5) Artificial Neural Network

FOSSEE (R Team)

28 January, 2020

What is an Artificial Neural Network?

An **Artificial Neural Network or ANN** is a statistical model capable of modelling non-linear relationships between inputs and outputs in parallel. Biological neural networks inspire it. An ANN consists of multiple layers of artificial neurons which are used to receive input, apply an activation function on the input, and omit output. These layers are named as the input layer, hidden layer, and output layer, respectively.

Introduction

The purpose of this experiment is to create an artificial neural network on "iris" dataset to predict species depending upon the input. The ANN will consist of three layers named as the input, hidden, and output layer, respectively.

Procedure

Step by step procedure to conduct the required experiment -

1. Splitting the dataset for training and testing
2. Normalizing data
3. Labelling species
4. Creating ANN for predicting species
5. Predictions using ANN
6. Calculating the error of the generated model
7. Plotting Actual v/s Predicted Species

Note : Please make sure that the following package is already installed -

- *neuralnet*

Code and Results

Data used for analysis

```
## R has a predefined dataset with the name "iris"
  # To know more about the dataset type "?iris" in the console
```

Splitting the dataset for training and testing

```
# 1) Creating Training and Testing data from "iris" dataset
  # 1.1) Generating Random Numbers
  set.seed(100)
  # 1.2) Creating sample for splitting dataset
  Sample <- sample(nrow(iris),0.7*nrow(iris))
  # 1.3) Training dataset containing 70% of data
  Training <- iris[Sample,]
  # 1.4) Testing dataset containing 30% of data
  Testing <- iris[-Sample,]
```

Normalizing data

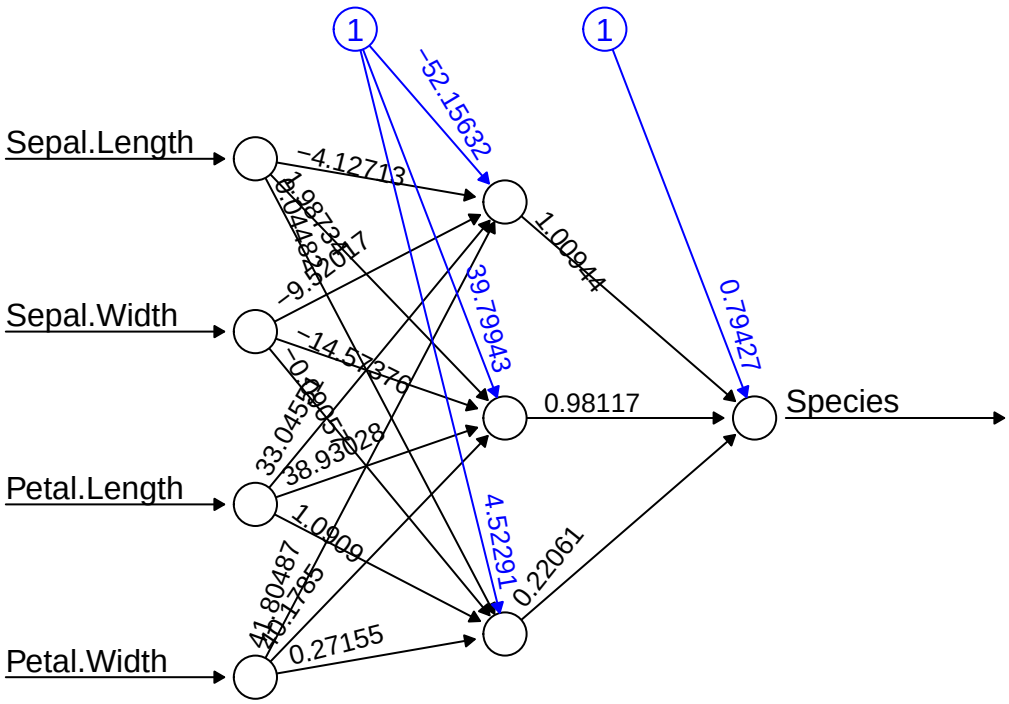
```
# 2) Normalizing data
  # 2.1) Normalizing training data
  Scaled_Training = Training
  Scaled_Training[,1:4] <- scale(Training[,1:4])
  # 2.2) Normalizing testing data
  Scaled_Testing = Testing
  Scaled_Testing[,1:4] <- scale(Testing[,1:4])
```

Labelling species

```
# 3) Labelling species
  # 3.1) Labelling species for training data
  Scaled_Training$Species <- as.integer(factor(Scaled_Training$Species))
  # 3.2) Labelling species for testing data
  Scaled_Testing$Species <- as.integer(factor(Scaled_Testing$Species))
```

Creating ANN for predicting species

```
# 4) Creating ANN and predicting values
  # 4.1) Fitting ANN
  # Run the following command after removing "#" if "neuralnet" library is not installed
  # install.packages("neuralnet")
  library(neuralnet)
  ANN = neuralnet(Species ~ Sepal.Length + Sepal.Width + Petal.Length +
        Petal.Width, Scaled_Training, hidden = 3, linear.output = T)
  # 4.2) Plotting ANN
  plot(ANN, rep = "best")
```



Error: 0.015972 Steps: 3101

Predictions using ANN

```
# 5) Predictions using ANN
  predict = compute(ANN, Scaled_Testing[,c(1:4)])
```

Calculating the error of the generated model

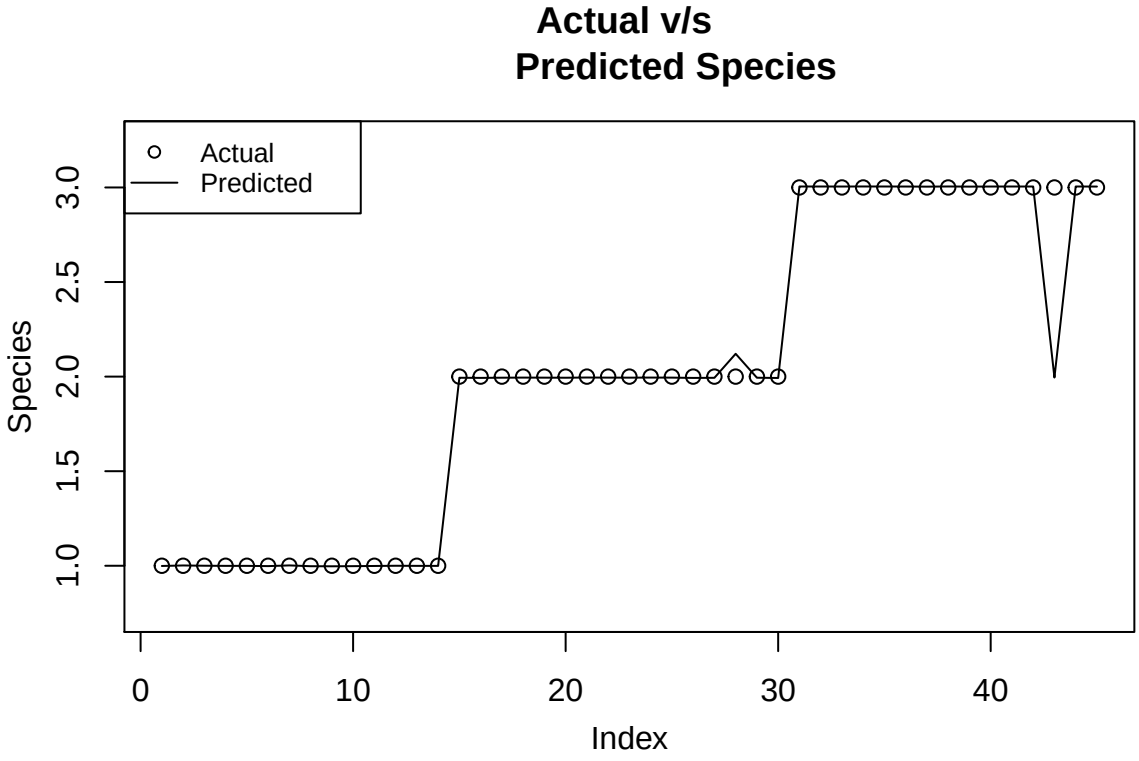
```
# 6) Calculate Root Mean Square Error (RMSE)
  RMSE = (sum((as.numeric(predict$net.result - Scaled_Testing$Species))^2)/
         nrow(Scaled_Testing)) ^ 0.5
  paste0("Root Mean Square Error is ", RMSE)
```

```
## [1] "Root Mean Square Error is 0.150944658699485"
```

Plotting Actual v/s Predicted Species

```
# 7) Plotting actual values v/s predicted values
  plot(as.integer(factor(Testing$Species)),main = "Actual v/s
       Predicted Species",ylab = "",xlab = "",ylim = c(0.75,3.25))
```

```
# 7.1) Adding a line to represent the predicted values
lines(predict$net.result,lty=1)
# 7.2) Clipping legend to the figure region instead of the plot region
par(xpd=TRUE)
# 7.3) Adding a legend
legend("topleft",text.width = 7,legend=c("Actual","Predicted"),
       pch = c(1,NA) ,lty = c(NA,1) ,cex=0.8)
# 7.4) Setting and adjusting labels of x and y axes
title(ylab="Species",xlab="Index", line=2.25, cex.lab=1)
```



Conclusion

We have successfully created an ANN with three layers as expected, over the "iris" dataset. The root mean square error over the predicted values and test dataset is small for the generated neural network. Hence, the predictions made by the ANN are mostly accurate. We have also plotted a graph containing the actual values of the test dataset and the predictions made using the same. We can observe the accuracy of the generated ANN from its effectiveness in predicting correct results.