Diagnosing Breast Cancer Using Artificial Network Models

By
Vanessa Fotso

Prepared for

University of Maryland Global Campus

DATA 630: Machine Learning

July 22, 2020

Introduction

Cancer is defined an uncontrolled growth and division of abnormal cells. There are several types of cancer, one of which is breast cancer. Breast cancer is the second leading cause of cancer death in women according to the American Cancer Society, and it has been widely studied around the world. A study from the National Cancer Institute (NCI) predicted the number of cases to grow from 283,000 cases in 2011 to 441,000 in 2030, representing more than 50% increase. The risk of developing the disease increases with age. Additional risk factors include family and medical history, obesity, physical activity, smoking, etc. prevention in this case is hard as breast cancer is caused by both the combination of controllable and non-controllable factors. Thus, the ability to properly diagnose breast cancer and early detection is essential to tailor treatment options and beat the disease.

Several machine learning approaches have been used over the years to help medical staff community provide accurate diagnostic. Al Mutaz et al in their study built a model using Support Vector Machine (SVM) that was able to better classify breast cancer, helping doctors detect the disease at it early stage. Additionally, Ferroni et al were able to build a model using ML-based decision support system (DSS) and random optimization that provided breast cancer prognostic with 86% accuracy. With early detection and accurate diagnosis being the key to increase survival rate, machine learning still needs to be further exploited to decrease the current statistics.

This paper seek to analyze various characteristics of cell nucleus and determine which combination of characteristics lead to the classification of cells as being cancerous malignant or benign. The predictive model will be build using Neural Networks method in R language on the

Breast Cancer Wisconsin (Diagnostic) Data Set with the goal to accurately predict breast cancer in an individual to help early detection.

Method

Artificial Neural Network (ANN) is a machine learning method composed of dense net of computational nodes or neurons performing a computation in a parallel and distributed manner (Haykin, 20100). It is an adaptive approach that learn by examples. The method gain knowledge through learning steps and the knowledge gained ais stored with interneuron connections. The algorithm here uses the neurons to learn hidden pattens in a training set then is evaluated on unknown data (test set) to check its performance. The most commonly type of neural nets is the multilayer feedforward networks, which is a layered structure of computational neurons. It takes an input signal, which propagates through the structured network layer by layer in a forward direction. The feedforward network exploit a backpropagation training algorithm to calculate the error and adjust the weight values to generate statistically correct output. The learning rules of neural nets include the least mean square, gradient descent, conjugated gradient, etc. These rules will be used to calcite the error at the output node.

Data and Preprocessing

The data was obtained from the University of California Irvine Machine Learning repository. This dataset describes the characteristics of cell nuclei of 569 patients with the purpose to classify tumorous cells as either malignant(M, dangerous) or benign(B, non dangerous) using feature values calculated from digitized image of breast mass (UCI). The data consists of 31 variables characterizing a patient's cell and there are no missing values which one requirement of neural nets. The summary of the data can be visualized in figure 2 in the appendix. The summary

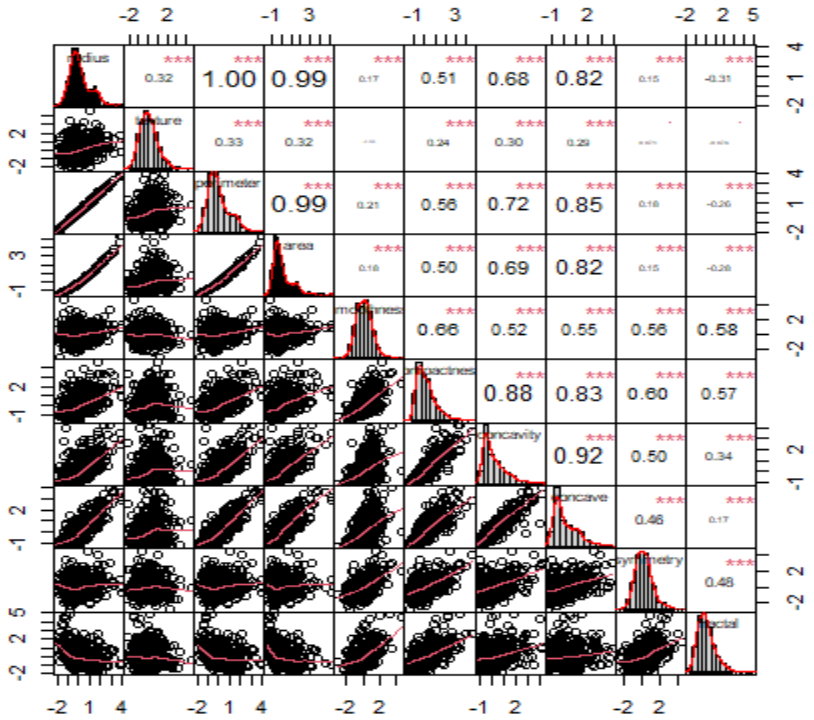
shows that all variables are numerical except for the diagnosis. There is also an ID column which is irrelevant to the study and will be dropped from the dataset. Next, the diagnosis variable being the dependent variable will also be transformed to binary digits assigning 1 to malignant (M) and 0 to benign (B) for the fitting step. The resulting summary can be observed in figure 3. As seen in figure 3, the dataset is unbalanced.

```
> prop.table(table(data$diagnosis))


        0         1

0.6274165 0.3725835
```

There is about 37% of malignant cells and 63% of benign cells. This can be fix by scaling the data to ensure that all dataset fall in the same range. The correlation matrix was then derived to check the correlation between variables.



The histograms on the above plot now follow a normal distribution and there's a great correlation between some variables in the dataset as the correlation coefficient is closed to 1 in most cases.

Neural Nets Modeling and Results

R neuralnets package was used to the neural nets model that will be used to diagnose breast cancer in an individual given a specific set of attributes. The data was first randomly divided into two sets, 75% for model training and 25% for training set that will be used to evaluate the performance. Neuralnets provide a great number of parameters for model optimization. The most commonly used include the formula, data, hidden, err.fct and linear.output

Parameters:

. Formula: given in the form of Y~ x1 + x2 + … where Y is the label and all the variables after ~ are the features. This is the formula definition that will be used to build the model.

. data: is the dataframe used to train the model

. hidden represents single layer with x neurons respectively. hidden layers optimize the weight of the input variables to improve the prediction performance

- err.fct : is used for the calculation of errors

- linear.ouput: this is the logic. Set FALSE for apply act.fct (classification) otherwise TRUE

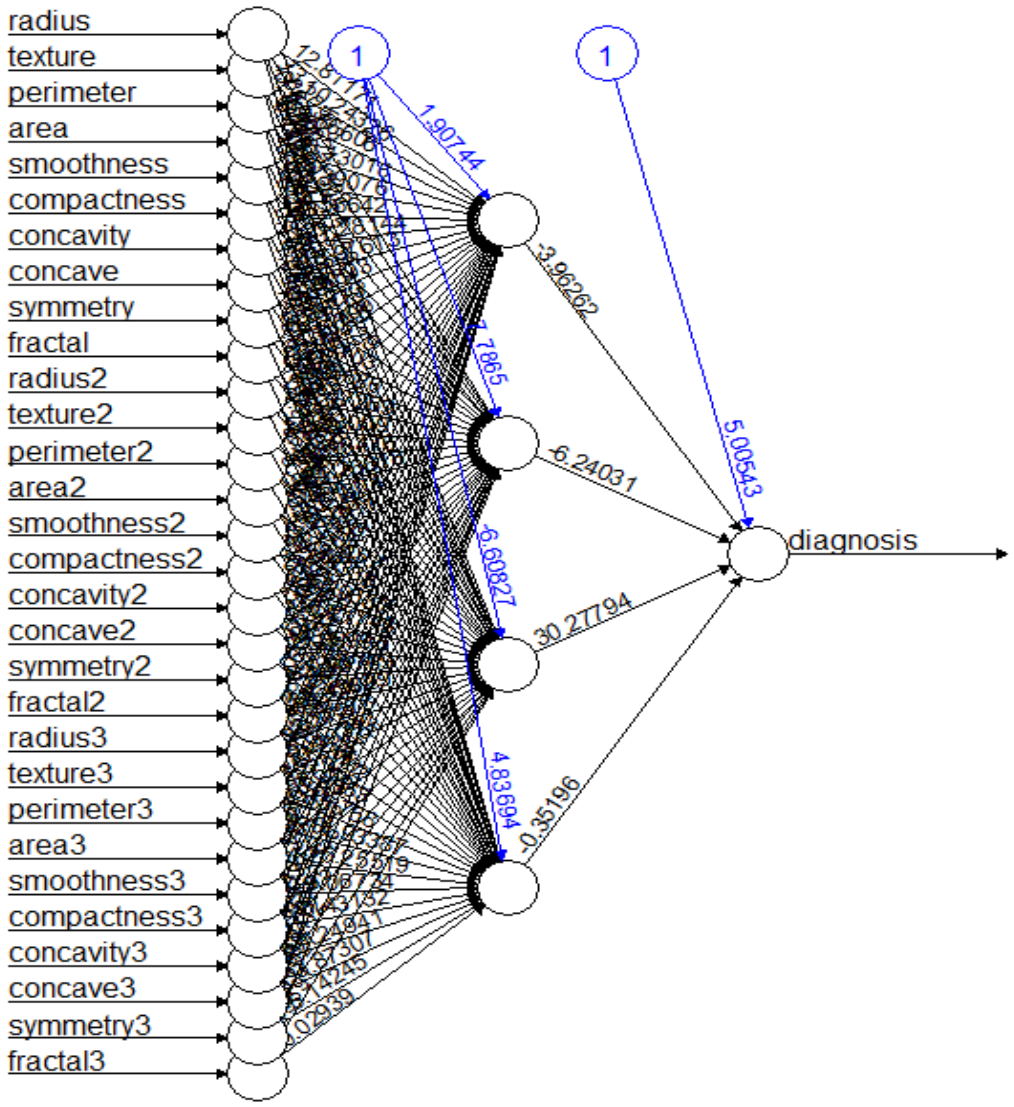
This is the parameter values used for the model:

```
nn<-neuralnet(formula = diagnosis~., data = train.data, hidden=4, err.fct="ce", linear.output = FALSE)
```

```
> nn$net.result[[1]][1:10] # display the first 10 predicted probabilities
```

```
 [1] 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000
1.0000000 0.9999996
```

The above result represents the first 10 patients in the training set. All the probabilities displayed here are above 0.5; thus, the cells for those patients will be labeled as malignant. Figure 5 shows the method had 319 training steps, with the calculated error for the model on the training set of 6.57. The picture below show the resulting network:



The model has 4 neurons in its hidden layer. The black lines represent the connections with weights, which are calculated with the backpropagation algorithm (propagate error at the output unit to all units in the way that each unit error is proportional to the contribution of the given unit towards total error at the output). The blue lines shows the bias term. No error was outputted with the resulting network. The picture below shows the resulted confusion matrix:

```
> # confusion matrix for the training set
> table(mypredict, train.data$diagnosis, dnn =c("Predicted", "Actual"))
         Actual
Predicted   0   1
        0 258   1
        1   0 150
> |
```

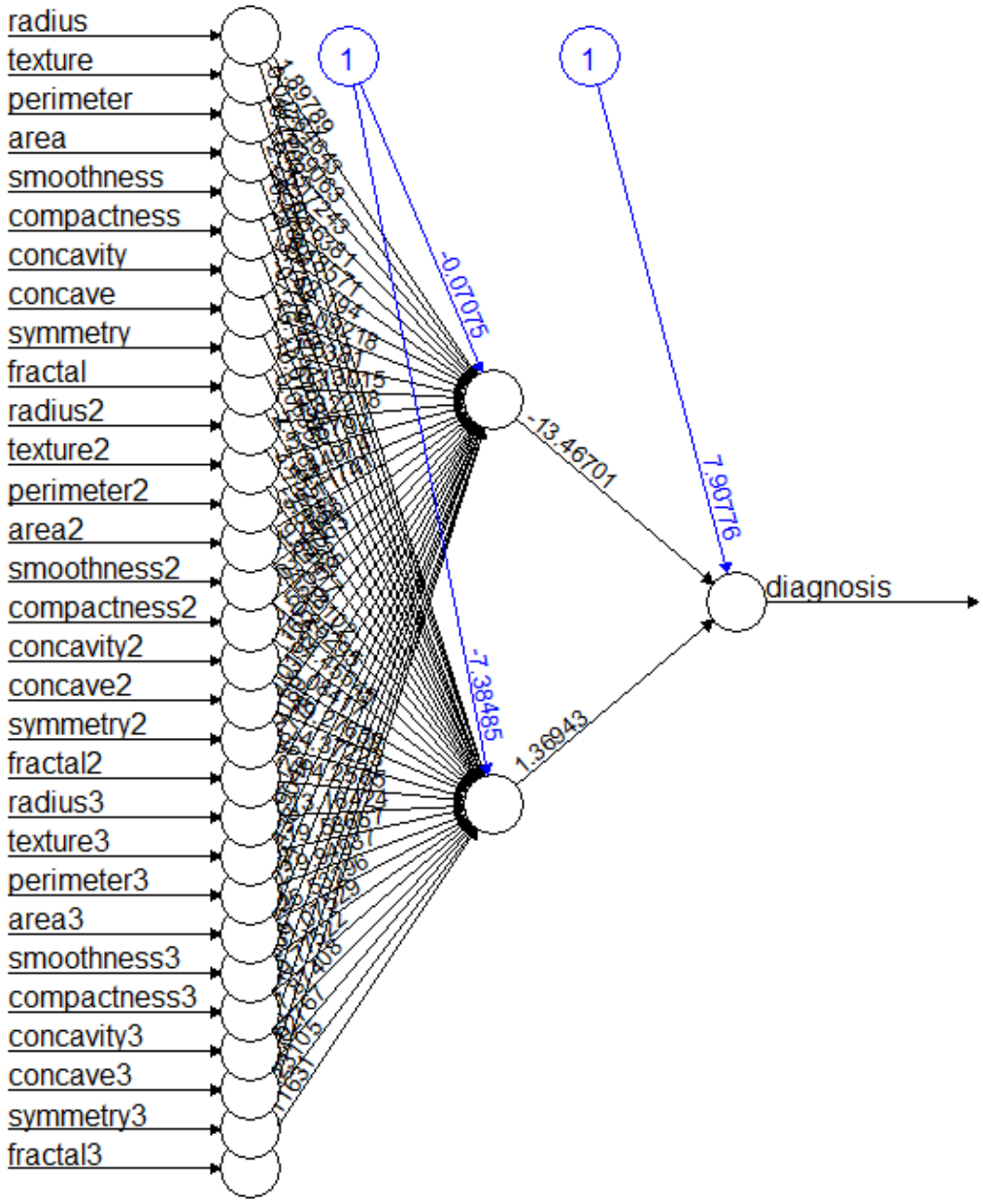
The training set has a total of 409 instances. 258 predictions are true negative while 150 predictions are true positive. This results in a total of 408 correct predictions out of 409; thus, the accuracy of 99.76% and precision of 99.34% (true positive divided by the total number predicted positive).

The model was the evaluated on the test set, resulting in the following matrix:

```
> # confusion matrix for the test set
> testPred <- compute(nn, test.data[, 0:31])$net.result
> testPred<-apply(testPred, c(1), round)
> table(testPred, test.data$diagnosis, dnn =c("Predicted", "Actual"))
         Actual
Predicted  0  1
        0 99  4
        1  0 57
```

The test set had 160 observations, 99 of which were true negative and 57 were true positive. Thus, the accuracy of 97.5% and the precision of 93.44%. The model have a high performance on both the training dataset and the unknown data. Thus the model has a good optimization making it capable to generalize unknown data.

For validation, the hidden parameter of the model was adjusted to 2 (from 4 to 2) while other params remained constant. Below are the generated neural network plot and confusion matrices.



Neural Network with hidden = 2

```
> # confusion matrix for the training set
> table(mypredict, train.data$diagnosis, dnn =c("Predicted", "Actual"))
         Actual
Predicted   0   1
        0 258   1
        1   0 150
>
> # confusion matrix for the test set
> testPred <- compute(nn, test.data[, 0:31])$net.result
> testPred<-apply(testPred, c(1), round)
> table(testPred, test.data$diagnosis, dnn =c("Predicted", "Actual"))
         Actual
Predicted  0  1
        0 99  5
        1  0 56
> |
```

Confusion Matrices on the Training and Test set

The second trial generated an accuracy of 99.76% and a precision of 99.34% on the training set while the accuracy was 96.9% and the precision was 91.8% on the test set. The accuracy and precision on the training set here is similar to the one obtained on the first run and the difference on the test set is negligible (1% difference). Decreasing the number of neurons in this case has not affected the model's capacity. Therefore, the neural network model built here is efficient in diagnosing breast cancer given the 32 cell characteristics provided in the dataset.

Conclusion

The analysis presented an artificial neural network approach for disease diagnosis. Two different configurations were used to train and test the model against the given Breast Cancer dataset. The results show the generated model's ability to properly classify an individual's cell as benign or malignant, given a certain cell's characteristics. The best configuration provided an accuracy of 97.5% on unknown data. The model shows a very good performance and can help medical providers with diagnosis of breast cancer. This is probably due to the fact the data was well processed before the fitting phase and there was enough data to feed the model during the training stage. However, the analysis conducted here is limited to a single source of data, which

may have limited attributes for breast cancer. Further data collection from different source with additional set of attributes is required to better assess the model performance.

References

Al Mutaz et al. “Breast Cancer Detection Based on Statistical Textural Features Classification. 2008

American Cancer Society. Cancer Facts and Figures 2020. Atlanta, Ga: American Cancer Society; 2020

Ferroni P. et al. Breast Cancer Prognosis Using a Machine learning Approach. Cancers, 11(3), 328. https://doi.org/10/3390/cancers11030328

Haykin, Simon, Neural Networks and Learning Machines. New Dehli: Pearson Education, 2010.

National Cancer Institute. Study Forecasts New Breast Cancer Cases by 2030. April 2015

Zurada M., Introduction to Artificial Neural Systems. Mumbai, India: Jaico Publishing House, 1994

Appendix

Data Variable

Attribute Information:

1) ID number

2) Diagnosis (M = malignant, B = benign)

3-32)

Ten real-valued features are computed for each cell nucleus:

a) radius (mean of distances from center to points on the perimeter)

b) texture (standard deviation of gray-scale values)

c) perimeter

d) area

e) smoothness (local variation in radius lengths)

f) compactness (perimeter^2 / area - 1.0)

g) concavity (severity of concave portions of the contour)

h) concave points (number of concave portions of the contour)

i) symmetry

j) fractal dimension ("coastline approximation" - 1)

```
> #Preview the first 6 rows
> head(data)
       ID diagnosis radius texture perimeter   area smoothness compactness concavity concave symmetry
1   842302         M  17.99   10.38    122.80 1001.0    0.11840     0.27760    0.3001 0.14710   0.2419
2   842517         M  20.57   17.77    132.90 1326.0    0.08474     0.07864    0.0869 0.07017   0.1812
3 84300903         M  19.69   21.25    130.00 1203.0    0.10960     0.15990    0.1974 0.12790   0.2069
4 84348301         M  11.42   20.38     77.58  386.1    0.14250     0.28390    0.2414 0.10520   0.2597
5 84358402         M  20.29   14.34    135.10 1297.0    0.10030     0.13280    0.1980 0.10430   0.1809
6   843786         M  12.45   15.70     82.57  477.1    0.12780     0.17000    0.1578 0.08089   0.2087
  fractal radius2 texture2 perimeter2  area2 smoothness2 compactness2 concavity2 concave2 symmetry2
1 0.07871  1.0950   0.9053      8.589 153.40    0.006399      0.04904    0.05373  0.01587   0.03003
2 0.05667  0.5435   0.7339      3.398  74.08    0.005225      0.01308    0.01860  0.01340   0.01389
3 0.05999  0.7456   0.7869      4.585  94.03    0.006150      0.04006    0.03832  0.02058   0.02250
4 0.09744  0.4956   1.1560      3.445  27.23    0.009110      0.07458    0.05661  0.01867   0.05963
5 0.05883  0.7572   0.7813      5.438  94.44    0.011490      0.02461    0.05688  0.01885   0.01756
6 0.07613  0.3345   0.8902      2.217  27.19    0.007510      0.03345    0.03672  0.01137   0.02165
  fractal2 radius3 texture3 perimeter3  area3 smoothness3 compactness3 concavity3 concave3 symmetry3
1 0.006193   25.38    17.33     184.60 2019.0      0.1622       0.6656     0.7119   0.2654     0.4601
2 0.003532   24.99    23.41     158.80 1956.0      0.1238       0.1866     0.2416   0.1860     0.2750
3 0.004571   23.57    25.53     152.50 1709.0      0.1444       0.4245     0.4504   0.2430     0.3613
4 0.009208   14.91    26.50      98.87  567.7      0.2098       0.8663     0.6869   0.2575     0.6638
5 0.005115   22.54    16.67     152.20 1575.0      0.1374       0.2050     0.4000   0.1625     0.2364
6 0.005082   15.47    23.75     103.40  741.6      0.1791       0.5249     0.5355   0.1741     0.3985
  fractal3
1  0.11890
2  0.08902
3  0.08758
4  0.17300
5  0.07678
6  0.12440
>
```

Figure1: Dataset Preview

```
> summary(data)
       ID            diagnosis             radius          texture         perimeter
 Min.   :     8670   Length:569         Min.   : 6.981   Min.   : 9.71   Min.   : 43.79
 1st Qu.:   869218   Class :character   1st Qu.:11.700   1st Qu.:16.17   1st Qu.: 75.17
 Median :   906024   Mode  :character   Median :13.370   Median :18.84   Median : 86.24
 Mean   : 30371831                      Mean   :14.127   Mean   :19.29   Mean   : 91.97
 3rd Qu.:  8813129                      3rd Qu.:15.780   3rd Qu.:21.80   3rd Qu.:104.10
 Max.   :911320502                      Max.   :28.110   Max.   :39.28   Max.   :188.50
      area          smoothness       compactness        concavity          concave
 Min.   : 143.5   Min.   :0.05263   Min.   :0.01938   Min.   :0.00000   Min.   :0.00000
 1st Qu.: 420.3   1st Qu.:0.08637   1st Qu.:0.06492   1st Qu.:0.02956   1st Qu.:0.02031
 Median : 551.1   Median :0.09587   Median :0.09263   Median :0.06154   Median :0.03350
 Mean   : 654.9   Mean   :0.09636   Mean   :0.10434   Mean   :0.08880   Mean   :0.04892
 3rd Qu.: 782.7   3rd Qu.:0.10530   3rd Qu.:0.13040   3rd Qu.:0.13070   3rd Qu.:0.07400
 Max.   :2501.0   Max.   :0.16340   Max.   :0.34540   Max.   :0.42680   Max.   :0.20120
    symmetry         fractal          radius2          texture2        perimeter2         area2
 Min.   :0.1060   Min.   :0.04996   Min.   :0.1115   Min.   :0.3602   Min.   : 0.757   Min.   :  6.802
 1st Qu.:0.1619   1st Qu.:0.05770   1st Qu.:0.2324   1st Qu.:0.8339   1st Qu.: 1.606   1st Qu.: 17.850
 Median :0.1792   Median :0.06154   Median :0.3242   Median :1.1080   Median : 2.287   Median : 24.530
 Mean   :0.1812   Mean   :0.06280   Mean   :0.4052   Mean   :1.2169   Mean   : 2.866   Mean   : 40.337
 3rd Qu.:0.1957   3rd Qu.:0.06612   3rd Qu.:0.4789   3rd Qu.:1.4740   3rd Qu.: 3.357   3rd Qu.: 45.190
 Max.   :0.3040   Max.   :0.09744   Max.   :2.8730   Max.   :4.8850   Max.   :21.980   Max.   :542.200
  smoothness2       compactness2       concavity2        concave2         symmetry2
 Min.   :0.001713   Min.   :0.002252   Min.   :0.00000   Min.   :0.000000   Min.   :0.007882
 1st Qu.:0.005169   1st Qu.:0.013080   1st Qu.:0.01509   1st Qu.:0.007638   1st Qu.:0.015160
 Median :0.006380   Median :0.020450   Median :0.02589   Median :0.010930   Median :0.018730
 Mean   :0.007041   Mean   :0.025478   Mean   :0.03189   Mean   :0.011796   Mean   :0.020542
 3rd Qu.:0.008146   3rd Qu.:0.032450   3rd Qu.:0.04205   3rd Qu.:0.014710   3rd Qu.:0.023480
 Max.   :0.031130   Max.   :0.135400   Max.   :0.39600   Max.   :0.052790   Max.   :0.078950
```

```
    fractal2          radius3          texture3        perimeter3         area3          smoothness3
 Min.   :0.0008948   Min.   : 7.93   Min.   :12.02   Min.   : 50.41   Min.   : 185.2   Min.   :0.07117
 1st Qu.:0.0022480   1st Qu.:13.01   1st Qu.:21.08   1st Qu.: 84.11   1st Qu.: 515.3   1st Qu.:0.11660
 Median :0.0031870   Median :14.97   Median :25.41   Median : 97.66   Median : 686.5   Median :0.13130
 Mean   :0.0037949   Mean   :16.27   Mean   :25.68   Mean   :107.26   Mean   : 880.6   Mean   :0.13237
 3rd Qu.:0.0045580   3rd Qu.:18.79   3rd Qu.:29.72   3rd Qu.:125.40   3rd Qu.:1084.0   3rd Qu.:0.14600
 Max.   :0.0298400   Max.   :36.04   Max.   :49.54   Max.   :251.20   Max.   :4254.0   Max.   :0.22260
  compactness3       concavity3        concave3        symmetry3        fractal3
 Min.   :0.02729   Min.   :0.0000   Min.   :0.00000   Min.   :0.1565   Min.   :0.05504
 1st Qu.:0.14720   1st Qu.:0.1145   1st Qu.:0.06493   1st Qu.:0.2504   1st Qu.:0.07146
 Median :0.21190   Median :0.2267   Median :0.09993   Median :0.2822   Median :0.08004
 Mean   :0.25427   Mean   :0.2722   Mean   :0.11461   Mean   :0.2901   Mean   :0.08395
 3rd Qu.:0.33910   3rd Qu.:0.3829   3rd Qu.:0.16140   3rd Qu.:0.3179   3rd Qu.:0.09208
 Max.   :1.05800   Max.   :1.2520   Max.   :0.29100   Max.   :0.6638   Max.   :0.20750
```

Figure 2: Dataset Summary

```
> summary(data)
   diagnosis          radius          texture         perimeter          area          smoothness
 Min.   :0.0000   Min.   : 6.981   Min.   : 9.71   Min.   : 43.79   Min.   : 143.5   Min.   :0.05263
 1st Qu.:0.0000   1st Qu.:11.700   1st Qu.:16.17   1st Qu.: 75.17   1st Qu.: 420.3   1st Qu.:0.08637
 Median :0.0000   Median :13.370   Median :18.84   Median : 86.24   Median : 551.1   Median :0.09587
 Mean   :0.3726   Mean   :14.127   Mean   :19.29   Mean   : 91.97   Mean   : 654.9   Mean   :0.09636
 3rd Qu.:1.0000   3rd Qu.:15.780   3rd Qu.:21.80   3rd Qu.:104.10   3rd Qu.: 782.7   3rd Qu.:0.10530
 Max.   :1.0000   Max.   :28.110   Max.   :39.28   Max.   :188.50   Max.   :2501.0   Max.   :0.16340
  compactness       concavity          concave          symmetry         fractal
 Min.   :0.01938   Min.   :0.00000   Min.   :0.00000   Min.   :0.1060   Min.   :0.04996
 1st Qu.:0.06492   1st Qu.:0.02956   1st Qu.:0.02031   1st Qu.:0.1619   1st Qu.:0.05770
 Median :0.09263   Median :0.06154   Median :0.03350   Median :0.1792   Median :0.06154
 Mean   :0.10434   Mean   :0.08880   Mean   :0.04892   Mean   :0.1812   Mean   :0.06280
 3rd Qu.:0.13040   3rd Qu.:0.13070   3rd Qu.:0.07400   3rd Qu.:0.1957   3rd Qu.:0.06612
 Max.   :0.34540   Max.   :0.42680   Max.   :0.20120   Max.   :0.3040   Max.   :0.09744
    radius2          texture2        perimeter2         area2          smoothness2
 Min.   :0.1115   Min.   :0.3602   Min.   : 0.757   Min.   :  6.802   Min.   :0.001713
 1st Qu.:0.2324   1st Qu.:0.8339   1st Qu.: 1.606   1st Qu.: 17.850   1st Qu.:0.005169
 Median :0.3242   Median :1.1080   Median : 2.287   Median : 24.530   Median :0.006380
 Mean   :0.4052   Mean   :1.2169   Mean   : 2.866   Mean   : 40.337   Mean   :0.007041
 3rd Qu.:0.4789   3rd Qu.:1.4740   3rd Qu.: 3.357   3rd Qu.: 45.190   3rd Qu.:0.008146
 Max.   :2.8730   Max.   :4.8850   Max.   :21.980   Max.   :542.200   Max.   :0.031130
  compactness2       concavity2         concave2         symmetry2          fractal2
 Min.   :0.002252   Min.   :0.00000   Min.   :0.000000   Min.   :0.007882   Min.   :0.0008948
 1st Qu.:0.013080   1st Qu.:0.01509   1st Qu.:0.007638   1st Qu.:0.015160   1st Qu.:0.0022480
 Median :0.020450   Median :0.02589   Median :0.010930   Median :0.018730   Median :0.0031870
 Mean   :0.025478   Mean   :0.03189   Mean   :0.011796   Mean   :0.020542   Mean   :0.0037949
 3rd Qu.:0.032450   3rd Qu.:0.04205   3rd Qu.:0.014710   3rd Qu.:0.023480   3rd Qu.:0.0045580
 Max.   :0.135400   Max.   :0.39600   Max.   :0.052790   Max.   :0.078950   Max.   :0.0298400
```

```
    radius3         texture3       perimeter3         area3          smoothness3       compactness3
 Min.   : 7.93   Min.   :12.02   Min.   : 50.41   Min.   : 185.2   Min.   :0.07117   Min.   :0.02729
 1st Qu.:13.01   1st Qu.:21.08   1st Qu.: 84.11   1st Qu.: 515.3   1st Qu.:0.11660   1st Qu.:0.14720
 Median :14.97   Median :25.41   Median : 97.66   Median : 686.5   Median :0.13130   Median :0.21190
 Mean   :16.27   Mean   :25.68   Mean   :107.26   Mean   : 880.6   Mean   :0.13237   Mean   :0.25427
 3rd Qu.:18.79   3rd Qu.:29.72   3rd Qu.:125.40   3rd Qu.:1084.0   3rd Qu.:0.14600   3rd Qu.:0.33910
 Max.   :36.04   Max.   :49.54   Max.   :251.20   Max.   :4254.0   Max.   :0.22260   Max.   :1.05800
   concavity3        concave3         symmetry3         fractal3
 Min.   :0.0000   Min.   :0.00000   Min.   :0.1565   Min.   :0.05504
 1st Qu.:0.1145   1st Qu.:0.06493   1st Qu.:0.2504   1st Qu.:0.07146
 Median :0.2267   Median :0.09993   Median :0.2822   Median :0.08004
 Mean   :0.2722   Mean   :0.11461   Mean   :0.2901   Mean   :0.08395
 3rd Qu.:0.3829   3rd Qu.:0.16140   3rd Qu.:0.3179   3rd Qu.:0.09208
 Max.   :1.2520   Max.   :0.29100   Max.   :0.6638   Max.   :0.20750
```

Figure 3: Data Summary After Transformation

```
> nn$weights                # network weights after the last method iteration
[[1]]
[[1]][[1]]
             [,1]          [,2]       [,3]          [,4]
 [1,]   1.9074423   7.78650064 -6.6082663   4.83694341
 [2,]  12.8117077  13.20902912  1.6678648  -0.15888321
 [3,] -10.2432641   8.79024414  1.3764595   0.14387164
 [4,]   9.6660608  15.51861867  1.2664743  -0.62472115
 [5,]  -0.1301617   0.37224404  0.2101033  -0.76014465
 [6,]   3.3907647  -1.44332789 -0.4115495   0.17537518
 [7,]   0.5664195  -0.46342537  0.3438495   1.06924287
 [8,] -21.2814358  -5.98109214  4.4284605 -16.94765220
 [9,] -10.3761312  -9.87529301  8.7177471  -5.69321827
[10,]   6.7589297   1.52102795  1.2037852  -0.53931602
[11,]   0.2117318   0.03685771  0.5179615  -4.34931014
[12,] -22.7320911  -2.17269026  2.7062066  -3.49591667
[13,]   9.3729938   1.83009184 -2.6554510   5.80137551
[14,]  -2.4138911  -0.96475589  2.8405620  -0.61126807
[15,]  -8.6226372 -21.26617683 19.5085721 -21.81368689
[16,] -10.3229938  -1.61803891  0.4933762   0.79329606
[17,]  16.4946766  15.04217098 -5.2963741  11.37338301
[18,]   2.3881344   0.06002572 -0.2174305  -3.21596547
[19,]  -4.1386985  -1.06974358  4.9999537 -10.43294943
[20,]  -3.4518926   5.62279918 -1.0076931   0.03137404
```

Figure 4: Network Weights After the Last Iteration

```
> nn$result.matrix          # number of trainings steps, the error, and the weights
                                    [,1]
error                         6.566064707
reached.threshold             0.009782281
steps                       319.000000000
Intercept.to.1layhid1         1.907442256
radius.to.1layhid1           12.811707654
texture.to.1layhid1         -10.243264137
perimeter.to.1layhid1         9.666060782
area.to.1layhid1             -0.130161672
smoothness.to.1layhid1        3.390764655
compactness.to.1layhid1       0.566419515
concavity.to.1layhid1       -21.281435752
concave.to.1layhid1         -10.376131224
symmetry.to.1layhid1          6.758929667
fractal.to.1layhid1           0.211731840
radius2.to.1layhid1         -22.732091077
texture2.to.1layhid1          9.372993848
perimeter2.to.1layhid1       -2.413891138
area2.to.1layhid1            -8.622637198
```

Figure 5: Total Number of Training Steps