Agenda
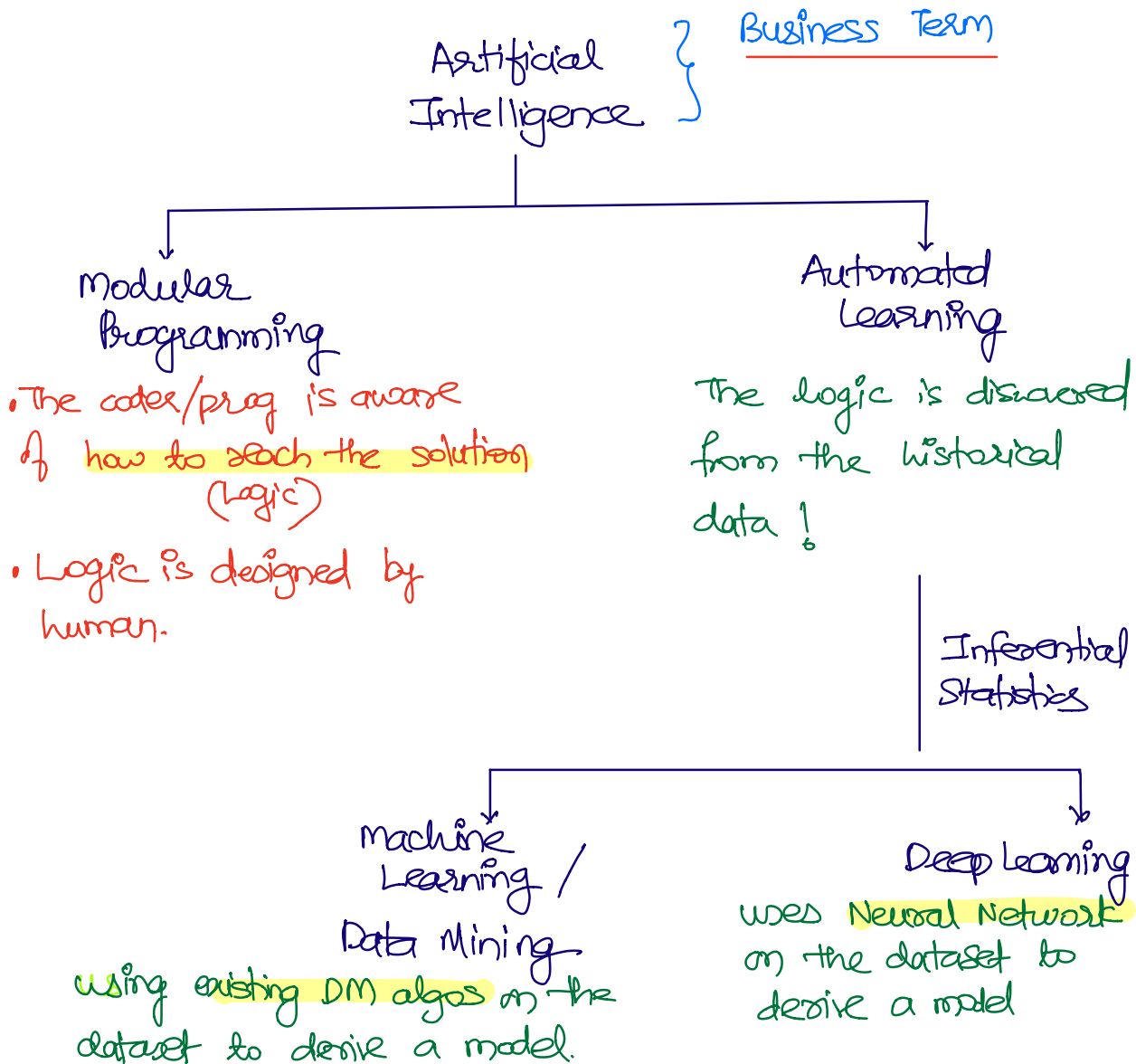
model Building and Evaluation

Inferential Statistics

Prediction, Discovery, Association, Grouping/Categorizing/Classification

Artificial Intelligence } **Business Term**

Artificial Intelligence splits into:

Modular Programming

- The coder/prog is aware of how to reach the solution (Logic)
- Logic is designed by human.

Automated Learning

The logic is discovered from the historical data !

Inferential Statistics

Automated Learning splits into:

Machine Learning / Data Mining

using existing DM algos on the dataset to derive a model.

Deep Learning

uses Neural Network on the dataset to derive a model

Algo refers to the standard data mining algorithms.

Machine Learning

uses algo. to get the logic from the historical data.

Supervised Learning
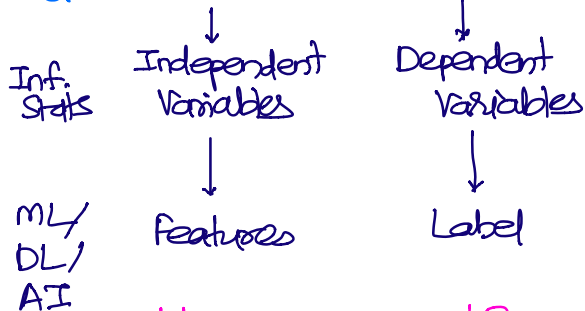
col1	col2	col3
2	2	4
3	2	6
12	1	12
13	2	26

Xtrain ytrain
Xtest ytest

50:50 split

Extracting Logic
↓
model

"Exploring all generated logics" ← Unsupervised Learning

Discovering
- Pattern
- Association

col1	col2	col3	col4
1	2	3	4
5	6	7	8
9	11	17	22

random dataset with hidden domain meaning

expects your dataset to have Question and Answer

Inf. Stats

Independent Variables
↓

Dependent Variables
↓

ML/ DL/ AI

Features

Label

col1, col2

col3

rdSpend, admSpend, markSpend, state

Profit

Phases of Machine Learning / Inferential Stats / DL

① Training Phase / Convergence Phase

converge the model (Learning)

Set [Features] ⟶
Label ⟶

ML algo
DL algo
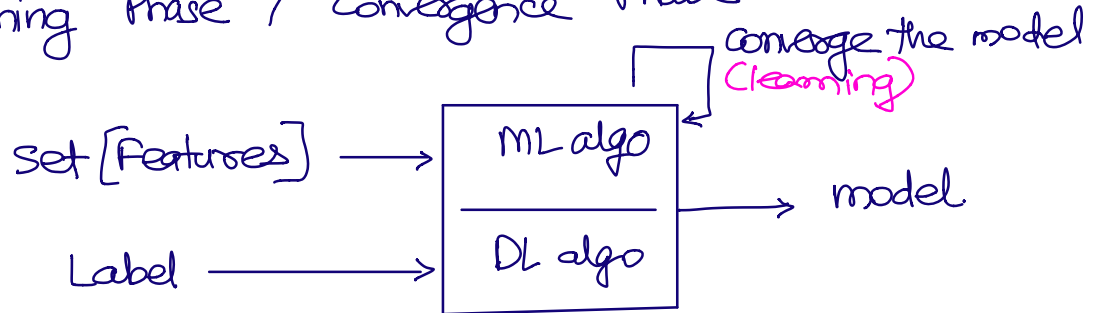
⟶ model

Goal of training phase is to create a model.

② Quality Phase / QA phase

Unknown data ↓

Set [features] ⟶

Trained model
(Converged model)

⟶ Predicted Label

Test Set QA which wasn't used during training phase.

Actual Label ⟶

Calc accuracy or Calc error

⟶ % accuracy
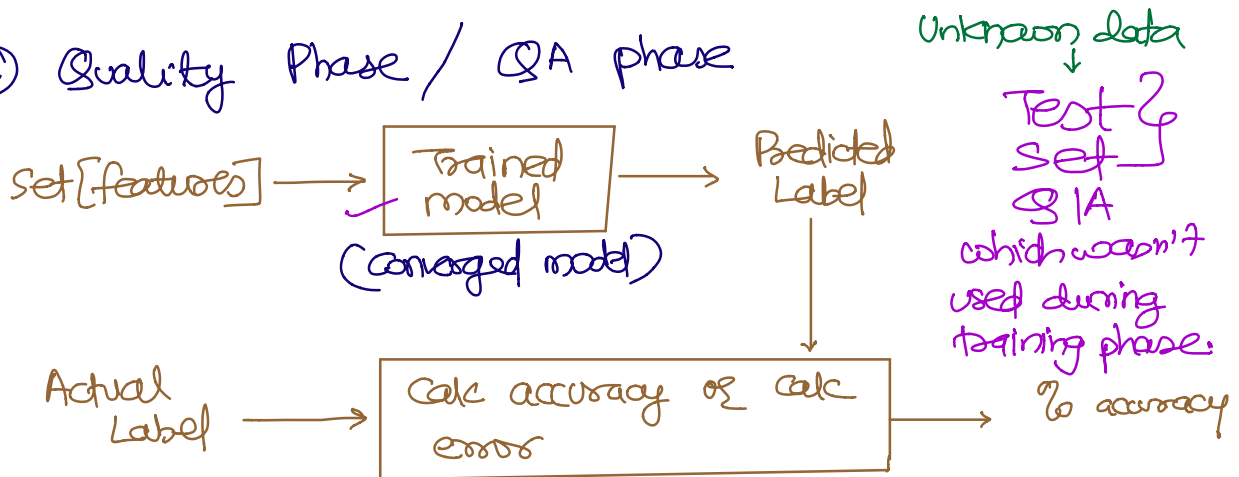
Goal of QA is to understand the reliability of the model's prediction.

③ Deployment :

* SUSPENSE *

RULE

ML
Supervised
Learning Algorithms

Observe Label
Column !

Regression
Algorithms

Label → Continuous
Numerical
Data.

Classification
Algorithms

Label → Categorical,
Ordinal,
Discrete