

Prediction of Wine by its range

K.Ajay

III MECH

Reg.No.: 950921114001

Wine is an [alcoholic drink](#) made from [fermented fruit](#). [Yeast](#) consumes the [sugar](#) in the fruits and converts it to [ethanol](#) and [carbon dioxide](#), releasing [heat](#) in the process. Though wine [can be made from a variety of fruit crops](#) such as [plum](#), [cherry](#), [pomegranate](#), [blueberry](#), [currant](#) and [elderberry](#), it is most often made from [grapes](#), and the term "wine" generally refers to grape wine when used without a qualifier.

our application's prediction, provide ideal solutions for the analysis of wine, which will make this whole process efficient and cheaper with less human interaction.

Objective

- Our Main objective is to predict the wine quality using machine learning through Python programming language
- A large dataset is considered and wine quality is modelled to analyse the quality of wine through different parameters like fixed acidity, volatile acidity etc.
- All these parameters will be analysed through Machine Learning algorithms like random forest classifier algorithm which will help to rate the wine on scale 1 - 10 or bad - good.
- Output obtained would further be checked for correctness and model will be optimized accordingly.
- It can support the wine expert evaluations and ultimately improve the Production.

Coding :

```

library('party') library('rpart')
library('rpart.plot') library('caret')
library('ROCR') library('randomForest')
set.seed(46)

wineData<-read.csv('../input/winequalityN.csv',sep=',',stringsAsFactors=F) str(wineData)
sum(is.na(wineData))

hist(wineData$quality, main= "Wine Quality", col= "blue") qualityFactor <- ifelse(wineData$quality >=
6, "high", "low") wineData <- data.frame(wineData, qualityFactor) table(wineData$qualityFactor)

index <- sample(2, nrow(wineData), replace=TRUE, prob = c(0.60, 0.40)) trainingData <-
wineData[index==1, ] validationData <- wineData[index==2, ]

control <- rpart.control(minsplit = 5L, maxdepth = 5L, minbucket = 5, cp =
0.002, maxsurrogate = 4)
rPartModel <- rpart(qualityFactor~., trainingData, method = "class", control = control)
predict_rpart <- predict(rPartModel, validationData[, -13], type = "class"
)
prp(rPartModel, type=2, extra=3, tweak=0.8, main = "The Quality of Wine", compress=TRUE)

```

Output :

```

Prediction high low    high
1299 380    low 305 595

```

```

Accuracy : 0.7344
 95% CI : (0.7169, 0.7514)
No Information Rate : 0.6219
P-Value [Acc > NIR] : < 2.2e-16

```

Kappa : 0.4265
McNemar's Test P-Value : 0.004693

Sensitivity : 0.8099
Specificity : 0.6103
Pos Pred Value : 0.7737
Neg Pred Value : 0.6611
Prevalence : 0.6219
Detection Rate : 0.5037
Detection Prevalence : 0.6510
Balanced Accuracy : 0.7101

'Positive' Class : high

Application

- It is used by the wine manufacturers to improve the quality of the future wines.
- Certification bodies can also use the result for quality control. • Results can be used to make wine selection guides for wine magazines.
- Results can be used by consumers for wine selection

Reference

- “The Classification of White Wine and Red Wine According to Their Physicochemical Qualities,” Int. J. Intell. Syst. Appl. Eng., vol. 4, no. SpecialIssue-1, pp. 23–26, 2016.
- [2] E. Summary, W. P. Monitoring, W. Quality, W. Safety, and W. Complexity, “Wine Analysis : from ‘Grape to Glass’ An analytical testing digest of the wine manufacturing process,” 2016.
- [3] A. Ghosh, “Project Report : -Red Wine Quality Analysis Final 3 . An empirical Red Wine Quality Analysis of the Portuguese ‘ Vinho Verde ’ wine,” no. December 2017, 2018.
- [4] Y. Gupta, “Selection of important features and predicting wine quality using machine learning techniques,” Procedia Comput. Sci., vol. 125, pp. 305–312, 2018.
- [5] P. Model, L. Regression, and R. Studio, “Building and Evaluating a Predictive Model w/ Linear Regression in RapidMiner Studio,” 2018.
- [6] B. Tajini and O. C. Paris, “Badr Tajini – On Campus Paris – DSTI 2017,” vol. 47, no. 4, pp. 547–553, 2017.

