50	0.7598790	0.5189398
##	0.3	3	0.6	0.625	100	0.7786173	0.5561995
##	0.3	3	0.6	0.625	150	0.7864731	0.5712056
##	0.3	3	0.6	0.625	200	0.7872788	0.5725230
##	0.3	3	0.6	0.625	250	0.7817974	0.5615035
##	0.3	3	0.6	0.750	50	0.7755534	0.5502194
##	0.3	3	0.6	0.750	100	0.7896348	0.5778411

##	0.3	3	0.6	0.750	150	0.7934497	0.5851513
##	0.3	3	0.6	0.750	200	0.7919052	0.5818144
##	0.3	3	0.6	0.750	250	0.7911424	0.5804949
##	0.3	3	0.6	0.875	50	0.7739541	0.5472262
##	0.3	3	0.6	0.875	100	0.7879866	0.5743419
##	0.3	3	0.6	0.875	150	0.7895799	0.5775571
##	0.3	3	0.6	0.875	200	0.7809307	0.5599835
##	0.3	3	0.6	0.875	250	0.7785502	0.5549837
##	0.3	3	0.6	1.000	50	0.7707983	0.5398264
##	0.3	3	0.6	1.000	100	0.7763224	0.5512796
##	0.3	3	0.6	1.000	150	0.7887802	0.5759604
##	0.3	3	0.6	1.000	200	0.7935165	0.5856769
##	0.3	3	0.6	1.000	250	0.7888232	0.5760199
##	0.3	3	0.8	0.500	50	0.7622411	0.5227151
##	0.3	3	0.8	0.500	100	0.7731915	0.5442609
##	0.3	3	0.8	0.500	150	0.7786541	0.5551835
##	0.3	3	0.8	0.500	200	0.7880663	0.5743062
##	0.3	3	0.8	0.500	250	0.7755475	0.5491527
##	0.3	3	0.8	0.625	50	0.7567048	0.5114848
##	0.3	3	0.8	0.625	100	0.7645057	0.5272179
##	0.3	3	0.8	0.625	150	0.7832439	0.5639043
##	0.3	3	0.8	0.625	200	0.7848188	0.5671119
##	0.3	3	0.8	0.625	250	0.7800824	0.5579814
##	0.3	3	0.8	0.750	50	0.7614844	0.5224570
##	0.3	3	0.8	0.750	100	0.7888659	0.5771488
##	0.3	3	0.8	0.750	150	0.7872485	0.5729713
##	0.3	3	0.8	0.750	200	0.7841171	0.5661450
##	0.3	3	0.8	0.750	250	0.7872115	0.5725945
##	0.3	3	0.8	0.875	50	0.7732032	0.5451414
##	0.3	3	0.8	0.875	100	0.7880109	0.5746701
##	0.3	3	0.8	0.875	150	0.7872175	0.5723508
##	0.3	3	0.8	0.875	200	0.7919419	0.5820184
##	0.3	3	0.8	0.875	250	0.7911301	0.5804708
##	0.3	3	0.8	1.000	50	0.7615086	0.5214383
##	0.3	3	0.8	1.000	100	0.7715739	0.5407831
##	0.3	3	0.8	1.000	150	0.7887682	0.5754822
##	0.3	3	0.8	1.000	200	0.7840621	0.5658800
##	0.3	3	0.8	1.000	250	0.7840497	0.5660374
##	0.3	4	0.6	0.500	50	0.7700236	0.5385152
##	0.3	4	0.6	0.500	100	0.7902878	0.5784652
##	0.3	4	0.6	0.500	150	0.7957815	0.5893390
##	0.3	4	0.6	0.500	200	0.7879748	0.5732948
##	0.3	4	0.6	0.500	250	0.7879563	0.5735625
##	0.3	4	0.6	0.625	50	0.7802591	0.5587225
##	0.3	4	0.6	0.625	100	0.7857039	0.5697863
##	0.3	4	0.6	0.625	150	0.7817793	0.5618455
##	0.3	4	0.6	0.625	200	0.7918809	0.5820712
##	0.3	4	0.6	0.625	250	0.7848185	0.5682002
##	0.3	4	0.6	0.750	50	0.7856673	0.5694546
##	0.3	4	0.6	0.750	100	0.7927233	0.5833754
##	0.3	4	0.6	0.750	150	0.7888290	0.5760581
##	0.3	4	0.6	0.750	200	0.7919419	0.5822275
##	0.3	4	0.6	0.750	250	0.7919479	0.5822724
##	0.3	4	0.6	0.875	50	0.7833174	0.5651369

##	0.3	4	0.6	0.875	100	0.7966236	0.5912135
##	0.3	4	0.6	0.875	150	0.7903734	0.5789977
##	0.3	4	0.6	0.875	200	0.7903977	0.5791135
##	0.3	4	0.6	0.875	250	0.7911971	0.5809806
##	0.3	4	0.6	1.000	50	0.7873399	0.5734610
##	0.3	4	0.6	1.000	100	0.7982229	0.5948509
##	0.3	4	0.6	1.000	150	0.7982351	0.5951877
##	0.3	4	0.6	1.000	200	0.8005421	0.5996593
##	0.3	4	0.6	1.000	250	0.7974476	0.5934838
##	0.3	4	0.8	0.500	50	0.7895979	0.5770846
##	0.3	4	0.8	0.500	100	0.7903793	0.5784700
##	0.3	4	0.8	0.500	150	0.7989489	0.5961828
##	0.3	4	0.8	0.500	200	0.7989735	0.5960480
##	0.3	4	0.8	0.500	250	0.7973862	0.5931361
##	0.3	4	0.8	0.625	50	0.7693028	0.5365449
##	0.3	4	0.8	0.625	100	0.7871750	0.5720375
##	0.3	4	0.8	0.625	150	0.7918931	0.5817755
##	0.3	4	0.8	0.625	200	0.7887434	0.5756207
##	0.3	4	0.8	0.625	250	0.7918747	0.5816409
##	0.3	4	0.8	0.750	50	0.7920027	0.5827138
##	0.3	4	0.8	0.750	100	0.7982959	0.5952875
##	0.3	4	0.8	0.750	150	0.7935534	0.5856912
##	0.3	4	0.8	0.750	200	0.7959156	0.5901755
##	0.3	4	0.8	0.750	250	0.7974293	0.5933849
##	0.3	4	0.8	0.875	50	0.7786357	0.5557589
##	0.3	4	0.8	0.875	100	0.7864058	0.5705646
##	0.3	4	0.8	0.875	150	0.7825175	0.5629837
##	0.3	4	0.8	0.875	200	0.7856733	0.5691224
##	0.3	4	0.8	0.875	250	0.7848981	0.5676694
##	0.3	4	0.8	1.000	50	0.7857040	0.5692326
##	0.3	4	0.8	1.000	100	0.7848799	0.5679847
##	0.3	4	0.8	1.000	150	0.7872234	0.5729158
##	0.3	4	0.8	1.000	200	0.7864119	0.5715685
##	0.3	4	0.8	1.000	250	0.7840680	0.5665247
##	0.3	5	0.6	0.500	50	0.7786298	0.5550211
##	0.3	5	0.6	0.500	100	0.7817488	0.5611171
##	0.3	5	0.6	0.500	150	0.7817490	0.5608772
##	0.3	5	0.6	0.500	200	0.7801679	0.5575508
##	0.3	5	0.6	0.500	250	0.7817978	0.5607806
##	0.3	5	0.6	0.625	50	0.7864364	0.5709254
##	0.3	5	0.6	0.625	100	0.8021110	0.6019869
##	0.3	5	0.6	0.625	150	0.8021046	0.6021478
##	0.3	5	0.6	0.625	200	0.7950793	0.5879980
##	0.3	5	0.6	0.625	250	0.7958482	0.5897298
##	0.3	5	0.6	0.750	50	0.7997670	0.5970532
##	0.3	5	0.6	0.750	100	0.8013175	0.6000909
##	0.3	5	0.6	0.750	150	0.7950613	0.5878309
##	0.3	5	0.6	0.750	200	0.7903737	0.5787964
##	0.3	5	0.6	0.750	250	0.7966362	0.5914354
##	0.3	5	0.6	0.875	50	0.7965930	0.5913410
##	0.3	5	0.6	0.875	100	0.7910933	0.5803737
##	0.3	5	0.6	0.875	150	0.7840008	0.5667238
##	0.3	5	0.6	0.875	200	0.7902878	0.5791474
##	0.3	5	0.6	0.875	250	0.7895185	0.5774600

##	0.3	5	0.6	1.000	50	0.7943777	0.5873877
##	0.3	5	0.6	1.000	100	0.7990223	0.5963210
##	0.3	5	0.6	1.000	150	0.7974355	0.5930620
##	0.3	5	0.6	1.000	200	0.7958363	0.5896841
##	0.3	5	0.6	1.000	250	0.7974110	0.5930618
##	0.3	5	0.8	0.500	50	0.7880783	0.5746859
##	0.3	5	0.8	0.500	100	0.7887988	0.5760596
##	0.3	5	0.8	0.500	150	0.7919178	0.5821667
##	0.3	5	0.8	0.500	200	0.7919362	0.5819467
##	0.3	5	0.8	0.500	250	0.7895495	0.5771379
##	0.3	5	0.8	0.625	50	0.7872114	0.5718813
##	0.3	5	0.8	0.625	100	0.7911118	0.5803600
##	0.3	5	0.8	0.625	150	0.7856487	0.5692072
##	0.3	5	0.8	0.625	200	0.7817180	0.5612785
##	0.3	5	0.8	0.625	250	0.7833112	0.5642046
##	0.3	5	0.8	0.750	50	0.7989124	0.5961500
##	0.3	5	0.8	0.750	100	0.7973744	0.5928058
##	0.3	5	0.8	0.750	150	0.7965871	0.5914180
##	0.3	5	0.8	0.750	200	0.7942676	0.5864553
##	0.3	5	0.8	0.750	250	0.7911487	0.5805282
##	0.3	5	0.8	0.875	50	0.7903121	0.5784018
##	0.3	5	0.8	0.875	100	0.8067679	0.6114529
##	0.3	5	0.8	0.875	150	0.8036062	0.6056872
##	0.3	5	0.8	0.875	200	0.8012684	0.6007662
##	0.3	5	0.8	0.875	250	0.8036245	0.6056663
##	0.3	5	0.8	1.000	50	0.8037041	0.6059577
##	0.3	5	0.8	1.000	100	0.8036489	0.6057016
##	0.3	5	0.8	1.000	150	0.8005115	0.5993210
##	0.3	5	0.8	1.000	200	0.8021108	0.6024151
##	0.3	5	0.8	1.000	250	0.8044364	0.6069584
##	0.4	1	0.6	0.500	50	0.7544035	0.5071352
##	0.4	1	0.6	0.500	100	0.7505402	0.4995513
##	0.4	1	0.6	0.500	150	0.7520113	0.5021209
##	0.4	1	0.6	0.500	200	0.7450039	0.4881901
##	0.4	1	0.6	0.500	250	0.7504545	0.4994772
##	0.4	1	0.6	0.625	50	0.7387353	0.4749176
##	0.4	1	0.6	0.625	100	0.7441798	0.4865656
##	0.4	1	0.6	0.625	150	0.7371971	0.4722709
##	0.4	1	0.6	0.625	200	0.7497465	0.4979282
##	0.4	1	0.6	0.625	250	0.7481599	0.4940845
##	0.4	1	0.6	0.750	50	0.7473235	0.4922806
##	0.4	1	0.6	0.750	100	0.7504607	0.4989944
##	0.4	1	0.6	0.750	150	0.7535982	0.5054597
##	0.4	1	0.6	0.750	200	0.7582860	0.5147886
##	0.4	1	0.6	0.750	250	0.7567662	0.5124845
##	0.4	1	0.6	0.875	50	0.7496794	0.4973059
##	0.4	1	0.6	0.875	100	0.7520477	0.5022188
##	0.4	1	0.6	0.875	150	0.7544220	0.5078791
##	0.4	1	0.6	0.875	200	0.7552031	0.5089141
##	0.4	1	0.6	0.875	250	0.7543912	0.5073940
##	0.4	1	0.6	1.000	50	0.7497037	0.4982648
##	0.4	1	0.6	1.000	100	0.7489652	0.4969808
##	0.4	1	0.6	1.000	150	0.7473904	0.4931745
##	0.4	1	0.6	1.000	200	0.7489225	0.4962205

##	0.4	1	0.6	1.000	250	0.7450037	0.4886313
##	0.4	1	0.8	0.500	50	0.7434779	0.4848276
##	0.4	1	0.8	0.500	100	0.7481961	0.4945737
##	0.4	1	0.8	0.500	150	0.7606296	0.5192571
##	0.4	1	0.8	0.500	200	0.7606786	0.5199217
##	0.4	1	0.8	0.500	250	0.7528111	0.5033873
##	0.4	1	0.8	0.625	50	0.7481780	0.4956870
##	0.4	1	0.8	0.625	100	0.7473843	0.4932674
##	0.4	1	0.8	0.625	150	0.7419339	0.4823654
##	0.4	1	0.8	0.625	200	0.7395841	0.4779688
##	0.4	1	0.8	0.625	250	0.7372524	0.4741210
##	0.4	1	0.8	0.750	50	0.7536102	0.5058429
##	0.4	1	0.8	0.750	100	0.7520479	0.5022340
##	0.4	1	0.8	0.750	150	0.7536286	0.5061993
##	0.4	1	0.8	0.750	200	0.7496978	0.4977901
##	0.4	1	0.8	0.750	250	0.7497526	0.4981361
##	0.4	1	0.8	0.875	50	0.7457486	0.4901998
##	0.4	1	0.8	0.875	100	0.7426724	0.4836051
##	0.4	1	0.8	0.875	150	0.7442351	0.4864002
##	0.4	1	0.8	0.875	200	0.7387660	0.4759706
##	0.4	1	0.8	0.875	250	0.7521028	0.5029076
##	0.4	1	0.8	1.000	50	0.7457791	0.4905375
##	0.4	1	0.8	1.000	100	0.7442408	0.4873412
##	0.4	1	0.8	1.000	150	0.7442530	0.4872146
##	0.4	1	0.8	1.000	200	0.7395165	0.4773847
##	0.4	1	0.8	1.000	250	0.7426724	0.4840272
##	0.4	2	0.6	0.500	50	0.7465725	0.4912467
##	0.4	2	0.6	0.500	100	0.7551666	0.5073021
##	0.4	2	0.6	0.500	150	0.7676369	0.5334014
##	0.4	2	0.6	0.500	200	0.7770733	0.5519552
##	0.4	2	0.6	0.500	250	0.7708595	0.5399003
##	0.4	2	0.6	0.625	50	0.7700111	0.5378942
##	0.4	2	0.6	0.625	100	0.7622475	0.5221865
##	0.4	2	0.6	0.625	150	0.7699930	0.5376726
##	0.4	2	0.6	0.625	200	0.7583105	0.5145056
##	0.4	2	0.6	0.625	250	0.7574803	0.5119854
##	0.4	2	0.6	0.750	50	0.7621554	0.5228829
##	0.4	2	0.6	0.750	100	0.7614475	0.5210616
##	0.4	2	0.6	0.750	150	0.7716103	0.5415674
##	0.4	2	0.6	0.750	200	0.7739357	0.5464037
##	0.4	2	0.6	0.750	250	0.7770854	0.5524646
##	0.4	2	0.6	0.875	50	0.7465665	0.4917192
##	0.4	2	0.6	0.875	100	0.7731910	0.5450811
##	0.4	2	0.6	0.875	150	0.7692845	0.5366366
##	0.4	2	0.6	0.875	200	0.7747352	0.5478288
##	0.4	2	0.6	0.875	250	0.7708288	0.5395609
##	0.4	2	0.6	1.000	50	0.7544219	0.5073866
##	0.4	2	0.6	1.000	100	0.7739173	0.5461292
##	0.4	2	0.6	1.000	150	0.7700294	0.5383173
##	0.4	2	0.6	1.000	200	0.7652745	0.5286024
##	0.4	2	0.6	1.000	250	0.7621553	0.5218868
##	0.4	2	0.8	0.500	50	0.7582488	0.5152597
##	0.4	2	0.8	0.500	100	0.7536161	0.5058953
##	0.4	2	0.8	0.500	150	0.7630590	0.5241268

##	0.4	2	0.8	0.500	200	0.7591159	0.5163147
##	0.4	2	0.8	0.500	250	0.7622720	0.5225858
##	0.4	2	0.8	0.625	50	0.7497098	0.4978863
##	0.4	2	0.8	0.625	100	0.7676914	0.5340577
##	0.4	2	0.8	0.625	150	0.7630406	0.5246388
##	0.4	2	0.8	0.625	200	0.7684667	0.5355369
##	0.4	2	0.8	0.625	250	0.7692784	0.5370459
##	0.4	2	0.8	0.750	50	0.7567477	0.5110766
##	0.4	2	0.8	0.750	100	0.7692299	0.5368833
##	0.4	2	0.8	0.750	150	0.7684670	0.5347091
##	0.4	2	0.8	0.750	200	0.7692359	0.5354977
##	0.4	2	0.8	0.750	250	0.7692116	0.5357602
##	0.4	2	0.8	0.875	50	0.7473480	0.4926257
##	0.4	2	0.8	0.875	100	0.7637667	0.5265237
##	0.4	2	0.8	0.875	150	0.7762308	0.5505359
##	0.4	2	0.8	0.875	200	0.7785930	0.5554309
##	0.4	2	0.8	0.875	250	0.7864243	0.5710801
##	0.4	2	0.8	1.000	50	0.7410914	0.4800247
##	0.4	2	0.8	1.000	100	0.7653297	0.5293098
##	0.4	2	0.8	1.000	150	0.7754985	0.5497422
##	0.4	2	0.8	1.000	200	0.7707986	0.5401593
##	0.4	2	0.8	1.000	250	0.7684362	0.5352763
##	0.4	3	0.6	0.500	50	0.7591218	0.5161441
##	0.4	3	0.6	0.500	100	0.7597811	0.5183969
##	0.4	3	0.6	0.500	150	0.7637184	0.5254795
##	0.4	3	0.6	0.500	200	0.7792953	0.5563617
##	0.4	3	0.6	0.500	250	0.7770002	0.5518933
##	0.4	3	0.6	0.625	50	0.7653725	0.5293958
##	0.4	3	0.6	0.625	100	0.7801803	0.5583352
##	0.4	3	0.6	0.625	150	0.7825671	0.5630908
##	0.4	3	0.6	0.625	200	0.7887990	0.5759191
##	0.4	3	0.6	0.625	250	0.7856616	0.5691764
##	0.4	3	0.6	0.750	50	0.7716467	0.5415899
##	0.4	3	0.6	0.750	100	0.7887678	0.5752194
##	0.4	3	0.6	0.750	150	0.7919602	0.5817174
##	0.4	3	0.6	0.750	200	0.7888108	0.5755013
##	0.4	3	0.6	0.750	250	0.7848677	0.5679389
##	0.4	3	0.6	0.875	50	0.7880597	0.5754887
##	0.4	3	0.6	0.875	100	0.7981737	0.5944221
##	0.4	3	0.6	0.875	150	0.7973437	0.5926874
##	0.4	3	0.6	0.875	200	0.8091423	0.6161452
##	0.4	3	0.6	0.875	250	0.8091300	0.6163746
##	0.4	3	0.6	1.000	50	0.7731604	0.5453262
##	0.4	3	0.6	1.000	100	0.7793865	0.5573107
##	0.4	3	0.6	1.000	150	0.7848555	0.5677303
##	0.4	3	0.6	1.000	200	0.7871931	0.5729692
##	0.4	3	0.6	1.000	250	0.7864302	0.5715116
##	0.4	3	0.8	0.500	50	0.7598483	0.5177128
##	0.4	3	0.8	0.500	100	0.7802101	0.5595100
##	0.4	3	0.8	0.500	150	0.7856976	0.5698723
##	0.4	3	0.8	0.500	200	0.7903548	0.5794422
##	0.4	3	0.8	0.500	250	0.7833111	0.5648440
##	0.4	3	0.8	0.625	50	0.7770611	0.5523875
##	0.4	3	0.8	0.625	100	0.7801374	0.5585931

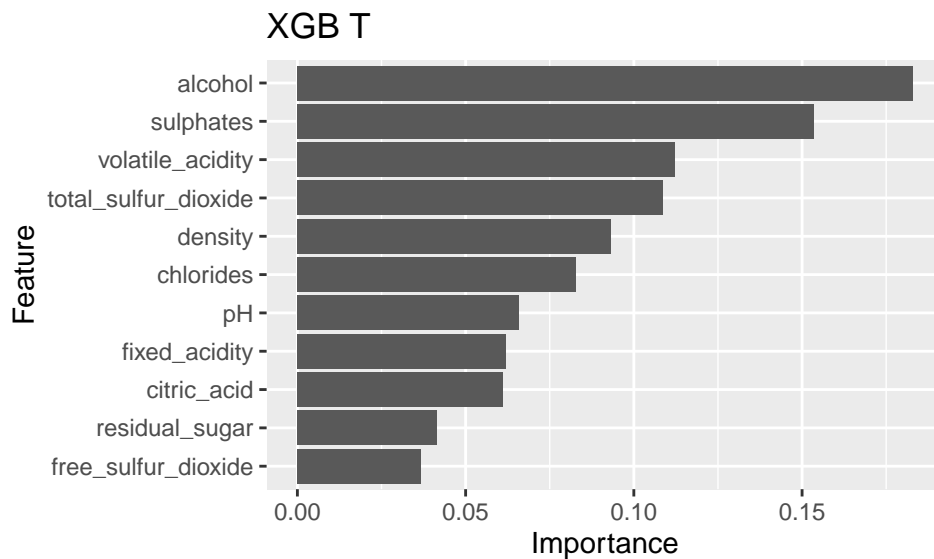
##	0.4	3	0.8	0.625	150	0.7808882	0.5603335
##	0.4	3	0.8	0.625	200	0.7824507	0.5631224
##	0.4	3	0.8	0.625	250	0.7934253	0.5852960
##	0.4	3	0.8	0.750	50	0.7614598	0.5214419
##	0.4	3	0.8	0.750	100	0.7715739	0.5405223
##	0.4	3	0.8	0.750	150	0.7793501	0.5561831
##	0.4	3	0.8	0.750	200	0.7824935	0.5626371
##	0.4	3	0.8	0.750	250	0.7817245	0.5613200
##	0.4	3	0.8	0.875	50	0.7638340	0.5262589
##	0.4	3	0.8	0.875	100	0.7786355	0.5555425
##	0.4	3	0.8	0.875	150	0.7817241	0.5620034
##	0.4	3	0.8	0.875	200	0.7746866	0.5477905
##	0.4	3	0.8	0.875	250	0.7770243	0.5526511
##	0.4	3	0.8	1.000	50	0.7699989	0.5388049
##	0.4	3	0.8	1.000	100	0.7817427	0.5614696
##	0.4	3	0.8	1.000	150	0.7793927	0.5566184
##	0.4	3	0.8	1.000	200	0.7770428	0.5517042
##	0.4	3	0.8	1.000	250	0.7832928	0.5645620
##	0.4	4	0.6	0.500	50	0.7724219	0.5436088
##	0.4	4	0.6	0.500	100	0.7755411	0.5499737
##	0.4	4	0.6	0.500	150	0.7833294	0.5659513
##	0.4	4	0.6	0.500	200	0.7849105	0.5687607
##	0.4	4	0.6	0.500	250	0.7919725	0.5830506
##	0.4	4	0.6	0.625	50	0.7888168	0.5762247
##	0.4	4	0.6	0.625	100	0.7895614	0.5774059
##	0.4	4	0.6	0.625	150	0.7770365	0.5527530
##	0.4	4	0.6	0.625	200	0.7707557	0.5403766
##	0.4	4	0.6	0.625	250	0.7770060	0.5522986
##	0.4	4	0.6	0.750	50	0.7958543	0.5899334
##	0.4	4	0.6	0.750	100	0.7934677	0.5848521
##	0.4	4	0.6	0.750	150	0.7903425	0.5782535
##	0.4	4	0.6	0.750	200	0.7958053	0.5894562
##	0.4	4	0.6	0.750	250	0.7973496	0.5923393
##	0.4	4	0.6	0.875	50	0.7863632	0.5710464
##	0.4	4	0.6	0.875	100	0.7949511	0.5880819
##	0.4	4	0.6	0.875	150	0.7934373	0.5851609
##	0.4	4	0.6	0.875	200	0.7911057	0.5807260
##	0.4	4	0.6	0.875	250	0.7871870	0.5726445
##	0.4	4	0.6	1.000	50	0.7911913	0.5803608
##	0.4	4	0.6	1.000	100	0.7864364	0.5705533
##	0.4	4	0.6	1.000	150	0.7864606	0.5706509
##	0.4	4	0.6	1.000	200	0.7895920	0.5776188
##	0.4	4	0.6	1.000	250	0.7927171	0.5838385
##	0.4	4	0.8	0.500	50	0.7935777	0.5859454
##	0.4	4	0.8	0.500	100	0.7966905	0.5915286
##	0.4	4	0.8	0.500	150	0.7958604	0.5897004
##	0.4	4	0.8	0.500	200	0.7973619	0.5930356
##	0.4	4	0.8	0.500	250	0.7895614	0.5775327
##	0.4	4	0.8	0.625	50	0.7865400	0.5711051
##	0.4	4	0.8	0.625	100	0.7801614	0.5577822
##	0.4	4	0.8	0.625	150	0.7817790	0.5618127
##	0.4	4	0.8	0.625	200	0.7786540	0.5554546
##	0.4	4	0.8	0.625	250	0.7771037	0.5526345
##	0.4	4	0.8	0.750	50	0.7824933	0.5628723



##	0.4	4	0.8	0.750	100	0.7762800	0.5505003
##	0.4	4	0.8	0.750	150	0.7848984	0.5677250
##	0.4	4	0.8	0.750	200	0.7895860	0.5772952
##	0.4	4	0.8	0.750	250	0.7840988	0.5663714
##	0.4	4	0.8	0.875	50	0.7887497	0.5759867
##	0.4	4	0.8	0.875	100	0.7856550	0.5694797
##	0.4	4	0.8	0.875	150	0.7848555	0.5679448
##	0.4	4	0.8	0.875	200	0.7926867	0.5840442
##	0.4	4	0.8	0.875	250	0.7934680	0.5856940
##	0.4	4	0.8	1.000	50	0.7801618	0.5585827
##	0.4	4	0.8	1.000	100	0.7918749	0.5816219
##	0.4	4	0.8	1.000	150	0.7895494	0.5770301
##	0.4	4	0.8	1.000	200	0.7895309	0.5770918
##	0.4	4	0.8	1.000	250	0.7934923	0.5849133
##	0.4	5	0.6	0.500	50	0.7723858	0.5432220
##	0.4	5	0.6	0.500	100	0.7856673	0.5684486
##	0.4	5	0.6	0.500	150	0.7739543	0.5451944
##	0.4	5	0.6	0.500	200	0.7755048	0.5478901
##	0.4	5	0.6	0.500	250	0.7786666	0.5548965
##	0.4	5	0.6	0.625	50	0.7982106	0.5940758
##	0.4	5	0.6	0.625	100	0.7974479	0.5925025
##	0.4	5	0.6	0.625	150	0.7903856	0.5785752
##	0.4	5	0.6	0.625	200	0.7911670	0.5802035
##	0.4	5	0.6	0.625	250	0.7864916	0.5712987
##	0.4	5	0.6	0.750	50	0.7880478	0.5741964
##	0.4	5	0.6	0.750	100	0.7942616	0.5870650
##	0.4	5	0.6	0.750	150	0.7864363	0.5716711
##	0.4	5	0.6	0.750	200	0.7926989	0.5841424
##	0.4	5	0.6	0.750	250	0.7903428	0.5792474
##	0.4	5	0.6	0.875	50	0.7888472	0.5754737
##	0.4	5	0.6	0.875	100	0.7865220	0.5710718
##	0.4	5	0.6	0.875	150	0.7919911	0.5821772
##	0.4	5	0.6	0.875	200	0.7911793	0.5805115
##	0.4	5	0.6	0.875	250	0.7896228	0.5774629
##	0.4	5	0.6	1.000	50	0.8044487	0.6069289
##	0.4	5	0.6	1.000	100	0.7966174	0.5910181
##	0.4	5	0.6	1.000	150	0.8013051	0.6006953
##	0.4	5	0.6	1.000	200	0.7989430	0.5953441
##	0.4	5	0.6	1.000	250	0.7981740	0.5941187
##	0.4	5	0.8	0.500	50	0.7856732	0.5696933
##	0.4	5	0.8	0.500	100	0.7895738	0.5772457
##	0.4	5	0.8	0.500	150	0.7864181	0.5706418
##	0.4	5	0.8	0.500	200	0.7887987	0.5754746
##	0.4	5	0.8	0.500	250	0.7958362	0.5896496
##	0.4	5	0.8	0.625	50	0.7949328	0.5874869
##	0.4	5	0.8	0.625	100	0.7941945	0.5860809
##	0.4	5	0.8	0.625	150	0.7887193	0.5752835
##	0.4	5	0.8	0.625	200	0.7926685	0.5830595
##	0.4	5	0.8	0.625	250	0.7903124	0.5784105
##	0.4	5	0.8	0.750	50	0.7934558	0.5846526
##	0.4	5	0.8	0.750	100	0.7997119	0.5973215
##	0.4	5	0.8	0.750	150	0.8075431	0.6134020
##	0.4	5	0.8	0.750	200	0.8044300	0.6071978
##	0.4	5	0.8	0.750	250	0.8028674	0.6041606

```
## 0.4 5 0.8 0.875 50 0.7919233 0.5816062
## 0.4 5 0.8 0.875 100 0.7919540 0.5814025
## 0.4 5 0.8 0.875 150 0.7903671 0.5785885
## 0.4 5 0.8 0.875 200 0.7919541 0.5823326
## 0.4 5 0.8 0.875 250 0.7943103 0.5869851
## 0.4 5 0.8 1.000 50 0.7981737 0.5945265
## 0.4 5 0.8 1.000 100 0.7958300 0.5894348
## 0.4 5 0.8 1.000 150 0.7958179 0.5896373
## 0.4 5 0.8 1.000 200 0.7942920 0.5864927
## 0.4 5 0.8 1.000 250 0.7927172 0.5834134
##
## Tuning parameter 'gamma' was held constant at a value of 0
## Tuning
## parameter 'min_child_weight' was held constant at a value of 1
## Accuracy was used to select the optimal model using the largest value.
## The final values used for the model were nrounds = 200, max_depth = 3, eta
## = 0.4, gamma = 0, colsample_bytree = 0.6, min_child_weight = 1 and subsample
## = 0.875.
```

```
p_xgbt <- varImp(wine_xgbt, scale = F) %>%
  ggplot() + ggtitle("XGB T")
p_xgbt
```



```
# Evaluate
confusionMatrix(predict(wine_xgbt, test), test$quality)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 115  25
##           1  34 147
##
```

```
##           Accuracy : 0.8162
##           95% CI : (0.7694, 0.857)
##      No Information Rate : 0.5358
##      P-Value [Acc > NIR] : <2e-16
##
##           Kappa : 0.629
##
##  McNemar's Test P-Value : 0.2976
##
##           Sensitivity : 0.7718
##           Specificity : 0.8547
##      Pos Pred Value : 0.8214
##      Neg Pred Value : 0.8122
##           Prevalence : 0.4642
##      Detection Rate : 0.3583
##      Detection Prevalence : 0.4361
##      Balanced Accuracy : 0.8132
##
##      'Positive' Class : 0
##
```

```
xgbt_acc <- confusionMatrix(predict(wine_xgbt, test), test$quality)$overall["Accuracy"]
xgbt_acc
```

```
## Accuracy
## 0.8161994
```

#### 4.1.9 torch

```
### Torch / Deep Learning -----

# Change the type from factor to numeric for DNN
train_torch <-
  train %>% mutate(
    quality = as.numeric(quality) - 1
  )

test_torch <-
  test %>% mutate(
    quality = as.numeric(quality) - 1
  )

# check the quality

table(train$quality, train_torch$quality) %>% kable()
```

	0	1
0	595	0
1	0	683

```

# Create {torch} dataset

df_dataset <- dataset(

  "wine",

  initialize = function(df, response_variable) {
    self$df <- df[, -which(names(df) == response_variable)]
    self$response_variable <- df[[response_variable]]
  },

  .getitem = function(index) {
    response <- torch_tensor(self$response_variable[index])
    x <- torch_tensor(as.numeric(self$df[index,]))

    list(x = x, y = response)
  },

  .length = function() {
    length(self$response_variable)
  }

)

# Create the data set train and test

train_torch_ds <- df_dataset(train_torch, "quality")
test_torch_ds <- df_dataset(test_torch, "quality")

# Create the data loader train and test

train_torch_dl <- dataloader(train_torch_ds, batch_size = 32, shuffle = T)
test_torch_dl <- dataloader(test_torch_ds, batch_size = 1, shuffle = F)

# Define a network / fc1 - 3 and dropout

net <- nn_module(

  "wine_DNN",

  initialize = function() {
    self$fc1 <- nn_linear(11, 66)
    self$fc2 <- nn_linear(66, 44)
    self$fc3 <- nn_linear(44, 1)
    self$dropout <- nn_dropout(0.5)
  },

  forward = function(x) {
    x %>%
      self$fc1() %>%
      nnf_relu() %>%
      self$fc2() %>%
      nnf_relu() %>%

```

```

        self$dropout() %>%
        self$fc3() %>%
        nnf_sigmoid()
    }
)

# Use CPU version

model <- net()
model$to(device = "cpu")

# Set the optimizer condition

optimizer <- optim_adam(model$parameters, lr = 0.01)

# Run a learning iteration

coro::loop(for (epoch in 1:20) {

    l <- c()

    coro::loop(for (b in enumerate(train_torch_dl)) {
        optimizer$zero_grad()
        output <- model(b[[1]]$to(device = "cpu"))
        loss <- nnf_binary_cross_entropy_with_logits(output, b[[2]]$to(device = "cpu"))
        loss$backward()
        optimizer$step()
        l <- c(l, loss$item())
    })

    cat(sprintf("Loss at epoch %d: %3f\n", epoch, mean(l)))
})

# Model prediction

model$eval()

i <- 1
pred_labels <- rep(0, nrow(test_torch))

coro::loop(for (b in enumerate(test_torch_dl)) {
    output <- model(b[[1]]$to(device = "cpu"))
    pred_labels[i] <- round(output$item(), 0)
    i <- i + 1
})

# Evaluate

table(test=test_torch$quality, pred_labels)

##      pred_labels
## test   0    1
##      0 125  24

```

```
##      1  74  98
```

```
torch_acc <- sum(diag(table(test_torch$quality, pred_labels))) / nrow(test_torch)
torch_acc
```

```
## [1] 0.694704
```

#### 4.1.10 Tabnet

```
### Tabnet -----
```

```
# splits the data (set vfold to 4 because it takes time)
```

```
splits <-
  vfold_cv(train, v = 4, strata = "quality") %>%
  with_seed(1, .)
```

```
# Create rule "classification"
```

```
rule <- tabnet(
  epochs = tune(),
  penalty = tune(),
  batch_size = tune(),
  decision_width = tune(),
  attention_width = tune(),
  num_steps = tune(),
  learn_rate = 0.08,
  momentum = 0.6) %>%
  set_engine("torch", verbose = TRUE) %>%
  set_mode("classification")
```

```
# making a recipe
```

```
rec <- recipe(quality ~ ., data = train) %>%
  step_nzv(all_predictors())
```

```
# making a work flow
```

```
wf <- workflow() %>%
  add_model(rule) %>%
  add_recipe(rec)
```

```
# range of hyper parameter
```

```
range_hypara <-
  wf %>%
  parameters() %>%
  update(
    epochs = epochs(c(50, 70)),
    decision_width = decision_width(range = c(20, 40)),
    attention_width = attention_width(range = c(20, 40)),
    num_steps = num_steps(range = c(4, 8))
```

```

) %>%
finalize(train)

# making a grid

grid <- range_hypara %>%
  grid_latin_hypercube(size = 1) %>%
  with_seed(1, .)

# tuning of hyper parameters (long time)

tune <-
  wf %>%
  tune_grid(
    resamples = splits,
    grid = grid,
    control = control_grid(save_pred = TRUE),
    metrics = metric_set(accuracy)
  )

# select the best hyper parameters

good_hypara <-
  tune %>%
  show_best() %>%
  dplyr::slice(1)

# update the rule with best parameters

upd_rule <-
  tabnet(
    epochs = good_hypara %>% pull(epochs),
    penalty = good_hypara %>% pull(penalty),
    batch_size = good_hypara %>% pull(batch_size),
    decision_width = good_hypara %>% pull(decision_width),
    attention_width = good_hypara %>% pull(attention_width),
    num_steps = good_hypara %>% pull(num_steps),
    learn_rate = 0.08,
    momentum = 0.6) %>%
  set_engine("torch", verbose = TRUE) %>%
  set_mode("classification")

# update the work flow with updated rule

upd_wf <-
  workflow() %>%
  add_model(upd_rule) %>%
  add_recipe(rec)

# build the model

model_tn <-

```

```

upd_wf %>%
  fit(train) %>%
  with_seed(1,.)

# Predict

pred_tabnet <-
  predict(model_tn,new_data = test, type = "class")

```

```

# Evaluate

table(test=test$quality, pred= pred_tabnet$.pred_class)

```

```

##      pred
## test  0   1
##      0 111  38
##      1  49 123

```

```

tabnet_acc <- sum(diag(table(test$quality, pred_tabnet$.pred_class))) /
  nrow(test)

tabnet_acc

```

```

## [1] 0.728972

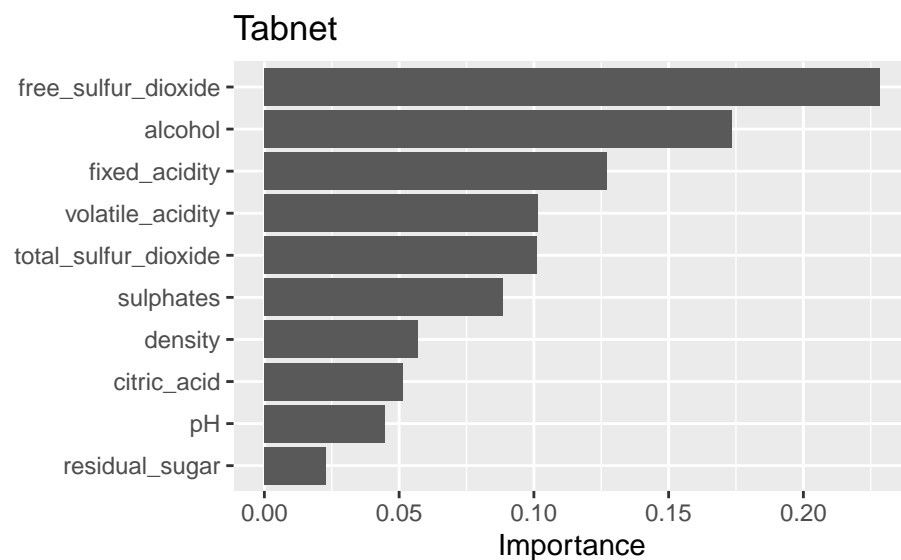
```

```

# Check the importance

fit <- extract_fit_parsnip(model_tn)
p_tabnet <- vip(fit) + ggtitle("Tabnet")
p_tabnet

```





## 4.2 Compare the models

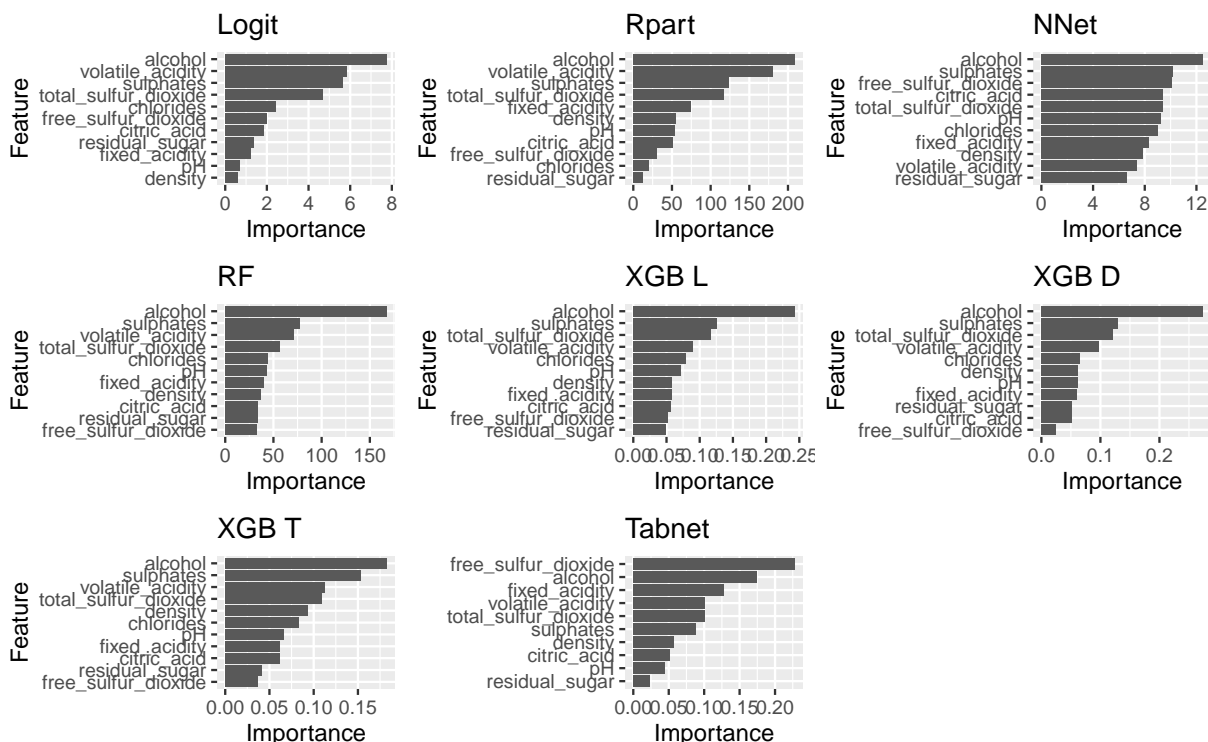
As a result of comparing each model using Accuracy, the DART model of XGBoost showed the highest score. Since we have not yet reached the optimal tuning for each model, we cannot judge which model is the best, but XGBoost was able to calculate a high score with a short code for all of them. In the Importance comparison, all models except Tabnet showed the highest score for Alcohol, while Tabnet showed the highest score for free sulfur dioxide. The ability to obtain multiple features in real work is attractive, and in competitions such as Kaggle, we believe it can be used to build ensembles.

```
### Compare the models -----
```

```
rbind(logit_acc, rpart_acc, rf_acc, svm_acc, nnet_acc,
      xgbl_acc, xgbd_acc, xgbt_acc, torch_acc, tabnet_acc) %>% kable()
```

	Accuracy
logit_acc	0.7258567
rpart_acc	0.7476636
rf_acc	0.7912773
svm_acc	0.7476636
nnet_acc	0.7819315
xgbl_acc	0.8099688
xgbd_acc	0.8193146
xgbt_acc	0.8161994
torch_acc	0.6947040
tabnet_acc	0.7289720

```
wrap_plots(p_logit, p_rpart, p_nnet, p_rf, p_xgbl, p_xgbd, p_xgbt, p_tabnet)
```



### 4.3 Additional: Google Colaboratory

In order to perform deep learning using GPU, we tried to calculate with R/torch using Google Colaboratory, where we can build a GPU environment for free. Link of the code and calculation results: **Google Colab: CYO\_colab**

Note: The library will automatically determine whether the environment is CPU or GPU and install the library, so you need to set the runtime to GPU first and then run the installation.

The number of epochs was set to 20 in this report, but 2000 in the Google Colaboratory calculation. 53 minutes was required in the CPU environment, but 35 minutes in the GPU environment, which reduced the calculation time.

## 5 conclusion

Following the previous MovieLens project, we worked on machine learning about classification using a small data table: Wine Quality dataset, assuming a BtoB business.

We were able to build models not only using the methods from the edx course and the widely used gradient boosting, but also using a library newly incorporated into R in 2020. In search of a better computing environment, we took advantage of the external environment and actually reduced the computation time.

Through this series of initiatives, we were able to learn how to work with data science, which will continue to grow in the future. By creating many models for the same data set, we were able to learn about the characteristics of each model. On the other hand, we have not yet calculated the optimal values of hyper-parameters for each model, so we think this is a future work to be done with better resources such as GPU environment.

## Reference

- Shunsuke1015.github (my previous work: Movie Lens PJ) <https://github.com/Shunsuke1015/homework-0>
- rafalab.github <https://rafalab.github.io/dsbook/large-datasets.html#recommendation-systems>
- UCI: Machine Learning Repository <https://archive.ics.uci.edu/ml/index.php>
- Pandoc <https://pandoc.org/index.html>
- The caret Package <https://topepo.github.io/caret/index.html>
- Tidymodels <https://www.tidymodels.org/>
- XGBoost.ai <https://xgboost.ai/>
- DART: Dropouts meet Multiple Additive Regression Trees (2015) <https://arxiv.org/pdf/1505.01866.pdf>
- torch for R <https://torch.mlverse.org/>
- Applied deep learning with torch from R [https://mlverse.github.io/torchbook\\_materials/](https://mlverse.github.io/torchbook_materials/)
- GitHub: mlverse/tabnet <https://github.com/mlverse/tabnet>
- TabNet: Attentive Interpretable Tabular Learning (2020) <https://arxiv.org/pdf/1908.07442.pdf>

- kaggle: Simple R - xgboost - caret kernel <https://www.kaggle.com/nagsdata/simple-r-xgboost-caret-kernel>
- towards data science: How to use R in Google Colab <https://towardsdatascience.com/how-to-use-r-in-google-colab-b6e02d736497>
- Wikipedia: Gradient boosting [https://en.wikipedia.org/wiki/Gradient\\_boosting](https://en.wikipedia.org/wiki/Gradient_boosting)
- Wikipedia: Deep learning [https://en.wikipedia.org/wiki/Deep\\_learning](https://en.wikipedia.org/wiki/Deep_learning)