Page 1

**Capstone Project**

**Preliminary Stage Assignment 2**

**Course code:** CSA1643

**Course:** Data warehousing and Data Mining for Data Science

**S. No**: 16

**Name**: Keshav Reddy

**Reg. No**: 192210248

**Slot**: c

**Title:** predictive maintenance for industrial equipment in data mining

**Assignment Release Date:** 14/02/2024

**Assignment** **Preliminary Stage (Assignment 2) submission Date**: 16/02/2024

**Mentor Name:** D r. D. Iranian

**Mentor Phone number:** 9994195096

**Mentor Department**: Department of mathematical sciences

**R PROGRAMMING:**# Load required libraries

library(caret) # for data preprocessing

library (random Forest) # for building the predictive model

# Load the dataset (replace 'data.csv' with your dataset)

data <- read.csv("data.csv")

# Explore the data (optional)

head(data)

summary(data)

# Preprocess the data (optional)

# Here, you may handle missing values, scale or normalize data, etc.

# For simplicity, we skip preprocessing in this example

# Split data into training and testing sets

set. Seed (123) # for reproducibility

train index <- create Data Partition (data $ target variable, p = 0.8, list = FALSE)

train data <- data [train index,]

test data <- data [train index,]

# Build a predictive model (Random Forest)

model <- random Forest (target variable ~., data = train data)

# Make predictions on the test set

predictions <- predict (model, test data)

# Evaluate model performance

confusion Matrix (predictions, test data $ target variable)

0.

# Visualize feature importance

vamplate(model)

**output:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **s.no** | **Sensor 1** | **Sensor 2** | **Sensor 3** | **Sensor 4** | **Target variable** |
| 1 | 0.1 | 0.2 | 0.3 | 0.4 | 0 |
| 2 | 0.2 | 0.3 | 0.4 | 0.5 | 1 |
| 3 | 0.3 | 0.4 | 0.5 | 0.6 | 0 |
| 4 | 0.4 | 0.5 | 0.6 | 0.7 | 0 |
| 5 | 0.5 | 0.6 | 0.7 | 0.8 | 1 |
| 6 | 0.6 | 0.7 | 0.8 | 0.9 | 0 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sensor 1** | **Sensor 2** | **Sensor 3** | **Sensor 4** | **Target variable** |
| Min. 0.1 | Min: 0.2 | Min: 0.3 | Min: 0.4 | Min: 0.00 |
| 1st Qu: 0.2 | 1st Qu: 0.3 | 1st Qu: 0.4 | 1st Qu: 0.5 | 1st Qu: 0.00 |
| Median: 0.3 | Median: 0.4 | Median: 0.5 | Median: 0.6 | Median: 0.00 |
| Mean: 0.4 | Mean: 0.5 | Mean: 0.6 | Mean: 0.7 | Mean: 0.33 |
| 3rd Qu: 0.5 | 3rd Qu: 0.7 | 3rd Qu: 0.7 | 3rd Qu: 0.8 | 3rd Qu: 1.00 |
| Max: 0.6 | Max: 0.7 | Max: 0.8 | Max: 0.9 | Max: 1.00 |