**Health monitoring and threat detection through digital processing**

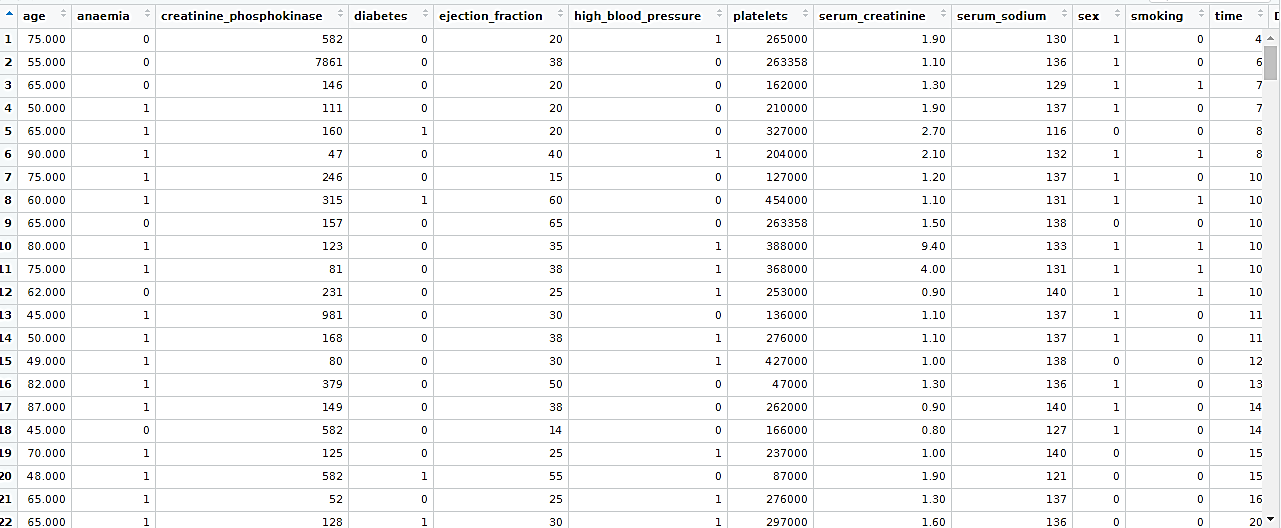


Figure 1: Dataset import in R Software

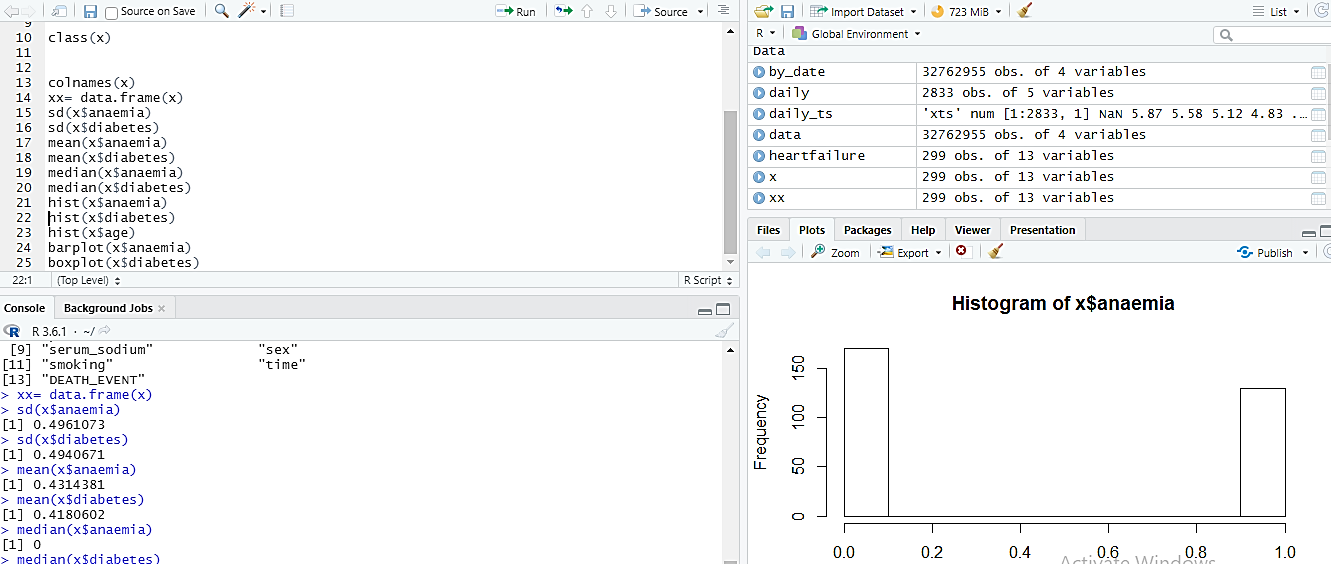


Figure 2: Descriptive Statistics on the dataset

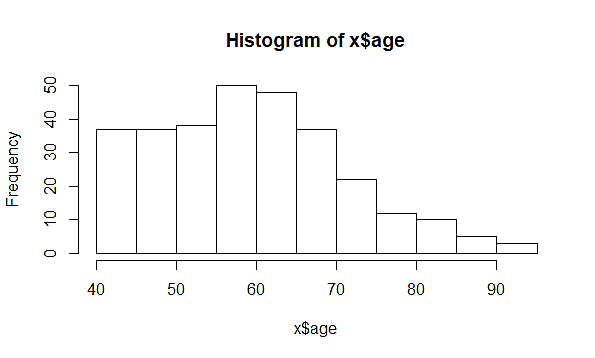


Figure 3: Histogram of Age variable

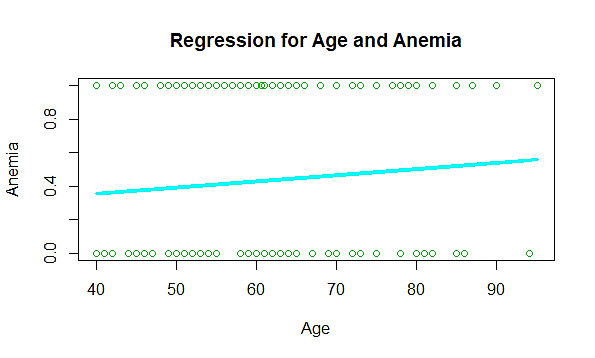


Figure 4: Regression Plot between Age and Anemia variables

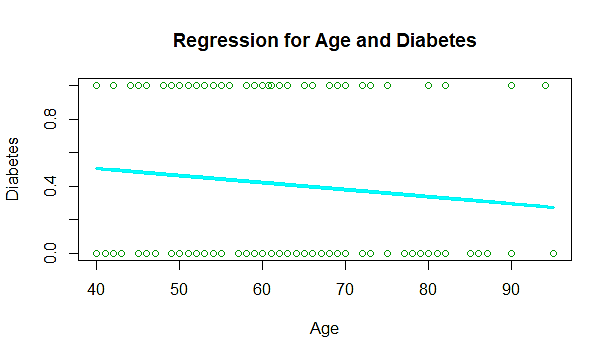


Figure 5: Regression between Age and Diabetes

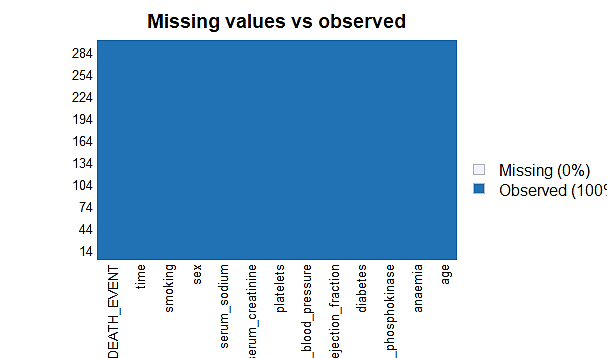


Figure 6: Missing Values Finding

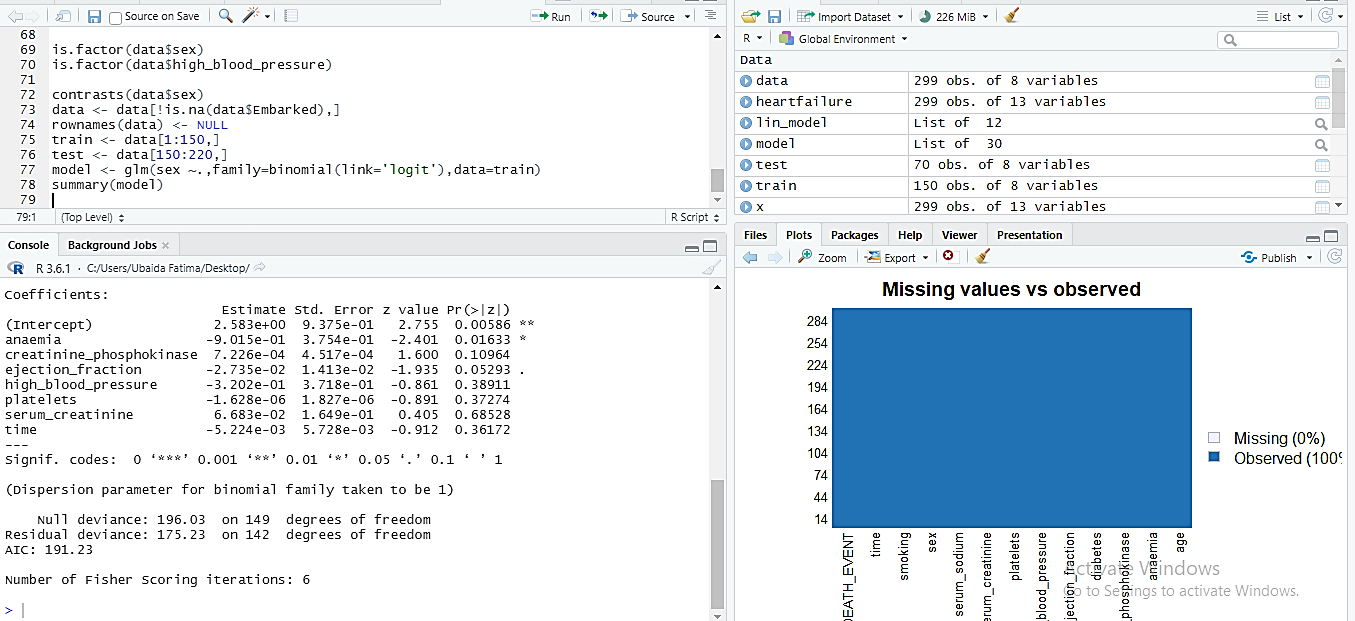


Figure 7: Training Dataset in R Software

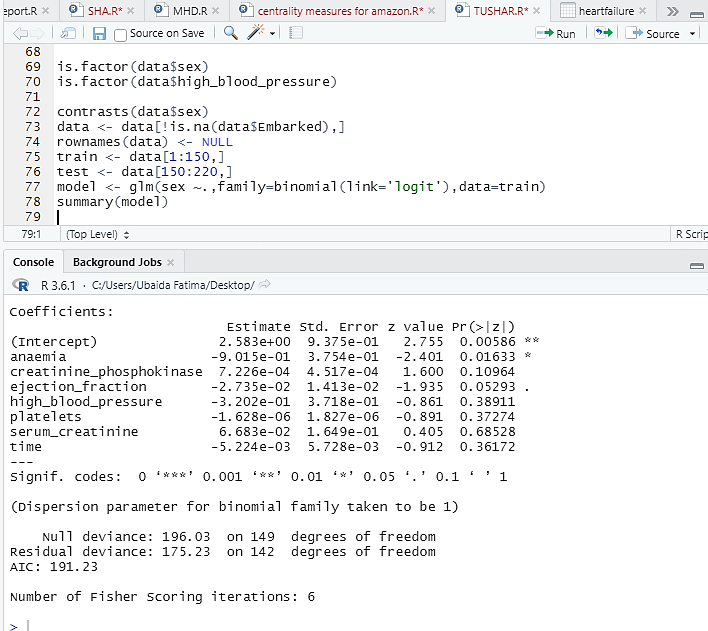


Figure 8: Logistic Regression Model Fitting on Dataset

**Interpreting the outcomes of dataset through logistic regression model:**

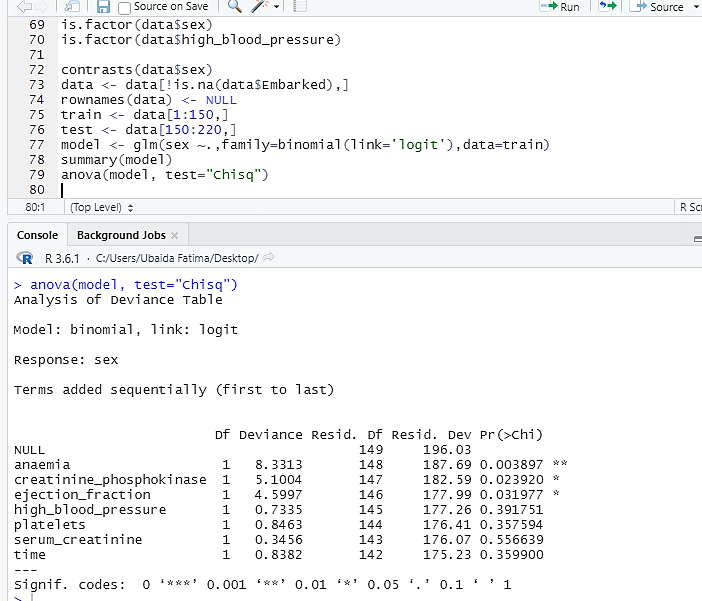


Figure 9: ANOVA on dataset through R Software

**Assessing the predictive ability of the model:**

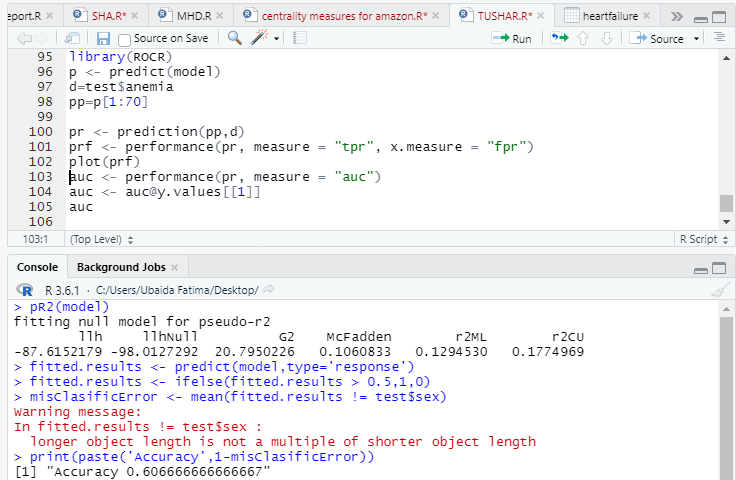


Figure 10: Accuracy of Fitted Model

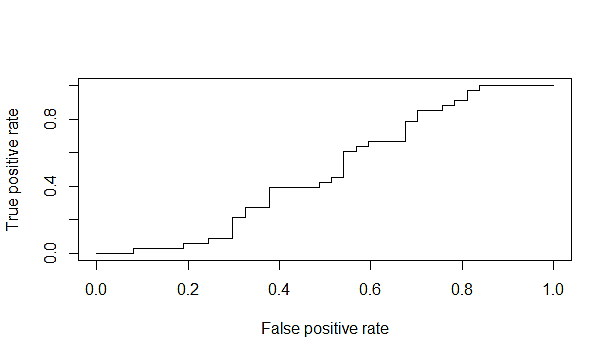


Figure 11: Prediction Plot for Anemia through Logistic Regression

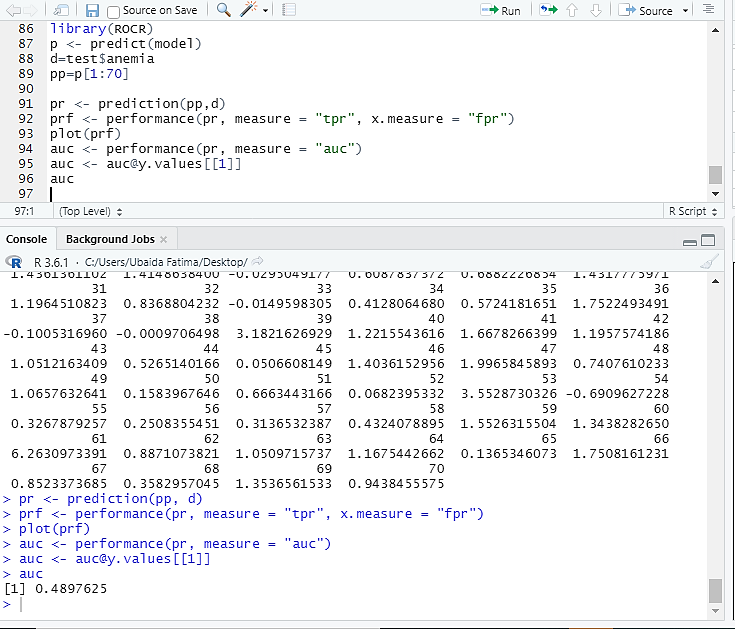


Figure 12: Prediction Model and Accuracy of Prediction Model on dataset through R Software

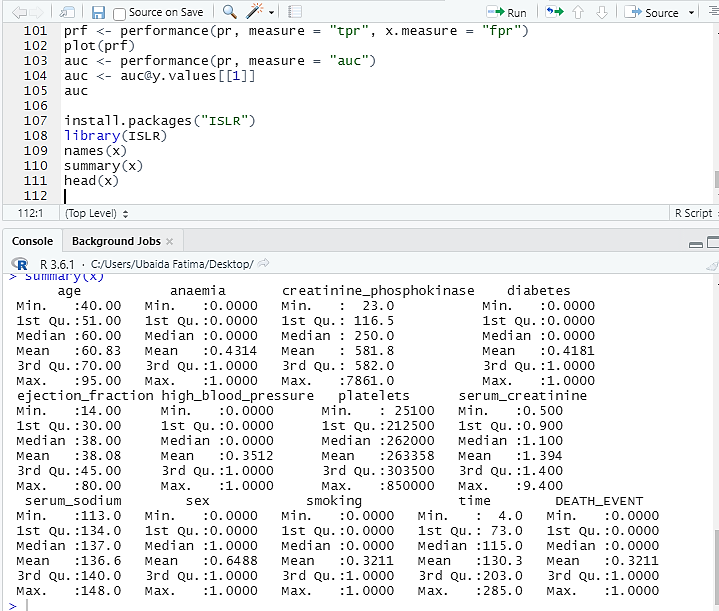


Figure 13: Summary of Heart Monitoring Dataset

Calculating the correlation between each pair of numeric variables. These pair-wise correlations can be plotted in a correlation matrix plot to given an idea of which variables change together.

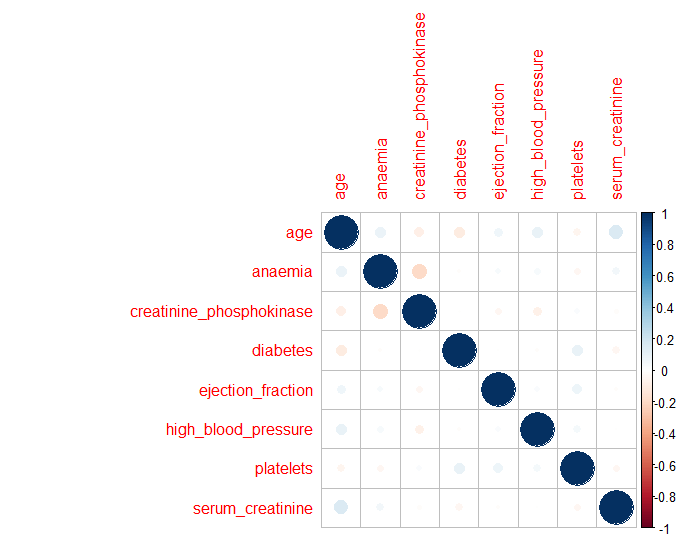


Figure 14: Correlation Plot for Dataset

A dot-representation was used where blue represents positive correlation and red negative. The larger the dot the larger the correlation. You can see that the matrix is symmetrical and that the diagonal are perfectly positively correlated because it shows the correlation of each variable with itself. Unfortunately, none of the variables are correlated with one another.

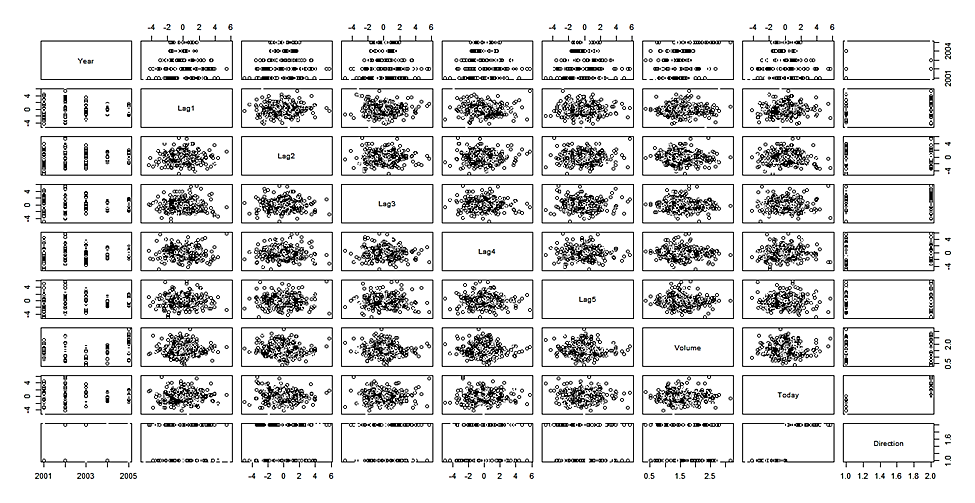


Figure 15: Density Distribution for every factor

The density distribution of each variable broken down by Sex factor in the data. Like the scatterplot matrix above, the density plot by Sex factor can help see the separation of Up and Down.

**Support Vector Machine on Heart Monitoring Dataset**

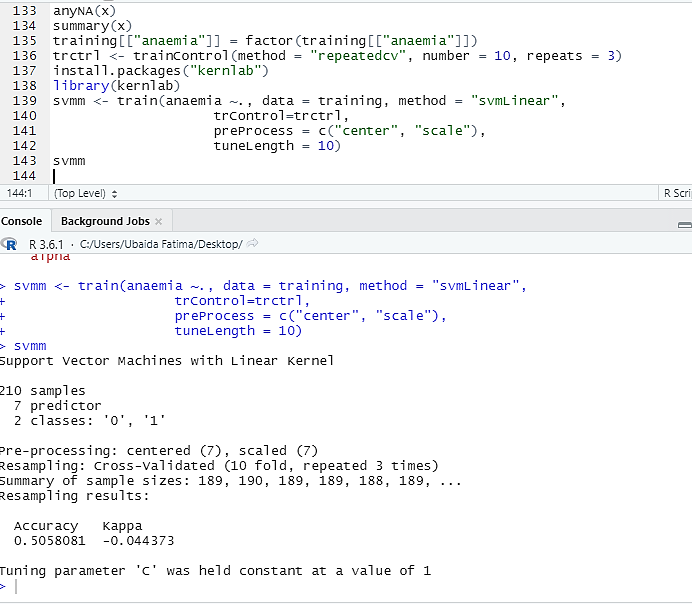


Figure 16: Support Vector Machine (SVM) results on R Software

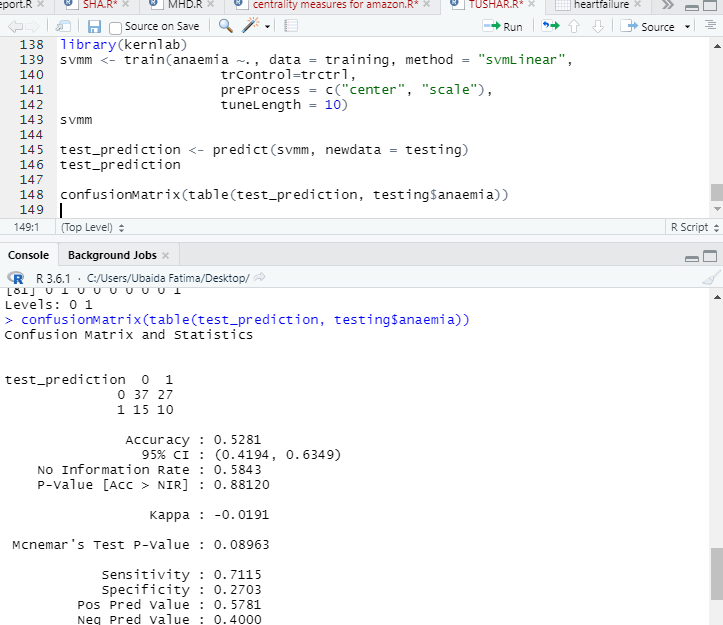


Figure 17: Test Prediction and Confusion Matrix Statistics

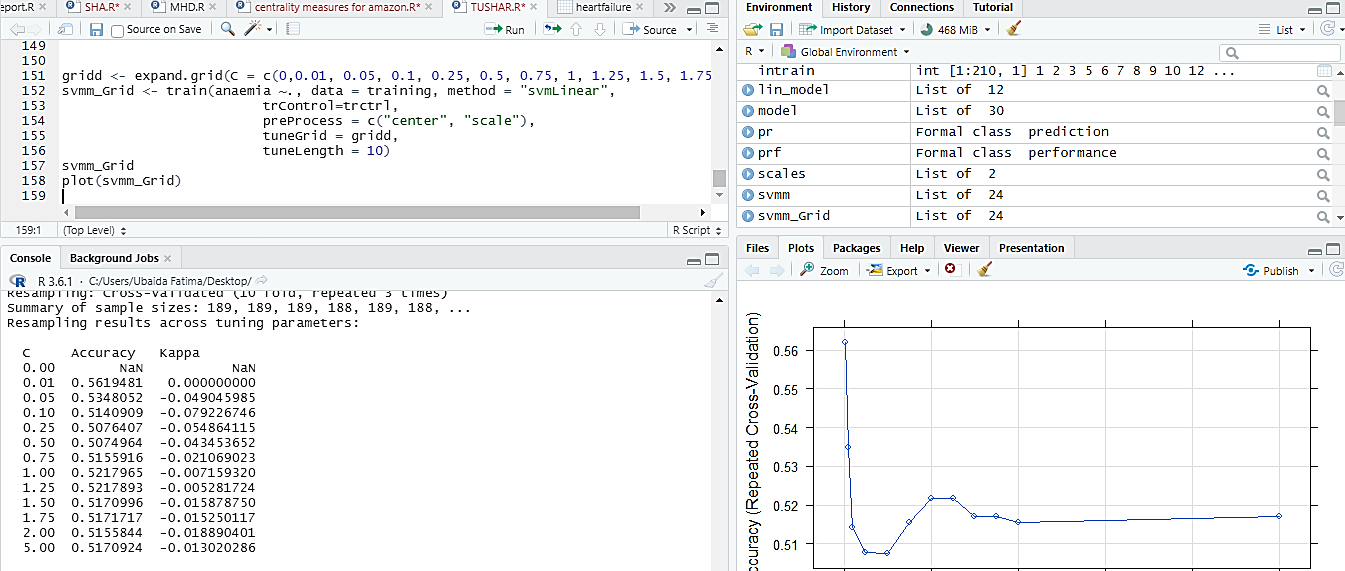


Figure 18: Support Vector Machine (SVM) Accuracy for dataset

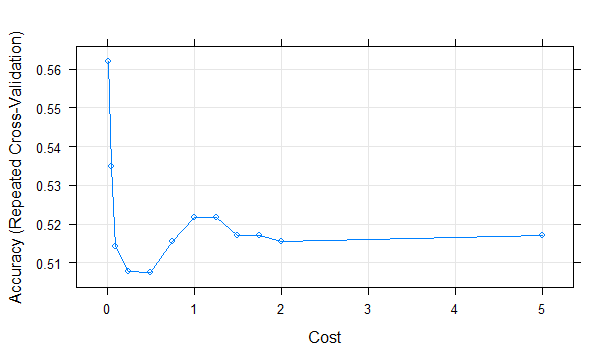


Figure 19: Support Vector Machine (SVM) plot for dataset

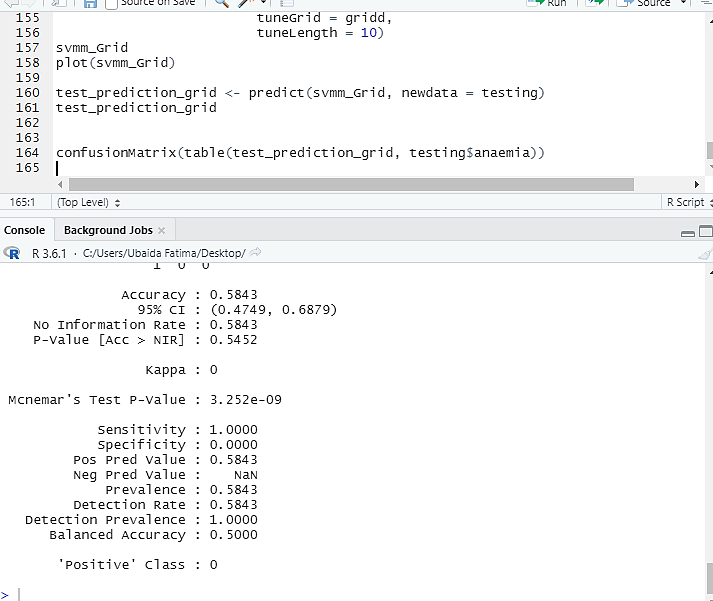


Figure 20: Confusion Matrix accuracy statistics for test prediction model

**Code on R software:**

install.packages("bnlearn")

install.packages("snow")

library('snow')

library(bnlearn)

library(readxl)

library(rlang)

x=heartfailure #Import data set heart failure

class(x)

colnames(x)

xx= data.frame(x)

sd(x$anaemia)

sd(x$diabetes)

mean(x$anaemia)

mean(x$diabetes)

median(x$anaemia)

median(x$diabetes)

hist(x$anaemia)

hist(x$diabetes)

hist(x$age)

barplot(x$anaemia)

boxplot(x$diabetes)

X=x$age

Y=x$anaemia

lin\_model <- lm(Y~X)

coef(lin\_model)

c <- 0.208143546

m <- 0.003670562

y\_cap <- c + (m\*X)

length(y\_cap)

plot(Y~X,

main='Regression for Age and Anemia',

xlab='Age',ylab='Anemia', col=3)

lines(X,y\_cap , col = 5 , lwd = 3)

Z=x$diabetes

lin\_model <- lm(Z~X)

coef(lin\_model)

c <- 0.673300159

m <- -0.004195687

y\_cap <- c + (m\*X)

length(y\_cap)

plot(Z~X,

main='Regression for Age and Diabetes',

xlab='Age',ylab='Diabetes', col=3)

lines(X,y\_cap , col = 5 , lwd = 3)

install.packages("Amelia")

library(Amelia)

missmap(x, main = "Missing values vs observed")

data <- subset(x,select=c(2,3,5,6,7,8,10,12))

zz=data$anaemia

zz[is.na(zz)] <- mean(zz,na.rm=T)

is.factor(data$sex)

is.factor(data$high\_blood\_pressure)

contrasts(data$sex)

data <- data[!is.na(data$Embarked),]

rownames(data) <- NULL

train <- data[1:150,]

test <- data[150:220,]

model <- glm(sex ~.,family=binomial(link='logit'),data=train)

summary(model)

anova(model, test="Chisq")

t= data$anaemia

install.packages("pscl")

library(pscl)

pR2(model)

fitted.results <- predict(model,type='response')

fitted.results <- ifelse(fitted.results > 0.5,1,0)

misClasificError <- mean(fitted.results != test$sex)

print(paste('Accuracy',1-misClasificError))

install.packages(ROCR)

library(ROCR)

p <- predict(model)

d=test$anemia

pp=p[1:70]

pr <- prediction(pp,d)

prf <- performance(pr, measure = "tpr", x.measure = "fpr")

plot(prf)

auc <- performance(pr, measure = "auc")

auc <- auc@y.values[[1]]

auc

install.packages("ISLR")

library(ISLR)

names(x)

summary(x)

head(x)

install.packages("corrplot")

library(corrplot)

correlations <- cor(x[,1:8])

corrplot(correlations, method="circle")

pairs(x, col=x$sex)

install.packages("caret")

library(caret)

x <- x[,1:8]

y <- x[,9]

scales <- list(x=list(relation="free"), y=list(relation="free"))

featurePlot(x=x, y=y, plot="density", scales=scales)

str(x)

head(x)

intrain <- createDataPartition(y = x$anaemia, p= 0.7, list = FALSE)

training <- x[intrain,]

testing <- x[-intrain,]

dim(training);

dim(testing);

anyNA(x)

summary(x)

training[["anaemia"]] = factor(training[["anaemia"]])

trctrl <- trainControl(method = "repeatedcv", number = 10, repeats = 3)

install.packages("kernlab")

library(kernlab)

svmm <- train(anaemia ~., data = training, method = "svmLinear",

trControl=trctrl,

preProcess = c("center", "scale"),

tuneLength = 10)

svmm

test\_prediction <- predict(svmm, newdata = testing)

test\_prediction

confusionMatrix(table(test\_prediction, testing$anaemia))

gridd <- expand.grid(C = c(0,0.01, 0.05, 0.1, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2,5))

svmm\_Grid <- train(anaemia ~., data = training, method = "svmLinear",

trControl=trctrl,

preProcess = c("center", "scale"),

tuneGrid = gridd,

tuneLength = 10)

svmm\_Grid

plot(svmm\_Grid)

test\_prediction\_grid <- predict(svmm\_Grid, newdata = testing)

test\_prediction\_grid

confusionMatrix(table(test\_prediction\_grid, testing$anaemia))