"Time Series" Analytics & Forecasting

LSTM
Deep learning
TSF

Past time data which would be in sequence/series (ascending)

Analyse

Future predictn

Python

ML

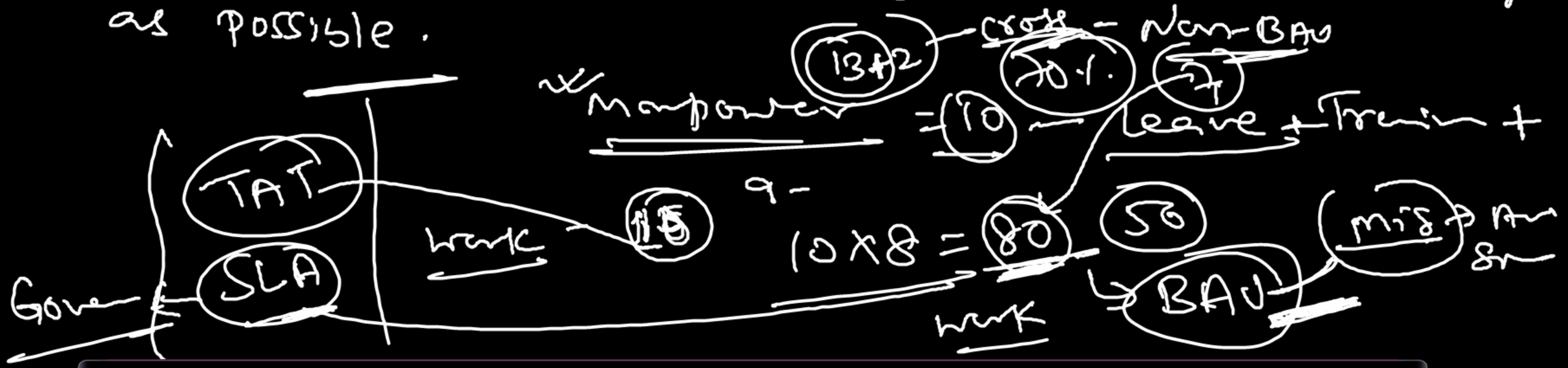
⇒ Introductn.

⇒ case study

→ Forecasting is one of the most Significant & frequently addressed problem in analytics. Inaccurate forecasting can have significant impact on both top line & bottom line of an organizatn..

Example :- Non-availability of a product in the market can result in customer dis-satisfn.

Whereas, too much inventory can erode the organization's Profit. Thus, it become necessary to forecast the demand for a product & services as accurately as possible.



SLA - 80 hrs - Manpower

1 Person = 45 hrs | 48 hrs

9 hrs

8 hrs

Emp :- $\frac{80}{8} = 10$ manpower

buff —

1 Leave + Training + 1

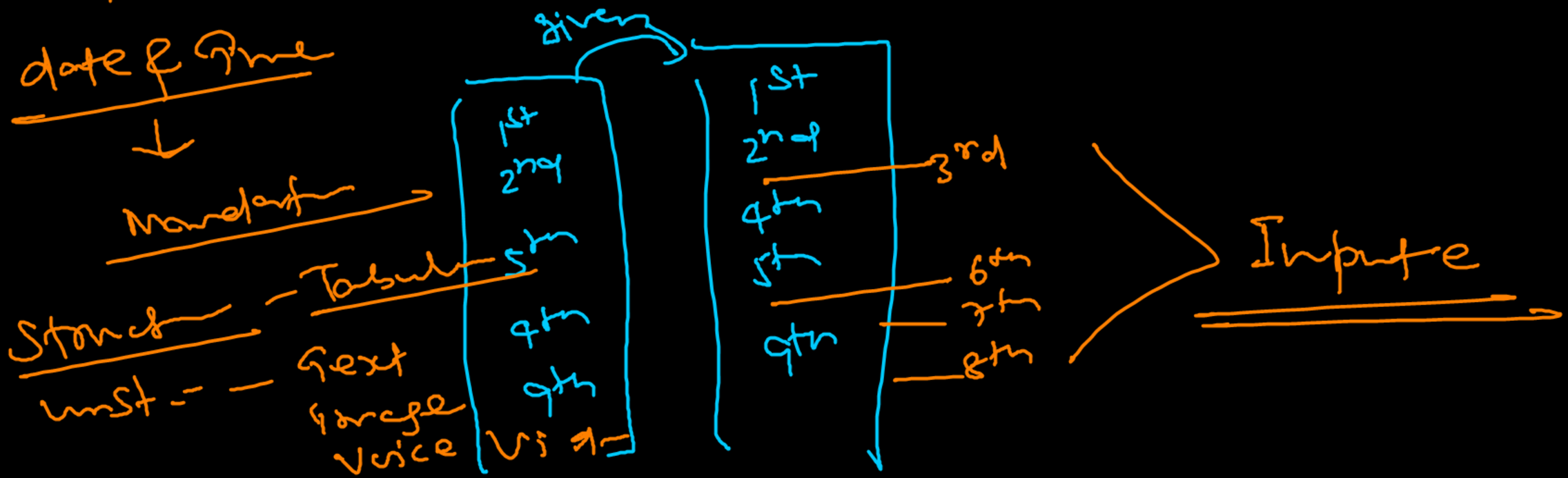
70 60

90

3 ≡ 13

Cross-sectional data :-

- Time series data - univariate or multivariate
  - Time in a sequential manner -



Why forecasting?

Control – failure of technology, inflation, recession, & change in Government law.

Pen ₹50 - ₹500

→ Every business operates under risk & uncertainty

→ Forecast is necessary to lessen the adverse effect of risk. – Mitigate

→ Regression, Classifier, Data Mining

future

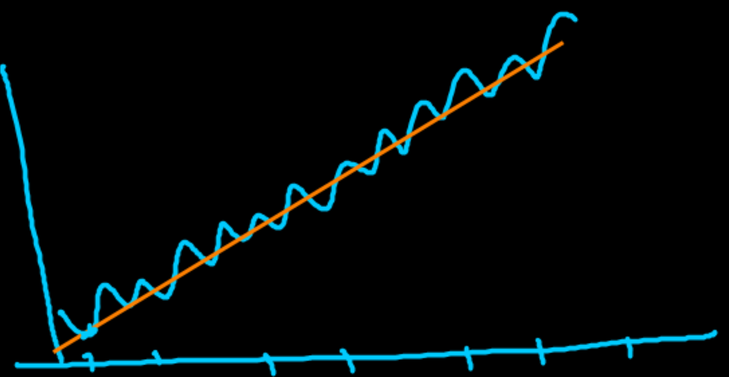
$$Y = f(x)$$

past

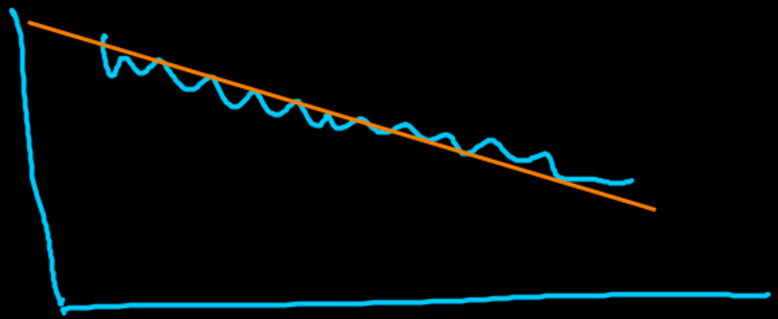
How ?

Components

① Trend (Long Term movement)
↳ more than one year



Trend (T_t) - When the series increases or decreases over the entire length of time.



45 days - 3 months

Appreciation of Dollar vs rupees

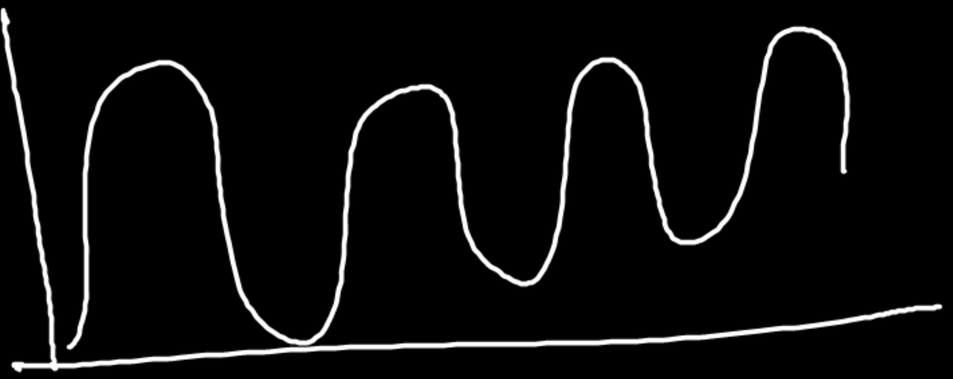
GDP

GDP - 7% - base year 2023
24 - 8%
25% - 65/26 = 2.5%
7%

NVIDIA

2012
$3
$3 in 2023
2023 → 2024, 25, 26

(2) Seasonality (S_t) – (Short term) – 1 year



Ac demand

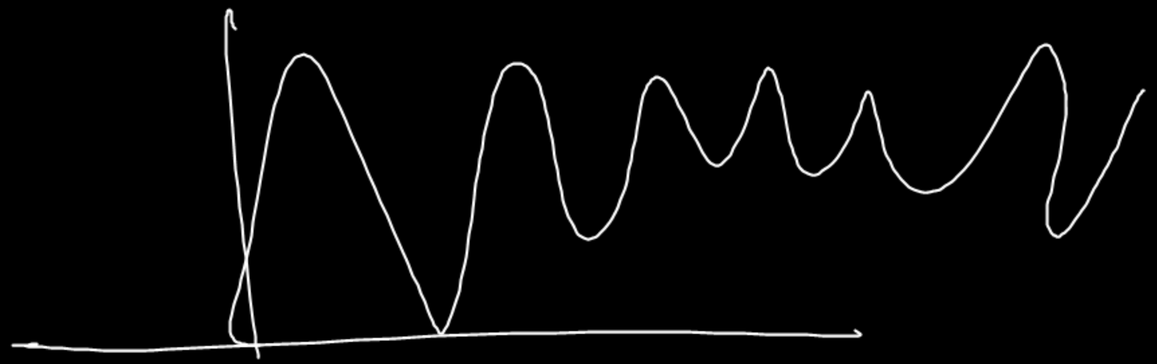
Ice cream

It reflects upward & downward movement over a time period

Stock Price

(3) Cyclicity (C_t) ⇒ rarely - it looks like the same as Seasonality only diff. is it is more than one year.

→ We can't predict

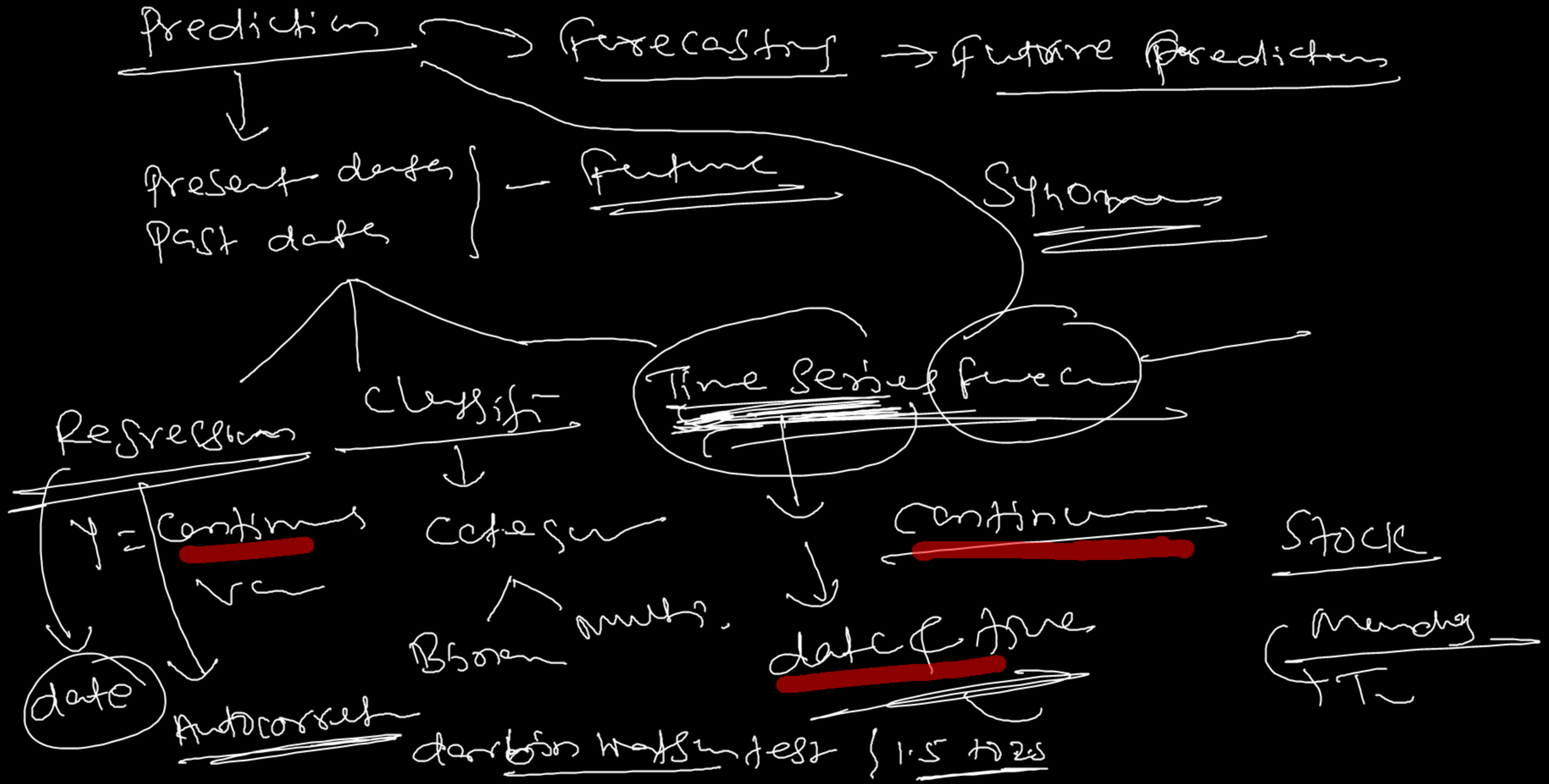


⇒ unemployment rate
⇒ Recession

EMI
$100 → $80

(4) Irregular / Random / Noisy ⇒ we can't predict

↳ error



① Additive method

$$Y_t = T_t + S_t + I_t$$

→ ①

② Multiplicative method

$$Y_t = T_t * S_t * I_t$$

→ ②

	Y_1	Y_2
T_1	10	$10 \times 6 + 50$
T_2	12 (+2)	60 (×5) + 240
T_3	15 (+3)	300 (×4)
T_4	20 (+5)	1200
1	1	1
1	1	1

T_n

* What are the measurements to find the forecast?

① $$MAE = \frac{\sum_{t=1}^{n} |Y_t - F_t|}{n}$$

Y_t = Actual

F_t = Forecast

② $$MAPE = \frac{1}{n} \sum_{t=1}^{n} \frac{|\text{Actual} - \text{forecast}|}{|\text{Actual}|} * 100$$

Month	Actual	Forecast	Absolute diff
1	112	124	11
2	108	103	4
3	198	116	23
4	117	78	33

$$\text{MAPE} = \frac{11 + 4 + 23 + 33}{4} = 17\%$$

$$\frac{117 - 78}{117} \times 100 \qquad \frac{3900}{117}$$

MAPE	Forecasting
< 10%	Highly accurate
10% – 20%	Good forecasting
20% – 50%	Reasonable forecasting
> 50%	Weak & inaccurate forecasting Not acceptable

Please remember

Param

(3) $$MSE = \frac{1}{n} \sum_{t=1}^{n} (Y_t - F_t)^2$$

(4) $$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (Y_t - F_t)^2}$$

(5) AIC (6) BIC

→ Akike Information Criterion

AIC : $(-2LL + 2K)$

Where, LL = log-likelihood

*** K = parameter

Parameter = p d q or P D Q or both

Trend Seasonality

p = Partial Auto correlation = Auto Regressive

d = difference value (Stationary) - Integrated

Rule :- Time Series Forecasting we can Predict only when data is Stationary

q = Autocorrelation = Moving Avg.

model Name

p d q = AutoRegressive Integrated Moving Avg

ARIMA → Traditional method

↓

Trend

PDQ → SARIMA

↓

Seasonality

(6) BIC – Bayesian Information Criterion

$$-2LL + K * Ln(n)$$

LL = Log likelihood

K = parameter (PdF or PDF or both)

Ln = Natural Log (Log$_e$ / Log$_2$ / Ln)

n = Total no. of Observation

Thumb rule

Application –