PROJECT-3 REPORT

DATA ANALYTICS - CS40003

TOPIC -3.3

Roll no: 16CS30033   Name: Shaikh Moin Dastagir

Mob no: 7709400624

- First to check whether the data is a linear fit or not I applied the Linear Regression Model to it

  $y = a + bx$

  $a = 5.854613$ $b = 1.004402$

- I calculated the R^2 (R square value) and found it to be 0.934 which is very close to 1 , so the linear regression is indeed a good fit and hence the ERP vs PRP is linearly related.

- The data plot for linear regression using the above method with the line is shown below.

Plot of PRP vs ERP
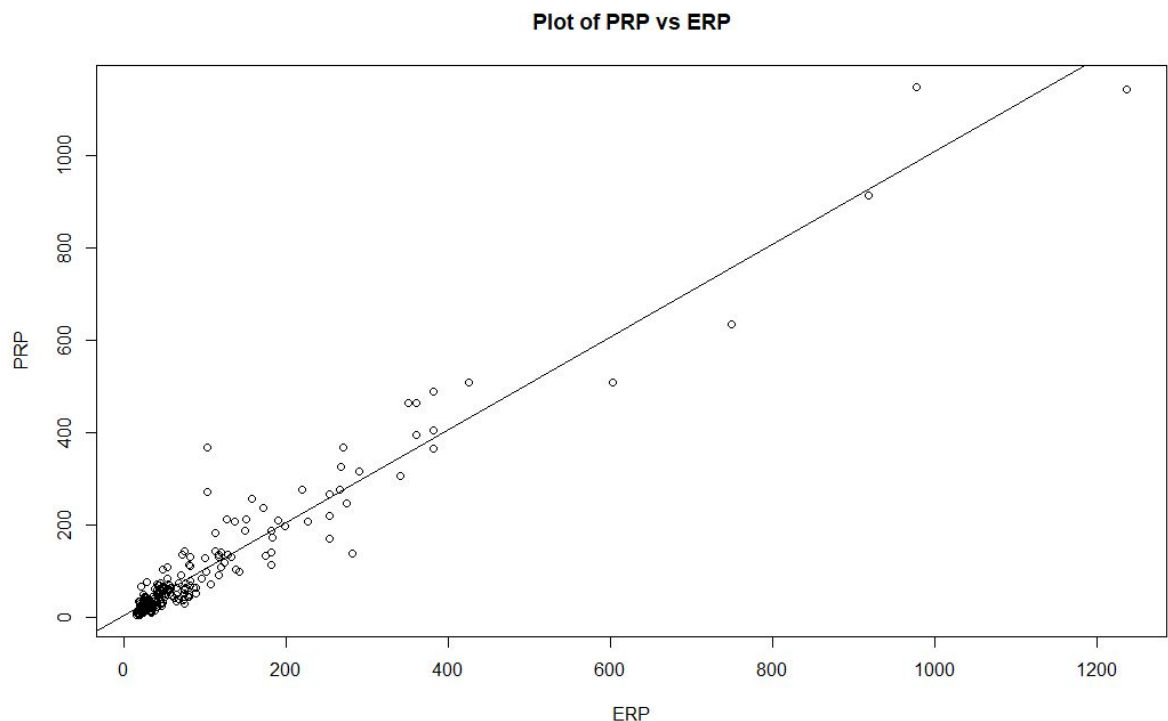
- For non linear regression I calculated using the inbulit function lm(), For degree 2 non linear regression : R2= 0.9354983

  For Degree 3 non linear regression : R2 = 0.9355019

Coefficients :   x =1.084284e+00   x2=-7.699776e-05
x3=-1.652821e-08 constant=9.083695e-01

- The plot for degree 3 non linear regression with the fitted smooth curve as per the regression model is shown below

Plot of PRP vs ERP
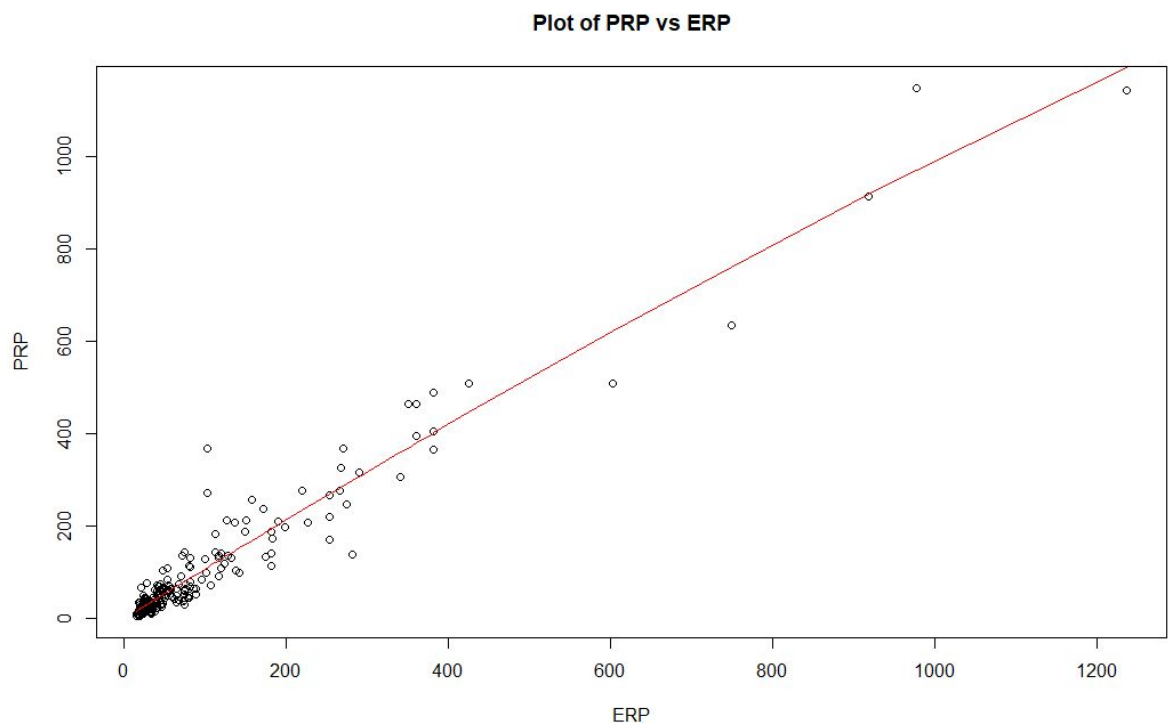
- Conclusion : The value of the coefficients obtained from the non linear fit is low and although the the value of R2 in Nonlinear fit obtained is slightly greater than that of linear fit , the plot of the regression line looks almost identical in both the cases , so the the non linear fit is trying to over fit the data, it is considering the noise in the data with it and overfitting the data .
- Finally , non linear fit is better , although it would be the best if the data has less noise with it .