CA Assignment

Neural Network Ensembles

Diabetes 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. [Probabilistic neural network(PNN):-](#_Probabilistic_neural_network(PNN):-)
  2. [Radial Basis Function(RBF):-](#_1.3_Radial_Basis)
  3. [Ensemble MLP Backpropagation, PNN and RBF using Majority voting :-](#_1.4_Ensemble_MLP)

1. [Using Weka:-](#_Using_Weka:-)
   1. [Multilayer Perceptron](#_2.1_Multilayer_Perceptron)
   2. [Using RBF Classifier:-](#_2.2_Using_RBF)
   3. [RBF Network](#_2.3_RBF_Network)
   4. [Ensemble MLP , RBF Classifier and RBF Network using majority voting :-](#_2.4_Ensemble_MLP)
2. [Observation Summary:-](#_Observation_Summary:-)
3. [Understanding:-](#_Understanding:-)

Diabetes data set

The data set contains 9 variable out of which first 8 columns are input and 9th column is outcome. This is a binary classification problem as outcome value is either 0 or 1 i.e, showing sign of diabetes is either yes or no. Used following implementation of neural networks. As this is a classification problem we used **majority voting** technique for ensemble.

# Using R:-

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

Using multilayer perceptron individually with 10 hidden nodes in one hidden layer we get accuracy of **61.97%.**

Package / library used – **nnet**

Total data set = 768 rows, Split = 75% training , 25% testing

Rows 1-576 = training

Rows 577-768 = testing

Time taken Model= 0.4329121 secs

Time taken Predict= 0.286922 secs

The details are given in snapshot below:-

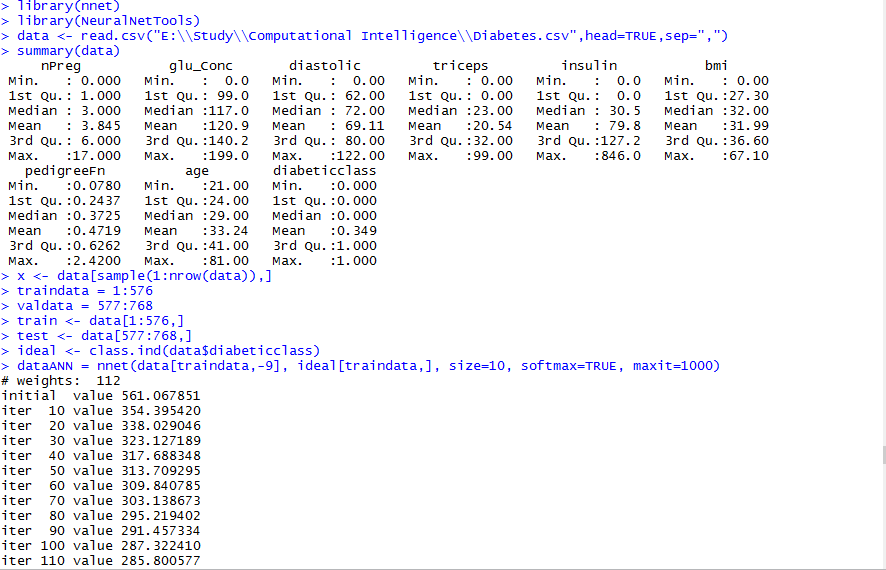


Figure Multi layer perceptron - part1

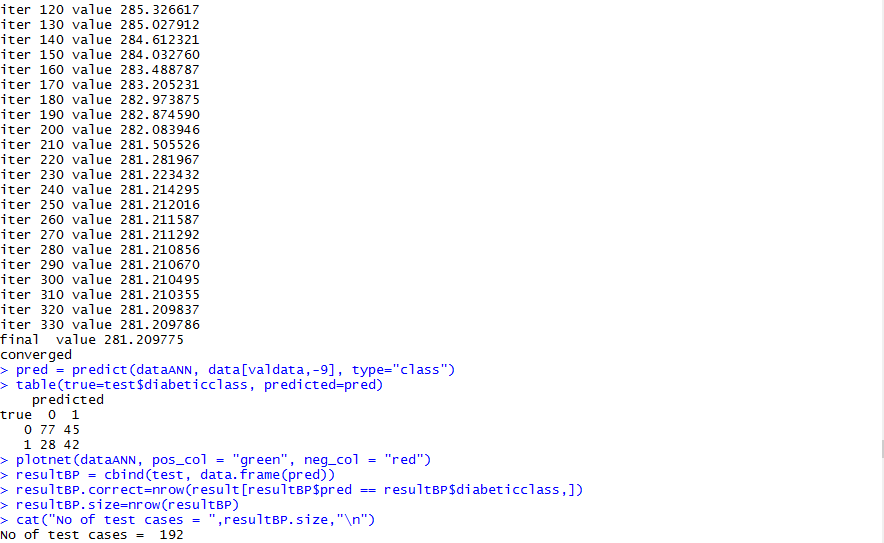


Figure Multi layer perceptron - part2

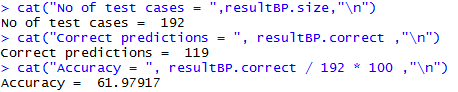


Figure Multi layer perceptron - part3

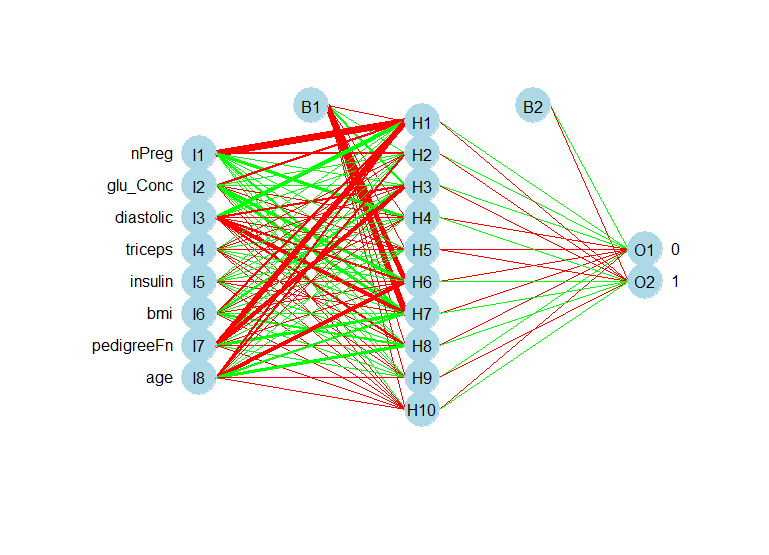


Figure . Plot - Diabetes Neural Network Multilayer Perceptron with Backpropagation

## Probabilistic neural network(PNN):-

Using probabilistic neural network individually gives 64.025% accuracy. As is it binary classification problem so we have chosen PNN over GRNN.

Package/library used – **pnn**

Total data set = 768 rows, Split = 75% training , 25% testing

Rows 1-576 = training

Rows 577-768 = testing

Time taken Model= 0.8680301 secs

Time taken Predict= 22.65863 secs

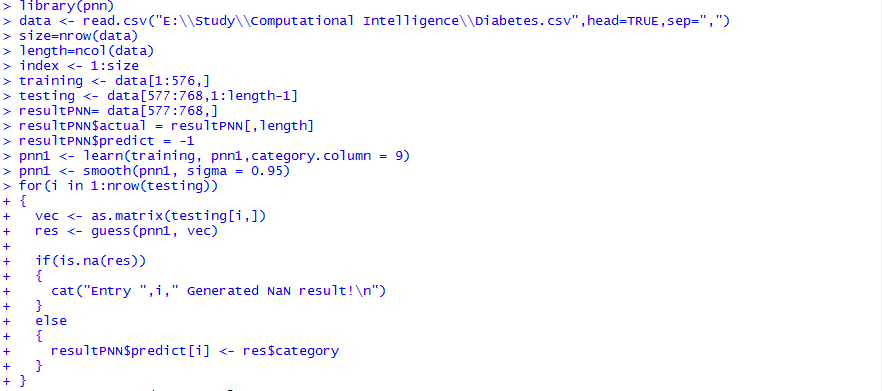


Figure Probabilistic neural network - part1

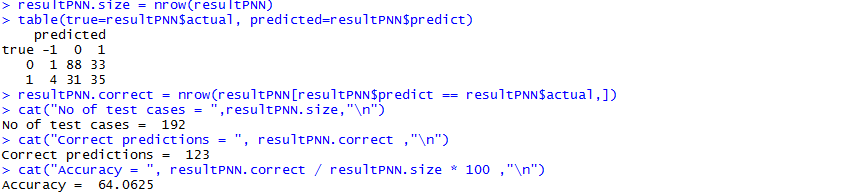


Figure Probabilistic neural network - part2

## 1.3 Radial Basis Function(RBF):-

Using RBF individually we get accuracy of **63.02%** . More details of RBF is given below:-

Package used – **RSNNS**

Time taken Model= 5.51747 secs

Time taken Predict= 0.04653406 secs

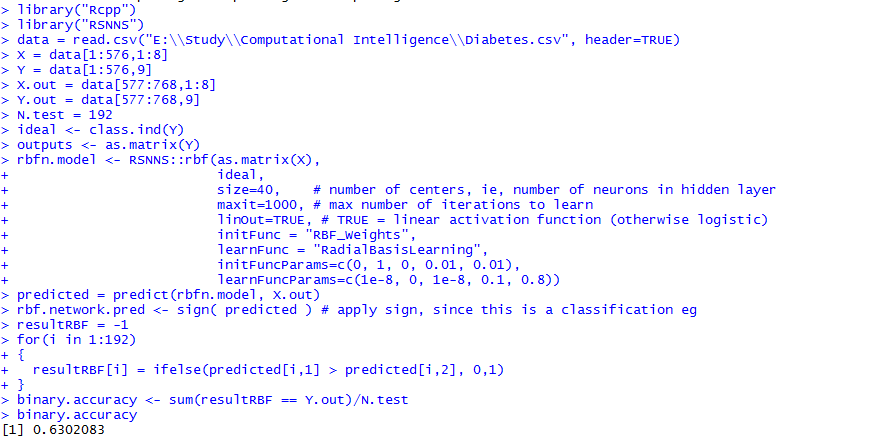


Figure Radial basis function NN

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

We combined the output of multi layer perceptron with backpropagation, probabilistic neural network and radial basis function. Then used voting technique to arrive at the final result. The accuracy of this combined model is more than any other model used individually. The accuracy of voting ensemble achieved is **66.67 %.**

Time taken for ensemble logic to execute is **1.995603 secs**.

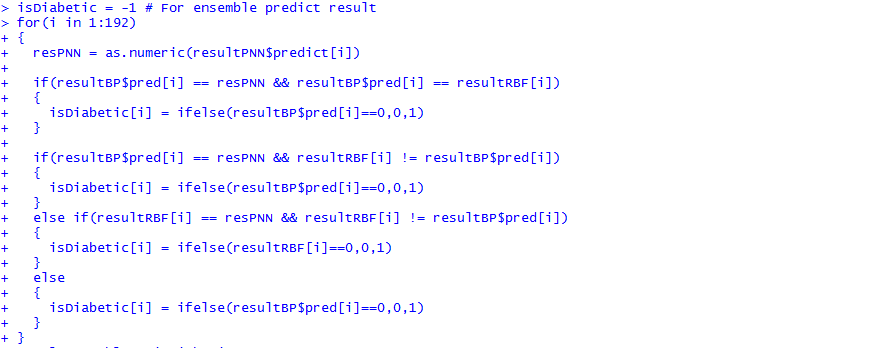


Figure Ensemble Voting logic

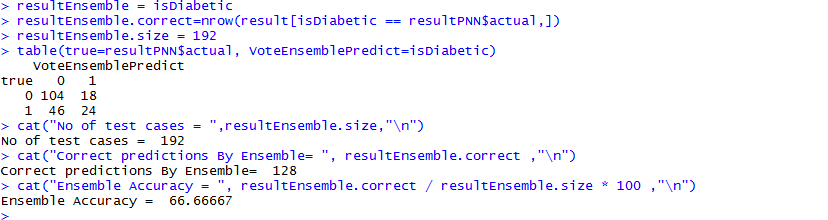


Figure Ensemble voting result

# Using Weka:-

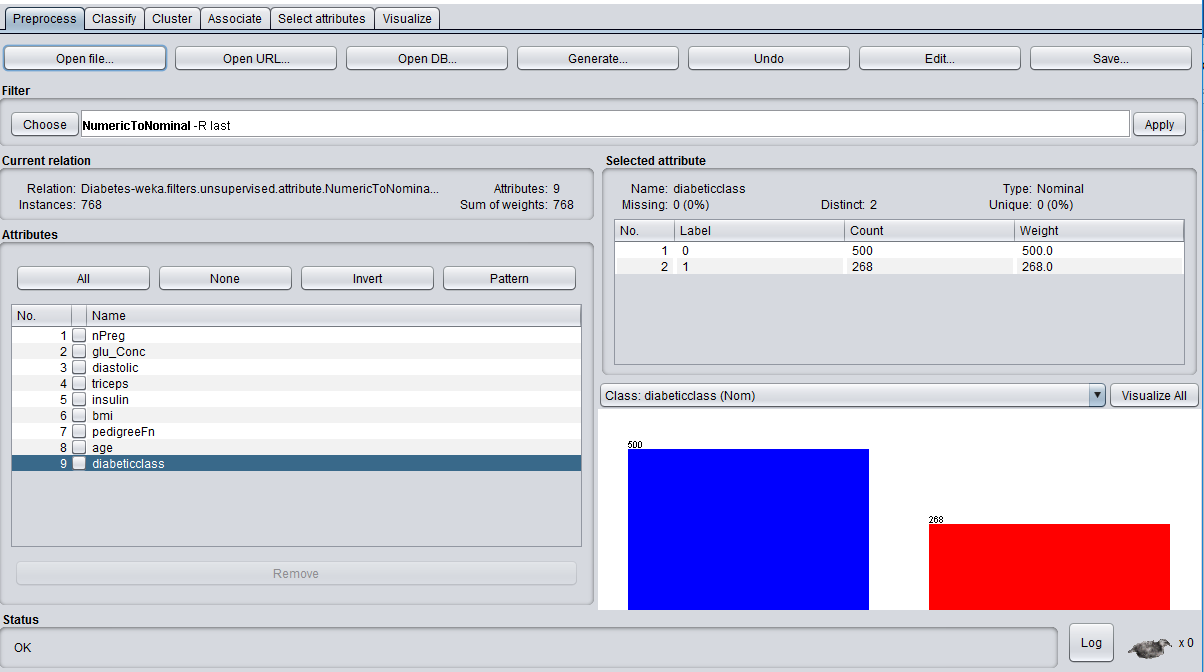


Figure DataSet for diabetic class

## 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 **75.5%.**

Hidden nodes = 10, Hidden layer = 1.

Training – 75%, Testing 25%

Time taken to build model: 2.83 seconds

Time taken to test model on test split: 0 seconds

The result is shown below:-

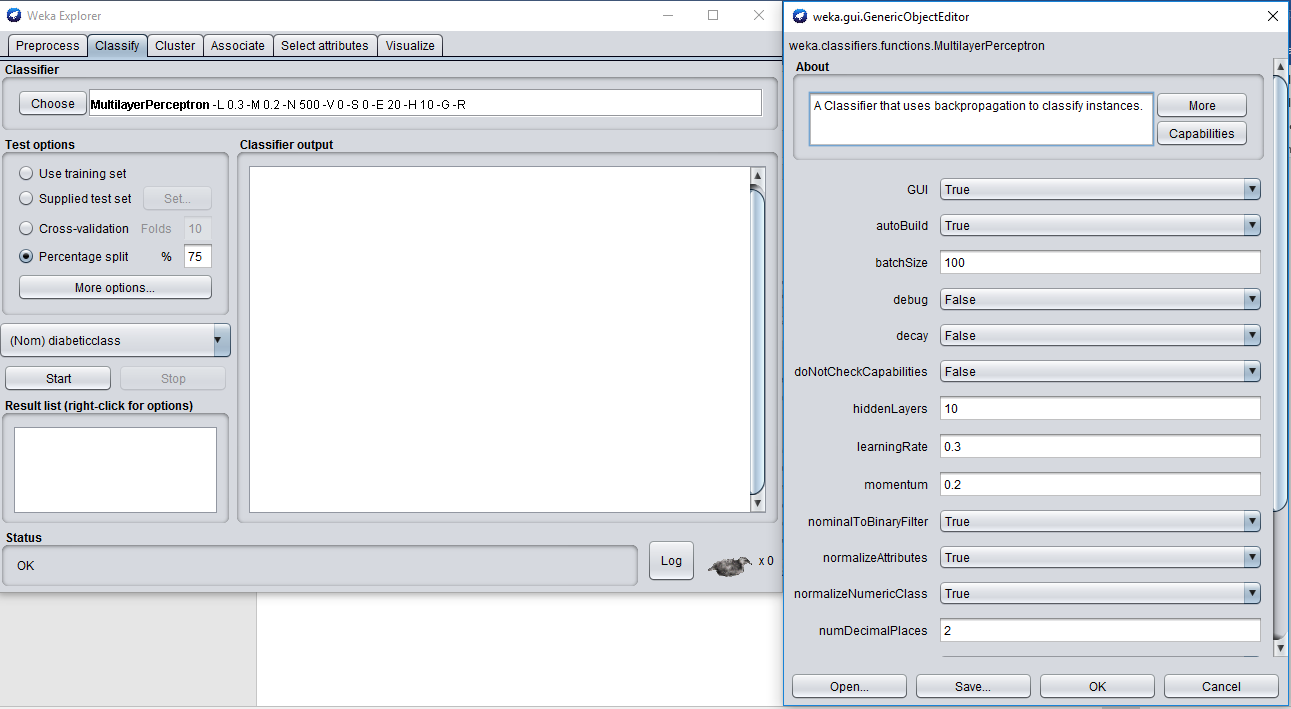


Figure Multilayer perceptron settings

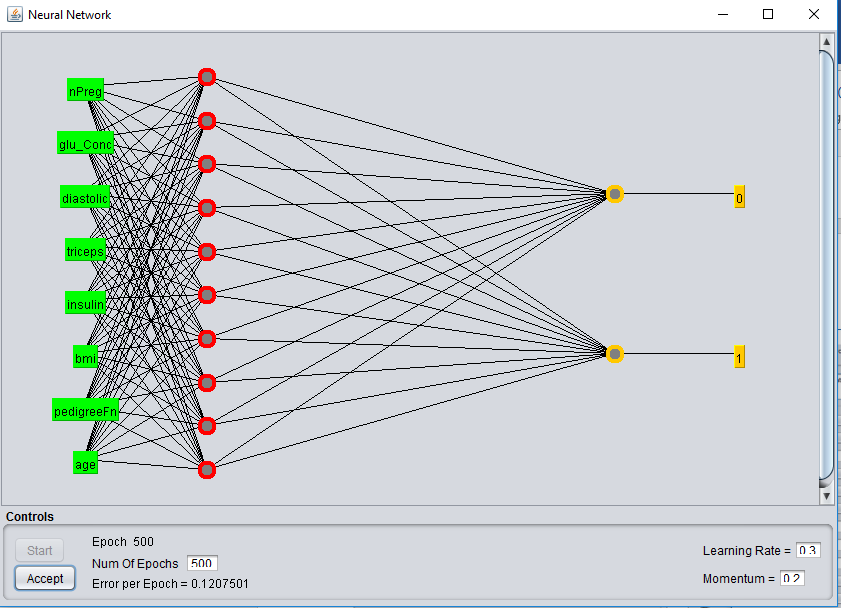


Figure MLP Plot

The report is as follows:-

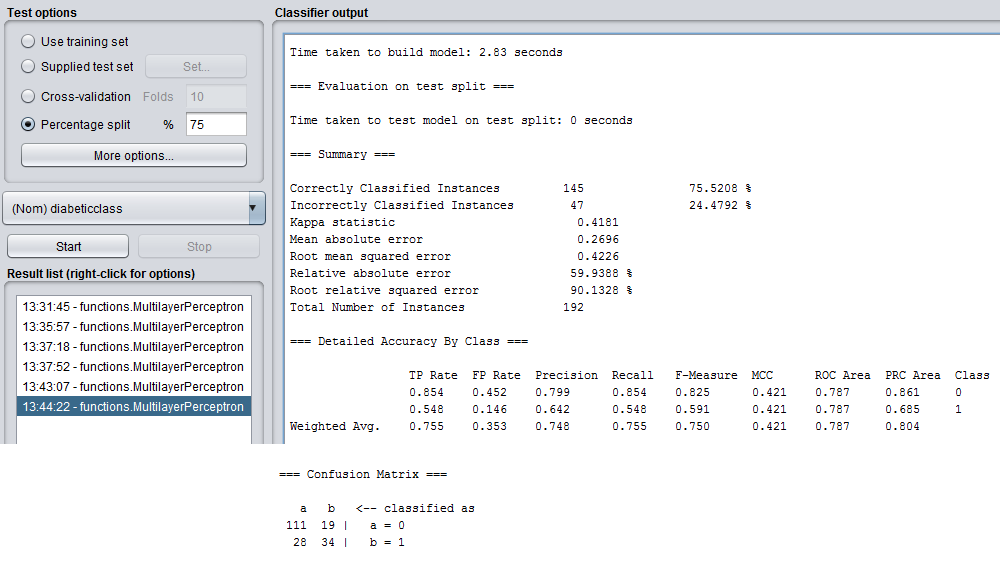


Figure MLP Result

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

## 2.2 Using RBF Classifier:-

Using individual RBF classifier we get the accuracy as **77.60%.**

Testing = 75%, Training =25%

Time taken to build model: 0.2 seconds

Time taken to test model on test split: 0.01 seconds

The settings for RBF Classifier is shown in fig 14.

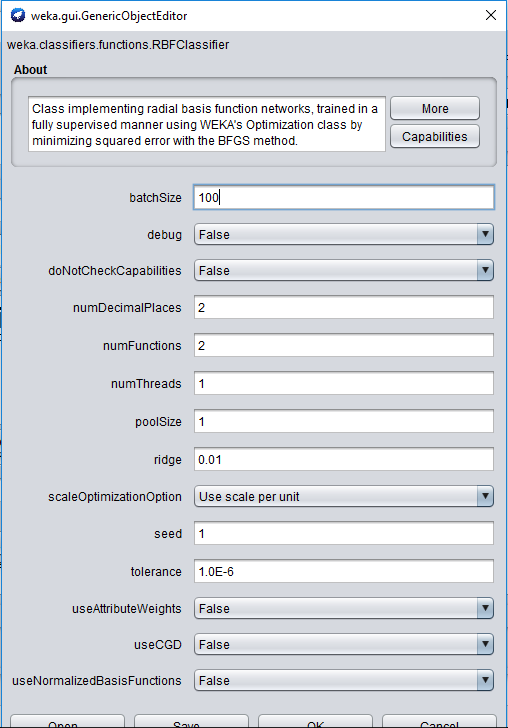


Figure RBFClassifier Settings

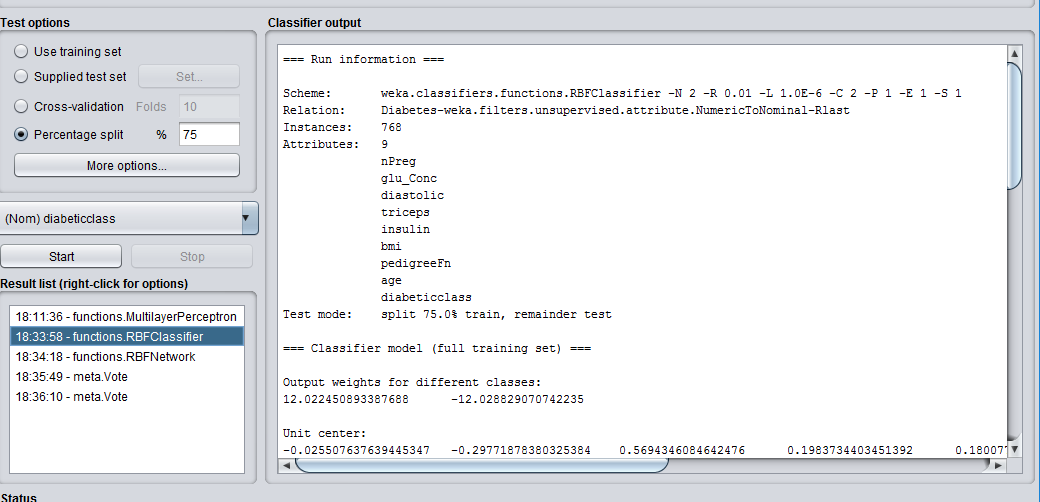


Figure RBFClassifier run info

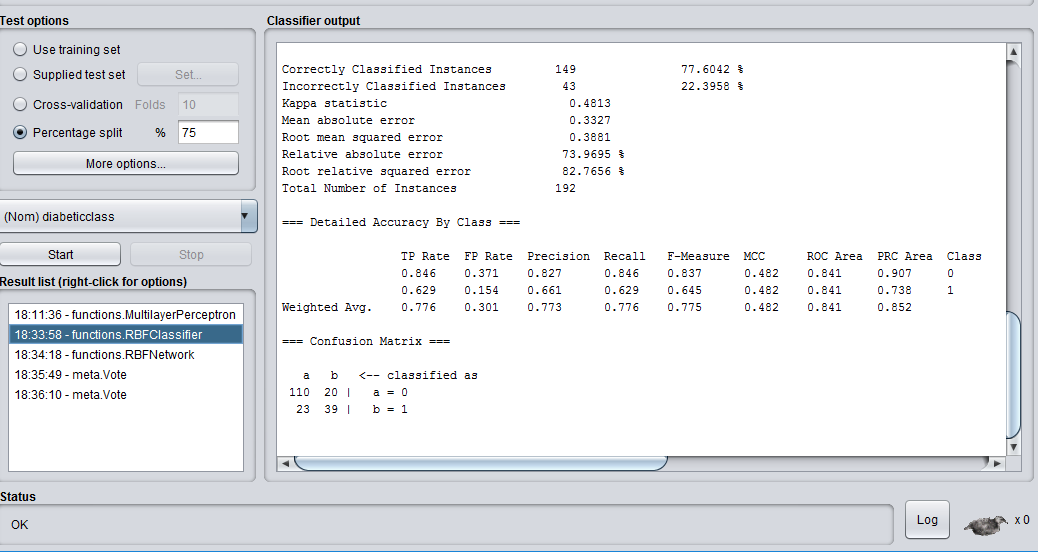


Figure RBF Classifier result

For whole log info check attached “**Weka-Diabetes\_RBF\_Classifier.txt**”.

## 2.3 RBF Network

Using RBFNetwork individually gives accuracy of **80.20%**.

Training = 75%, Testing = 25%

Time taken to build model: 0.1 seconds

Time taken to test model on test split: 0 seconds

The settings is shown in fig.17.

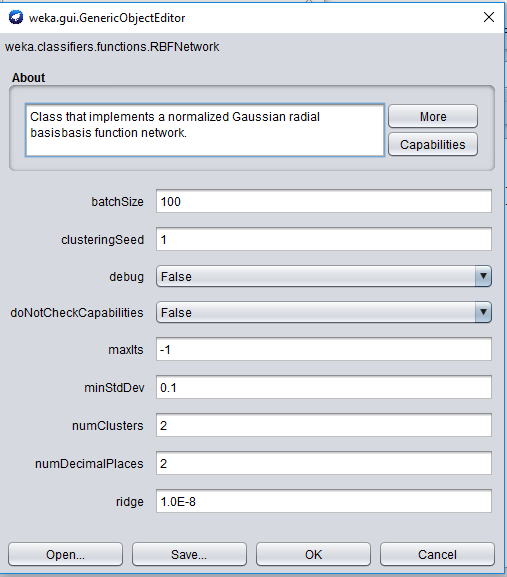


Figure RBFNetwork Settings

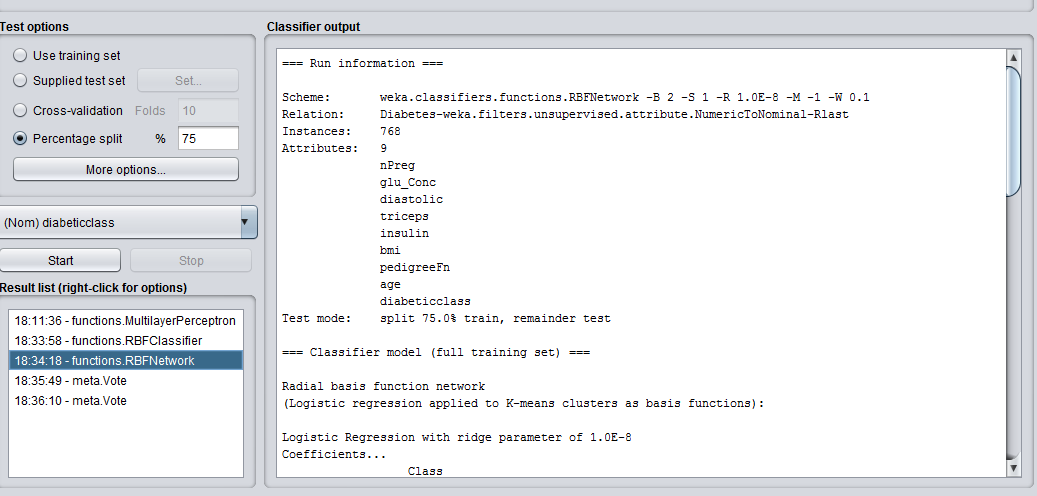


Figure RBF Network Run info

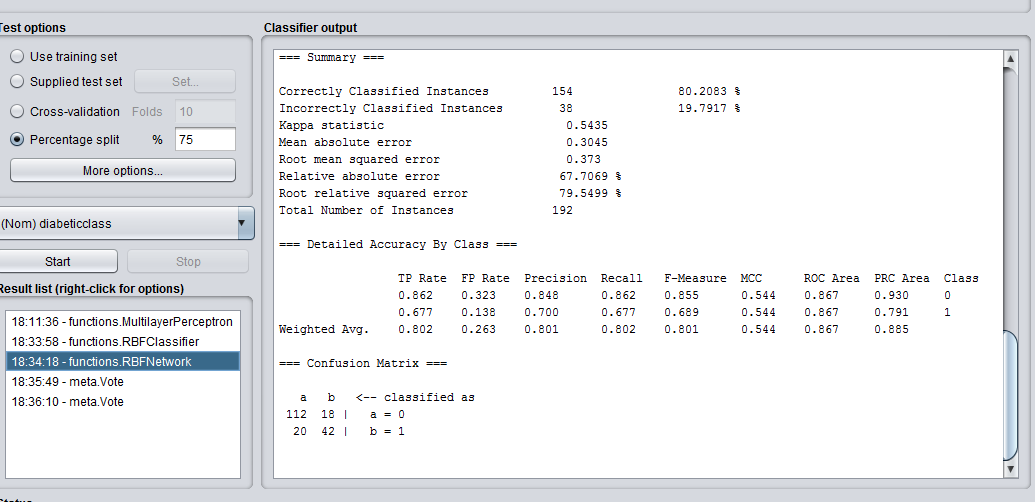


Figure RBF Network result

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

## 2.4 Ensemble MLP , RBF Classifier and RBF Network using majority voting :-

Accuracy = **80.7292 %**

Time taken to build model: 3.33 seconds

Time taken to test model on test split: 0.02 seconds

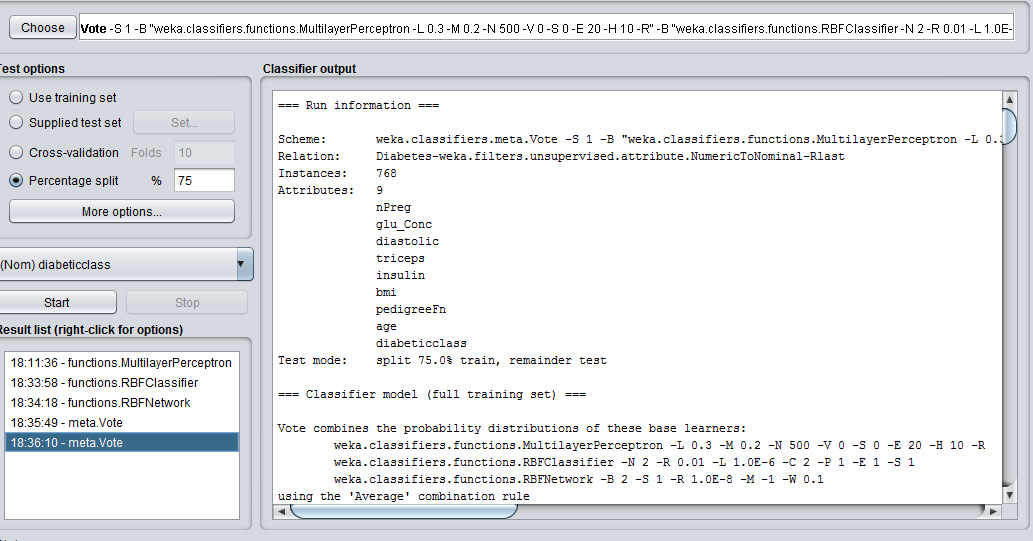


Figure Ensemble (ML, RBFClassifier, RBF Network) Voting

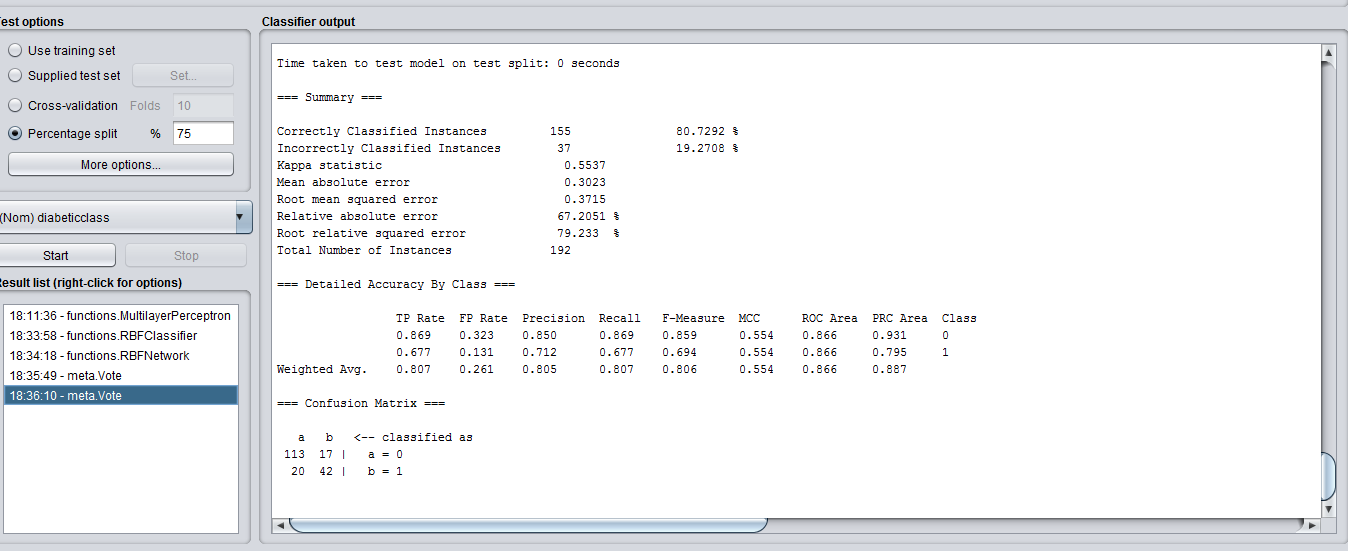


Figure Ensemble result

For log info check “**Weka-Ensemble\_Voting\_ML\_RBFClassifier\_RBFNetwork.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) | 0.4329121 | 0.286922 | 61.97%. |
| Probabilistic Neural Network(PNN) | 0.8680301 | 22.65863 | 64.025% |
| Radial Basis Function(RBF) | 5.51747 | 0.04653406 | 63.02% |
| Ensemble Majority Voting (ML BP + PNN + RBF) | Total time = Sum of all individual model times = 29.81049826 sec | 1.995603 | 66.67 %. |

Table Summary Neural Network using R

### Using Weka:-

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **Time taken in modelling(sec).** | **Time taken in testing/predicting(sec).** | **Accuracy** |
| Multi layer perceptron | 2.83 | 0 | 75.5%. |
| Radial Basis Function Classifier(RBFClassifier) | 0.2 | 0.01 | 77.60%. |
| Radial Basis Function Network(RBFNetwork) | 0.1 | 0 | 80.20%. |
| Ensemble Majority Voting (ML + RBFClassifier + RBFNetwork) | 3.33 | 0.02 | 80.73% |

Table Summary Neural Network using Weka

# **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 PNN gives best accuracy of 64.025% but when we combine all the three networks we get the accuracy of 66.67% which is 2.645% more. In weka although the increment in accuracy is not very much significant but still there is increment of 0.53% accuracy by using ensemble rather than using RBFNetwork alone which gave 80.20% accuracy. Total accuracy of ensemble using majority voting in weka is 80.73%.

Although there is improvement in result using ensemble but if we look at performance we see a delay. It is obvious as computation of one model alone would be faster than combining it with other models. The observation for time taken in ensemble is shown in table1 and table2 for R and weka respectively.

So there is always tradeoff between performance time and accuracy if we use neural network models in an ensemble.