INFORMATICS INSTITUTE OF TECHNOLOGY

In collaboration with

UNIVERSITY OF WESTMINSTER

BEng (Hons) Software Engineering

Machine Learning and Data Mining

5DATA001C.2

Course Work Report

Name- Ravihara De Silva

UoW ID-w1913374

IIT ID- 20211620

Partitioning Clustering Objective

Load the Dataset:

```
vehicle_dataset <- read_excel("vehicles.xlsx")
vehicle_dataset <- vehicle_dataset[,-20]
```

Initially load the vehicle dataset with the help of read_excel method. This method reads the excel file and converts the data into the dataframe.

Question 1:

Preprocessing Dataset (Removing Outliers and Scaling)?

Our first task is to remove the outliers and scaling the Dataset. The advantage of the removing outlier and scaling, in dataset we have multiple features and each features hold the different

Range of data scaling converts the data into common range like (0-1) and remove outliers help to remove the data in the features because in the feature some data is not useful, and we remove this data.

```
for (i in 1:length(vehicle_dataset)) {
  # Calculating the q1/q3
  q1 <- quantile(vehicle_dataset[,i] , 0.25)
  q3 <- quantile(vehicle_dataset[,i] , 0.75)

  # calculate the IQR and Lower and Upper Values
  iqr <- q3 - q1
  lowerValue <- q1 - 1.5 * iqr
  upperValue <- q3 + 1.5 * iqr

  # Find out the Outliers
  outliers <- vehicle_dataset[,i] < lowerValue | vehicle_dataset[,i] > upperValue

  # Remove the Outliers from Dataset
  vehicle_dataset  <- vehicle_dataset[!outliers,]
}
```

In this figure, I am removing the outliers in the dataset, and I have used the Quartile and inter-quartiles range technique.

```
vehicle_dataset <- scale(vehicle_dataset)
```

In this figure, I am converting the data into the common range. In which I have used the scale method.

Question No 2:

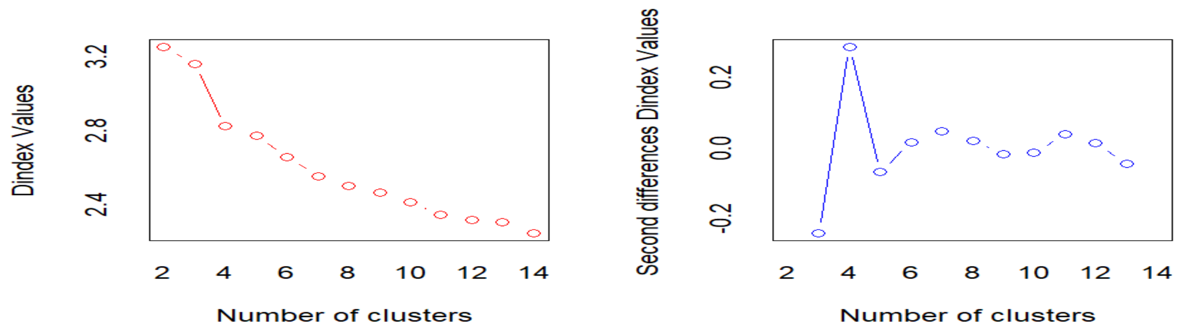
Find out Optimal Number of Clusters of Each Method?

Second task is to determine the number of clusters centers, in this task we have multiple type of methods like (NBclust, Elbow, Gap Statistic and silhouette). In this task I have displayed the Code of each method and displayed the best optimal cluster by using graphs too.

Method NBclust:

```
nb <- NbClust(vehicle_dataset , distance = "euclidean", min.nc = 2, max.nc = 14, method = "kmeans")
k <- nb$Best.nc[1]
```

In this Figure, I have used the NBclust method to determine the optimal number of clusters.

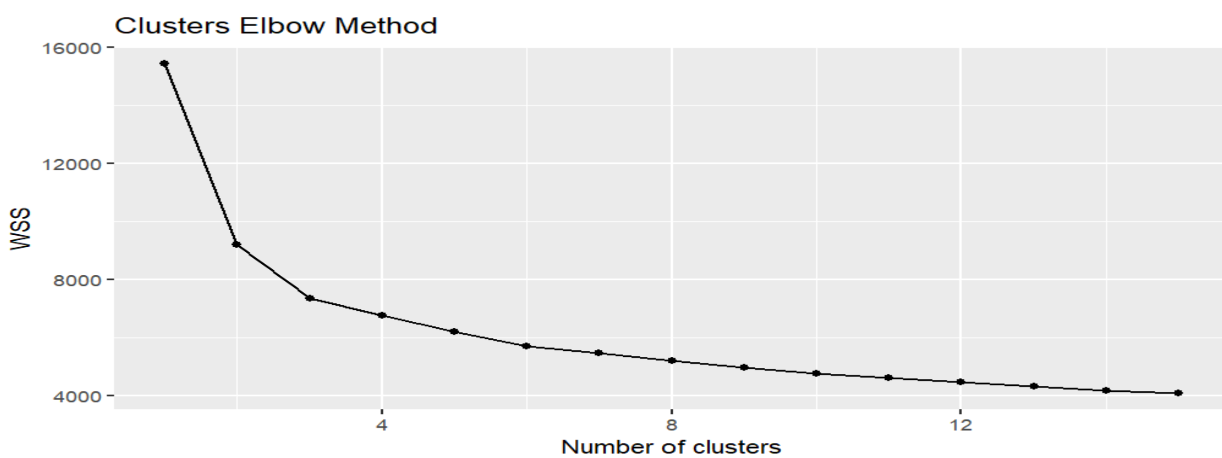


This is the graphical representation to describe the optimal number of clusters for NBclust method.

Method Elbow:

```
elbow <- sapply(1:15, function(k){kmeans(vehicle_dataset, k, nstart=10, iter.max = 150)$tot.withinss})
ggplot(data.frame(k=1:15, WSS=elbow), aes(x=k, y=WSS)) + geom_line() + geom_point() +
  labs(x = "Number of clusters", y = "WSS") + ggtitle("Clusters Elbow Method")
```

In this Figure, I have used the Elbow method to determine the optimal number of clusters.



This is the graphical representation to describe the optimal number of clusters for Elbow method.

Method Gap Statistics:

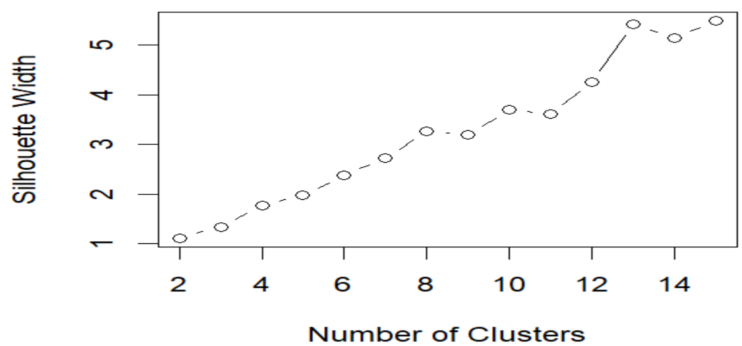
```
Gap <- clusGap(vehicle_dataset, FUN = kmeans, nstart = 20, K.max = 15, B = 50)
k <- maxSE(Gap$Tab[, "gap"], Gap$Tab[, "SE.sim"], method = "Tibs2001SEmax")
```

In this Figure, I have used the Gap Statistics method to determine the optimal number of clusters.

Method silhouette:

```
Matrix <- as.matrix(vehicle_dataset)
width <- sapply(2:15, function(k){
  kmeans_w <- kmeans(Matrix, k)
  avg <- mean(silhouette(kmeans_w$cluster, dist(Matrix)))
})
plot(2:15, width, type="b", xlab="Number of Clusters", ylab="Silhouette Width")
```

In this Figure, I have used the Silhouette method to determine the optimal number of clusters.



This is the graphical representation to describe the optimal number of clusters for Silhouette method.

Question No 3:

Kmeans Clustering Appling each Method?

In this task, our main goal is to apply the Kmeans clustering on each method and generate the results. Each method generates the K values and Kmeans clustering must use this k value and generate the results on the Dataset and calculating the Results like (BSS, WSS, TSS and Ratio of the BSS/TSS)

Method NBclust:

```
NB_kmeans <- kmeans(vehicle_dataset, centers = k)
```

In this Figure, I have applied the Kmeans clustering with NBclust method and generate the results on the Dataset.

```
> cat("Best Cluster for NB Clust-->",k,"\n")
Best Cluster for NB Clust--> 12
> cat("NB Clust BSS-->", BSS,"\n")
NB Clust BSS--> 158.0612
> cat("NB Clust WSS-->", WSS,"\n")
NB Clust WSS--> 4494.226
> cat("NB Clust TSS-->", TSS,"\n")
NB Clust TSS--> 15447
> cat("NB Clust BCC TSS Ratio-->", BSS_TSS_ratio,"\n")
NB Clust BCC TSS Ratio--> 0.01023249
```

In this figure, I have calculated all evaluation and displayed the results.

Method Elbow:

```
Elbow_Kmeans <- kmeans(vehicle_dataset, k, nstart=10, iter.max=300)
```

In this Figure, I have applied the Kmeans clustering with Elbow method and generate the results on the Dataset.

```
> cat("Best Cluster for Elbow-->",k,"\n")
Best Cluster for Elbow--> 2
> cat("Elbow BSS-->", BSS,"\n")
Elbow BSS--> 27720.6
> cat("Elbow Wss-->", WSS,"\n")
Elbow Wss--> -12273.6
> cat("Elbow TSS-->", TSS,"\n")
Elbow TSS--> 15447
> cat("Elbow BCC TSS Ratio-->", BSS_TSS_ratio,"\n")
Elbow BCC TSS Ratio--> 1.794562
```

In this figure, I have calculated all evaluation and displayed the results.

Method Gap Statistics:

```
Gap_Kmeans <- kmeans(vehicle_dataset, centers = k, nstart = 25)
```

In this Figure, I have applied the Kmeans clustering with Gap Statistics method and generate the results on the Dataset.

```
> cat("Best Cluster for GAP-->",k,"\n")
Best Cluster for GAP--> 3
> cat("GAP BSS-->", BSS,"\n")
GAP BSS--> 8073.767
> cat("GAP Wss-->", WSS,"\n")
GAP Wss--> -700.5346
> cat("GAP BSS Ratio-->", BSS_ratio,"\n")
GAP BSS Ratio--> 0.5226754
> cat("GAP WSS Ratio-->", WSS_ratio,"\n")
GAP WSS Ratio--> -0.04535085
```

In this figure, I have calculated all evaluation and displayed the results.

Method silhouette:

```
silhouette_Kmeans<- kmeans(vehicle_dataset, k, nstart = 20)
```

In this Figure, I have applied the Kmeans clustering with Silhouette method and generate the results on the Dataset.

```
> cat("Best Cluster for Silhouette -->",k,"\n")
Best Cluster for Silhouette --> 3
> cat("Silhouette  BSS-->", BSS,"\n")
Silhouette  BSS--> 27444.25
> cat("Silhouette  Wss-->", WSS,"\n")
Silhouette  Wss--> 7373.233
> cat("Silhouette BSS and TSS Ratio-->", BSS_TSS_ratio,"\n")
Silhouette BSS and TSS Ratio--> 0.7882319
```

In this figure, I have calculated all evaluation and displayed the results.

Question No 4:

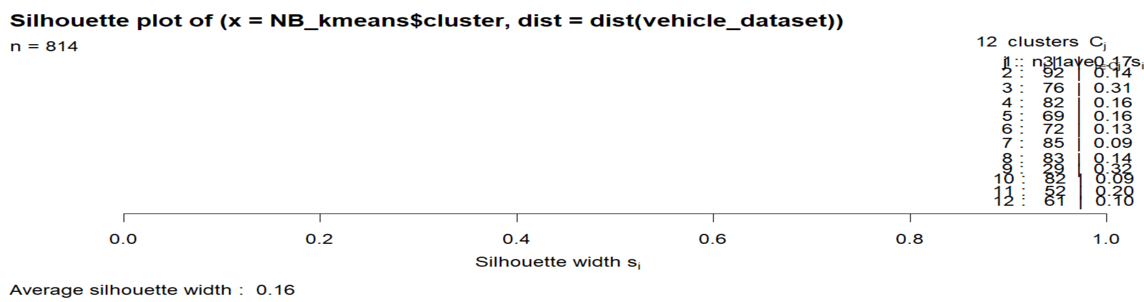
silhouette plot?

Silhouette plots generate the graphs and display the results of the Kmeans cluster and generate the average silhouette width with the help of average score of the cluster. In the silhouette plot accept the only basics two parameters like (value of k and the dataset) with the base of these data silhouette plot display the graph.

<u>Method NBclust:</u>

```
plot(silhouette(NB_kmeans$cluster, dist(vehicle_dataset)))
```

In this Figure, I have applied the Silhouette plots for NBclust Method.

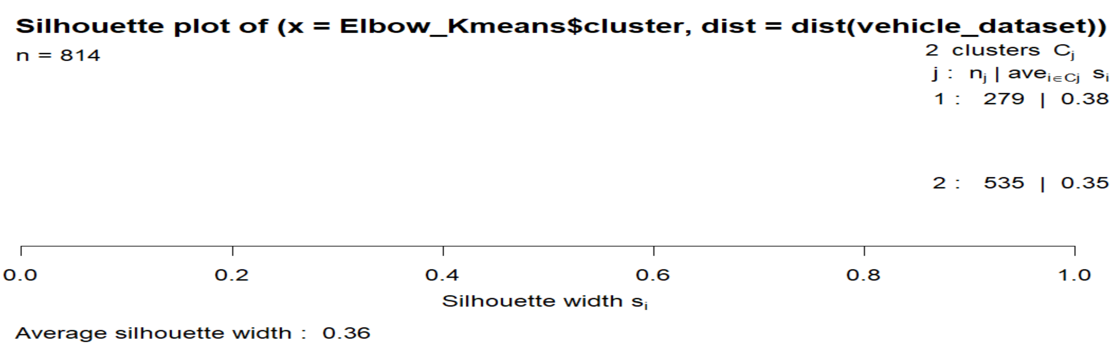


This is the graphically representation of Silhouette plot for NBclust method. In the Silhouette plot display the results of the Cluster.

<u>Method Elbow:</u>

```
plot(silhouette(Elbow_Kmeans$cluster, dist(vehicle_dataset)))
```

In this Figure, I have applied the Silhouette plots for Elbow Method.

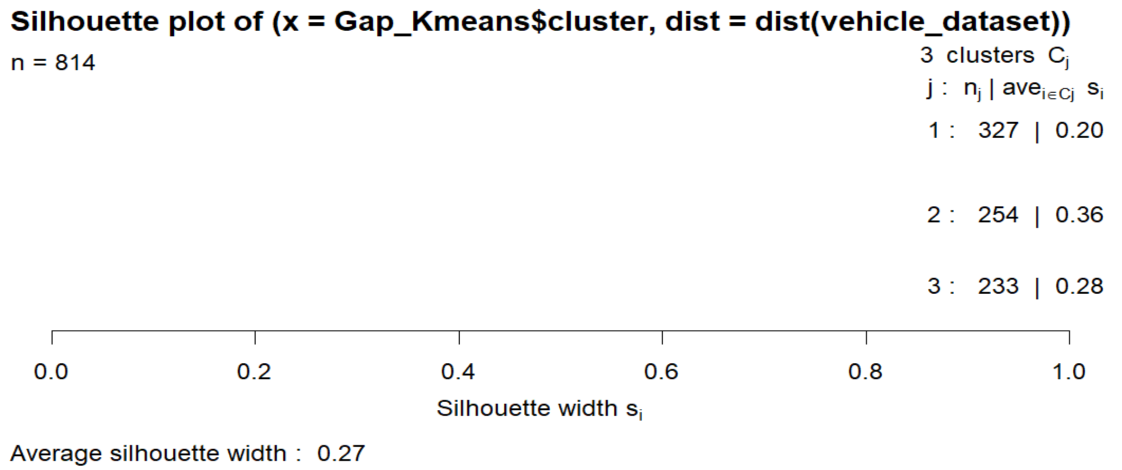


This is the graphically representation of Silhouette plot for Elbow method. In the Silhouette plot display the results of the Cluster.

<u>Method Gap Statistics:</u>

```
plot(silhouette(Gap_Kmeans$cluster, dist(vehicle_dataset)))
```

In this Figure, I have applied the Silhouette plots for Gap Statistics Method.

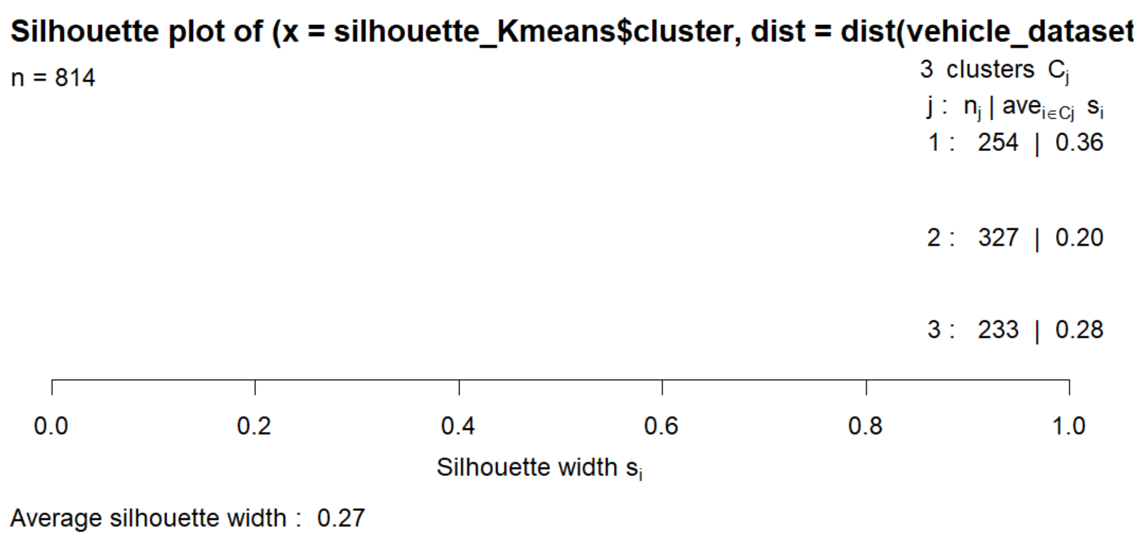


This is the graphically representation of Silhouette plot for Gap Statistics method. In the Silhouette plot display the results of the Cluster.

Method silhouette:

```
plot(silhouette(silhouette_Kmeans$cluster, dist(vehicle_dataset)))
```

In this Figure, I have applied the Silhouette plots for silhouette Method.



This is the graphically representation of Silhouette plot for Silhouette method. In the Silhouette plot display the results of the Cluster.

Question No 5:

PCA based New Dataset?

In this task, our main goal is to Apply the PCA in the Dataset and generate the new dataset with less features. Basically, PCA is a (Principal Component Analysis) this method converts the dataset into the useful dataset. This method accepts the dataset and generate the useful features with the help of cumulative score.

```
pca_dataset <- PCA(vehicle_dataset, graph = FALSE)

# Create a Attribute of Databaset with PCA
data_trans <- as.data.frame(pca_dataset$ind$coord)
actul_data <- prcomp(data_trans)

# Display the Eign Vectors/Values
cat(pca_dataset$eig)
cat(pca_dataset$var$coord)

# Create the New Dataset with PCA Analysis # Cumulatice Score
cumulative_scores <- cumsum(actul_data$sdev^2 / sum(actul_data$sdev^2))
new_dataset <- data.frame(actul_data$x[, cumulative_scores > 0.92])
```

In this figure, I have applied the PCA in the Dataset. PCA return the new dataset now we applied the cumulative score to generate the Dataset also base on the cumulative score after tha our dataset is ready to the further wrok.

Question No 6:

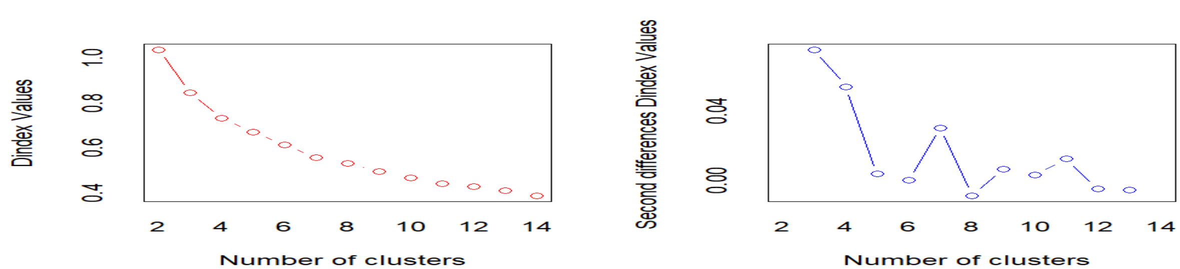
Determine the Number of Clusters of Each Method?

This task is to determine the number of clusters centers with PCA Dataset, in this task we have multiple type of methods like (NBclust, Elbow, Gap Statistic and silhouette). In this task I have displayed the Code of each method and displayed the best optimal cluster by using graphs too.

<u>Method NBclust:</u>

```
nb <- NbClust(new_dataset , distance = "euclidean", min.nc = 2, max.nc = 14, method = "kmeans")
k <- nb$Best.nc[1]
```

In this Figure, I have used the NBclust method to determine the optimal number of clusters with PCA Dataset.

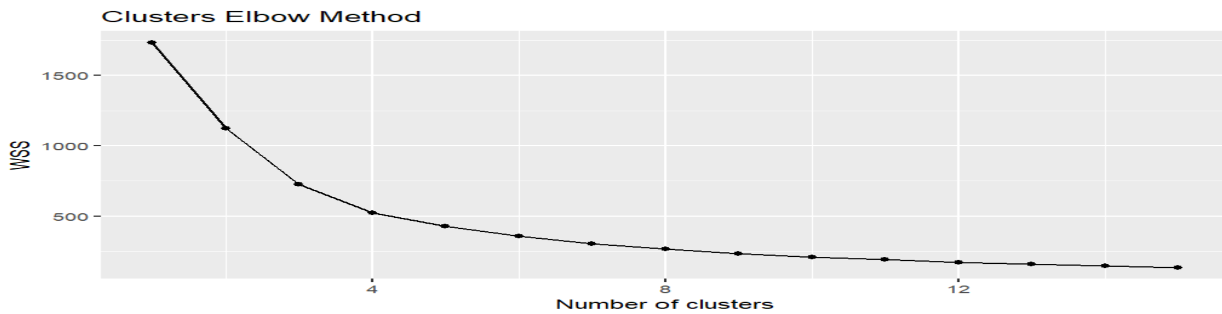


This is the graphical representation to describe the optimal number of clusters for NBclust method with PCA Dataset.

<u>Method Elbow:</u>

```
elbow <- sapply(1:15, function(k){kmeans(new_dataset, k, nstart=10, iter.max = 150)$tot.withinss})
ggplot(data.frame(k=1:15, WSS=elbow), aes(x=k, y=WSS)) + geom_line() + geom_point() +
  labs(x = "Number of clusters", y = "WSS") + ggtitle("Clusters Elbow Method")
```

In this Figure, I have used the Elbow method to determine the optimal number of clusters with PCA Dataset.



This is the graphical representation to describe the optimal number of clusters for Elbow method with PCA Dataset.

<u>Method Gap Statistics:</u>

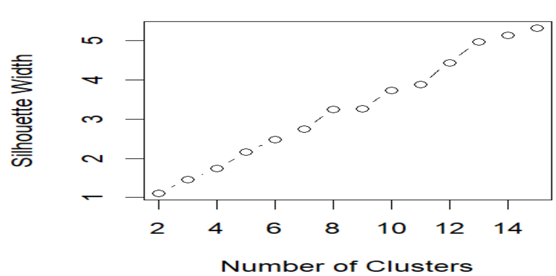
```
Gap <- clusGap(new_dataset, FUN = kmeans, nstart = 20, K.max = 15, B = 50)
k <- maxSE(Gap$Tab[, "gap"], Gap$Tab[, "SE.sim"], method = "Tibs2001SEmax")
```

In this Figure, I have used the Gap Statistics method to determine the optimal number of clusters with PCA Dataset.

<u>Method silhouette:</u>

```
Matrix <- as.matrix(new_dataset)
width <- sapply(2:15, function(k){
  kmeans_w <- kmeans(Matrix, k)
  avg <- mean(silhouette(kmeans_w$cluster, dist(Matrix)))
})
plot(2:15, width, type="b", xlab="Number of Clusters", ylab="Silhouette Width")
```

In this Figure, I have used the Silhouette method to determine the optimal number of clusters with PCA Dataset.



This is the graphical representation to describe the optimal number of clusters for Silhouette method with PCA Dataset.

Question No 7:

Kmeans Clusters Appling each Method?

In this task, our main goal is to apply the Kmeans clustering on each method and generate the results in the PCA Dataset. Each method generates the K values and Kmeans clustering must use this k value and generate the results on the Dataset and calculating the Results like (BSS, WSS, TSS and Ratio of the BSS/TSS)

Method NBclust:

```
NB_kmeans <- kmeans(new_dataset, centers = k)
```

In this Figure, I have applied the Kmeans clustering with Gap Statistics method and generate the results on the PCA Dataset.

```
> cat("Best Cluster for NB Clust-->",k,"\n")
Best Cluster for NB Clust--> 2
> cat("NB Clust BSS-->", BSS,"\n")
NB Clust BSS--> 2.662155
> cat("NB Clust Wss-->", WSS,"\n")
NB Clust Wss--> 1130.948
> cat("NB Clust TSS-->", TSS,"\n")
NB Clust TSS--> 1733.201
> cat("NB Clust BCC TSS Ratio-->", BSS_TSS_ratio,"\n")
NB Clust BCC TSS Ratio--> 0.001535976
```

In this figure, I have calculated all evaluation and displayed the results.

Method Elbow:

```
Elbow_Kmeans <- kmeans(new_dataset, k, nstart=10, iter.max=300)
```

In this Figure, I have applied the Kmeans clustering with Gap Statistics method and generate the results on the PCA Dataset.

```
> cat("Best Cluster for Elbow-->",k,"\n")
Best Cluster for Elbow--> 2
> cat("Elbow BSS-->", BSS,"\n")
Elbow BSS--> 2446.628
> cat("Elbow Wss-->", WSS,"\n")
Elbow Wss--> -713.4272
> cat("Elbow TSS-->", TSS,"\n")
Elbow TSS--> 1733.201
> cat("Elbow BCC TSS Ratio-->", BSS_TSS_ratio,"\n")
Elbow BCC TSS Ratio--> 1.411624
```

In this figure, I have calculated all evaluation and displayed the results.

Method Gap Statistics:

```
Gap_Kmeans <- kmeans(new_dataset, centers = k+3, nstart = 25)
```

In this Figure, I have applied the Kmeans clustering with Gap Statistics method and generate the results on the PCA Dataset.

```
> cat("Best Cluster for GAP-->",k,"\n")
Best Cluster for GAP--> 1
> cat("GAP BSS-->", BSS,"\n")
GAP BSS--> 1207.042
> cat("GAP WSS-->", WSS,"\n")
GAP WSS--> -680.883
> cat("GAP BSS Ratio-->", BSS_ratio,"\n")
GAP BSS Ratio--> 0.6964235
> cat("GAP WSS Ratio-->", WSS_ratio,"\n")
GAP WSS Ratio--> -0.392847
```

In this figure, I have calculated all evaluation and displayed the results.

<u>Method silhouette:</u>

```
silhouette_Kmeans<- kmeans(new_dataset, k, nstart = 20)
```

In this Figure, I have applied the Kmeans clustering with Gap Statistics method and generate the results on the PCA Dataset.

```
> cat("Best Cluster for Silhouette -->",k,"\n")
Best Cluster for Silhouette --> 4
> cat("Silhouette  BSS-->", BSS,"\n")
Silhouette  BSS--> 5257.135
> cat("Silhouette  WSS-->", WSS,"\n")
Silhouette  WSS--> 526.1698
> cat("Silhouette BSS and TSS Ratio-->", BSS_TSS_ratio,"\n")
Silhouette BSS and TSS Ratio--> 0.9090192
```

In this figure, I have calculated all evaluation and displayed the results.

Question No 8:

silhouette plot?

Silhouette plots generate the graphs and display the results of the Kmeans cluster in the PCA Dataset and generate the average silhouette width with the help of average score of the cluster. In the silhouette plot accept the only basics two parameters like (value of k and the dataset) with the base of these data silhouette plot display the graph.

<u>Method NBclust:</u>

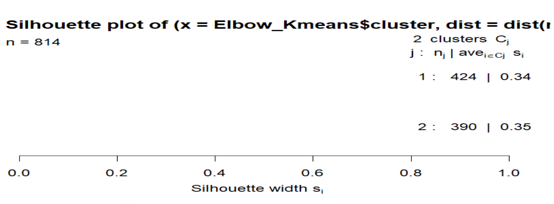
```
plot(silhouette(NB_kmeans$clsuster, dist(new_dataset)))
```

In this Figure, I have applied the Silhouette plots for NBclust Method in the PCA Dataset.

<u>Method Elbow:</u>

```
plot(silhouette(Elbow_Kmeans$cluster, dist(new_dataset)))
```

In this Figure, I have applied the Silhouette plots for Elbow Method in the PCA Dataset.

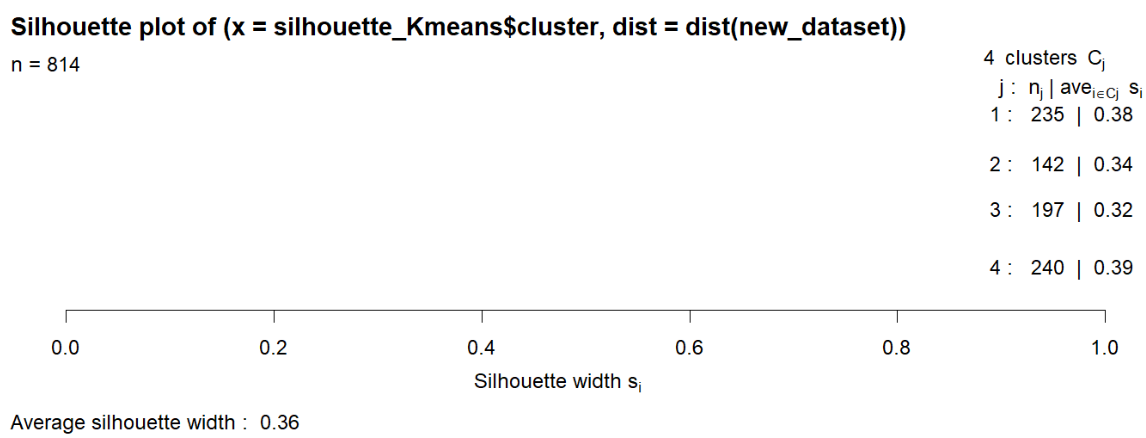


This is the graphically representation of Silhouette plot for Elbow method. In the Silhouette plot display the results of the Cluster.

Method Gap Statistics:

```
plot(silhouette(silhouette_Kmeans$cluster, dist(new_dataset)))
```

In this Figure, I have applied the Silhouette plots for Gap Statistics Method in the PCA Dataset.

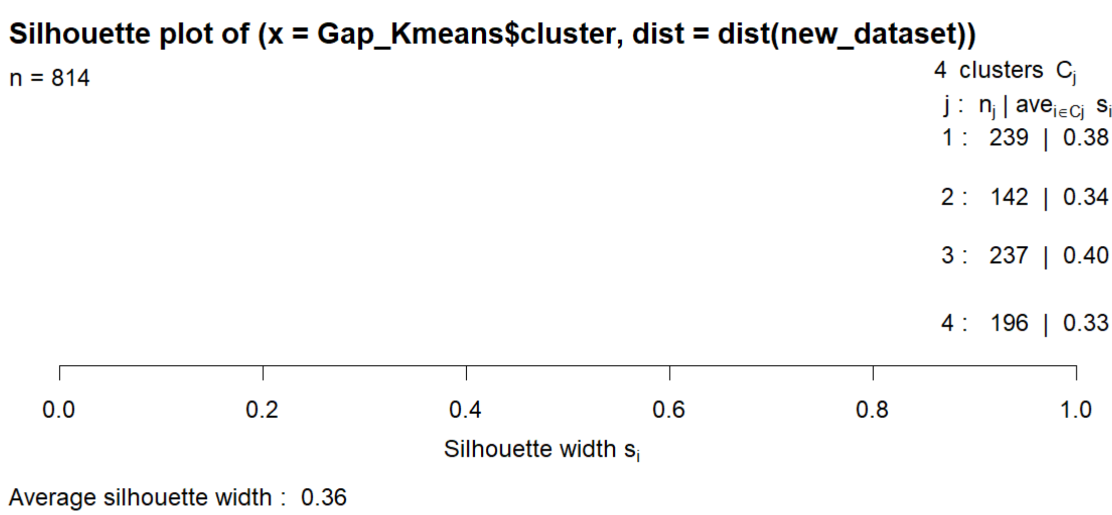


This is the graphically representation of Silhouette plot for Gap Statistics method. In the Silhouette plot display the results of the Cluster.

Method silhouette:

```
plot(silhouette(Gap_Kmeans$cluster, dist(new_dataset)))
```

In this Figure, I have applied the Silhouette plots for Silhouette Method in the PCA Dataset.



This is the graphically representation of Silhouette plot for Silhouette method. In the Silhouette plot display the results of the Cluster.

Question No 9:

Calinski-Harabasz index?

Calinski-Harabasz Index also known as (Variance Ratio Criterion) is the method with the help of we evaluate the cluster model. In which we used the predicted and Actual data. Calinski-Harabasz index evaluate the model with predicted and actual data and return the number of the index and this method validate the cluster by quantities and the features of the Dataset.

<u>Method NBclust:</u>

```
NB_index <- (SSB / (k - 1)) / (SSW / (nrow(new_dataset) - k))
```

In this figure, I have generated the Calinski-Harabasz Index of the PCA Dataset.

```
> NB_index
[1] 872.4591
```

This is the value of Calinski-Harabasz Index of the NBclust method.

<u>Method Elbow:</u>

```
Elbow_index <- (SSB / (k - 1)) / (SSW / (nrow(new_dataset) - k))
```

In this figure, I have generated the Calinski-Harabasz Index of the PCA Dataset.

```
> Elbow_index
[1] 886.3835
```

This is the value of Calinski-Harabasz Index of the Elbow method.

<u>Method Gap Statistics:</u>

```
Gap_index <- (SSB / (k - 1)) / (SSW / (nrow(new_dataset) - k))
```

In this figure, I have generated the Calinski-Harabasz Index of the PCA Dataset.

```
> gap_index
[1] 500.9017
```

This is the value of Calinski-Harabasz Index of the Gap Statistics method.

<u>Method silhouette:</u>

```
silhouette_index <- (SSB / (k - 1)) / (SSW / (nrow(new_dataset) - k))
```

In this figure, I have generated the Calinski-Harabasz Index of the PCA Dataset.

```
> silhouette_index
[1] 2653.588
```

This is the value of Calinski-Harabasz Index of the Silhouette method

Forecasting Problem

Question No 1:

Input/Output vector for forecasting problem?

There are many methods to generate the Input/output matrix for the Forecasting problem. In this project I have used electricity dataset and this problem is very famous problem in the energy sector. That's why I have generated the good result of this forecasting Problem.

Calendar based Feature:

In this method we store the data in the daily basis like holiday day, public holiday days and special day these data very useful for when we create the input vector for the Electricity dataset these data very helpful when we generate the accuracy of the model.

Seasonal based Variables:

In this method we store the data of the electricity base on the season like (winter season and summer season) based on these seasons we generate the interested pattern in the Electricity dataset, and we generate the Input Matrix based on this method.

Time Base Variables:

In this method we store the data of the electricity based on the time like (hours data, day data, weeks data. Months data and years data) based on these data we generate the input vector for the Forecasting Problem like (Electricity Problem).

Question No 2:

AR and NARX NN Models Input and Output Matrix?

AR NN Model Input and Output Matrix:

```
input_matrix <- matrix(NA, nrow = nrow(new_dataset) - time_delaye, ncol = time_delaye)
output_matrix <- as.matrix(new_dataset[(time_delaye + 1):nrow(new_dataset),])

for (i in 1:time_delaye) {
  input_matrix[,i] <- as.matrix(new_dataset[(time_delaye - i + 1):(nrow(new_dataset) - i),1])
}
```

In this figure, I have generated the Input and output matrix for the Electricity Dataset and this problem is forecasting and used AR NN Model. In this problem I have created the input output matrix with the previous time data and in this task, I have used only one feature of the electricity dataset.

NARX NN Model Input and Output Matrix:

```
Ninput_matrix <- as.matrix(new_dataset[1:465,c(1,2)])
output_matrix <- as.matrix(new_dataset[,c(3)])

# concatenation Current input matrix with Previous input matrix
input_matrix <- cbind(input_matrix , Ninput_matrix)
```

In this figure, I have generated the Input and output matrix for the Electricity Dataset and this problem is forecasting and use the NARX NN Model. In this problem I have created the input output matrix with the previous time data and in this task, I have used only one feature of the electricity dataset.

Question No 3:

Normalization of AR and NARX NN Models?

Normalization is the most popular technique; we know that our dataset is the combination of multiple features and each features hold multiple data with different range. Basically, Normalization converts each features data into the common range like (0-1). After normalizing each feature data than we are splitting the dataset into Training and Testing part.

Normalization of AR Input and Output Matrix:

```
input_matrix  <- apply(input_matrix , 2, function(x) (x - min(x)) / (max(x) - min(x)))
output_matrix <- apply(output_matrix, 2, function(x) (x - min(x)) / (max(x) - min(x)))
```

In this figure, I have converted the data into the common range with the help of normalization.

Normalization of NARX Input and Output Matrix:

```
Ninput_matrix  <- apply(Ninput_matrix , 2, function(x) (x - min(x)) / (max(x) - min(x)))
output_matrix <- apply(output_matrix, 2, function(x) (x - min(x)) / (max(x) - min(x)))
```

In this figure, I have converted the data into the common range with the help of normalization.

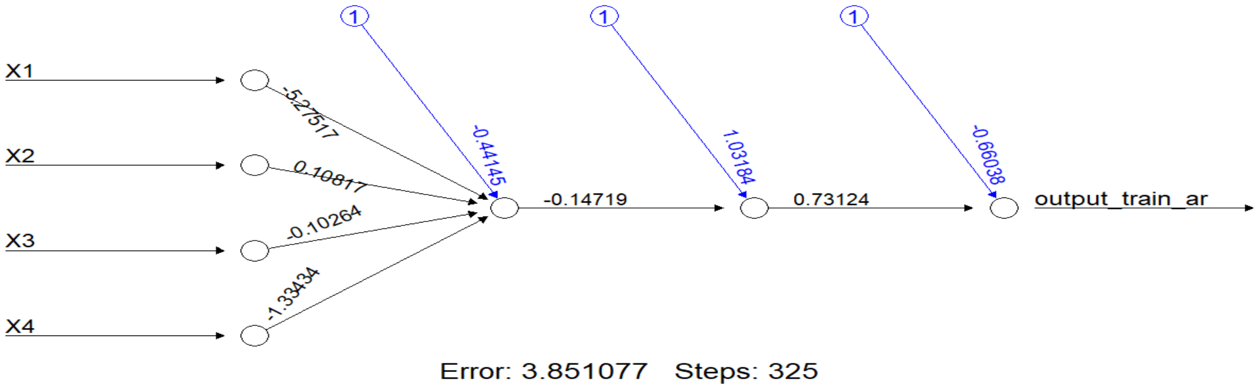
Question No 4:

AR NN Model Implementation with 12 Models?

In this task, I have generated the 12 models with the AR NN Model Approach with the setting the parameters of AR NN Model and each model display the Computational graphs and summary of the AR NN Model too.

Model No 1:

In the This Graph, I have showed you how this model is predicting.

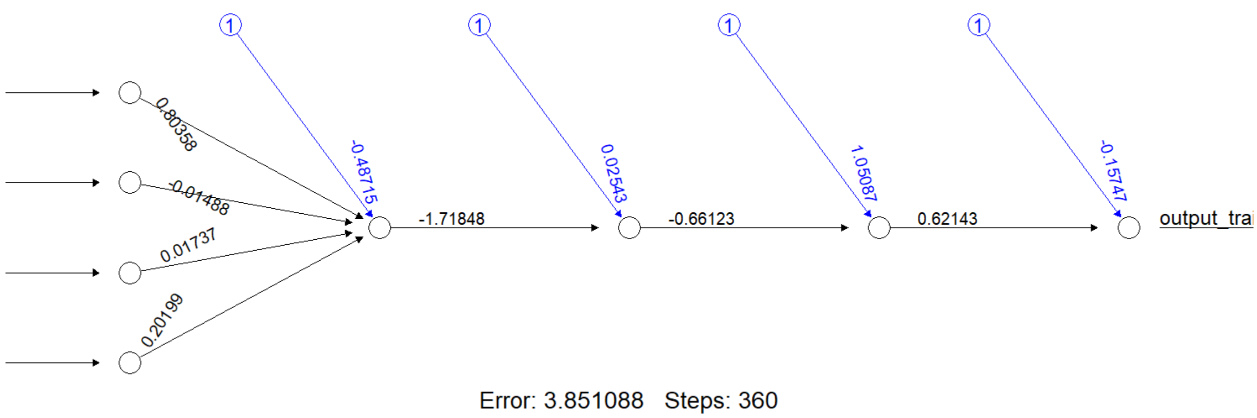


Here is the Summary of this Model.

```
                     Length Class      Mode
call                   6    -none-     call
response             376    -none-     numeric
covariate           1504    -none-     numeric
model.list             2    -none-     list
err.fct                1    -none-     function
act.fct                1    -none-     function
linear.output          1    -none-     logical
data                   5    data.frame list
exclude                0    -none-     NULL
net.result             1    -none-     list
weights                1    -none-     list
generalized.weights    1    -none-     list
startweights           1    -none-     list
result.matrix         12    -none-     numeric
```

Model No 2:

In the This Graph, I have showed you how this model is predicting.

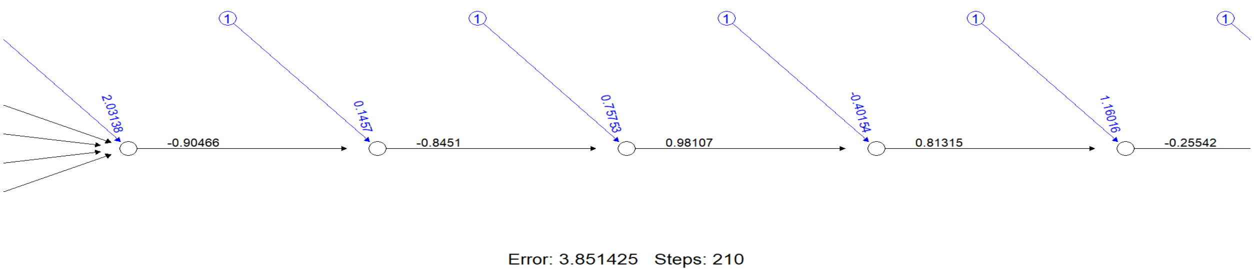


Here is the summary of this model.

```
                      Length Class      Mode
call                     6   -none-     call
response               376   -none-     numeric
covariate             1504   -none-     numeric
model.list               2   -none-     list
err.fct                  1   -none-     function
act.fct                  1   -none-     function
linear.output            1   -none-     logical
data                     5   data.frame list
exclude                  0   -none-     NULL
net.result               1   -none-     list
weights                  1   -none-     list
generalized.weights      1   -none-     list
startweights             1   -none-     list
result.matrix           14   -none-     numeric
```

Model No 3:

In the This Graph, I have showed you how this model is predicting.

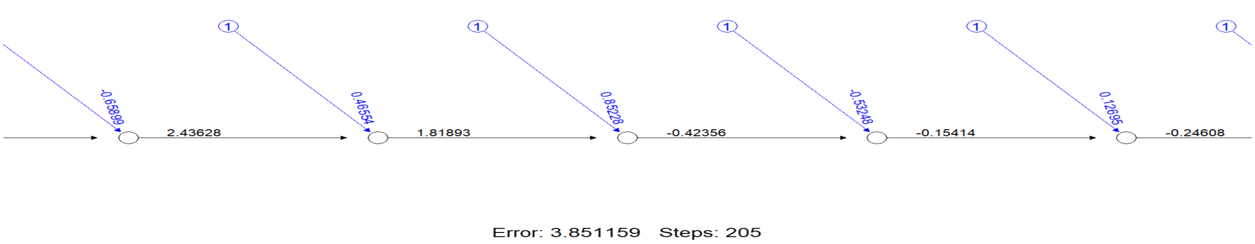


Here is the summary of this model.

```
                      Length Class      Mode
call                     6   -none-     call
response               376   -none-     numeric
covariate             1504   -none-     numeric
model.list               2   -none-     list
err.fct                  1   -none-     function
act.fct                  1   -none-     function
linear.output            1   -none-     logical
data                     5   data.frame list
exclude                  0   -none-     NULL
net.result               1   -none-     list
weights                  1   -none-     list
generalized.weights      1   -none-     list
startweights             1   -none-     list
result.matrix           18   -none-     numeric
```

Model No 4:

In the This Graph, I have showed you how this model is predicting.

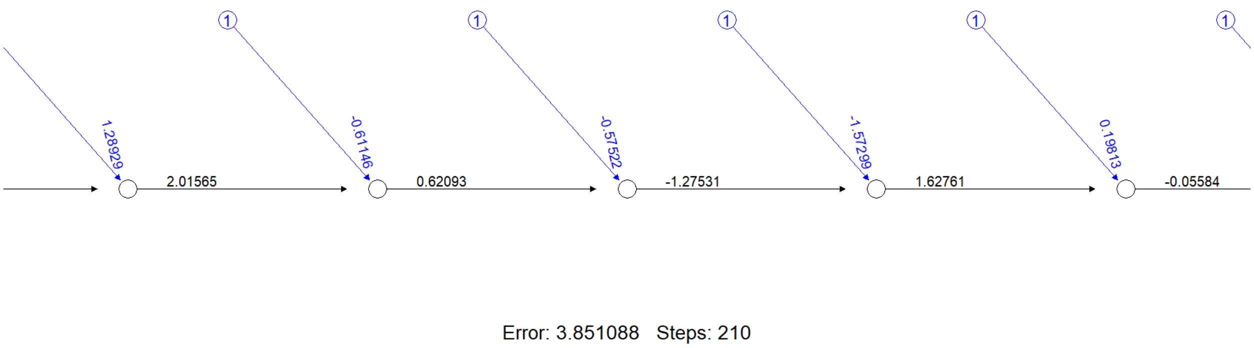


Here is the summary of this model.

```
                    Length Class      Mode
call                   6   -none-     call
response             376   -none-     numeric
covariate           1504   -none-     numeric
model.list             2   -none-     list
err.fct                1   -none-     function
act.fct                1   -none-     function
linear.output          1   -none-     logical
data                   5   data.frame list
exclude                0   -none-     NULL
net.result             1   -none-     list
weights                1   -none-     list
generalized.weights    1   -none-     list
startweights           1   -none-     list
result.matrix         22   -none-     numeric
```

Model No 5:

In the This Graph, I have showed you how this model is predicting.

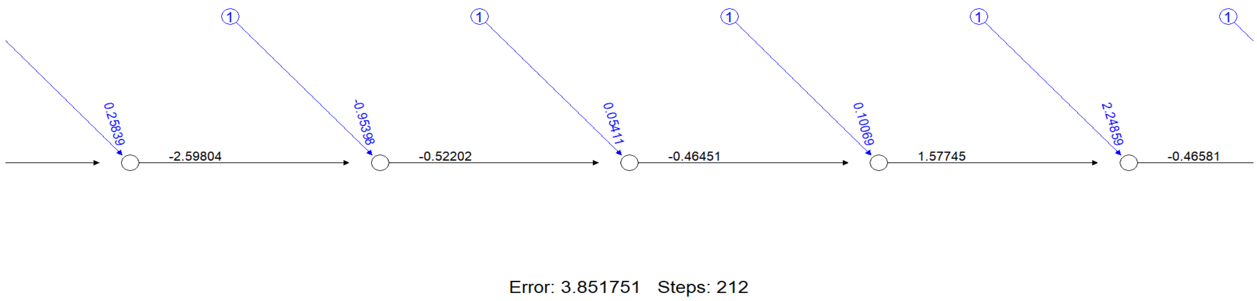


Here is the summary of this model.

```
                    Length Class      Mode
call                   6   -none-     call
response             376   -none-     numeric
covariate           1504   -none-     numeric
model.list             2   -none-     list
err.fct                1   -none-     function
act.fct                1   -none-     function
linear.output          1   -none-     logical
data                   5   data.frame list
exclude                0   -none-     NULL
net.result             1   -none-     list
weights                1   -none-     list
generalized.weights    1   -none-     list
startweights           1   -none-     list
result.matrix         26   -none-     numeric
```

Model No 6:

In the This Graph, I have showed you how this model is predicting.

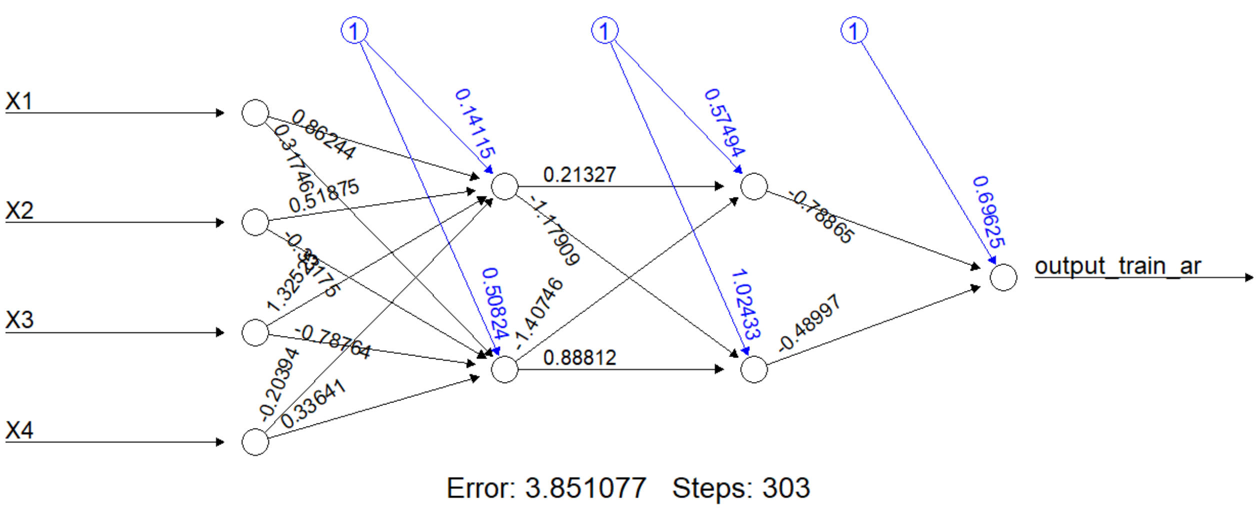


Here is the summary of this model.

```
                    Length Class      Mode
call                   6   -none-     call
response             376   -none-     numeric
covariate           1504   -none-     numeric
model.list             2   -none-     list
err.fct                1   -none-     function
act.fct                1   -none-     function
linear.output          1   -none-     logical
data                   5   data.frame list
exclude                0   -none-     NULL
net.result             1   -none-     list
weights                1   -none-     list
generalized.weights    1   -none-     list
startweights           1   -none-     list
result.matrix         30   -none-     numeric
```

Model No 7:

In the This Graph, I have showed you how this model is predicting.

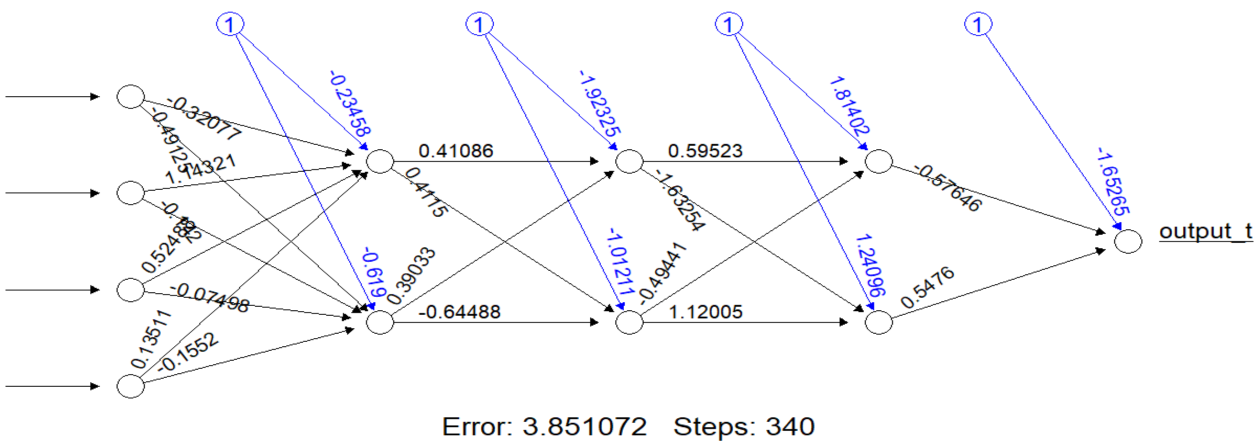


Here is the summary of this model.

```
                      Length  Class       Mode
call                    6    -none-      call
response              376    -none-      numeric
covariate            1504    -none-      numeric
model.list              2    -none-      list
err.fct                 1    -none-      function
act.fct                 1    -none-      function
linear.output           1    -none-      logical
data                    5    data.frame  list
exclude                 0    -none-      NULL
net.result              1    -none-      list
weights                 1    -none-      list
generalized.weights     1    -none-      list
startweights            1    -none-      list
result.matrix          22    -none-      numeric
```

Model No 8:

In the This Graph, I have showed you how this model is predicting.

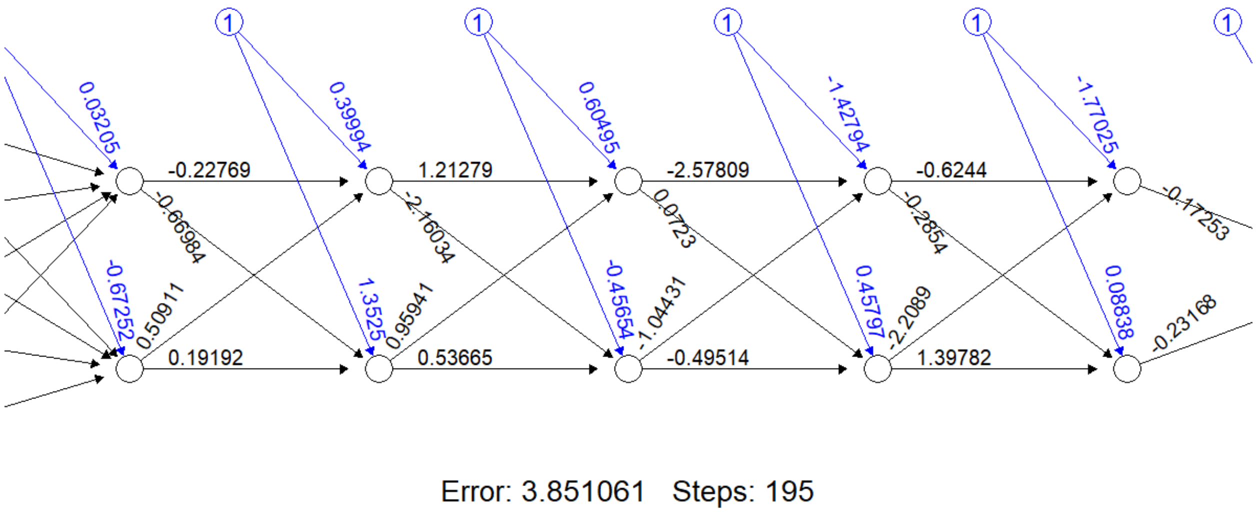


Here is the summary of this model.

```
                      Length  Class       Mode
call                    6    -none-      call
response              376    -none-      numeric
covariate            1504    -none-      numeric
model.list              2    -none-      list
err.fct                 1    -none-      function
act.fct                 1    -none-      function
linear.output           1    -none-      logical
data                    5    data.frame  list
exclude                 0    -none-      NULL
net.result              1    -none-      list
weights                 1    -none-      list
generalized.weights     1    -none-      list
startweights            1    -none-      list
result.matrix          28    -none-      numeric
```

Model No 9:

In the This Graph, I have showed you how this model is predicting.

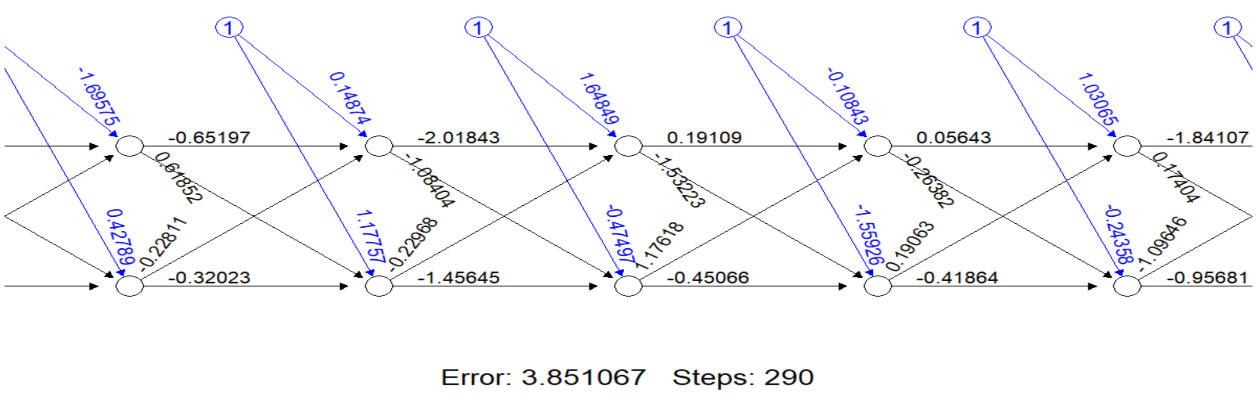


Here is the summary of this model.

```
                      Length Class      Mode
call                    6    -none-     call
response              376    -none-     numeric
covariate            1504    -none-     numeric
model.list              2    -none-     list
err.fct                 1    -none-     function
act.fct                 1    -none-     function
linear.output           1    -none-     logical
data                    5    data.frame list
exclude                 0    -none-     NULL
net.result              1    -none-     list
weights                 1    -none-     list
generalized.weights     1    -none-     list
startweights            1    -none-     list
result.matrix          40    -none-     numeric
```

Model No 10:

In the This Graph, I have showed you how this model is predicting.

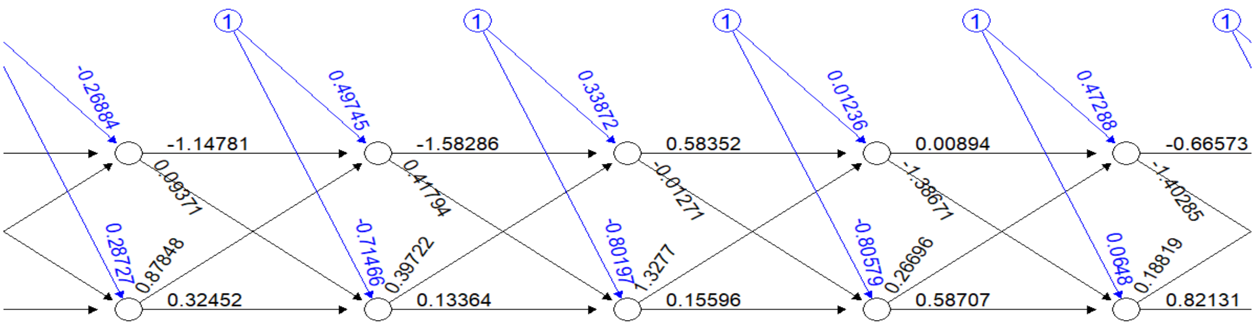


Here is the summary of this model.

	Length	Class	Mode
call	6	-none-	call
response	376	-none-	numeric
covariate	1504	-none-	numeric
model.list	2	-none-	list
err.fct	1	-none-	function
act.fct	1	-none-	function
linear.output	1	-none-	logical
data	5	data.frame	list
exclude	0	-none-	NULL
net.result	1	-none-	list
weights	1	-none-	list
generalized.weights	1	-none-	list
startweights	1	-none-	list
result.matrix	52	-none-	numeric

Model No 11:

In the This Graph, I have showed you how this model is predicting.



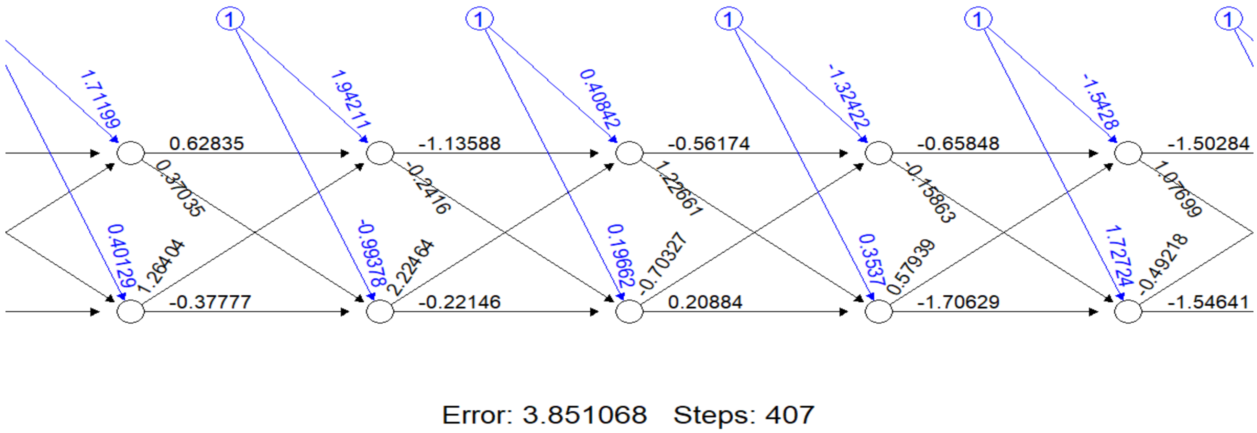
Error: 3.851113 Steps: 219

Here is the summary of this model.

	Length	Class	Mode
call	6	-none-	call
response	376	-none-	numeric
covariate	1504	-none-	numeric
model.list	2	-none-	list
err.fct	1	-none-	function
act.fct	1	-none-	function
linear.output	1	-none-	logical
data	5	data.frame	list
exclude	0	-none-	NULL
net.result	1	-none-	list
weights	1	-none-	list
generalized.weights	1	-none-	list
startweights	1	-none-	list
result.matrix	64	-none-	numeric

Model No 12:

In the This Graph, I have showed you how this model is predicting.



Here is the summary of this model.

```
                  Length Class      Mode
call                 6   -none-     call
response           376   -none-     numeric
covariate         1504   -none-     numeric
model.list           2   -none-     list
err.fct              1   -none-     function
act.fct              1   -none-     function
linear.output        1   -none-     logical
data                 5   data.frame list
exclude              0   -none-     NULL
net.result           1   -none-     list
weights              1   -none-     list
generalized.weights  1   -none-     list
startweights         1   -none-     list
result.matrix       76   -none-     numeric
```

Question No 5:

Statistics Indices for the evaluation of NN Model?

RMSE:

RMSE (Root Mean Square Error) in which we measure the performance of the nn model the RMSE calculate the value base on the average of the difference between the Prediction and Target values.

MAE:

MAE (Mean Absolute Error) in which we also measure the performance of the nn model the MAE calculate the value base on the absolute difference between the Prediction and Target values.

MAPE:

MAPE (Mean Absolute Percentage Error) in which we also measure the Performance of the nn model The MAPE calculate the value base on the percentage between the Prediction and Target value.

SMAPE:

SMAPE (Symmetrical Mean Absolute Percentage Error) in which we also measure the Performance of the nn model The SMAPE calculate the value base on the average of the absolute percentage error between the Prediction and Target Values.

Question No 6:

ARR NN Model Comparison Table?

```
> print(comparison_armodel)
        RMSE        MAE     MAPE    SMAPE
1  0.1116392 0.08936264 18.61858 17.54545
2  0.1116778 0.08941183 18.62578 17.55413
3  0.1116471 0.08935056 18.61662 17.54466
4  0.1116709 0.08940341 18.62486 17.55227
5  0.1116956 0.08943782 18.62923 17.55884
6  0.1116598 0.08937627 18.62054 17.54848
7  0.1116701 0.08940488 18.62460 17.55290
8  0.1116463 0.08935200 18.61699 17.54472
9  0.1116616 0.08938470 18.62133 17.54978
10 0.1116377 0.08935763 18.61767 17.54470
11 0.1116544 0.08938533 18.62058 17.54974
12 0.1117649 0.08959009 18.65121 17.58262
```

In this Figure, I have displayed the All four statistical Indices with the help of AR NN Model.

Question No 7:

AR NN Best two Models?

We know that many, we have many methods that help to select the Best Model with the models. Basically, we know that AR Approach based on the MLP NN Model and LSTM NN Model but I which I have used the MLP NN Model. In this task I have selected the model no 4 and model no 7 because as the statistically indices these models perform well and these models generate the good results in the Test Dataset.

Question No 8:

NARX Models Results?

In this task, I have generated the 4 models with the NARX NN Model Approach with the setting the parameters of NARX NN Model and each model display THE summary of the NARX NN Model too.

Model No 1:

Here is the Summary of the NARX first Model.

```
> summary(model)
          Length Class        Mode
x         376    ts           numeric
m           1    -none-       numeric
p           1    -none-       numeric
P           1    -none-       numeric
scalex      2    -none-       list
size        1    -none-       numeric
subset    376    -none-       numeric
model      50    nnetarmodels list
nnetargs    0    -none-       list
fitted    376    ts           numeric
residuals 376    ts           numeric
lags       15    -none-       numeric
series      1    -none-       character
method      1    -none-       character
call        4    -none-       call
```

Model No 2:

Here is the Summary of the NARX second Model.

```
> summary(model)
          Length Class        Mode
x         376    ts           numeric
m           1    -none-       numeric
p           1    -none-       numeric
P           1    -none-       numeric
scalex      2    -none-       list
size        1    -none-       numeric
subset    376    -none-       numeric
model      50    nnetarmodels list
nnetargs    0    -none-       list
fitted    376    ts           numeric
residuals 376    ts           numeric
lags       15    -none-       numeric
series      1    -none-       character
method      1    -none-       character
call        4    -none-       call
```

Model No 3:

Here is the Summary of the NARX third Model.

```
> summary(model)
          Length Class        Mode
x         376    ts           numeric
m           1    -none-       numeric
p           1    -none-       numeric
P           1    -none-       numeric
scalex      2    -none-       list
size        1    -none-       numeric
subset    376    -none-       numeric
model      50    nnetarmodels list
nnetargs    0    -none-       list
fitted    376    ts           numeric
residuals 376    ts           numeric
lags       15    -none-       numeric
series      1    -none-       character
method      1    -none-       character
call        4    -none-       call
```

Model No 4:

Here is the Summary of the NARX fifth Model.

```
> summary(model)
          Length Class        Mode
x         376    ts           numeric
m           1    -none-       numeric
p           1    -none-       numeric
P           1    -none-       numeric
scalex      2    -none-       list
size        1    -none-       numeric
subset    376    -none-       numeric
model      50    nnetarmodels list
nnetargs    0    -none-       list
fitted    376    ts           numeric
residuals 376    ts           numeric
lags       15    -none-       numeric
series      1    -none-       character
method      1    -none-       character
call        4    -none-       call
```

Comparison Table:

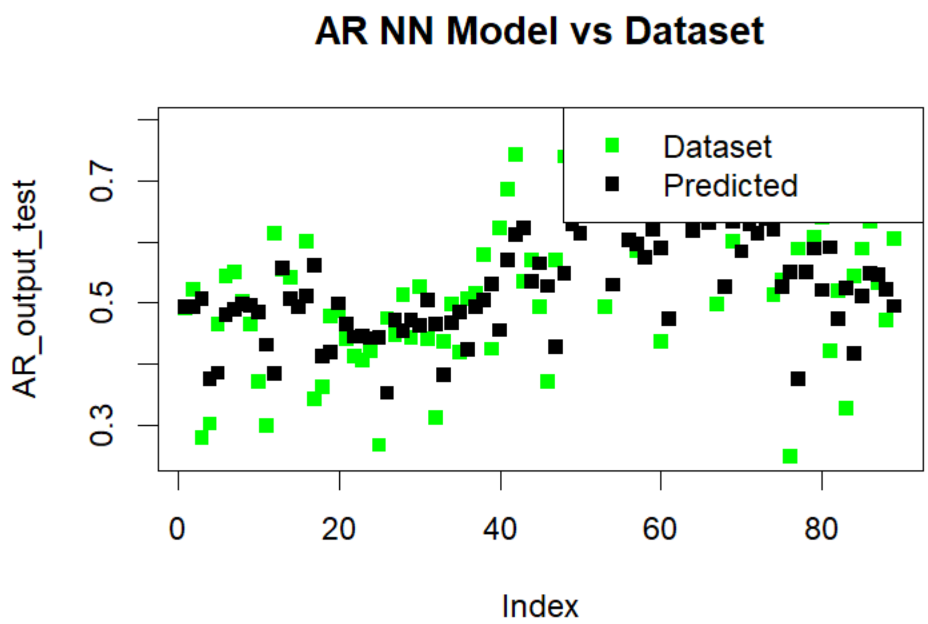
```
> print(comparison_narxmodel)
       RMSE       MAE     MAPE    SMAPE
1 0.1346346 0.1052800 20.51510 21.49111
2 0.1516405 0.1194104 23.52491 23.81580
3 0.1356060 0.1091150 21.74667 21.90841
4 0.1362271 0.1098729 21.54932 22.41349
```

In this Figure, I have displayed the All four statistical Indices with the help of NARX NN Model.

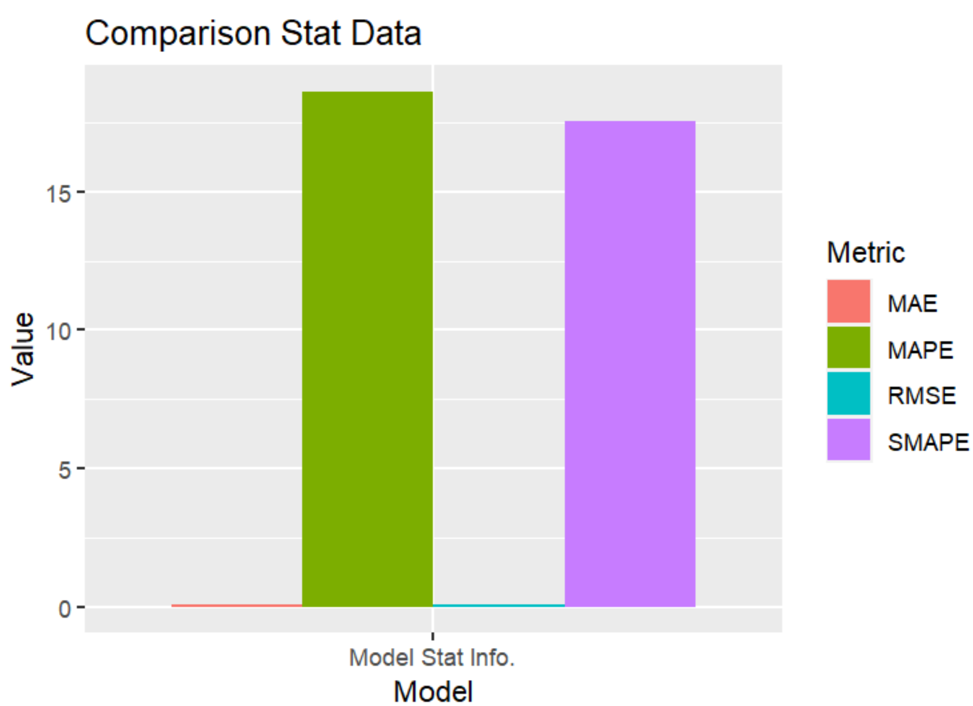
Question No 9:

Graphically Representation of Output and desire Output Data?

AR NN Model:

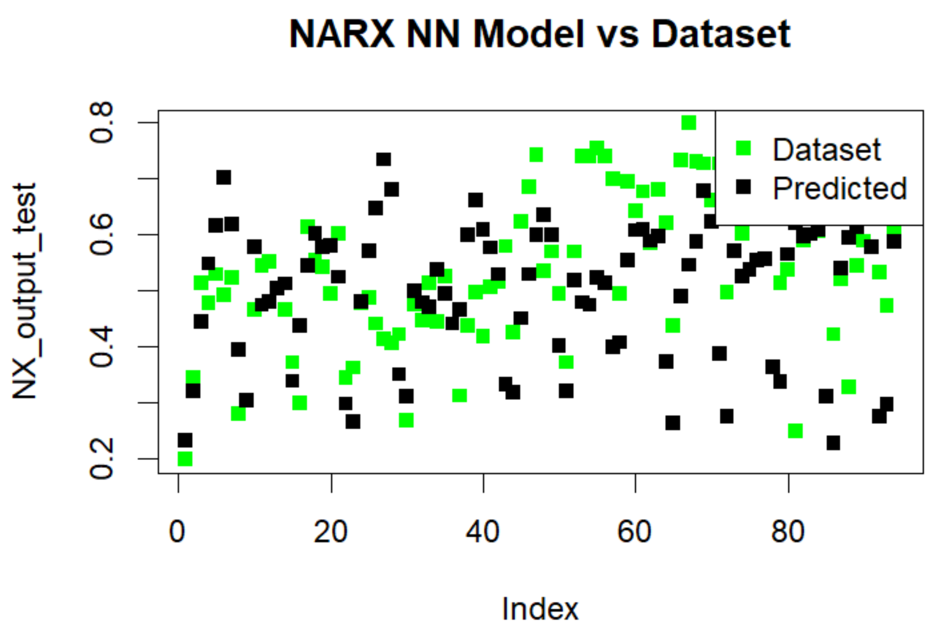


In this figure, I have displayed the AR model predicted data and actual data.

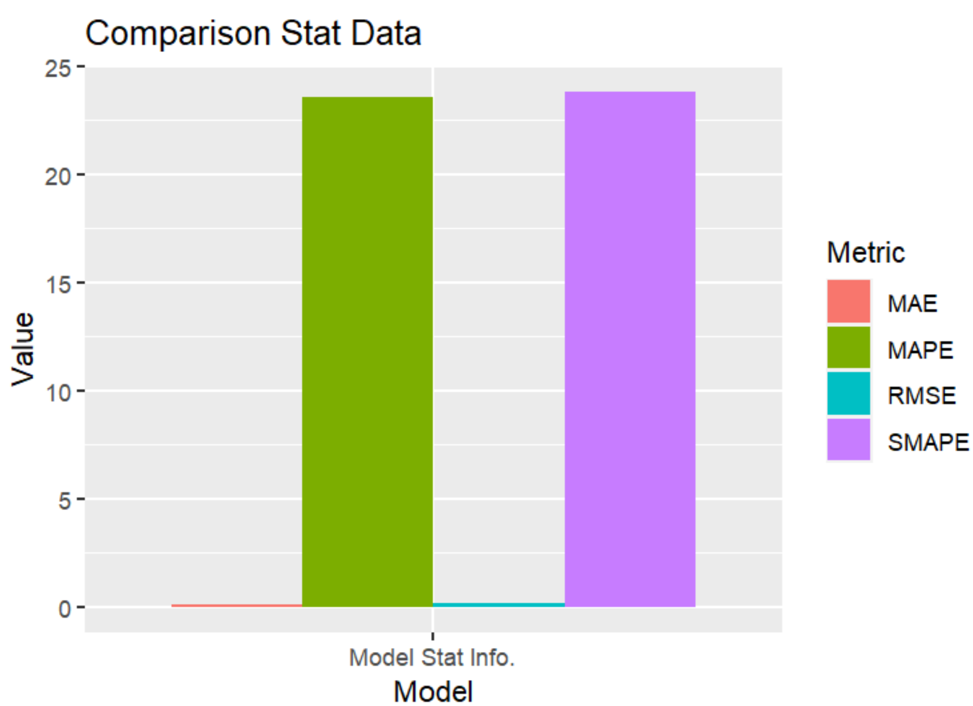


In this Figure, I have displayed the AR Model Statistical indic.

NARX Model:



In this figure, I have displayed the NARX model predicted data and actual data.



In this Figure, I have displayed the NARX Model Statistical Indic.