Machine Learning

AI

ML
DL
DS
DA
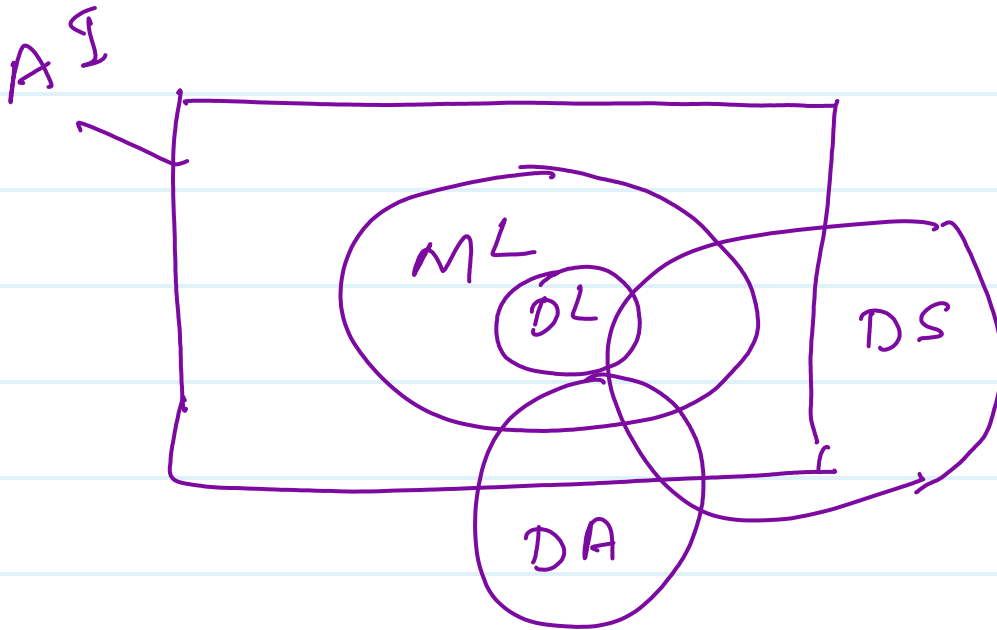
Types of machine learning

① Supervised ML
② unsupervised ML

① Supervised ML
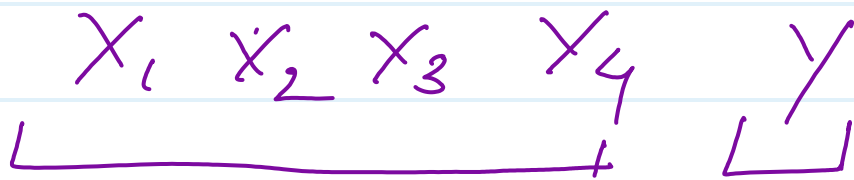
It required Lable data.

Age	cibil	Job	Salary	loan Amount

Independent variable / features

Dependent variable / Target feature

$$X_1 \quad X_2 \quad X_3 \quad X_4 \quad \boxed{Y}$$

Different Types of supervised learning

① Regression

② Classification

① Regression Algorithms

(a) Linear Regr.
(b) L_1 (Lasso)
(c) L_2 (Ridge)

(d) Elasticnet Regularization

(e) KNN

(f) SVR

(h) Decision Tree regre.

(i) Random forest reg.] bagging

(j) Ada boost

(k) Gradient boosting] Boosting]

(l) XG boost

(m) Polynomial Linear regression.   ensemble tech.

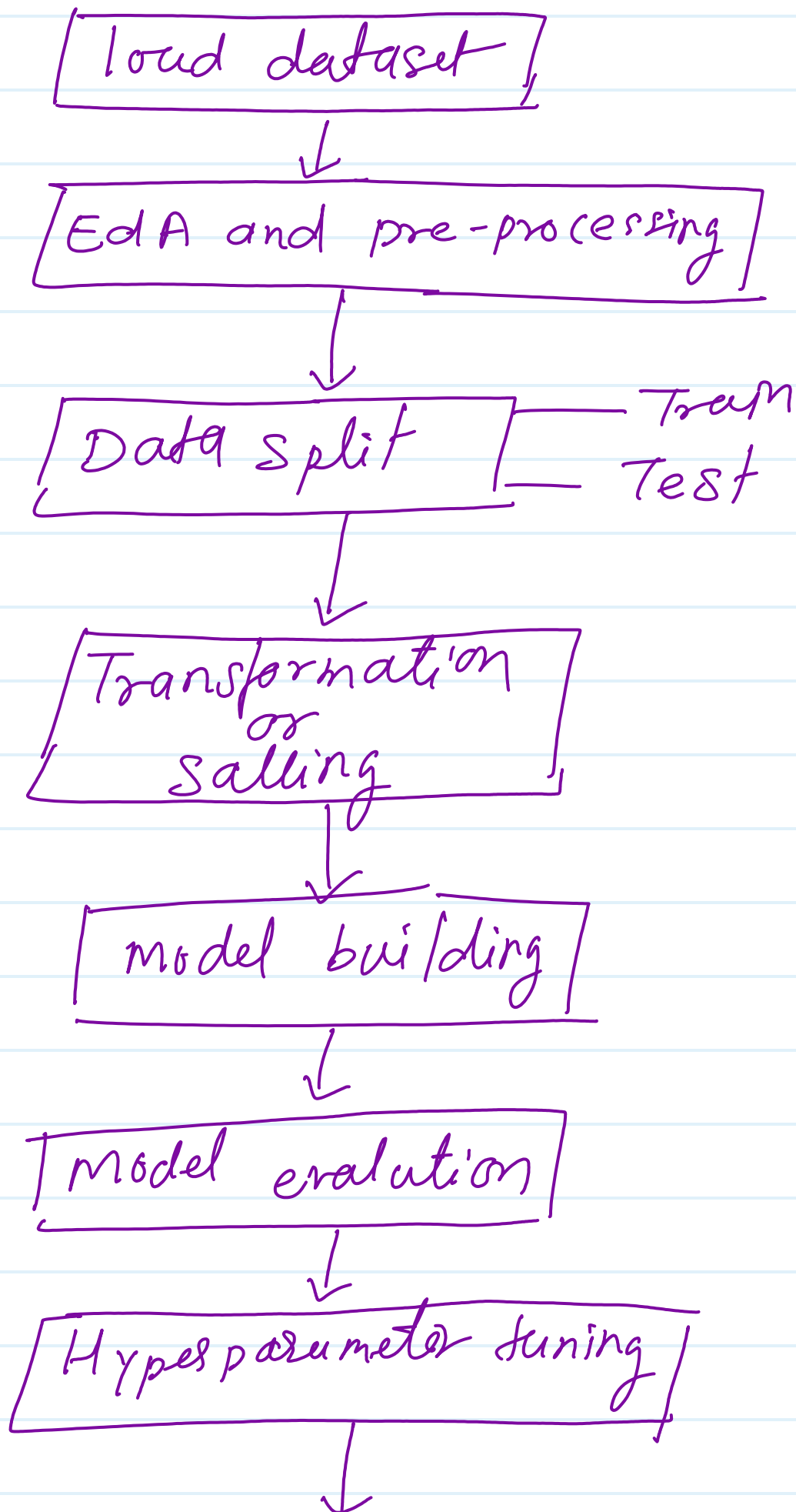(2) classification

(a) logistic regression

(b) Naive bay's

(c) KNN

(d) SVC

(e) DTC

(f) Random forest class.

(g) Ada boost class.

(h) Gradient boost class.

(i) X G boost classification

* Usupervised learning

Clustering / Grouping

(I) k-mean

(II) k-mean + f

(III) DB Scan

(IV) Hirerchical clustering.

```
┌─────────────────────┐
│   loud dataset      │
└─────────────────────┘
          ↓
┌─────────────────────────────┐
│  EdA and pre-procesing      │
└─────────────────────────────┘
          ↓
┌─────────────────────┐      ── Train
│   Data split        │──
└─────────────────────┘      ── Test
          ↓
┌─────────────────────┐
│  Transformation     │
│       or            │
│     salling         │
└─────────────────────┘
          ↓
┌─────────────────────┐
│  model building     │
└─────────────────────┘
          ↓
┌─────────────────────┐
│  model evalution    │
└─────────────────────┘
          ↓
┌─────────────────────────────┐
│  Hyperparameter tuning      │
└─────────────────────────────┘
          ↓
```

Re-training of model

↓

Deployment of model

overfitting –

age	height
5	3.5
10	4
15	4.5
20	5
25	5.5
40	7.5

11.8

→ low bias

→ High variance.

Underfitting -

training efficiency - low

Testin efficieney - High
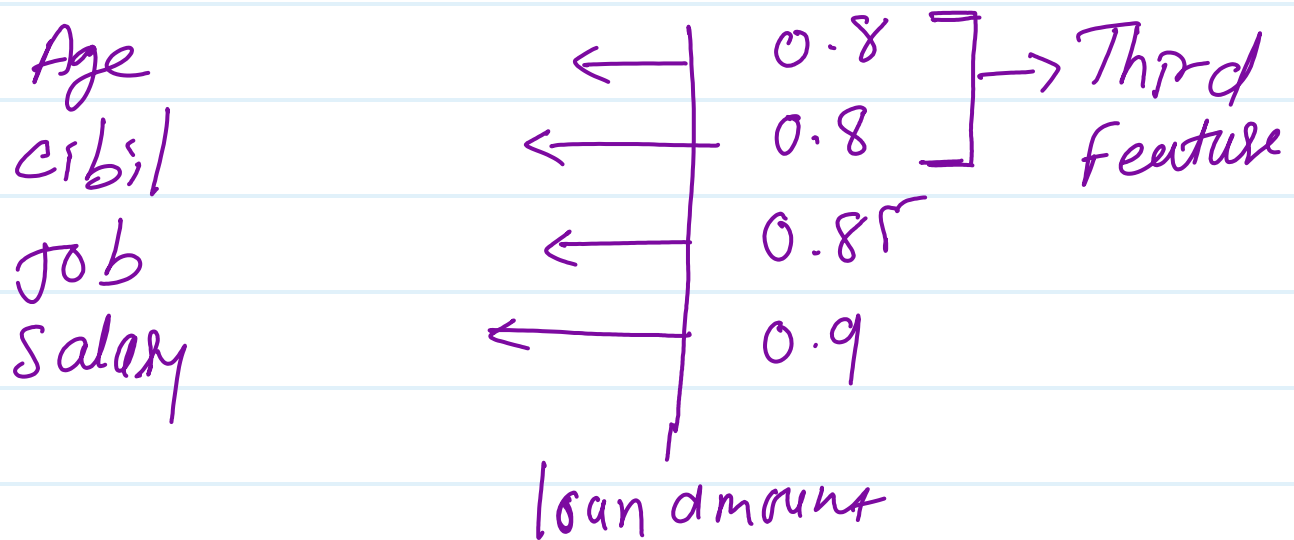
→ High bias
→ low variance

Best fitting model

training efficieney = moderat
testing efficieng = moderat

→ low bias
→ low variance

Correlation →

Age $\leftarrow$ 0.8 $\left.\right]$ → Third
cibil $\leftarrow$ 0.8 Feature
Job $\leftarrow$ 0.85
Salary $\leftarrow$ 0.9

loan amount

feature Selection or feature Reduction

$X_1, X_2 ------ X_{90} \quad Y$

① KalN
② PCA (Principle component Analysis)

Data transformation

⇒ Data leakage

Train dataset — fit transfer

test dataset — transform