Dimensionality

- refers to the number of features AKA Dimentions, variables or attributes in a data set

Ex: in this ex the data set has 6 featurs So dimentionality = 6

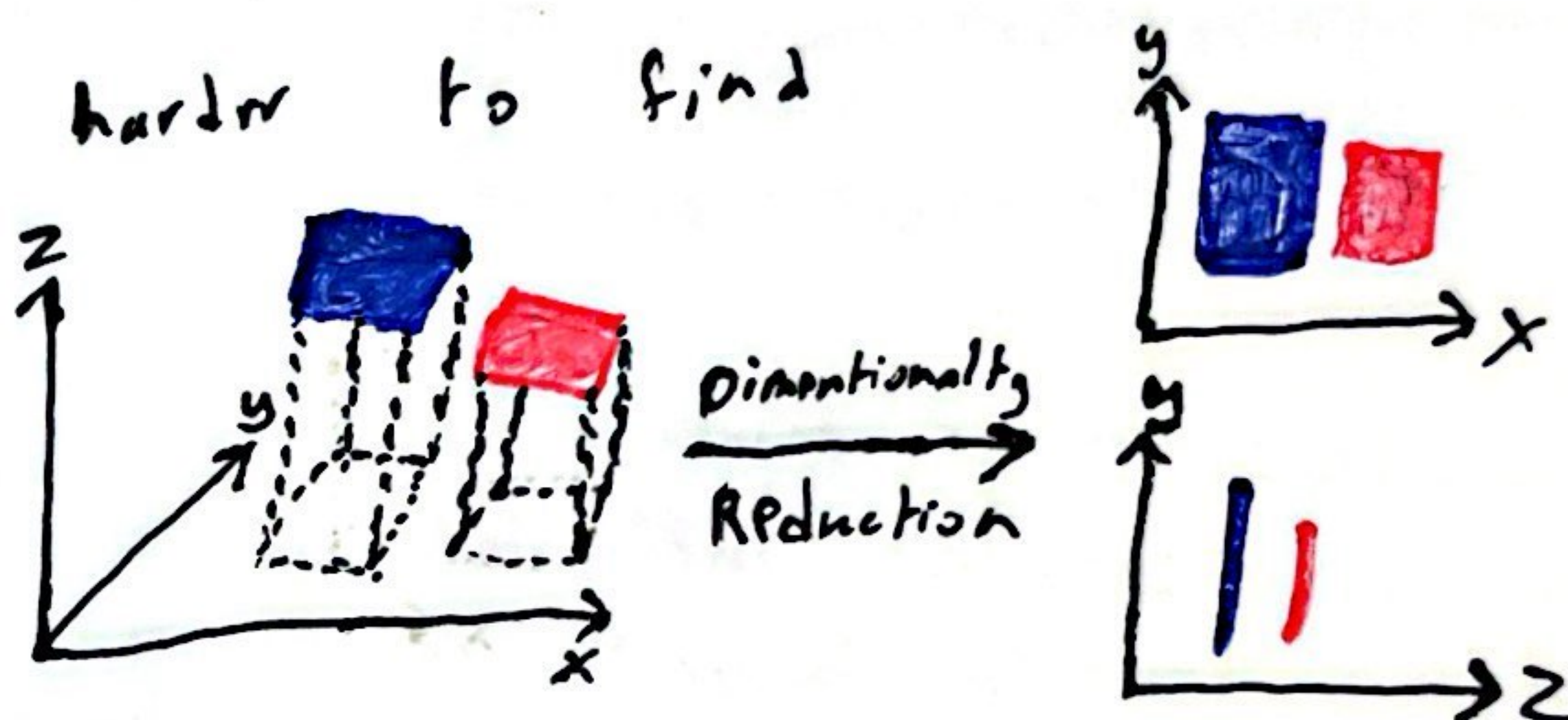• the observations = entries/instance /Sample / Example

index col is, Not a Feature

6 Features

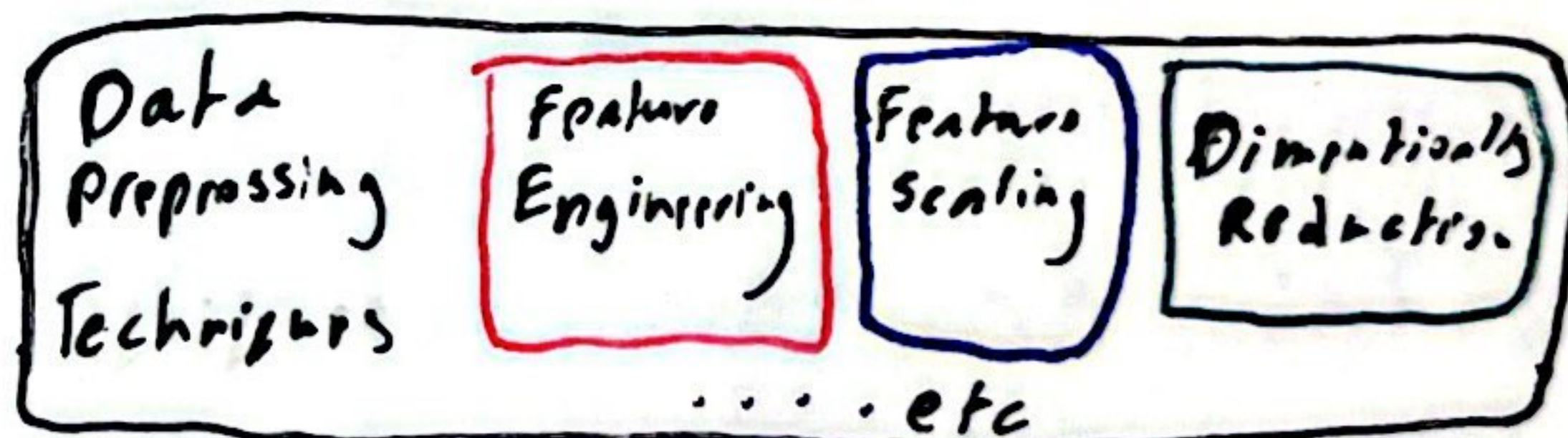|	Name	age	Height	weight	Gender	city
0	Alice	25	165	60	Female	toronto
1	BoB	30	180	70	Male	NYC
2	Ali	20	168	58	Male	toronto

observations

* High Dimentional Data, having many Features can pose unique Challenges, called "The curse of Dimentionality"

• as Dimentions increse → Data Becomes more sparse and patterns Become harder to find

• this is why Dimentionality Reduction is important "see Dimentionality reduction" Topic

Dimentionality Reduction

• the past few topics and more all all part of Data preprossessing

Data Preprossing Techniques | Feature Engineering | Feature Scaling | Dimentionally Reduction ....etc

Target (output) • only for supervised learning

- is what a ML model is trying to predict based on the features (input). For Ex in the house price prediction model from "Model" chapter the target is the price of the house while sqft is input or in email detector its target : spam or not spam

• Breakdown of terms

* in labeled data, traing data must include Both Features and corresponding target val

input, Feature (Var 1) → [Model] → output, target (Var2)
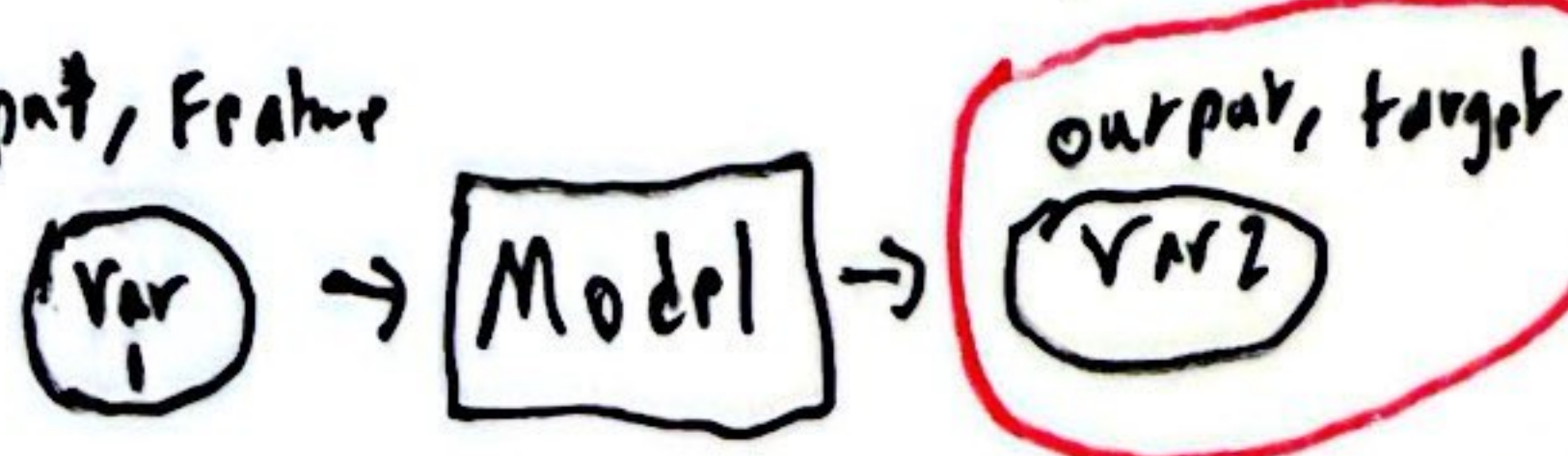
Term	used in	Meaning
target/lable	General ML	Value model is trying to predict
Dependent var	Stats/regression	Var that depends on input Features
Class	Classification tasks	cat or group input Belongs to eg: cat /Dog
output	NN/DL	final result of Model
Response var	stats	AKA Dependent Var

Scanned with CamScanner