1. data ingestion.
2. EDA
3. Pre Processing
4. model building
5. model evaluation.

1. Handle missing
2. outlier
3. Scale
4. encoding
5. feature selection
6. Dim red. [PCA]

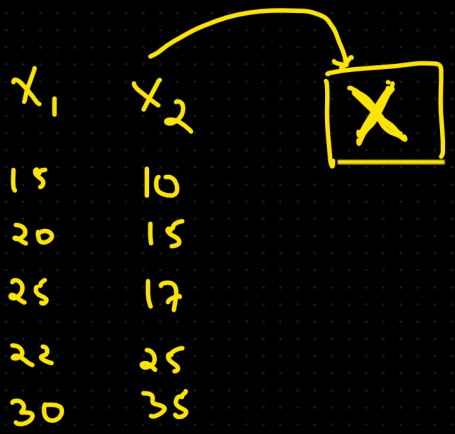
Sensor fault detection. → 170 (80-1)

250 (85-1.)

500 (75-1.)

$f_1, f_2, f_3 \dots f_{100}$ ⇒ feature selection.

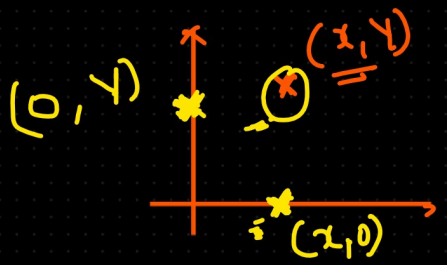
10
20
25

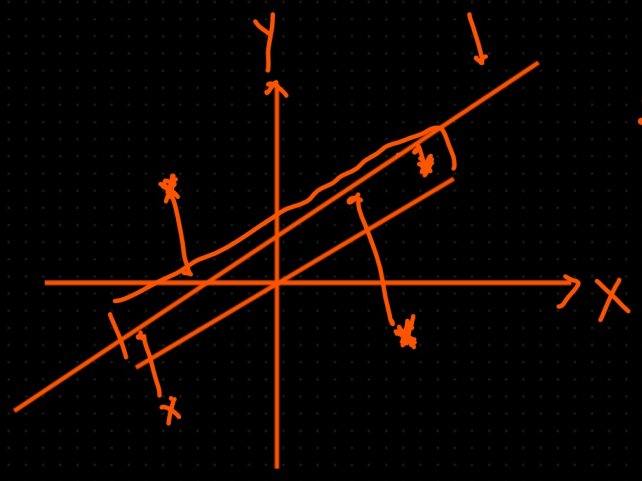


PCA

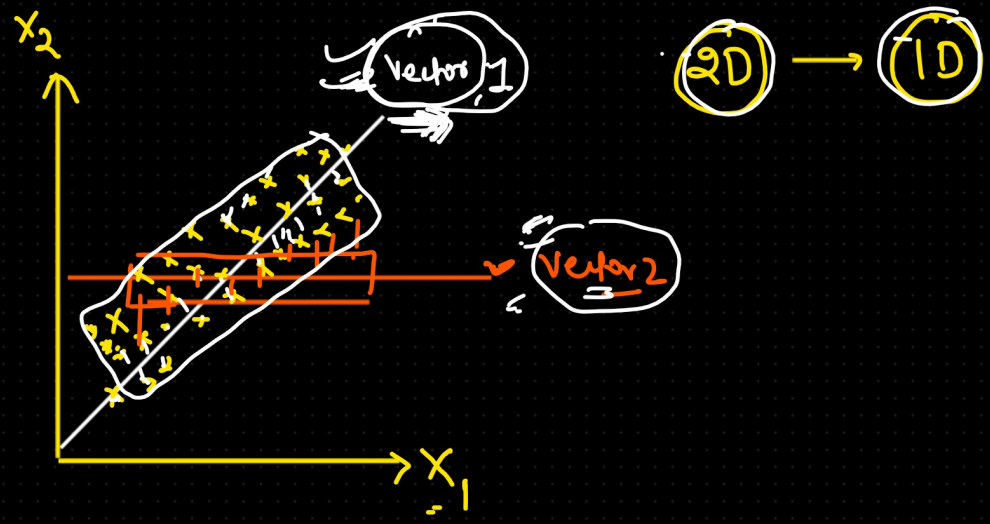
1. Standardized.
2. Cov-matrix
3. eigen value and eigen vector
4. PC

PCA





2D → 1D.


X_2
Vector 1
2D
1D
Vector 2
X_1