

**M.Sc. DATA ANALYTICS 2018-2019**

**MACHINE LEARNING**

**CA - 1**

**A large brick building

Description automatically generated**

**Boston Housing Price Prediction**

**using**

**Neural Network**

**In**

**R Programming Language**

**BY:**

**Poonam Dhoot – 10399137@mydbs.ie**

**Sunmeet Thapar – 10506082@mydbs.ie**

**Machine Learning CA 1 – R Programming**

**Prepared by:**

|  |  |
| --- | --- |
| Student Name(s) | Student ID number (email) |
| 1. Poonam Dhoot | 10399137@mydbs.ie |
| 1. Sunmeet Thapar | 10506082@mydbs.ie |

**Topic: Artificial Neural Network(ANN)**

Artificial neural networks are one of the main tools used in machine learning. As the “neural” part of their name suggests, they are brain-inspired systems which are intended to replicate the way that we humans learn. Neural networks consist of input and output layers, as well as (in most cases) a hidden layer consisting of units that transform the input into something that the output layer can use. They are excellent tools for finding patterns which are far too complex or numerous for a human programmer to extract and teach the machine to recognize.

While neural networks (also called “perceptrons”) have been around since the 1940s, it is only in the last several decades where they have become a major part of artificial intelligence. This is due to the arrival of a technique called “backpropagation,” which allows networks to adjust their hidden layers of neurons in situations where the outcome doesn’t match what the creator is hoping for — like a network designed to recognize dogs, which misidentifies a cat, for example.

Another important advance has been the arrival of deep learning neural networks, in which different layers of a multilayer network extract different features until it can recognize what it is looking for.

**Brief Explanation of the dataset:**

1. Boston Housing Price Prediction – It is a very famous dataset available in package MASS in RStudio
2. In this project, we will evaluate the performance and predictive power of a model that has been trained and tested on data collected from homes in suburbs of Boston, Massachusetts. A model trained on this data that is seen as a good fit could then be used to make certain predictions about a home — in particular, its monetary value. This model would prove to be invaluable for someone like a real estate agent who could make use of such information on a daily basis.
3. It is Regression type of prediction. Since the values are numerical.

**R Code:**

#Boston Housing Price Prediction using Neural Network

#install.packages("neuralnet")

library(Metrics)

library(MASS)

data <- Boston

# Set a seed

set.seed(1000)

# Check that no data is missing

apply(data,2,function(x) sum(is.na(x)))

# Train-test random splitting

index <- sample(1:nrow(data),round(0.75\*nrow(data)))

train <- data[index,]

test <- data[-index,]

# Neural net fitting

# Scaling data for the NN

maxs <- apply(data, 2, max)

maxs

mins <- apply(data, 2, min)

mins

scaled <- as.data.frame(scale(data, center = mins, scale = maxs - mins))

head(scaled)

# Train-test split

train\_set <- scaled[index,]

test\_set <- scaled[-index,]

# NN training

library(neuralnet)

n <- names(train\_set)

f <- as.formula(paste("medv ~", paste(n[!n %in% "medv"], collapse = " + ")))

nn <- neuralnet(f,data=train\_set,hidden=c(5,3),linear.output=T, act.fct = 'logistic')

# Predict

pr.nn <- compute(nn,test\_set[,1:13])

# Results from NN are normalized (scaled)

# Descaling for comparison

pr.nn\_ <- pr.nn$net.result\*(max(data$medv)-min(data$medv))+min(data$medv)

test.r <- (test\_$medv)\*(max(data$medv)-min(data$medv))+min(data$medv)

# Calculating MSE

MSE.nn<-mse(test.r, pr.nn\_)

MSE.nn

# Improve Performance - Change Parameter

nn\_new <- neuralnet(f,data=train\_set,hidden=c(10,5,3),linear.output=T, act.fct = 'logistic')

pr.nn\_new <- compute(nn\_new,test\_set[,1:13])

pr.nn\_new <- pr.nn\_new$net.result\*(max(data$medv)-min(data$medv))+min(data$medv)

# Re-calculate MSE

MSE.nn\_new<-mse(test.r, pr.nn\_new)

MSE.nn\_new

# Compare Performance

print(paste('Old MSE :',MSE.nn, 'New MSE :', MSE.nn\_new))

# Plot predictions - Old MSE

plot(test$medv,pr.nn\_,col='red',main='Real vs predicted NN',pch=18,cex=0.7, xlab = "Actual", ylab = "Predicted")

abline(0,1,lwd=2)

legend('bottomright',legend='NN',pch=18,col='red', bty='n')

# Plot predictions - New MSE

plot(test$medv,pr.nn\_new,col='red',main='Real vs predicted NN',pch=18,cex=0.7, xlab = "Actual", ylab = "Predicted")

abline(0,1,lwd=2)

legend('bottomright',legend='NN',pch=18,col='red', bty='n')

**Attachments:**

Boston Housing using Neural Network

1. R code
2. RMD code file
3. R code in Notepad
4. HTML file
5. PowerPoint Presentation