multiple linear regression

features: 'n' number of features

label: I label -> numeric

Ventual Capitalist Use-case:

RLDSpend -> numeric 41

Administration -> numeric f2

marketing -> numeric f3

State - Categolical f4

Brofit ______ numeric label

After handling categorical data,

California

Plorida

New York

RSD Spend

Administration

marketing

Bobit

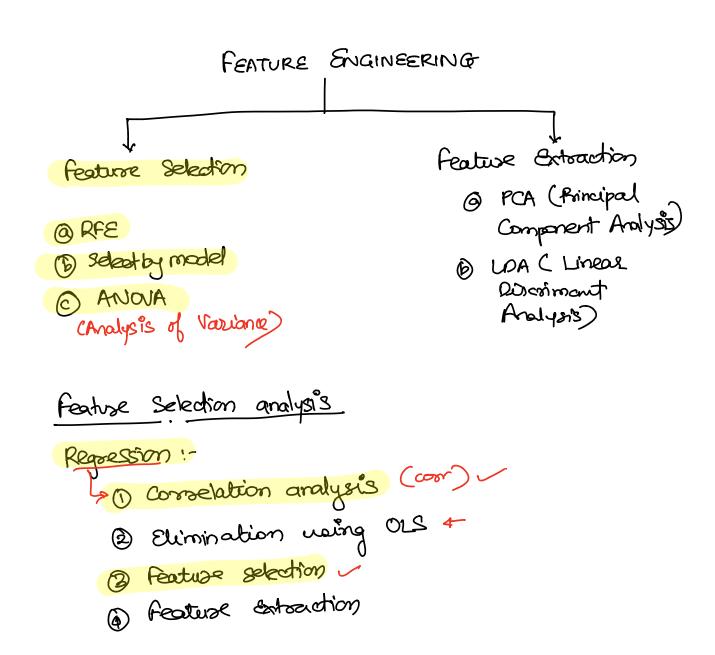
Equation of line goes:

Profit = bo + bi [(alifornia]

+ bz [Florida] + b3 [NY]

+ 64 [R&D Spend] + 65 [Adm]

4 b6 [morres]



Backward elimination using OLS. p-value - 340.054 essar hon con tolerate) confidence > 0.95+ (the trust you can have on the model) P Z = 0.05 -> Ral"ship enists so keep the - Diminde kette my significance level is

| P = 0.05 | Data Scientist Rde of ML engg: a Eliminate the fatures whose praluc is greater than 0.05

Ideal steps us perform in feature elimination using OLS! AIC

(1) Perform All-IN (erosse you provide all coeff and intexcept coeff)

partit = 10 bo + bo (California) + bo (Florida) + bo (NY) +

(RRO) + bo (adm) + bo (marok)

- @ Decide the Significance level (0.05)
- @ sdeet the feature in summary which how highest p-value.
- @ If proble > SL(0.05), eliminate that feature.
- 5 repeat step 3 & 4 till condition (1) Is false, go to step 6.
- © consider all semaining features as selected feature for model building.