



# International Journal for Innovative Engineering and Management Research

A Peer Reviewed Open Access International Journal

www.ijiemr.org

**COPY RIGHT**



**ELSEVIER**  
**SSRN**

**2019IJIEMR**. Personal use of this material is permitted. Permission from IJIEMR must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. No Reprint should be done to this paper, all copy right is authenticated to Paper Authors

IJIEMR Transactions, online available on 1<sup>st</sup> Jun 2019. Link

[:http://www.ijiemr.org/downloads.php?vol=Volume-08&issue=ISSUE-06](http://www.ijiemr.org/downloads.php?vol=Volume-08&issue=ISSUE-06)

Title: **WINEQUALITY ASSESSMENT USING MACHINE LEARNING AND PYTHON**

Volume 08, Issue 06, Pages: 220–225.

Paper Authors

**Dr. Venugeetha Y, Prakruthi Gopal Krishna ,Gagan Deep S**



USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per **UGC Guidelines** We Are Providing A Electronic Bar Code

## WINEQUALITY ASSESSMENT USING MACHINE LEARNING AND PYTHON

<sup>1</sup>Dr. Venugeetha Y, <sup>2</sup>Prakruthi Gopal Krishna, <sup>3</sup>Gagan Deep S

<sup>1</sup>Assistant professor (ph.D), Dept of CSE-GAT, Bengaluru

<sup>2</sup>UG Scholar, Dept of CSE-GAT, Bengaluru.

<sup>3</sup>UG Scholar, Dept of ISE-DSATM, Bengaluru.

### ABSTRACT

A fermented grape drink, treated anaerobically, to produce alcohol is known as Wine. The support vector machine model is used under the regression approach. The results are the predicted output from the machine which is calculated based on the values given for the wine attributes. Each wine quality is graded between 0 (very bad) and 10 (very excellent).

### 1. INTRODUCTION

As Artificial Intelligence (AI) allows the machine to learn and improve from the experience and data provided. Machine learning can access data and uses it learn and improve. The process of learning begins with a survey for data, such as examples, direct experience, or instruction, based on an example that is provided patterns in data make a better decision in the future. The primary purpose is to let the computers on their own isolated from human intervention and adjust actions accordingly. Wine is an alcoholic beverage made from anaerobically treated, juiced up grapes. Yeast consumes the sugar in the grapes and converts it to ethanol, carbon dioxide and heat. Just by reading the label, it is very impossible to postulate the quality of wine. Quality is rated or certified best by tasting. Usually, you won't be allowed to taste a wine before buying, but details are available on wine labels to help you contemplate. Any good wine sellers will have abundant knowledge to provide the details of wine manufacturers

### 2. REQUIREMENTS

First, we collect the software and hardware requirements based on, based on the initial requirement that are been demanded from the problem statement, there twqe of predicament that we face during the design and integration phase.

#### A. Software Requirements

Operating System: Windows 7 or above,  
Programing Language: Python 3.6.8, IDE Used: JetBrains PyCharm Community Edition 2018.3.3 x64.

#### B. Hardware Requirements

Processor-Pentium IV or above, RAM-2GB or more, Hard disk- 3 GB or more

### 3. IMPLEMENTATION

Python community has developed modules to help programmers implement Machine Learning. Here numpy, scipy and scikit-learn modules are used by installing them. One of the commands for installation is: pip install numpy scipy scikit.

#### A. Machine Learning overview

A dataset is given to the machine for learning and it then gets trained using the same and then this training is used to predict the new output for a new, given input. For example, training a computer by feeding 1000 photos of spiders and 1000 more photos which are not of a spider, and make a computer understand whether a photo is spider or not. Then computer is tested with a new photo, after, computer should be able to tell whether new photo is spider or not. We use specialized algorithms, to illustrate the process of training and prediction. To predict on a new test data, algorithm is trained with data set.

#### B. Algorithm used: SVR (Support Vector Regression)

Most of the people who are in Machine

Learning or Data Science are up to knowledge with the term SVM or Support Vector Machine. However, SVR (Support Vector Regression) is a little off-track from SVM. SVR is a regression algorithm and can work with contiguous values rather than that like Classification which is SVM. In other words, by using regression we try to eradicate or attenuate the error rate, Meanwhile, in SVR we try to fit the error within a particular range. The main focus when we are using SVR is to take in to consideration, the points that are inside the boundary line. Our best fit line is the line hwqerplane that has utmost points.

What we are trying to achieve here is simply coding to decide a decision boundary at 'e' distance from the genuine hwqer plane such that data points most in vicinity to the hwqer plane or the support vectors are within that same vicinity.

Hence, the decision boundary is our margin of tolerance that is, we are going to consider those points only, which are adhering to this boundary. Or in other words / terms, we are going to consider those points only, which have least error rate. This will serve us in the way i.e, by giving us a better model.

### C.Attributes Used

The Here out of two choice one can be selected i.e. either red wine or white wine. There are 11main attributes that help us assess the quality of the wine and they are as listed below: Equations are an exception to the prescribed specifications -

- Fixed Acidity
- Volatile Acidity
- Citric Acid
- Residual Sugar
- Chlorides
- Free Sulphur Di-Oxide
- Total Sulphur Di-Oxide
- Density
- pH
- Sulphates
- Amount of alcohol

### D. Advantages of Machine Learning are:

- Many industrial applications like banks and financial sector, health-care, retail, publishing and social media, and others have

taken advantage of this technology already and will continue to do so.

- Google and Facebook use these applications to push informed advertisements on the basis of the users search history.
- There is a reduction in the Time cycle and there is a very efficient use of resources.
- There are tools available that provide continuous quality enhancement in complex processing environments by the help of Machine Learning

### E.Code

```
from sklearn.svm import SVR
from sklearn.datasets import load_boston
from sklearn import model_selection
from sklearn.metrics import accuracy_score

from pandas import read_csv
fromsklearn.model_selectionimporttrain_test_split
choice = int (input("enter 1 for red wine quality checking and enter 2 for white wine quality checking"))
if choice==1:
    a1 = float(input( "Enter a list of the following values of the wine while entering return for every value : \n1. fixed acidity\n2. volatile acidity\n3.citric acid\n4.residual sugar\n5.chlorides\n6.free sulphur dioxide\n7. total sulfur dioxide\n8. density\n9. pH\n10. sulfates\n11.alcohol\n"))
    a2 = float(input())
    a3 = float(input())
    a4 = float(input())
    a5 = float(input())
    a6 = float(input())
    a7 = float(input())
    a8 = float(input())
    a9 = float(input())
    a10 = float(input())
    a11 = float(input())
attlist = [a1, a2, a3, a4, a5, a6, a7, a8, a9, a10, a11]
print(attlist)
import pandas as pd
a = pd.DataFrame([attlist])
dset=read_csv(("C:\\Users\\DELL\\Documents\\loan_prediction\\winewhite.csv"))
X = dset.drop("quality", axis=1)
```

```

Y = dset["quality"]
X_train, X_test, Y_train, Y_test =
train_test_split(X, Y, test_size=0.1)
svm_model = SVR()
svm_model.fit(X_train, Y_train)
Y_predicted = svm_model.predict(X_test)
wq = svm_model.predict(a)
    if (int(wq) >= 9.5):
print(10)
    if (int(wq) < 9.5 and int(wq) >= 8.5):
print(9)
    if (int(wq) < 8.5 and int(wq) >= 7.5):
print(8)
    if (int(wq) < 7.5 and int(wq) >= 6.5):
print(7)
    if (int(wq) < 6.5 and int(wq) >= 5.5):
print(6)
    if (int(wq) < 5.5 and int(wq) >= 4.5):
print(5)
    if (int(wq) < 4.5 and int(wq) >= 3.5):
print(4)
    if (int(wq) < 3.5 and int(wq) >= 2.5):
print(3)
    if (int(wq) < 2.5 and int(wq) >= 1.5):
print(2)
    if (int(wq) < 1.5 and int(wq) >= 0.5):
print(1)
    if (int(wq) < 0.5):
print(0)
elif choice==2:

    a1 = float(input("Enter a list of the
following values of the wine while entering
return for every value : \n1. fixed acidity\n2.
volatile acidity\n3.citric acid\n4.residual
sugar\n5.chlorides\n6.free sulphur dioxide\n7.
total sulfur dioxide\n8. density\n9. pH\n10.
sulfates\n11.alcohol\n "))
    a2 = float(input())
    a3 = float(input())
    a4 = float(input())
    a5 = float(input())
    a6 = float(input())
    a7 = float(input())
    a8 = float(input())
    a9 = float(input())
    a10 =float(input())
    a11= float(input())
attlist=[a1,a2,a3,a4,a5,a6,a7,a8,a9,a10,a11]
print (attlist)

```

```

import pandas as pd
a=pd.DataFrame([attlist])

dset=read_csv(("C:\\Users\\DELL\\Documen
ts\\loan_prediction\\winewhite.csv"))
X=dset.drop("quality",axis=1)
Y=dset["quality"]

X_train,X_test,Y_train,Y_test=train_test_spli
t(X,Y,test_size=0.1)
svm_model=SVR()
svm_model.fit(X_train,Y_train)
Y_predicted=svm_model.predict(X_test)
wq=svm_model.predict(a)
    if(int(wq)>=9.5):
        print (10)
    if (int(wq)<9.5 and int(wq)>=8.5):
print(9)
    if (int(wq)<8.5 and int(wq)>=7.5):
print(8)
    if (int(wq)<7.5 and int(wq)>=6.5):
print(7)
    if (int(wq)<6.5 and int(wq)>=5.5):
print(6)
    if (int(wq)<5.5 and int(wq)>=4.5):
print(5)
    if (int(wq)<4.5 and int(wq)>=3.5):
print(4)
    if (int(wq)<3.5 and int(wq)>=2.5):
print(3)
    if (int(wq)<2.5 and int(wq)>=1.5):
print(2)
    if (int(wq)<1.5 and int(wq)>=0.5):
print(1)
    if (int(wq)<0.5):
print(0)
else:
    print("INVALIDCHOICE")

```

## F. DESIGN

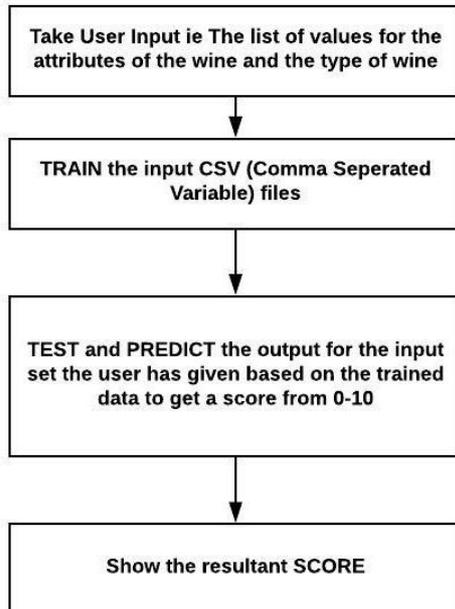


Figure 1: Flowchart of Wine Quality Determination

First the user input is taken which is the list of values for the attributes of the wine and the twqe of wine. Using the csv train the input. Then the input is tested and the output is predicted for the input set that the user has given based on the trained data to get a score from 0-10. Then the resultant score is shown.

### G.Snapshots

```

enter 1 for red wine quality checking and enter 2 for white wine quality checking!
Enter a list of the following values of the wine while entering return for every valu
1. fixed acidity
2. volatile acidity
3.citric acid
4.residual sugar&l.chlorides
6.free sulphur dioxide
7. total sulfur dioxide
8. density
9. pH
10. sulfates
11.alcohol
  
```

Snapshot 1: The prompt for the input which asks if we want to test the red or white wine and after selecting the type of Wine, the attributes are displayed.

```

10. sulfates
11.alcohol
7.8
0.88
0
2.6
0.098
25
67
0.9968
3.2
0.68
9.8
[7.8, 0.88, 0.0, 2.6, 0.098, 25.0, 67.0, 0.9968, 3.2, 0.68, 9.8]
C:\Users\DELL\PycharmProjects\untitled\venvo\lib\site-packages\sklearn\svm\base.py:196: FutureWarning: "avoid this warning.", FutureWarning)
5
Process finished with exit code 0
  
```

Snapshot 2: The inputs are given and after testing the output is predicted to “5”

fixed acidity	volatile acidity	citric acid	residual sugar	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH	sulphates	alcohol	quality
7.4	0.7	0	1.8	0.076	11	24	0.9878	3.51	0.56	9.4	5
7.8	0.88	0	2.6	0.098	25	67	0.9968	3.2	0.68	9.8	5
7.8	0.70	0.04	2.3	0.032	13	34	0.937	3.20	0.53	9.6	5
11.2	0.28	0.56	1.9	0.075	17	60	0.998	3.16	0.58	9.8	6

Snapshot 3: The output which is predicted is been verified with the trained data output

## 4. CONCLUSION

Machine Learning is one of those technological concepts that is just endless. There is a whole new world of opportunities due to this and many people embracing this and owning it as well. Applying these technologies to the real world and the daily life, especially to a normal citizen will make this a better world and a better place to live in. Manufacturers can mainly benefit from this. Hence, due to this project, the mass producers of wine can make sure they can save time in testing their products manually, reducing the power dissipated by the links of

## 5. REFERENCE

[1] Machine Learning by Authors [PanosLouridas](#) Athens University of Economics and Business and [Christof Ebert](#) Vector Consulting Services (IEEE Paper)

[2] Textbook – Python for Everybody and Python Documentation website <https://docs.python.org>

[3] A Study of Machine Learning in Healthcare IEEE Paper by [Rohan Bhardwaj](#) ; [Ankita R.](#)



[Nambiar ; Debojyoti Dutta](#)

[4][Machine Learning 2015](#) by Tom Mitchell and Maria-Florina Balcan, Carnegie Mellon University

[5][Introduction to Machine Learning 2018](#) by Maria-Florina Balcan, Carnegie Mellon University.



# International Journal for Innovative Engineering and Management Research

PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL

[www.ijiemr.org](http://www.ijiemr.org)