Multicollinear and VIF and bias

~~1) diff bet Multicollinearity vs Correlation~~ — { 0.7, 0.8

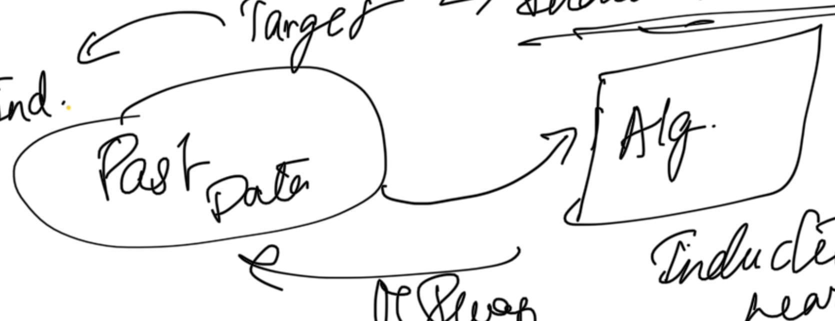
~~2) future data~~ | we have some cols — No. New cols — VIF | In future if we see no. M.C how to adjust

3) Multicollinear → Feature Engineering
- Combination of cols
- Ratios
- How M.C.

4) Productionizing → handling saving of models when we Extra data

---

Class - Linear of algorithms → Type of learning
↳ Inductive learning



N Images

Task Classification
Male
female.

Data to see

Male Data

90 : 10

90%

female Data → None of
90:10 → 90% images
Male Female ?
90
90

Metric / Comparison Parameter

We will need More data → Make the predict effectively.

Some assumptions / Inductive Bias !!!

① Multicollinearity ② Homoscedasticity ③ Linearity
④ Normal Distribution of Residuals

VIF

$$* \quad y = m_1 x_1 + m_2 x_2 + c \; [+E]$$

Slope Intercept - y

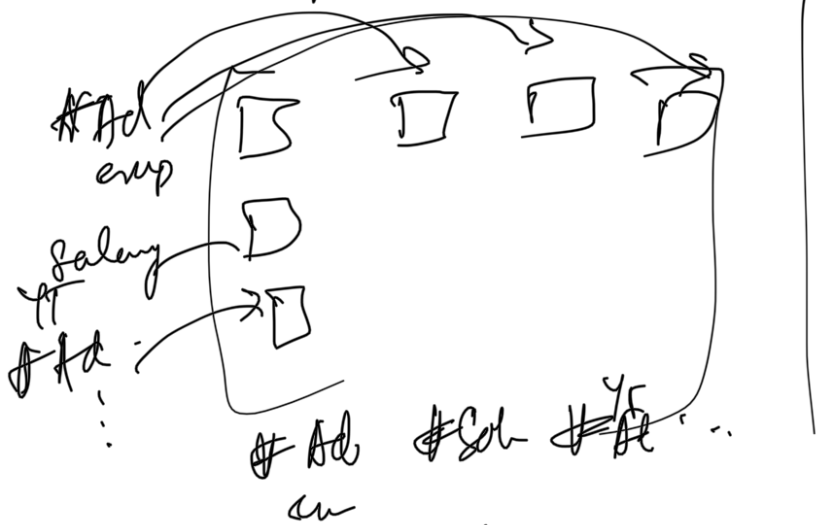
How much with that variable affect the Rise/fall of any target.

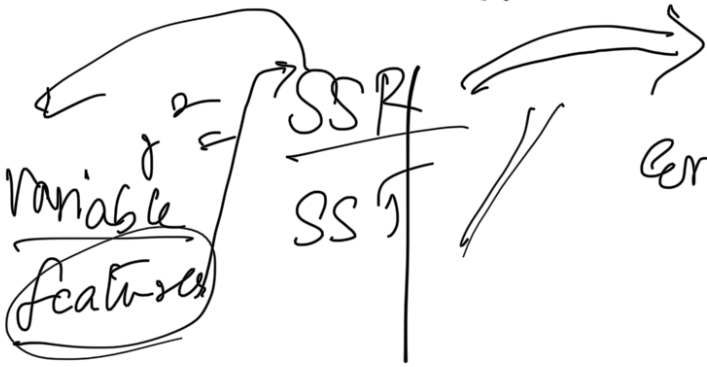
#AD expenditure   # Youtube AD expenditure x_1

$$y = m_1 x_1 + m_2 x_2 + m_3 x_3 + \ldots$$

0.86↑   # Salary   0.39↑

* Where there is multicollinearity the scale of the coefficients and it changes & the computation fn becomes unreliable

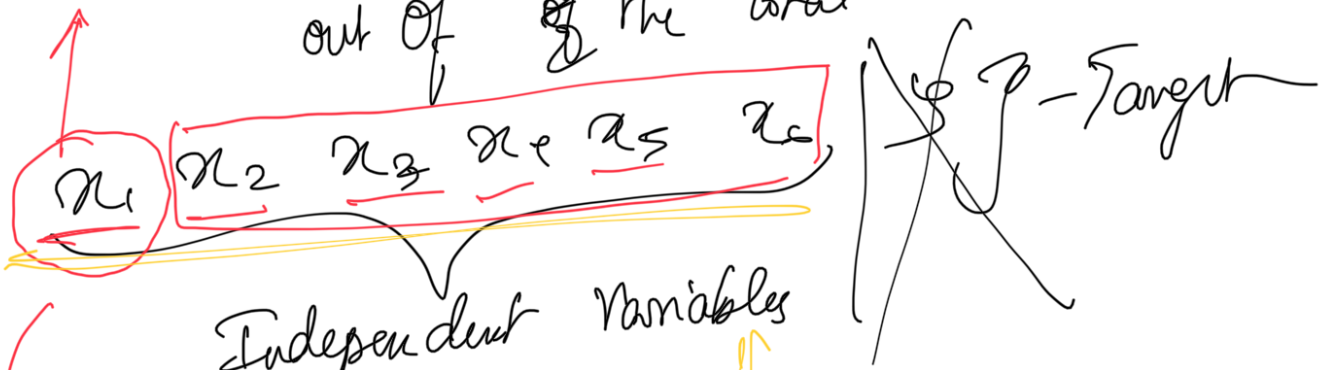




How much of my error is explained by the Relationship

out of the total

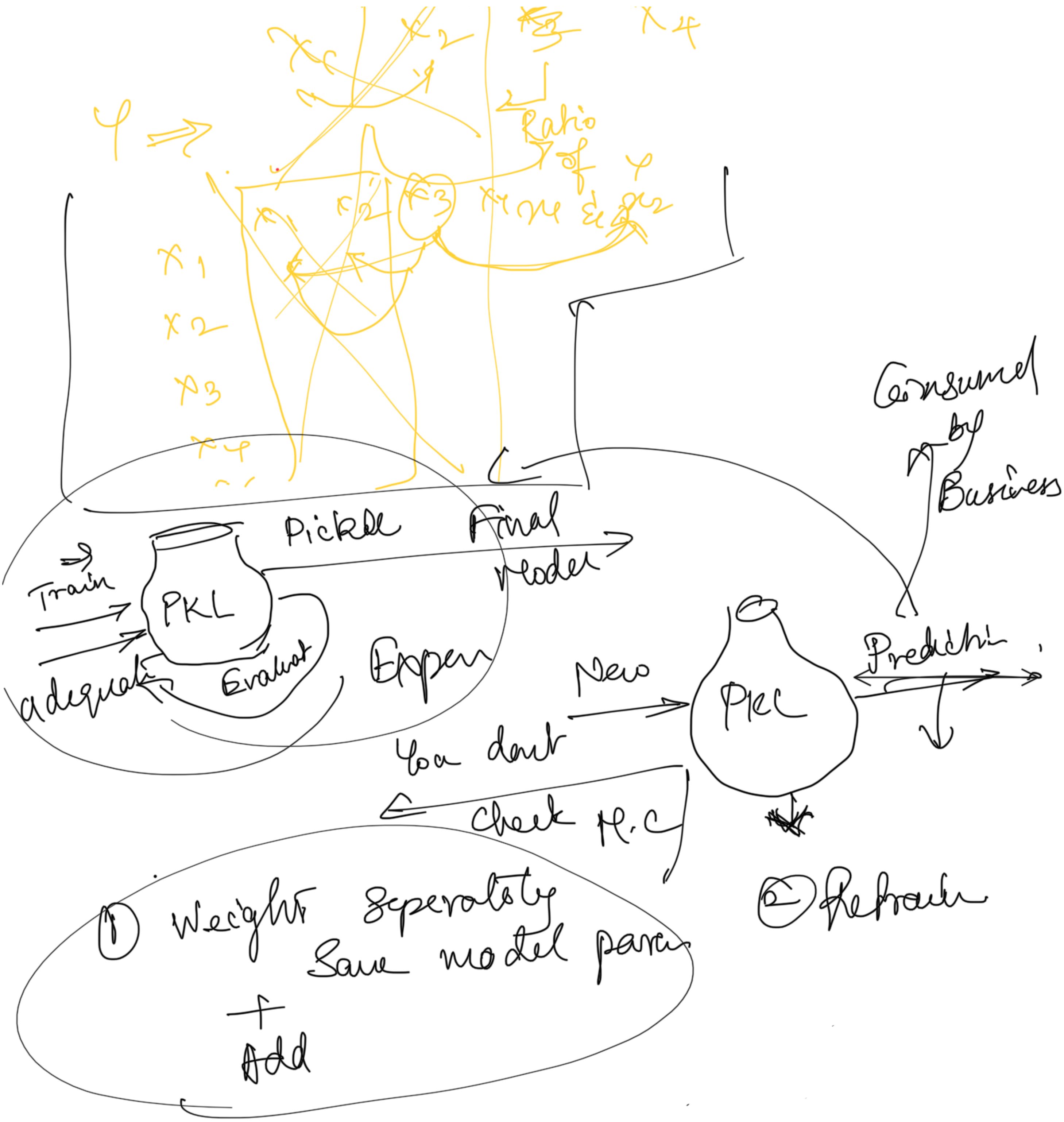
Reg line



Independent Variables

How much inflation of Variance in x_1 can be explained by other variables.

R^2



Y
Ratio of Y
Pickle
Final model
Consumed by Business
Train
PKL
Evaluat
adequat
Expen
New
You don't
PKL
Prediction
Check M.C
① Weight seperately
Save model param
+ Add
② Retrain