SUPPORT VECTOR MACHINES

Assignment 1

BP: PREDICT THE BURNED AREA OF FOREST FIRES WITH SUPPORT VECTOR MACHINES.

**PROCEDURE:**

STEP 1: First we have to Exploratory Data Analysis which can be done by plotting scattered plot, box plots and summary.

summary(fireforests)

FFMC DMC DC ISI temp

Min. :18.70 Min. : 1.1 Min. : 7.9 Min. : 0.000 Min. : 2.20

1st Qu.:90.20 1st Qu.: 68.6 1st Qu.:437.7 1st Qu.: 6.500 1st Qu.:15.50

Median :91.60 Median :108.3 Median :664.2 Median : 8.400 Median :19.30

Mean :90.64 Mean :110.9 Mean :547.9 Mean : 9.022 Mean :18.89

3rd Qu.:92.90 3rd Qu.:142.4 3rd Qu.:713.9 3rd Qu.:10.800 3rd Qu.:22.80

Max. :96.20 Max. :291.3 Max. :860.6 Max. :56.100 Max. :33.30

RH wind rain area

Min. : 15.00 Min. :0.400 Min. :0.00000 Min. : 0.00

1st Qu.: 33.00 1st Qu.:2.700 1st Qu.:0.00000 1st Qu.: 0.00

Median : 42.00 Median :4.000 Median :0.00000 Median : 0.52

Mean : 44.29 Mean :4.018 Mean :0.02166 Mean : 12.85

3rd Qu.: 53.00 3rd Qu.:4.900 3rd Qu.:0.00000 3rd Qu.: 6.57

Max. :100.00 Max. :9.400 Max. :6.40000 Max. :1090.84

dayfri daymon daysat daysun

Min. :0.0000 Min. :0.0000 Min. :0.0000 Min. :0.0000

1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000

Median :0.0000 Median :0.0000 Median :0.0000 Median :0.0000

Mean :0.1644 Mean :0.1431 Mean :0.1625 Mean :0.1838

3rd Qu.:0.0000 3rd Qu.:0.0000 3rd Qu.:0.0000 3rd Qu.:0.0000

Max. :1.0000 Max. :1.0000 Max. :1.0000 Max. :1.0000

daythu daytue daywed monthapr

Min. :0.000 Min. :0.0000 Min. :0.0000 Min. :0.00000

1st Qu.:0.000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.00000

Median :0.000 Median :0.0000 Median :0.0000 Median :0.00000

Mean :0.118 Mean :0.1238 Mean :0.1044 Mean :0.01741

3rd Qu.:0.000 3rd Qu.:0.0000 3rd Qu.:0.0000 3rd Qu.:0.00000

Max. :1.000 Max. :1.0000 Max. :1.0000 Max. :1.00000

monthaug monthdec monthfeb monthjan

Min. :0.0000 Min. :0.00000 Min. :0.00000 Min. :0.000000

1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.000000

Median :0.0000 Median :0.00000 Median :0.00000 Median :0.000000

Mean :0.3559 Mean :0.01741 Mean :0.03868 Mean :0.003868

3rd Qu.:1.0000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.000000

Max. :1.0000 Max. :1.00000 Max. :1.00000 Max. :1.000000

monthjul monthjun monthmar monthmay

Min. :0.0000 Min. :0.00000 Min. :0.0000 Min. :0.000000

1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.0000 1st Qu.:0.000000

Median :0.0000 Median :0.00000 Median :0.0000 Median :0.000000

Mean :0.0619 Mean :0.03288 Mean :0.1044 Mean :0.003868

3rd Qu.:0.0000 3rd Qu.:0.00000 3rd Qu.:0.0000 3rd Qu.:0.000000

Max. :1.0000 Max. :1.00000 Max. :1.0000 Max. :1.000000

monthnov monthoct monthsep

Min. :0.000000 Min. :0.00000 Min. :0.0000

1st Qu.:0.000000 1st Qu.:0.00000 1st Qu.:0.0000

Median :0.000000 Median :0.00000 Median :0.0000

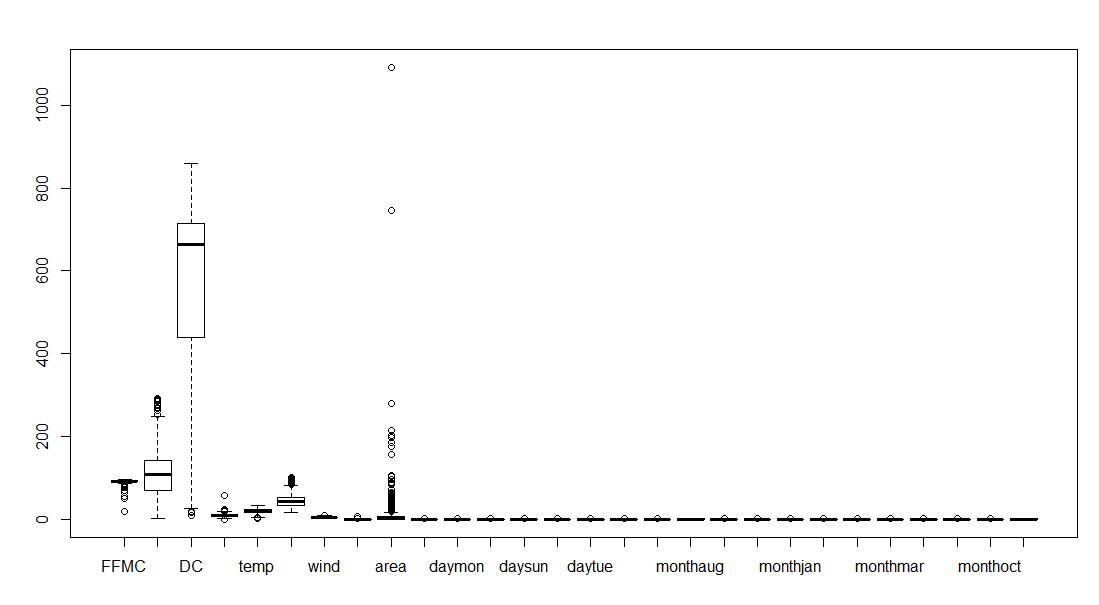
Mean :0.001934 Mean :0.02901 Mean :0.3327

3rd Qu.:0.000000 3rd Qu.:0.00000 3rd Qu.:1.0000

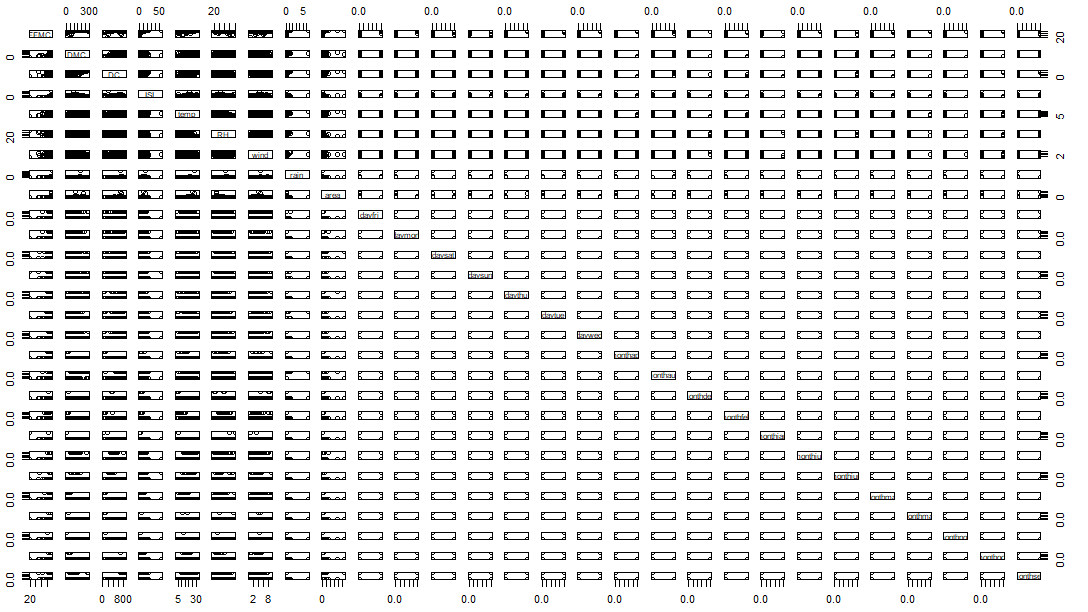
Max. :1.000000 Max. :1.00000 Max. :1.0000

From the summary we can see that in FFMC, DC and Area the difference between mean and max is large so it may be right skewed it can my confirmed by using Box Plot .

It can be confirmed that FFMC, DC and Area is right Skewed.



Scatter Plot for the Following data frame is:



STEP 2: Predicting the data using SVM techniques

Normalizing the data:

|  |
| --- |
| normalize<-function(x){  + return ( (x-min(x))/(max(x)-min(x)))  + }  >  fireforests\_norm<-as.data.frame(lapply(fireforests,FUN=normalize))  After normalization we can check by using summary :  > summary(fireforests\_norm$area)  Min. 1st Qu. Median Mean 3rd Qu. Max.  0.0000000000 0.0000000000 0.0004766969 0.0117774303 0.0060228814 1.0000000000 |
|  |
| |  | | --- | |  | |

|  |
| --- |
| Now splitting the data into test and train :  > fireforests\_train <- fireforestsfinal[1:400,]  > fireforests\_test <- fireforestsfinal[401:517,] |
|  |
| |  | | --- | |  | |

Before building the model we need to install some packages required for SVM formation:

|  |
| --- |
| install.packages("kernlab") |
| library(kernlab) |

Now after splitting the data into test and train , we can Build model on training data which is given by ( here we use vanilla dot mode ):

fireforests\_model <- ksvm(area ~ ., data=fireforests\_train , kernel="vanilladot")

Now predicting the model using some kernel techniques :

fireforests\_prediction <- predict(fireforests\_model , fireforests\_test)

> head(fireforests\_prediction)

[,1]

1 -0.8668609

2 0.8181923

3 2.3311731

4 2.2311499

5 1.1117748

6 1.7673233

> table(fireforests\_prediction , fireforests\_test$area)

fireforests\_prediction 0 0.33 0.43 0.54 0.72 1.23 1.29 1.63 1.76 1.94 1.95 2 2.07

-5.4132100369708 0 0 0 0 0 0 0 0 0 0 0 0 0

-1.09789679863193 1 0 0 0 0 0 0 0 0 0 0 0 0

-1.06394827265261 0 0 0 0 0 0 0 0 0 0 0 0 0

-0.961487598625538 0 0 0 0 0 0 0 0 0 0 0 0 0

-0.86686086294424 0 0 0 0 0 0 0 0 0 0 0 0 0

-0.739449440749524 1 0 0 0 0 0 0 0 0 0 0 0 0

-0.731193010562424 0 0 0 0 0 0 0 0 0 0 0 0 0

-0.578269129140542 0 0 0 0 0 0 0 0 0 0 0 0 0

-0.279746920667455 0 0 0 0 0 0 0 0 0 0 0 0 0

-0.0505272657091105 0 0 0 0 0 0 0 0 0 0 0 0 0

-0.0490100666529445 0 0 0 0 0 0 0 0 0 0 0 0 1

0.000879275337150265 1 0 0 0 0 0 0 0 0 0 0 0 0

0.180829873875901 0 0 0 0 0 0 0 0 0 0 0 0 0

0.198522295210461 1 0 0 0 0 0 0 0 0 0 0 0 0

fireforests\_prediction 2.14 2.17 2.21 2.35 2.44 2.75 2.77 3.05 3.18 3.19 3.2 3.32

-5.4132100369708 0 0 0 0 0 0 0 0 0 0 0 0

-1.09789679863193 0 0 0 0 0 0 0 0 0 0 0 0

-1.06394827265261 0 0 0 0 0 0 0 0 0 0 0 0

-0.961487598625538 0 0 0 0 0 0 0 0 0 0 0 0

-0.86686086294424 0 0 0 0 0 0 0 0 0 0 0 0

-0.739449440749524 0 0 0 0 0 0 0 0 0 0 0 0

-0.731193010562424 0 0 0 0 0 0 0 0 0 0 0 0

-0.578269129140542 0 0 0 0 0 0 0 0 0 0 0 0

-0.279746920667455 0 0 0 0 0 0 0 0 0 0 0 0

-0.0505272657091105 0 0 0 0 0 0 0 0 0 0 0 0

-0.0490100666529445 0 0 0 0 0 0 0 0 0 0 0 0

0.000879275337150265 0 0 0 0 0 0 0 0 0 0 0 0

0.180829873875901 0 0 0 0 0 0 0 0 0 0 0 0

0.198522295210461 0 0 0 0 0 0 0 0 0 0 0 0

fireforests\_prediction 3.35 3.71 3.95 4.62 4.95 4.96 5.39 5.55 5.8 6.04 6.3 6.36

-5.4132100369708 0 0 0 0 0 0 0 0 0 0 0 0

-1.09789679863193 0 0 0 0 0 0 0 0 0 0 0 0

-1.06394827265261 0 0 0 0 0 0 0 0 0 0 0 0

-0.961487598625538 0 0 0 0 0 0 0 0 0 0 0 0

-0.86686086294424 0 0 0 0 0 0 0 0 0 0 0 0

-0.739449440749524 0 0 0 0 0 0 0 0 0 0 0 0

-0.731193010562424 0 0 0 0 0 0 0 0 0 0 0 0

-0.578269129140542 0 0 0 0 0 0 0 0 0 0 0 0

-0.279746920667455 0 0 0 0 0 0 0 0 0 0 0 0

-0.0505272657091105 0 0 0 0 0 0 0 0 0 0 0 0

-0.0490100666529445 0 0 0 0 0 0 0 0 0 0 0 0

0.000879275337150265 0 0 0 0 0 0 0 0 0 0 0 0

0.180829873875901 0 0 0 0 0 0 0 0 0 0 0 0

0.198522295210461 0 0 0 0 0 0 0 0 0 0 0 0

fireforests\_prediction 6.43 6.44 6.61 6.84 7.02 7.36 7.8 8.16 8.59 9.71 9.96 10.08

-5.4132100369708 0 0 0 0 0 0 0 0 0 0 0 0

-1.09789679863193 0 0 0 0 0 0 0 0 0 0 0 0

-1.06394827265261 0 1 0 0 0 0 0 0 0 0 0 0

-0.961487598625538 0 0 0 0 0 0 0 0 0 0 0 0

-0.86686086294424 0 0 0 0 0 0 0 1 0 0 0 0

-0.739449440749524 0 0 0 0 0 0 0 0 0 0 0 0

-0.731193010562424 0 0 0 0 0 0 0 0 0 0 0 0

-0.578269129140542 0 0 0 0 0 0 0 0 0 0 0 0

-0.279746920667455 0 0 0 0 0 0 0 0 0 0 0 0

-0.0505272657091105 0 0 0 0 0 0 0 0 0 0 0 1

-0.0490100666529445 0 0 0 0 0 0 0 0 0 0 0 0

0.000879275337150265 0 0 0 0 0 0 0 0 0 0 0 0

0.180829873875901 0 0 0 0 0 0 0 0 0 0 0 0

0.198522295210461 0 0 0 0 0 0 0 0 0 0 0 0

fireforests\_prediction 10.82 11.16 14.68 15.34 16.4 26.43 38.48 40.54 43.32 46.7

-5.4132100369708 1 0 0 0 0 0 0 0 0 0

-1.09789679863193 0 0 0 0 0 0 0 0 0 0

-1.06394827265261 0 0 0 0 0 0 0 0 0 0

-0.961487598625538 0 0 0 0 0 0 0 0 0 0

-0.86686086294424 0 0 0 0 0 0 0 0 0 0

-0.739449440749524 0 0 0 0 0 0 0 0 0 0

-0.731193010562424 0 0 0 0 0 0 0 0 0 0

-0.578269129140542 0 1 0 0 0 0 0 0 0 0

-0.279746920667455 0 0 0 0 0 1 0 0 0 0

-0.0505272657091105 0 0 0 0 0 0 0 0 0 0

-0.0490100666529445 0 0 0 0 0 0 0 0 0 0

0.000879275337150265 0 0 0 0 0 0 0 0 0 0

0.180829873875901 0 0 0 0 0 0 1 0 0 0

0.198522295210461 0 0 0 0 0 0 0 0 0 0

fireforests\_prediction 49.59 54.29 61.13 70.32 82.75 185.76 278.53 746.28

-5.4132100369708 0 0 0 0 0 0 0 0

-1.09789679863193 0 0 0 0 0 0 0 0

-1.06394827265261 0 0 0 0 0 0 0 0

-0.961487598625538 0 0 0 1 0 0 0 0

-0.86686086294424 0 0 0 0 0 0 0 0

-0.739449440749524 0 0 0 0 0 0 0 0

-0.731193010562424 0 1 0 0 0 0 0 0

-0.578269129140542 0 0 0 0 0 0 0 0

-0.279746920667455 0 0 0 0 0 0 0 0

-0.0505272657091105 0 0 0 0 0 0 0 0

-0.0490100666529445 0 0 0 0 0 0 0 0

0.000879275337150265 0 0 0 0 0 0 0 0

0.180829873875901 0 0 0 0 0 0 0 0

0.198522295210461 0 0 0 0 0 0 0 0

agreement <- fireforests\_prediction == fireforests\_test$area

> table(agreement)

agreement

FALSE

117

> prop.table(table(agreement))

agreement

FALSE

1

The Above prediction shows only FALSE results , so it not accurate we can check with other models .