Linear regression

To predict continuous outcome ariable based in one or more predictable variables

Supervised/ unsupervised

Classification/ regression

Clustering/ classification

Table interpretation from lr model output// rsquared

Linear regression table output

Mlr model/ multiple logistic model

--

Attach for internal read for external

Head prints first six records

Has a response variable ==1/0

Dot means all variable are considered

Check p value for significance if closer to 0 more significant and has \* sign

If coefficient is positive then the significance increases

Annova model/ regression sumsquare/ residual sumsquare/r square value

Y= mx+b y= dependent variable, m= slope

Split data into train and test for cross validation

For linear algorithm we use algorithms like naïve bayes, svm, dt

Use independent variables for testing

Model accuracy is calculated through actual and predicted value/ the residual value

Overfitting= train data with more training input like 90/10

Categorical data—yes/no

Confusion matrix for categorical data

Ture positive/ negative, false positive/negative

Calculate average accuracy for final accurate prediction

Y= mx+b y= mpg, mx= hp, b= intercept

Higer the degree of polynomial the lesser the mse/ residual standard error