

DS – 510 Project 2

**TO STUDY AND ATTEMPT TO PREDICT IF, BASED ON A
NUMBER OF RISK FACTORS, A GIVEN PATIENT WILL
HAVE HEART DISEASE.**

Team Members:

Aditi Agarwal

Lovepreet Kaur

Mohit Supe

Table of Contents

Our Analysis:	3
Reading Data File and Arranging for Analysis.....	8
Reading the data file Heart.csv	8
Selecting first 250 rows for Training data	8
Selecting last 53 rows for Test data	8
Logistic Regression.....	8
Model 1: AHD vs all Explanatory Variables.	8
Model 2: AHD vs Sex, Ca, Chest pain and RestBP	9
Model 3: AHD vs Sex, Ca and Chest pain.....	10
Model 4: AHD vs Ca and Chest pain.....	10
Model 5: AHD vs ChestPain and Sex.....	11
Model 6: AHD vs Ca and Sex.....	11
Model 7: AHD vs Sex.....	12
Model 8: AHD vs RestBP.....	12
Model 9: AHD vs Ca	13
Model 10: AHD vs ChestPain.....	13
Predicting the Best Model:.....	14
Support Vector Machine	15
Model 11: AHD vs all Explanatory Variables.	15
Model 12: AHD vs Sex, Ca, Chest pain and RestBP.....	15
Model 13: AHD vs Sex, Ca and Chest pain.....	16
Model 14: AHD vs Ca and Chest pain.....	16
Model 15: AHD vs ChestPain and Sex.....	16
Model 16: AHD vs Ca and Sex.....	17
Model 17: AHD vs Sex.....	17
Model 18: AHD vs RestBP.....	18
Model 19: AHD vs Ca	18
Model 20: AHD vs ChestPain.....	18

Our Analysis:

Log Odds

```
> exp(cbind(OR=coef(Model1),confint(Model1)))  
Waiting for profiling to be done...  
              OR      2.5 %      97.5 %  
(Intercept)  0.004060436 4.311547e-06 2.5384421  
Age          1.003546094 9.474646e-01 1.0624847  
Sex          7.848956904 2.356256e+00 30.1421557  
ChestPainnonanginal 0.135133782 4.425756e-02 0.3755143  
ChestPainnontypical 0.184762682 3.639128e-02 0.7268512  
ChestPainypical  0.080139725 1.587923e-02 0.3394161  
RestBP       1.026710901 1.002147e+00 1.0539300  
Chol         1.006961031 9.976412e-01 1.0164869  
Fbs          0.333405944 8.125178e-02 1.2651664  
RestECG      1.364393077 8.851698e-01 2.1378434  
MaxHR        0.977052786 9.528498e-01 0.9999128  
ExAng        2.353518734 9.062628e-01 6.1573254  
Oldpeak      1.552667497 9.330640e-01 2.6607787  
Slope        1.905810142 7.851695e-01 4.6138495  
Ca           3.083886165 1.786751e+00 5.6744888  
Thalnormal   0.820720590 9.363980e-02 7.4240971  
Thalreversible 2.975987592 3.424086e-01 27.0215838
```

Confidence interval should not contain 1. If it contains 1, that means that Predictor/Explanatory variable is not making any changes in Logit/log odds (Ratio of occurrence of event/not occurrence of that event).

- We can exclude ChestPainnontypical2 from our study, as it contains 1 in its confidence interval. So it is not an impacting factor for AHD.
- MaxHR is also not an impacting factor AHD study as it contains 1 in its confidence interval.

Intercept (0.004060436) is a constant.

- One-unit increase in Sex will increase your odds by 7.85 units.
- One-unit increase in ExAng will increase your odds by 2.35 units.
- One-unit increase in Ca will increase your odds by 3.08 units.
- One-unit increase in Thalreversible will increase your odds by 2.97 units.
- One-unit increase in RestECG will increase your odds by 1.36 units.
- One-unit increase in Oldspeak, slope will increase your odds by 1.55 & 1.9 units respectively.
- ChestPainnontypical has four values {1,2,3,4}. If you change from 1 to 2, then its not significant.
But when you change from 1 to 3 and 4, it's significant and Intercepts <1: means its against heart disease. Changing from ChestPainnontypical1 to 2 and 3 decreases odds.
- Fbs and Thalnormal have slope Intercepts <1, so increasing them one-unit will decrease odds.

Age, RestBP, Chol have Slope intercept close to 1, so you have to change these variables by large number to make any change in odds.

Backward Approach:

Logistic Model of AHD v/s all other predictor variables:

```
> summary(Model1) #values of logit

Call:
glm(formula = AHD ~ ., family = "binomial", data = Training)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.4548  -0.4445  -0.1111   0.3157   2.7754

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  -5.506465   3.364858  -1.636 0.101743
Age           0.003540   0.029019   0.122 0.902912
Sex           2.060381   0.644627   3.196 0.001392 **
ChestPainnonanginal -2.001490   0.541224  -3.698 0.000217 ***
ChestPainnontypical -1.688683   0.750084  -2.251 0.024365 *
ChestPaintypical  -2.523984   0.771765  -3.270 0.001074 **
RestBP        0.026360   0.012719   2.073 0.038214 *
Chol          0.006937   0.004669   1.486 0.137322
Fbs          -1.098394   0.696691  -1.577 0.114890
RestECG       0.310710   0.223293   1.391 0.164077
MaxHR        -0.023215   0.012178  -1.906 0.056615 .
ExAng         0.855912   0.485137   1.764 0.077687 .
Oldpeak       0.439974   0.265335   1.658 0.097280 .
Slope         0.644907   0.446312   1.445 0.148466
Ca            1.126191   0.292823   3.846 0.000120 ***
Thalnormal   -0.197573   1.093361  -0.181 0.856601
Thalreversible 1.090576   1.090946   1.000 0.317475
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 339.88  on 246  degrees of freedom
Residual deviance: 146.09  on 230  degrees of freedom
(3 observations deleted due to missingness)
AIC: 180.09

Number of Fisher Scoring iterations: 6
```

- Sex, ChestPain, RestBP and Ca are four Significant variables. All other variables are not significant.

$P\text{value} < \text{Alpha (Significance level)}$

Logistic model (Model2) of AHD v/s Sex, ChestPain, RestBP and Ca

```
> summary(Model2)
```

Call:

```
glm(formula = AHD ~ Sex + Ca + RestBP + ChestPain, family = "binomial",
    data = Training)
```

Deviance Residuals:

```
      Min       1Q   Median       3Q      Max
-2.6522  -0.6091  -0.2104   0.5977   2.2877
```

Coefficients:

```
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  -4.58748    1.44995  -3.164  0.00156 **
Sex           1.86544    0.43589   4.280  1.87e-05 ***
Ca            1.09083    0.21863   4.989  6.06e-07 ***
RestBP        0.02719    0.01008   2.698  0.00698 **
ChestPainnonanginal -2.38591    0.43147  -5.530  3.21e-08 ***
ChestPainnontypical -2.80997    0.61184  -4.593  4.38e-06 ***
ChestPainatypical -2.59897    0.65819  -3.949  7.86e-05 ***
---
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

```
Null deviance: 341.07 on 247 degrees of freedom
Residual deviance: 201.51 on 241 degrees of freedom
(2 observations deleted due to missingness)
AIC: 215.51
```

Number of Fisher Scoring iterations: 5

- All variables are significant in this model.
- RestBP seems to be least important here as it is significant but at .001 level (Alpha=.001).(**)
- All other variables **Sex, ChestPain and Ca** are significant at 0 level. (***)
- Model 3 is the best Model

Error Rate: Error rate calculated is the percentage value of incorrect predictions of AHD that resulted from the models.

Model	AHD	VS	Independent Variables	LR Error (%)	SVM Error (%)
Model_1	AHD	VS	All	35.60	52.00
Model_2	AHD	VS	Sex + Ca + Chestpain + RestBP	32.40	30.00
Model_3	AHD	VS	Sex + Ca + Chestpain	32.40	28.00
Model_4	AHD	VS	Ca +ChestPain	34.00	30.00
Model_5	AHD	VS	ChestPain+Sex	24.40	30.00

Model_6	AHD	VS	Ca +Sex	32.40	24.00
Model_7	AHD	VS	Sex	38.40	38.00
Model_8	AHD	VS	RestBP	41.20	44.00
Model_9	AHD	VS	Ca	34.00	24.00
Model_10	AHD	VS	ChestPain	23.20	30.00

Model 3 is the Best Model

Backward approach

- In Model1, Sex, ChestPain, RestBP and Ca are four Significant variables. All other variables are not significant.
Pvalue < Alpha (Significance level)
- In Model2, all variables are significant. But RestBP seems to be least important here as it is significant but at .001 level (Alpha=.001) and its individual error rate is also very high 41.2%.
All other variables Sex, ChestPain and Ca are significant at 0 level. (***)

In Model 3, Sex, Ca and chest Pain are most significant explanatory variables among others.

Forward Approach

AHD v/s Sex, Ca, ChestPain, RestBP (Model 7, 8, 9 and 10)

- ChestPain and Ca are the most significant explanatory variables as error rate for these two is lowest in Logistic and SVM models.
- RestBP is the least significant explanatory variable as error rate is highest.

AHD v/s Ca + ChestPain, ChestPain + Sex, Ca + Se.

- Ca +Sex and ChestPain+Sex (Model5 and 6) are best models among these three as error rate for them is lowest for logistic and SVM Models.

So Model 3 (Sex + Ca + ChestPain v/s AHD) is the best Model because its error rate is low (LR=32.40% and SVM=28%).

Comparison among Logistic regression and Support Vector Machine

```
> summary(SVM2)
```

Call:

```
svm(formula = AHD ~ Sex + Ca + ChestPain, data = Training, Kernel = "linear", scale = FALSE)
```

Parameters:

```
SVM-Type: C-classification  
SVM-Kernel: radial  
cost: 1  
gamma: 0.1666667
```

Number of Support Vectors: 121

```
( 61 60 )
```

Number of Classes: 2

Levels:

```
No Yes
```

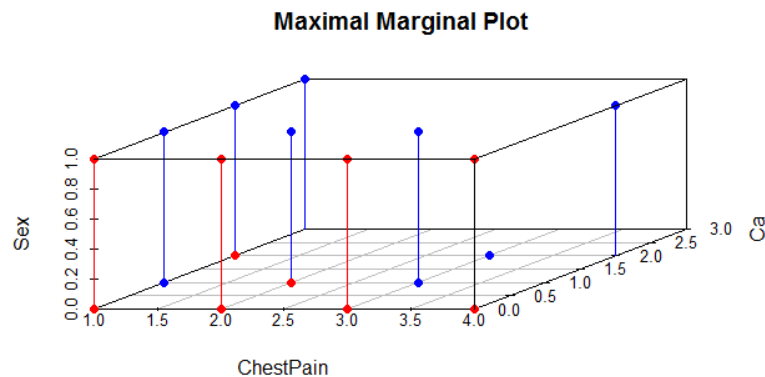
Model 3 is the best Model. SVM better predicts our model than logistic regression.

- Error rate is lesser for SVM than logistic regression because:
Logistic regression focuses on maximizing the probability of the data. The farther the data lies from the separating hyperplane (on the correct side), the highest accuracy of the LR model is.
- An SVM tries to find the separating hyperplane that maximizes the distance of the closest points to the margin (the support vectors). If a point is not a support vector, it doesn't really matter. In this model support vectors are only 121.

Error rate depends on prediction of only these 121 data points only.

Testing Best Model with Test Data:

- When we used the best model to predict the test data, the error rate observed was lower than that of the training data with difference of 4% approximately.



Scatterplot for Test Data

Hence, our model ChestPain + Sex + Ca predicts the AHD with relatively best accuracy.

Reading Data File and Arranging for Analysis

Reading the data file Heart.csv

```
Project<-read.csv("C://Users//Mohit//Desktop//ds510_project//Project_2//Heart.csv")
View(Project)
```

Selecting first 250 rows for Training data

```
Training<-Project[1:250,-1]
rownames(Training) <- seq(length=nrow(Training))
View(Training)
summary(Training)
```

Selecting last 53 rows for Test data

```
Test<-Project[251:303,-1]
View(Test)
rownames(Test) <- seq(length=nrow(Test))
```

Logistic Regression

Model 1: AHD vs all Explanatory Variables.

```
Model_1<-glm(AHD~.,data=Training,family = "binomial")
summary(Model_1)
```

#Probability of AHD=1(Yes) given values of predictor variables

```
ProbModel_1<-predict(Model_1, data= Training, type="response")
ProbModel_0[1:10]
contrasts(Training$AHD) #To check values of Binary Response variable
```


#Checking Error rate

```
PredModel_1=rep("0",53)
PredModel_1[ProbModel_1 > 0.5]="1"
head(PredModel_1)
Tab_1<-table(PredModel_1,ahd_test) #Confusion Matrix
print(Tab_1)
mean(PredModel_1 !=ahd_test)
Error_1<-1-sum(diag(Tab_1))/sum(Tab_1) #Error rate
print(Error_1)
```

#Confidence Interval

```
exp(cbind(OR=coef(Model_1),confint(Model_1)))
```

Model 2: AHD vs Sex, Ca, Chest pain and RestBP .

```
Model_2<-glm(AHD~Sex+Ca+ChestPain+RestBP,data=Training,family = "binomial")
summary(Model_2)
```

#Probability of AHD=1(Yes) given values of predictor variables

```
ProbModel_2<-predict(Model_2, data=Training, type="response")
ProbModel_2[1:10]
```

#Checking Error rate

```
PredModel_2=rep("0",250)
PredModel_2[ProbModel_2 > 0.5]="1"
head(PredModel_2)
Tab_2<-table(PredModel_2,Training$AHD) #Confusion Matrix
print(Tab_2)
Error_2<-1-sum(diag(Tab_2))/sum(Tab_2) #Training Error rate
print(Error_2)
```

Model 3: AHD vs Sex, Ca and Chest pain.

```
Model_3<-glm(AHD~Sex+Ca+ChestPain,data=Training,family = "binomial")
summary(Model_3)
```

#Probability of AHD=1(Yes) given values of predictor variables

```
ProbModel_3<-predict(Model_3, Training, type="response")
ProbModel_3[1:10]
```

#Checking Error rate

```
PredModel_3=rep("0",250)
PredModel_3[ProbModel_2 > 0.5]="1"
head(PredModel_3)
Tab_3<-table(PredModel_3,Training$AHD) #Confusion Matrix
print(Tab_3)
Error_3<-1-sum(diag(Tab_3))/sum(Tab_3) #Training Error rate
print(Error_3)
```

#Confidence Interval

```
exp(cbind(OR=coef(Model_3),confint(Model_3)))
```

Model 4: AHD vs Ca and Chest pain.

AHD vs Ca + Chestpain

```
Model_4<-glm(AHD~Ca+ChestPain,data=Training,family = "binomial")
summary(Model_4)
```

#Probability of AHD=1(Yes) given values of predictor variables

```
ProbModel_4<-predict(Model_4, Training, type="response")
ProbModel_4[1:10]
```

#Checking Error rate

```
PredModel_4=rep("0",250)
PredModel_4[ProbModel_4 > 0.5]="1"
head(PredModel_4)
Tab_4<-table(PredModel_4,Training$AHD) #Confusion Matrix
print(Tab_4)
Error_4<-1-sum(diag(Tab_4))/sum(Tab_4) #Training Error rate
print(Error_4)
```

Model 5: AHD vs ChestPain and Sex

AHD vs Chestpain + Sex

```
Model_5<-glm(AHD~ChestPain+Sex,data=Training,family = "binomial")
summary(Model_5)
```

#Probability of AHD=1(Yes) given values of predictor variables

```
ProbModel_5<-predict(Model_5,type="response")
ProbModel_5[1:10]
```

#Checking Error rate

```
PredModel_5=rep("0",250)
PredModel_5[ProbModel_5 > 0.5]="1"
head(PredModel_5)
Tab_5<-table(PredModel_5,Training$AHD) #Confusion Matrix
print(Tab_5)
Error_5<-1-sum(diag(Tab_5))/sum(Tab_5) #Training Error rate
print(Error_5)
```

Model 6: AHD vs Ca and Sex

AHD vs Ca+Sex

```
Model_6<-glm(AHD~Ca+Sex,data=Training,family = "binomial")
summary(Model_6)
```

#Probability of AHD=1(Yes) given values of predictor variables

```
ProbModel_6<-predict(Model_6,type="response")
ProbModel_6[1:10]
```

#Checking Error rate

```
PredModel_6=rep("0",250)
PredModel_6[ProbModel_6 > 0.5]="1"
head(PredModel_6)
Tab_6<-table(PredModel_6,Training$AHD) #Confusion Matrix
print(Tab_6)
Error_6<-1-sum(diag(Tab_6))/sum(Tab_6) #Training Error rate
print(Error_6)
```

Model 7: AHD vs Sex

AHD vs Sex

```
Model_7<-glm(AHD~Sex,data=Training,family = "binomial")
summary(Model_7)
```

#Probability of AHD=1(Yes) given values of predictor variables

```
ProbModel_7<-predict(Model_7,type="response")
ProbModel_7[1:10]
```

#Checking Error rate

```
PredModel_7=rep("0",250)
PredModel_7[ProbModel_7 > 0.5]="1"
head(PredModel_7)
Tab_7<-table(PredModel_7,Training$AHD) #Confusion Matrix
print(Tab_7)
Error_7<-1-sum(diag(Tab_7))/sum(Tab_7) #Training Error rate
print(Error_7)
```

Model 8: AHD vs RestBP

```
Model_8<-glm(AHD~RestBP,data=Training,family = "binomial")
summary(Model_8)
```

#Probability of AHD=1(Yes) given values of predictor variables

```
ProbModel_8<-predict(Model_8,type="response")
ProbModel_8[1:10]
```

#Checking Error rate

```
PredModel_8=rep("0",250)
PredModel_8[ProbModel_8 > 0.5]="1"
head(PredModel_8)
Tab_8<-table(PredModel_8,Training$AHD) #Confusion Matrix
print(Tab_8)
Error_8<-1-sum(diag(Tab_8))/sum(Tab_8) #Training Error rate
print(Error_8)
```

Model 9: AHD vs Ca

```
Model_9<-glm(AHD~Ca,data=Training,family = "binomial")
summary(Model_9)
```

#Probability of AHD=1(Yes) given values of predictor variables

```
ProbModel_9<-predict(Model_9,type="response")
ProbModel_9[1:10]
```

#Checking Error rate

```
PredModel_9=rep("0",250)
PredModel_9[ProbModel_9 > 0.5]="1"
head(PredModel_9)
Tab_9<-table(PredModel_9,Training$AHD) #Confusion Matrix
print(Tab_9)
Error_9<-1-sum(diag(Tab_9))/sum(Tab_9) #Training Error rate
print(Error_9)
```

Model 10: AHD vs ChestPain

```
Model_10<-glm(AHD~ChestPain,data=Training,family = "binomial")
summary(Model_10)
```

#Probability of AHD=1(Yes) given values of predictor variables

```
ProbModel_10<-predict(Model_10,type="response")
ProbModel_10[1:10]
```

#Checking Error rate

```
PredModel_10=rep("0",250)
PredModel_10[ProbModel_10 > 0.5]="1"
head(PredModel_10)
Tab_10<-table(PredModel_10,Training$AHD) #Confusion Matrix
print(Tab_10)
Error_10<-1-sum(diag(Tab_10))/sum(Tab_10) #Training Error rate
print(Error_10)
```

Predicting the Best Model:

```
Model_21<-glm(AHD~Sex+Ca+ChestPain,data=Test,family = "binomial")
summary(Model_21)
```

#Probability of AHD=1(Yes) given values of predictor variables

```
ProbModel_21<-predict(Model_21, data=Test, type="response")
ProbModel_21[1:10]
```

#Checking Error rate

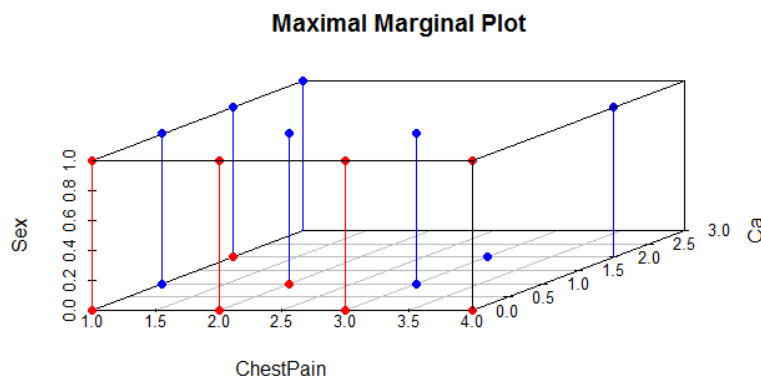
```
PredModel_21=rep("0",53)
PredModel_21[ProbModel_21 > 0.5]="1"
head(PredModel_21)
Tab_21<-table(PredModel_21,Test$AHD) #Confusion Matrix
print(Tab_21)
Error_21<-1-sum(diag(Tab_21))/sum(Tab_21) #Training Error rate
print(Error_21)
```

#Confidence Interval

```
exp(cbind(OR=coef(Model_21),confint(Model_21)))
```

Scatterplot for test data

```
colors <- c("red", "blue", "green")
colors <- colors[as.numeric(Test$AHD)]
scatterplot3d(Test[,c(3,12,2)], pch = 16, type = "h", color = colors, grid = TRUE,
main="Maximal Marginal Plot", sub = "Scatterplot for Test Data")
```



Support Vector Machine

Model 11: AHD vs all Explanatory Variables.

```
SVM_11<-svm(AHD~.,data=Training,Kernel="linear",scale=FALSE)
summary(SVM_11)
```

#Predicting values for Test AHD based on SVM1

```
Predict_11<-predict(SVM_11,Test[!rowSums(is.na(Test)),-14])
head(Predict_11,n=100)
tab_11 <- table(Predict_11,Test[!rowSums(is.na(Test)),14]) #Confusion Matrix
print(tab_11)
Error_11<-1-sum(diag(tab_11))/sum(tab_11) #Test Error rate
print(Error_11)
```

Model 12: AHD vs Sex, Ca, Chest pain and RestBP.

```
SVM_12<-
svm(AHD~Sex+Ca+ChestPain+RestBP,data=Training,Kernel="linear",scale=FALSE)
summary(SVM_12)
```

#Predicting values for Test AHD based on SVM1

```
Predict_12<-predict(SVM_12,Test[!rowSums(is.na(Test)),-14])
head(Predict_12,n=100)
tab_12 <- table(Predict_12,Test[!rowSums(is.na(Test)),14]) #Confusion Matrix
print(tab_12)
Error_12<-1-sum(diag(tab_12))/sum(tab_12) #Test Error rate
print(Error_12)
```

Model 13: AHD vs Sex, Ca and Chest pain.

```
SVM_13<-svm(AHD~Sex+Ca+ChestPain,data=Training,Kernel="linear",scale=FALSE)
summary(SVM_13)
```

#Predicting values for Test AHD based on SVM1

```
Predict_13<-predict(SVM_13,Test[!rowSums(is.na(Test)),-14])
head(Predict_13,n=100)
tab_13 <- table(Predict_13,Test[!rowSums(is.na(Test)),14]) #Confusion Matrix
print(tab_13)
Error_13<-1-sum(diag(tab_13))/sum(tab_13) #Test Error rate
print(Error_13)
```

Model 14: AHD vs Ca and Chest pain.

```
SVM_14<-svm(AHD~Ca+ChestPain,data=Training,Kernel="linear",scale=FALSE)
summary(SVM_14)
```

#Predicting values for Test AHD based on SVM1

```
Predict_14<-predict(SVM_14,Test[!rowSums(is.na(Test)),-14])
head(Predict_14,n=100)
tab_14 <- table(Predict_14,Test[!rowSums(is.na(Test)),14]) #Confusion Matrix
print(tab_14)
Error_14<-1-sum(diag(tab_14))/sum(tab_14) #Test Error rate
print(Error_14)
```

Model 15: AHD vs ChestPain and Sex

```
SVM_15<-svm(AHD~Sex+ChestPain,data=Training,Kernel="linear",scale=FALSE)
summary(SVM_15)
```

#Predicting values for Test AHD based on SVM1

```
Predict_15<-predict(SVM_15,Test[!rowSums(is.na(Test)),-14])
head(Predict_15,n=100)
tab_15 <- table(Predict_15,Test[!rowSums(is.na(Test)),14]) #Confusion Matrix
print(tab_15)
```



```
Error_15<-1-sum(diag(tab_15))/sum(tab_15)          #Test Error rate
print(Error_15)
```

Model 16: AHD vs Ca and Sex

```
SVM_16<-svm(AHD~Sex+Ca,data=Training,Kernel="linear",scale=FALSE)
summary(SVM_16)
```

#Predicting values for Test AHD based on SVM1

```
Predict_16<-predict(SVM_16,Test[!rowSums(is.na(Test)),-14])
head(Predict_16,n=100)
tab_16 <- table(Predict_16,Test[!rowSums(is.na(Test)),14])  #Confusion Matrix
print(tab_16)
Error_16<-1-sum(diag(tab_16))/sum(tab_16)          #Test Error rate
print(Error_16)
```

Model 17: AHD vs Sex

```
SVM_17<-svm(AHD~Sex,data=Training,Kernel="linear",scale=FALSE)
summary(SVM_17)
```

#Predicting values for Test AHD based on SVM1

```
Predict_17<-predict(SVM_17,Test[!rowSums(is.na(Test)),-14])
head(Predict_17,n=100)
tab_17 <- table(Predict_17,Test[!rowSums(is.na(Test)),14])  #Confusion Matrix
print(tab_17)
Error_17<-1-sum(diag(tab_17))/sum(tab_17)          #Test Error rate
print(Error_17)
```

Model 18: AHD vs RestBP

```
SVM_18<-svm(AHD~RestBP,data=Training,Kernel="linear",scale=FALSE)
summary(SVM_18)
```

#Predicting values for Test AHD based on SVM1

```
Predict_18<-predict(SVM_18,Test[!rowSums(is.na(Test)),-14])
head(Predict_18,n=100)
tab_18 <- table(Predict_18,Test[!rowSums(is.na(Test)),14]) #Confusion Matrix
print(tab_18)
Error_18<-1-sum(diag(tab_18))/sum(tab_18) #Test Error rate
print(Error_18)
```

Model 19: AHD vs Ca

```
SVM_19<-svm(AHD~Ca,data=Training,Kernel="linear",scale=FALSE)
summary(SVM_19)
```

#Predicting values for Test AHD based on SVM1

```
Predict_19<-predict(SVM_19,Test[!rowSums(is.na(Test)),-14])
head(Predict_19,n=100)
tab_19 <- table(Predict_19,Test[!rowSums(is.na(Test)),14]) #Confusion Matrix
print(tab_19)
Error_19<-1-sum(diag(tab_19))/sum(tab_19) #Test Error rate
print(Error_19)
```

Model 20: AHD vs ChestPain

```
SVM_20<-svm(AHD~ChestPain,data=Training,Kernel="linear",scale=FALSE)
summary(SVM_20)
```

#Predicting values for Test AHD based on SVM1

```
Predict_20<-predict(SVM_20,Test[!rowSums(is.na(Test)),-14])
head(Predict_20,n=100)
tab_20 <- table(Predict_20,Test[!rowSums(is.na(Test)),14]) #Confusion Matrix
print(tab_20)
Error_20<-1-sum(diag(tab_20))/sum(tab_20) #Test Error rate
print(Error_20)
```