CA Assignment

Neural Network Ensembles

Wine Quality white data set

By:-

1. Rahul Singh Chauhan
2. Tran Khanh Trung

Contents

1. [Using R:-](#_Using_R:-_1)

## [Multilayer perceptron with back propagation:-](#_Multilayer_perceptron_with)

* 1. [General Regression neural network(GRNN):-](#_General_Regression_neural)
  2. [Radial Basis Function(RBF):-](#_1.3_Radial_Basis)
  3. [Ensemble MLP Backpropagation, GRNN and RBF using weighted average :-](#_1.4_Ensemble_MLP)

1. [Using Weka:-](#_Using_Weka:-)
   1. [Multilayer Perceptron](#_2.1_Multilayer_Perceptron)
   2. [Using RBF Network:-](#_2.2_Using_RBF)
   3. [Ensemble](#_Ensemble_Stacking/Blending_MLP+RBFN) Stacking
   4. [Ensemble](#_2.4_Ensemble_MLP) Voting-Average of Probabilities
2. [Observation Summary:-](#_Observation_Summary:-)
3. [Understanding:-](#_Understanding:-)

Wine Quality data set

The data set contains 12 variable out of which first 11 columns are input and 12th column is outcome. This is a multi class classification problem as outcome value is one of 3,4,5,6,7 or 8 which shows score of the quality of wine. Used following implementation of neural networks. For ensemble we used weighted average.

# Using R:-

## Multilayer perceptron with back propagation**:-**

Using multilayer perceptron individually with 11 hidden nodes in one hidden layer we get accuracy of **54.04%.**

Package / library used – **nnet**

Total data set = 4898rows, Split = 75% training , 25% testing

Rows 1-3673 = training

Rows 3673-4898 = testing

Time taken Model= 5.248149 secs

Time taken Predict= 0.01562595 secs

The details are given in snapshot below:-

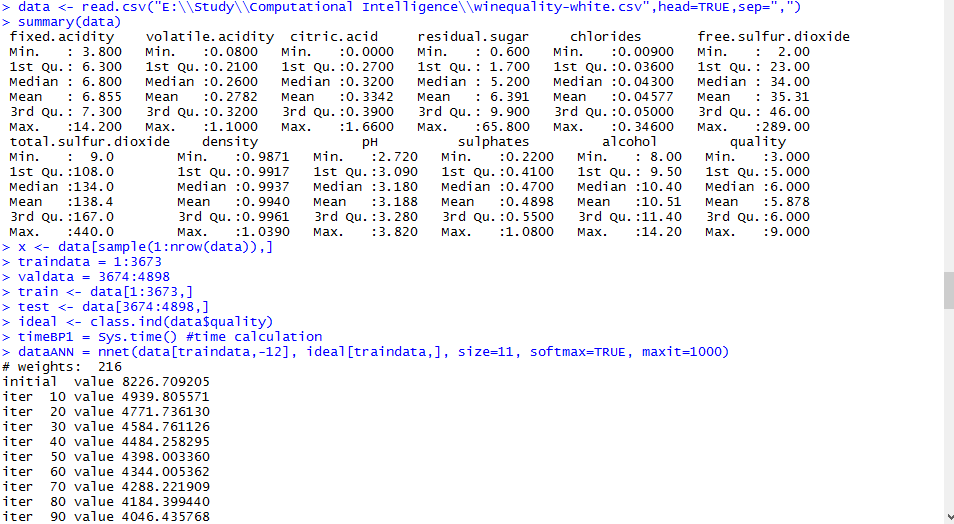


Figure 1 Multi layer perceptron - part1



Figure 2 Multi layer perceptron - part2

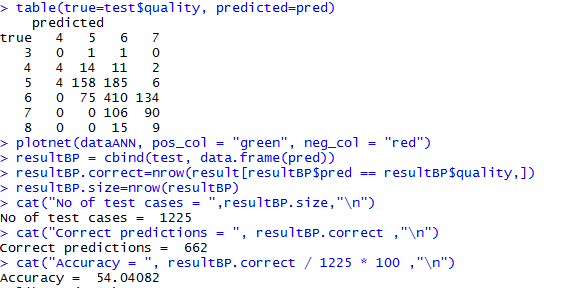


Figure 3 Multi layer perceptron - Result

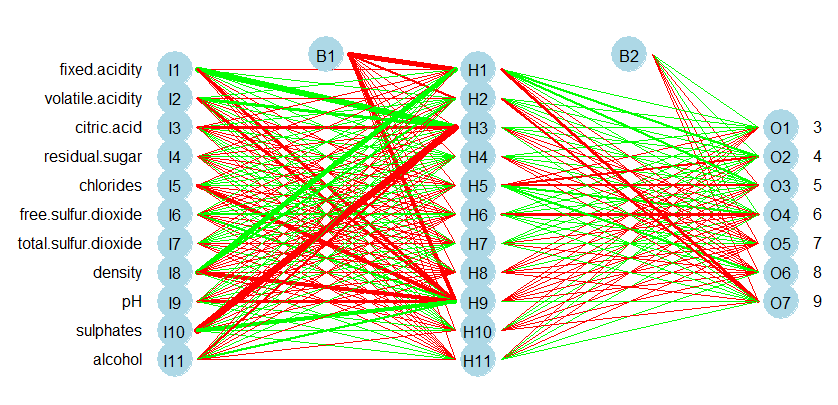


Figure 4 Plot – Wine Quality Neural Network Multilayer Perceptron with Backpropagation

## General Regression neural network(GRNN):-

Using General Regression neural network individually gives 39.91% accuracy. For this problem GRNN doesn’t look like a suitable model.

Package/library used – **grnn**

Total data set = 4898rows, Split = 75% training , 25% testing

Rows 1-3673 = training

Rows 3673-4898 = testing

Time taken Model= 1.19393 secs

Time taken Predict= 71.72 secs



Figure 5 General Regression neural network - Modelling

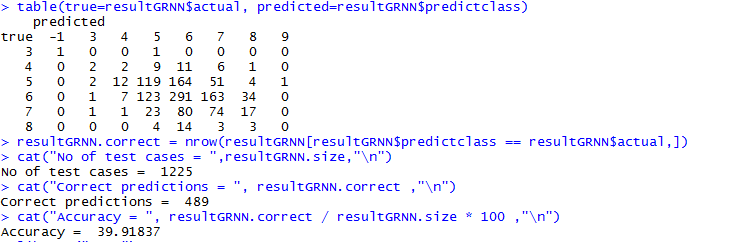


Figure 6 General Regression neural network - Result

## 1.3 Radial Basis Function(RBF):-

Using RBF individually we get accuracy of **50.36%** . More details of RBF is given below:-

Package used – **RSNNS**

Time taken Model= 9.397148 secs

Time taken Predict= 0.02213788 secs

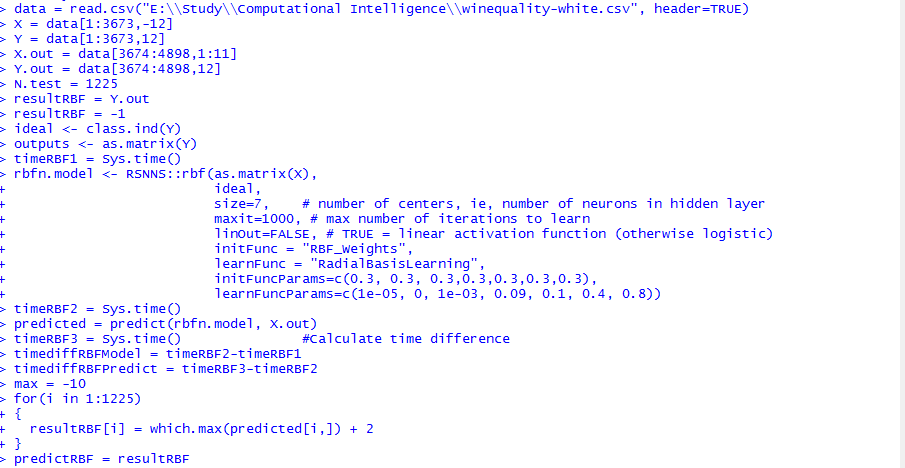


Figure 7 Radial basis function NN

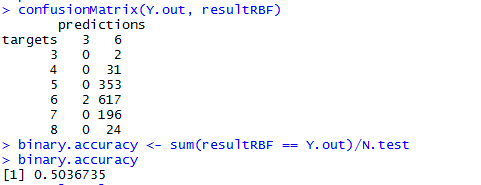


Figure 8 RBF result

## 1.4 Ensemble MLP Backpropagation, GRNN and RBF using Majority voting :-

We combined the output of multi layer perceptron with backpropagation, general regression neural network and radial basis function. Then used weighted average technique to arrive at the final result. Multi layer perceptron is given 60% weightage while GRNN and RBF are given 20%,20% weightage respectively. The result is rounded off to get final output classification. The accuracy of this combined model is more than any other model used individually. The accuracy of voting ensemble achieved is **54.85 %.**

Time taken for ensemble logic to execute is 1.44199 secs.

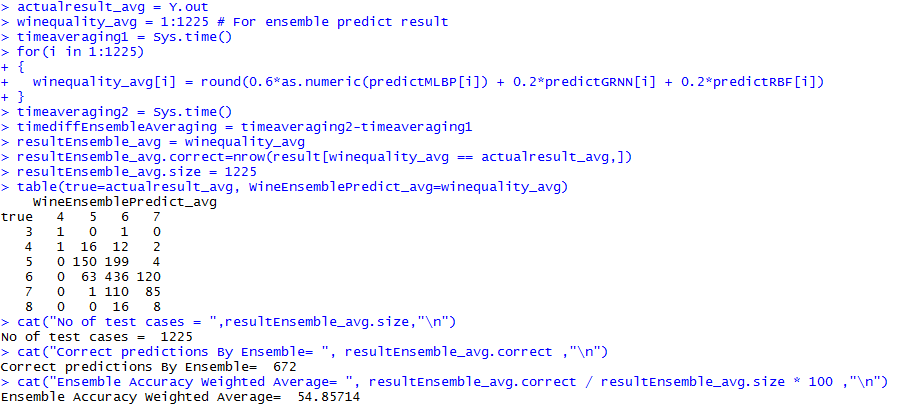


Figure 9 Ensemble weighted average result

# Using Weka:-

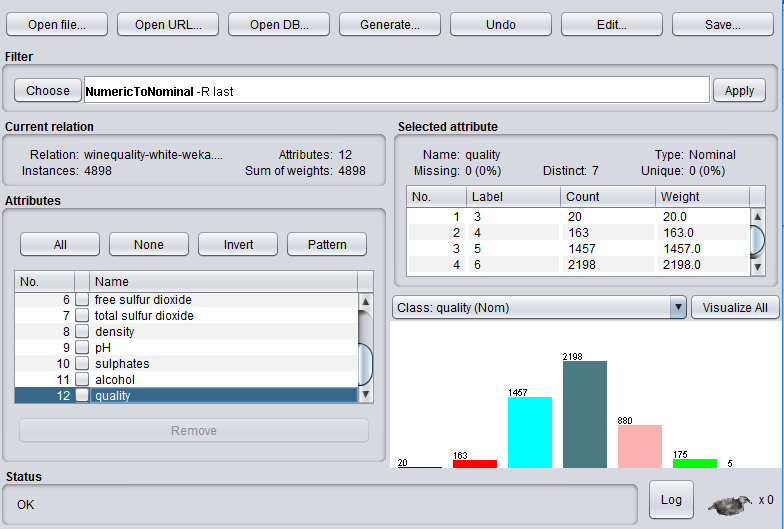


Figure 10 Dataset wine quality classification

## 2.1 Multilayer Perceptron

Using individual multi layer perceptron model. In weka tool with the settings as shown in fig. 11 it gives accuracy of **54.902%.**

Hidden nodes = 11, Hidden layer = 1.

Training – 75%, Testing 25%

Time taken to build model: 13.39 seconds

Time taken to test model on test split: 0 seconds

The process is shown below:-

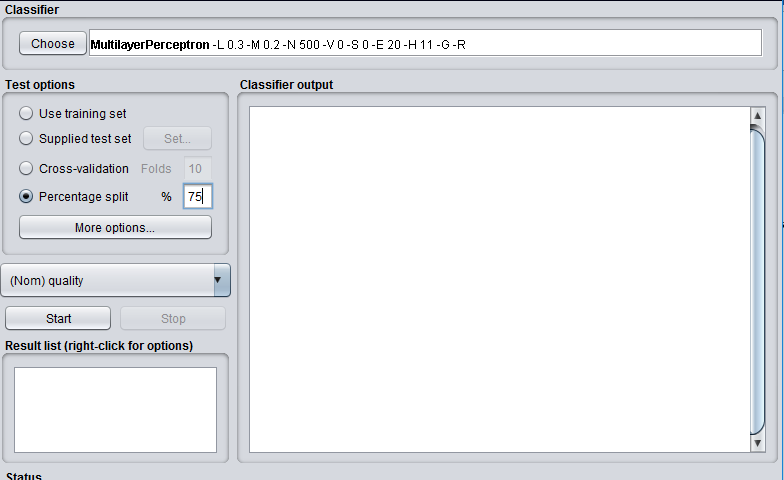


Figure 11 MLP settings

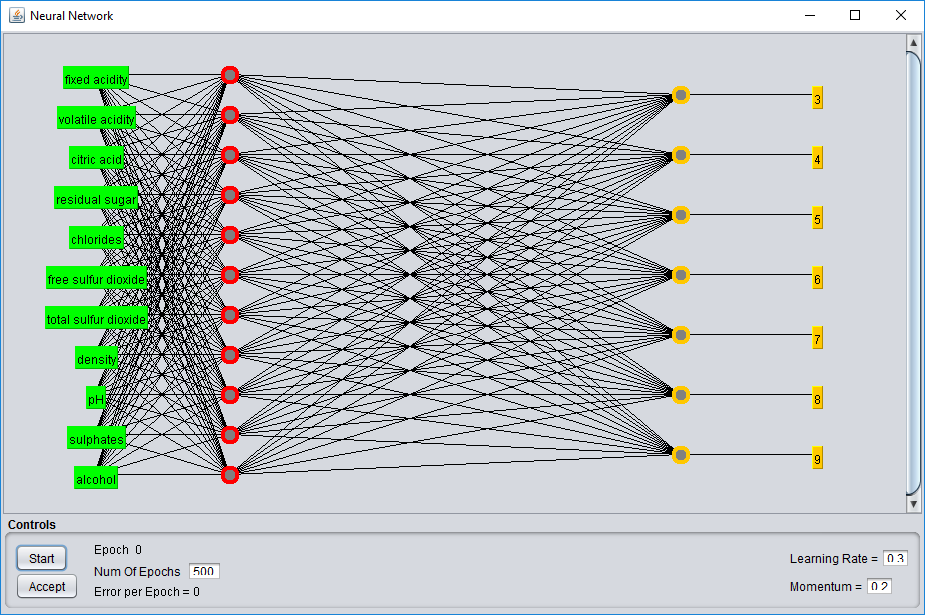


Figure 12 MLP Plot

The report is as follows:-

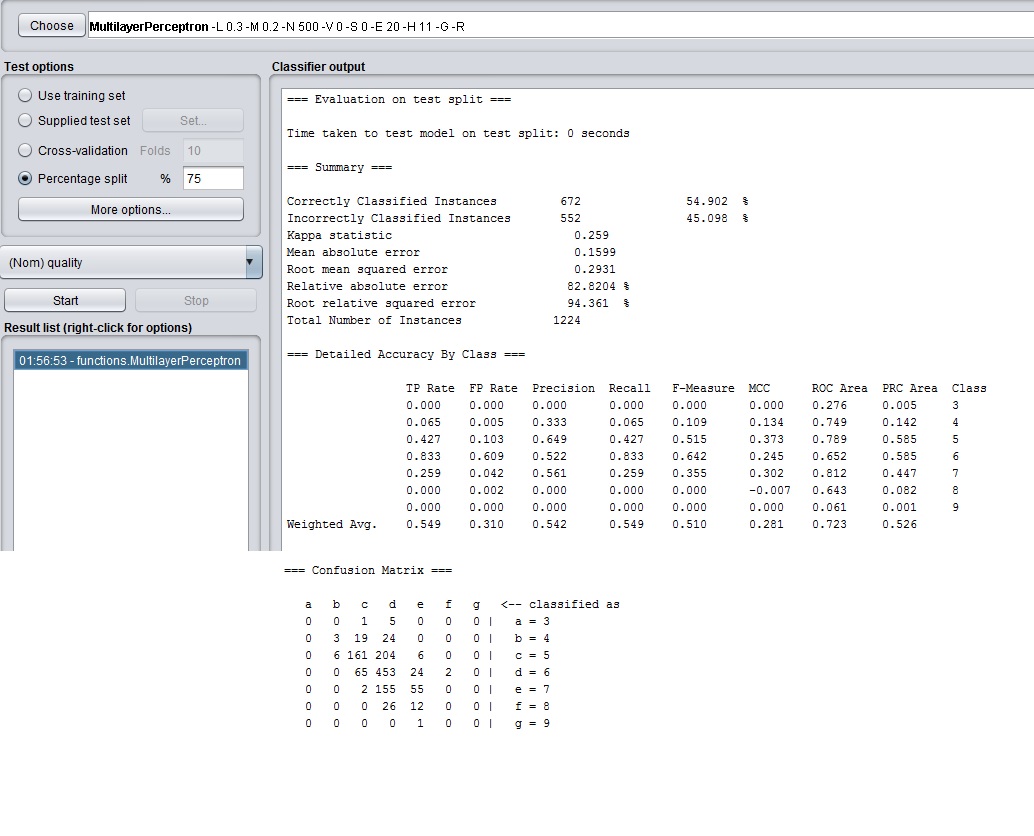


Figure 13 MLP Result

For the whole log detail check attached log file “**Weka-WineQuality\_MLP\_Log.txt**”.

## 2.2 Using RBF Network:-

Using individual RBF network we get the accuracy as **50.4902 %.**

Testing = 75%, Training =25%

Time taken to build model: 6.95 seconds

Time taken to test model on test split: 0.02 seconds

The settings for RBF network is shown in fig 14.

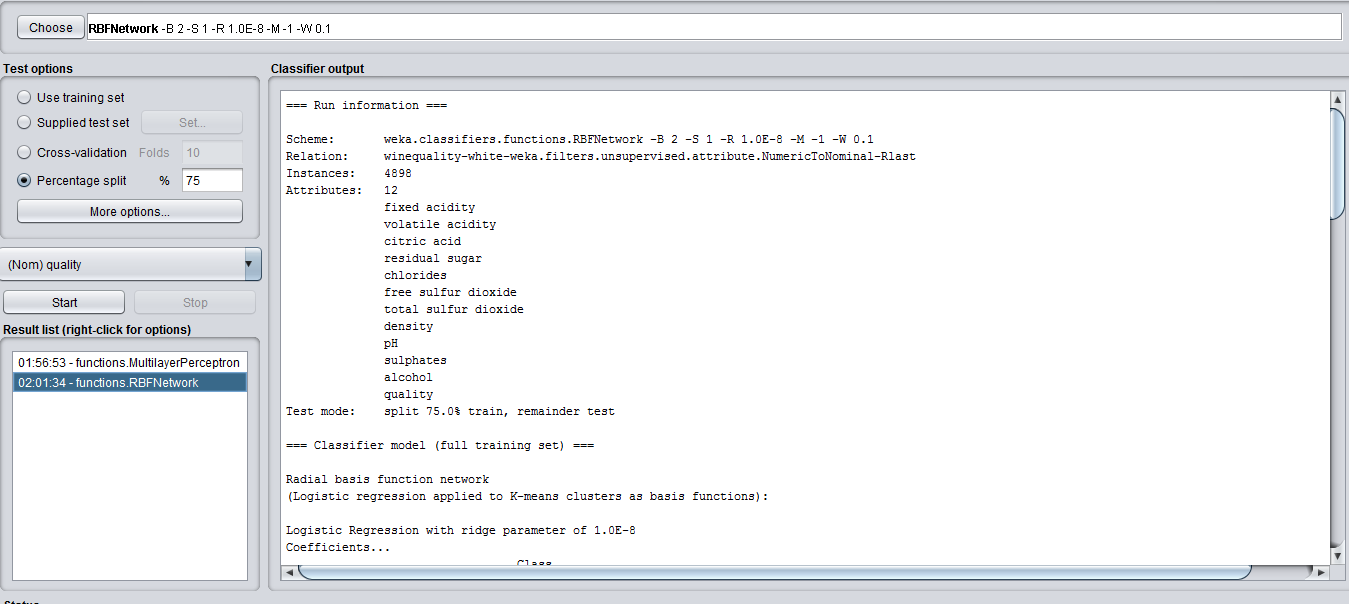


Figure 14 RBF Network Settings

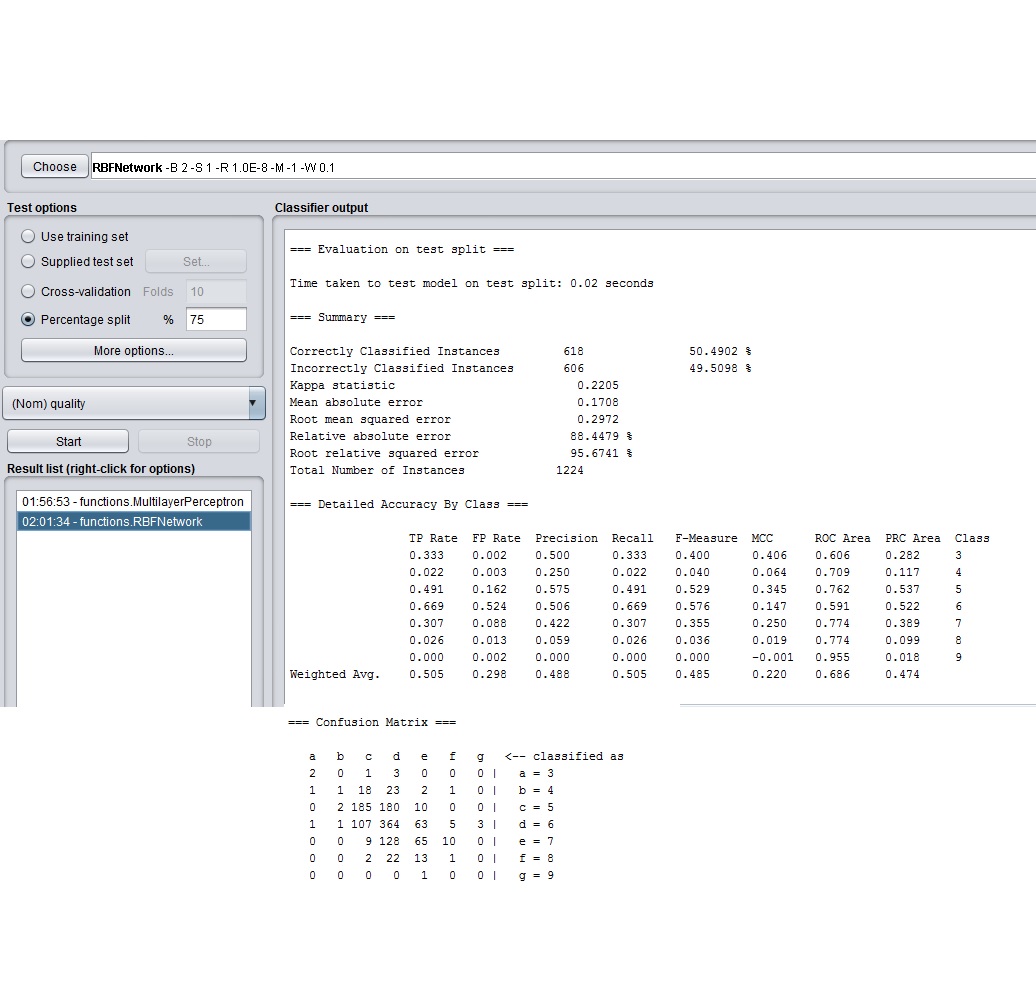


Figure 15 RBF Network result

Log info “**Weka-WineQuality\_RBFNetwork.txt**”

## Ensemble Stacking/Blending MLP+RBFNetwork

The accuracy of stacking ensemble is **56.1275** %.

Time taken to build model: 118.64 seconds

Time taken to test model on test split: 0.05 seconds

The settings is shown in fig.16.

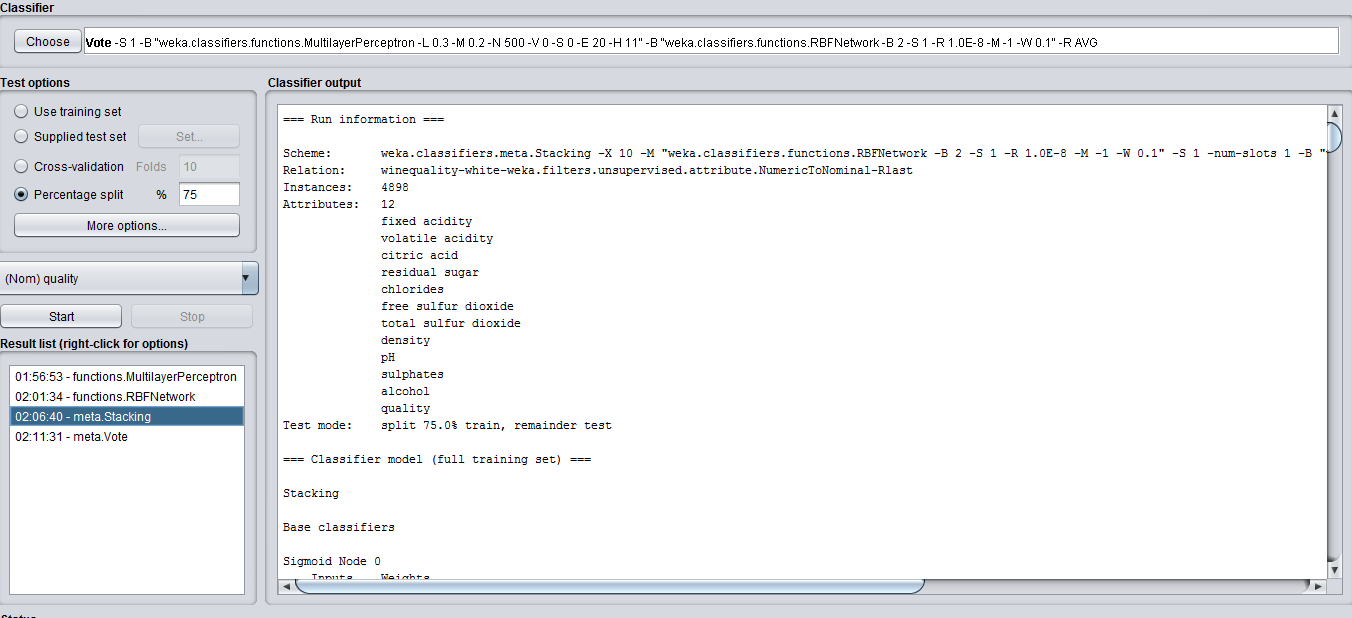


Figure 16 Ensemble stacking settings

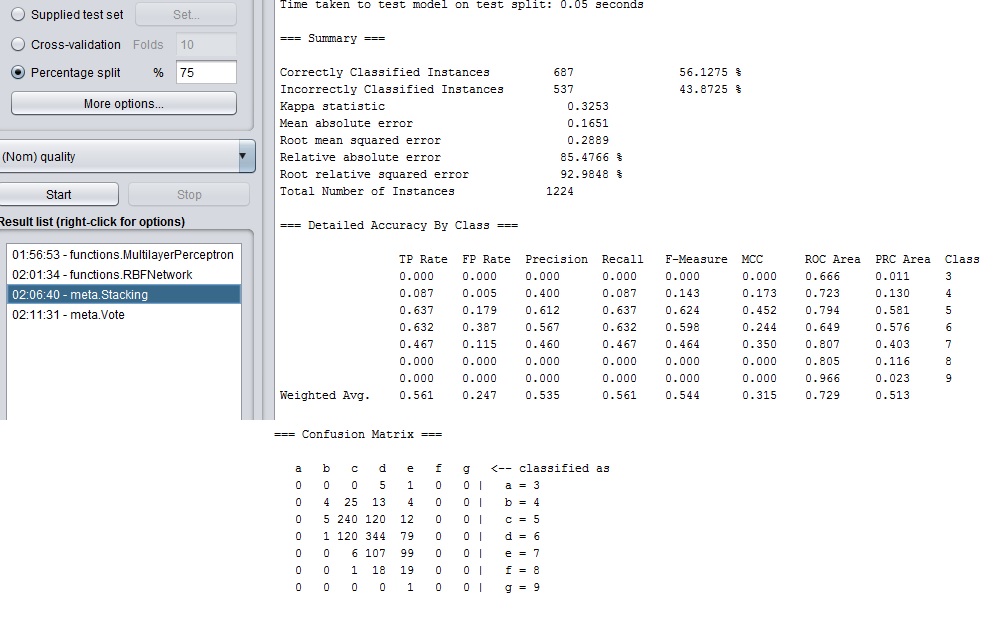


Figure 17 Ensemble Stacking result

Log info: “**Weka-WineQuality\_EnsembleStacking.txt**”

## 2.4 Ensemble MLP and RBF Network using Voting with average of probabilities :-

Accuracy = **55.3105%**

Time taken to build model: 18.36 seconds

Time taken to test model on test split: 0.02 seconds

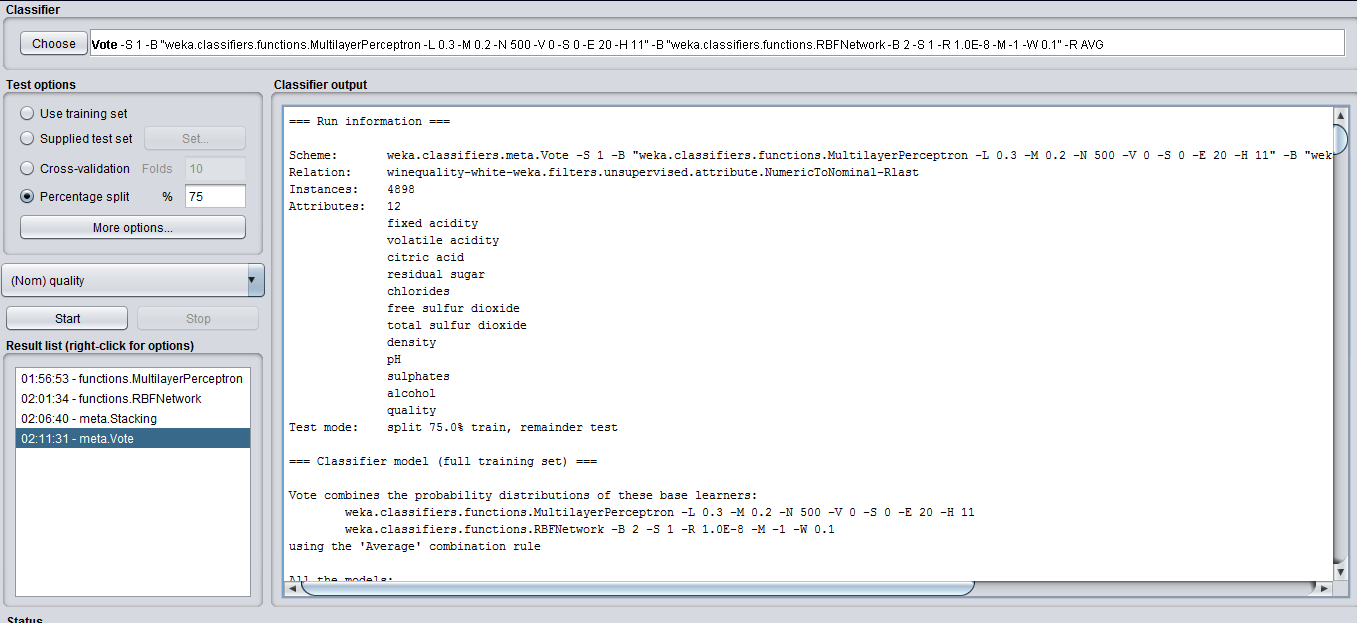


Figure 18 Ensemble Vote Average of probabilities settings

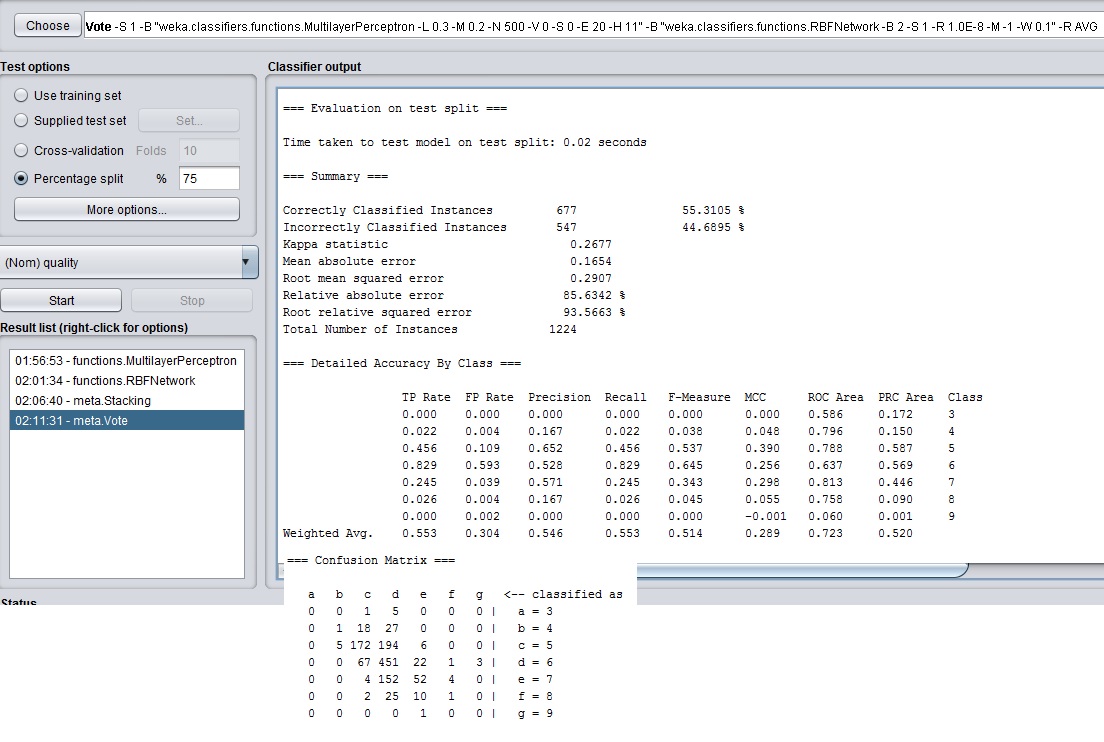


Figure 19 Ensemble Voting- Average of probabilities result

For log info check “**Weka-WineQuality\_EnsembleVoting-AvgOfProbabilities.txt”**

# **Observation Summary:-**

### Using R:-

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **Time taken in modelling(sec).** | **Time taken in testing/predicting(sec).** | **Accuracy** |
| Multi layer perceptron with backpropagation(ML BP) | 5.248149 | 0.01562595 | 54.04% |
| General Regression Neural Network(GRNN) | 1.19393 | 71.72 | 39.91% |
| Radial Basis Function(RBF) | 9.397148 | 0.02213788 | 50.36% |
| Ensemble Majority Voting (ML BP + GRNN + RBF) | Total time = Sum of all individual model times = 87.577 sec | 1.44199 | 54.85%. |

Table 1: Summary Neural Network using R on Wine data set

### Using Weka:-

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **Time taken in modelling(sec).** | **Time taken in testing/predicting(sec).** | **Accuracy** |
| Multi layer perceptron | 13.39 | 0 | 54.902%. |
| Radial Basis Function Network(RBFNetwork) | 6.95 | 0.02 | 50.490% |
| Ensemble Stacking MLP+RBFNetwork | 118.64 | 0.05 | 56.127%. |
| Ensemble Average of probabilities Voting (MLP + RBFNetwork) | 18.36 | 0.02 | 55.3105% |

Table 2: Summary using Weka on Wine Quality Data set

# **Understanding:-**

The table1 and table2 in observations shows that a neural network ensemble gives higher accuracy than individual network alone. While using R tool out of three individual models MLP with BP gives best accuracy of 54.04% but when we combine all the three networks we get the accuracy of 54.85% which is 0.81% more. In weka there is increment of 1.225% accuracy by using stacking ensemble and 0.4085% by using voting(averaging of probabilities) ensemble technique.

Just like diabetes data set example there is always tradeoff between performance time and accuracy if we use neural network models in an ensemble which is evident from the table 1 and table 2.