Machine Learning
Introduction

Definition of Learning systems

- Tom Mitchell: A Computer learns a Task T from experience E, if its performance P improves with E.

E: Experience of past election results

T: Predicting outcome of next elections for given Party Win, Loss - Classification

P: True +ve : No. of correctly predicted Wins

True –ve: No. of correctly predicted Losses

False +ve: No. of incorrectly predicted Wins

False –ve: No. of incorrectly predicted Losses

OR

P: Probability of win/loss - Regression

Purview

- ML – **Data does all the work** to solve problems
- As a programmer, you have to **develop an ML model – find best values of its parameters**
- **Great Applications**: Image recognition, sales prediction, medical diagnosis, Recommender systems
- **And more power when combined with AI**: Robotics, space exploration, game playing, humanoid voice recognition and synthesis, machine translation

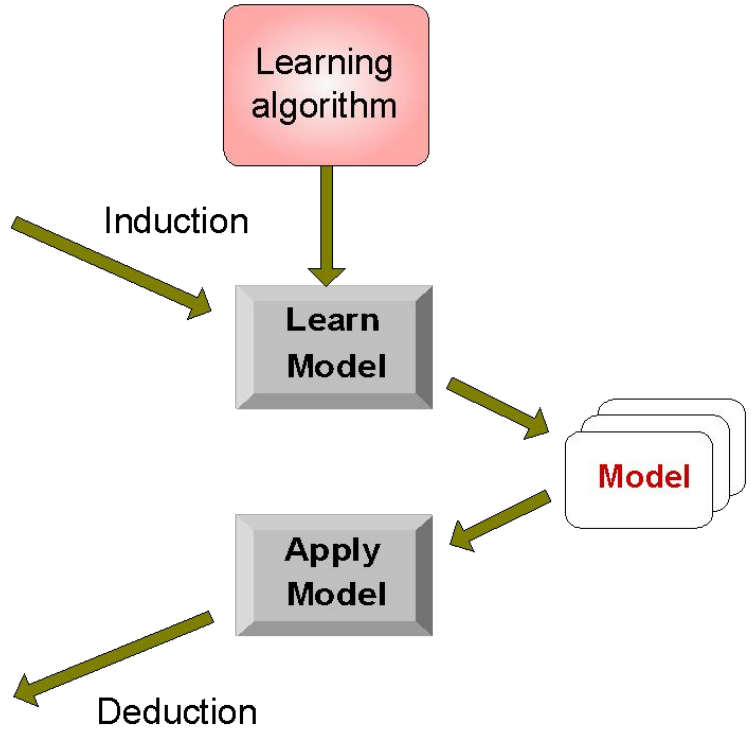
Interaction of Data and Learning Algorithm

Tid	Attrib1	Attrib2	Attrib3	Class
1	Yes	Large	125K	No
2	No	Medium	100K	No
3	No	Small	70K	No
4	Yes	Medium	120K	No
5	No	Large	95K	Yes
6	No	Medium	60K	No
7	Yes	Large	220K	No
8	No	Small	85K	Yes
9	No	Medium	75K	No
10	No	Small	90K	Yes

Training Set

Tid	Attrib1	Attrib2	Attrib3	Class
11	No	Small	55K	?
12	Yes	Medium	80K	?
13	Yes	Large	110K	?
14	No	Small	95K	?
15	No	Large	67K	?

Test Set



Features

- Regression – Volume of sales. Features??

Price, quality, time of sale, model of sale….

- Can ML deal with infinite features?

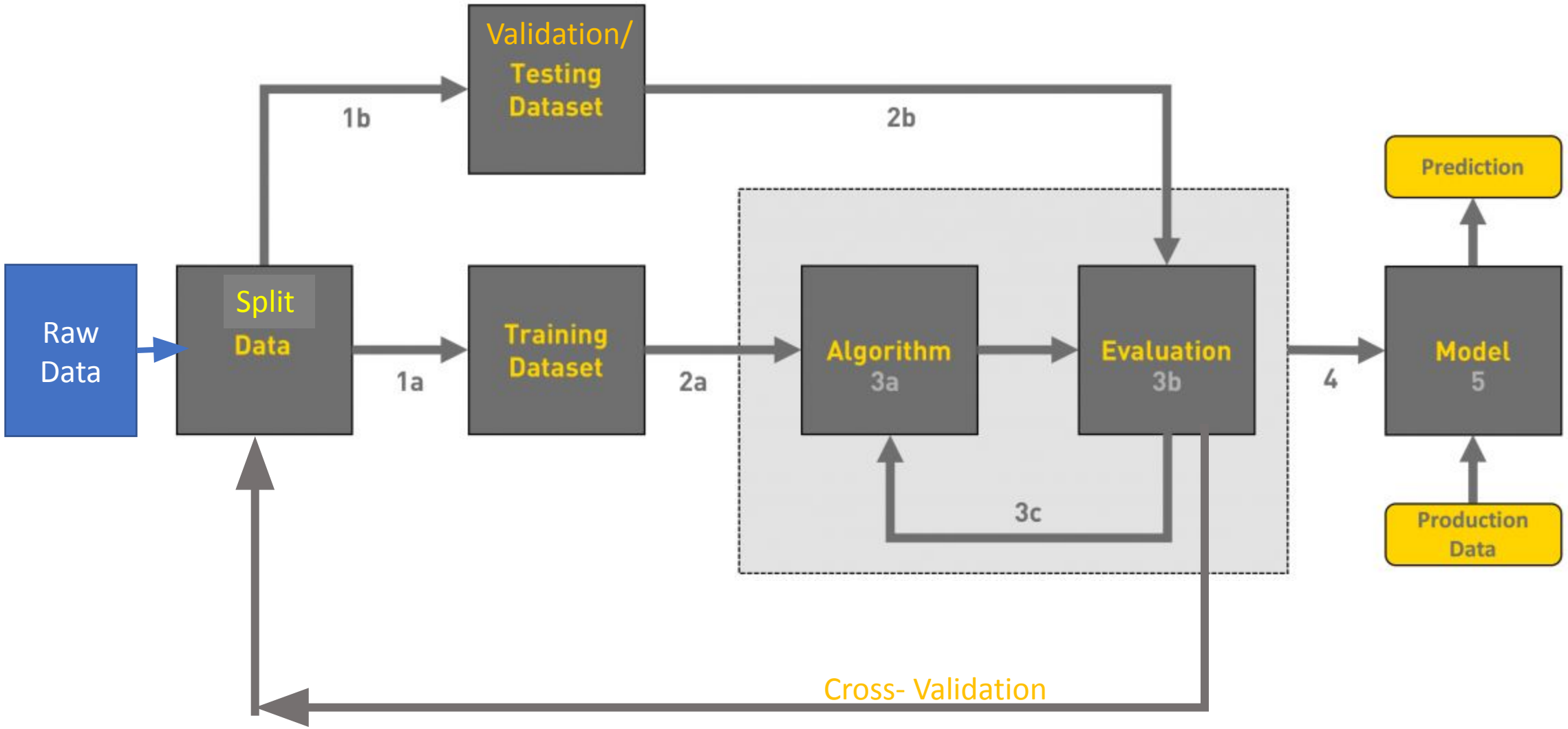
- Classification –Detection of breast cancer?

Size of tumour, degree of malignancy, medical history, age, thickness, type of tumour, clumpiness

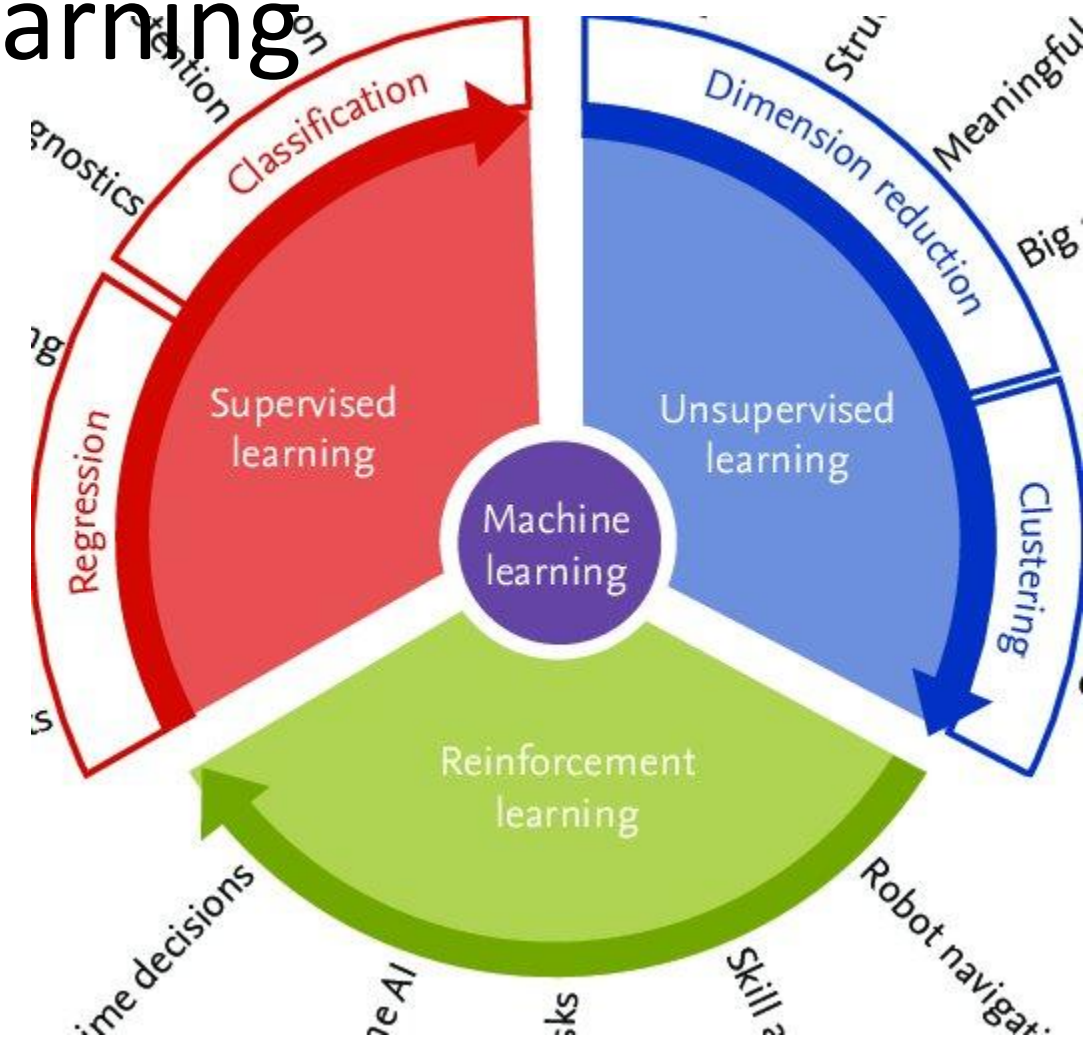
Detection of Depression?

Age,history, suicidal thoughts, motor skills, speech, interest in activities, anxiety level, psychological parameters

Work-Flow of ML



Types of Learning



Basic types of learning

- **Supervised:** Given training dataset {features, response}, predict response for new features

Bayesian, Decision Trees, Linear and Logistic Regression, SVM, KNN, ANN, LDA, HMM, CRF

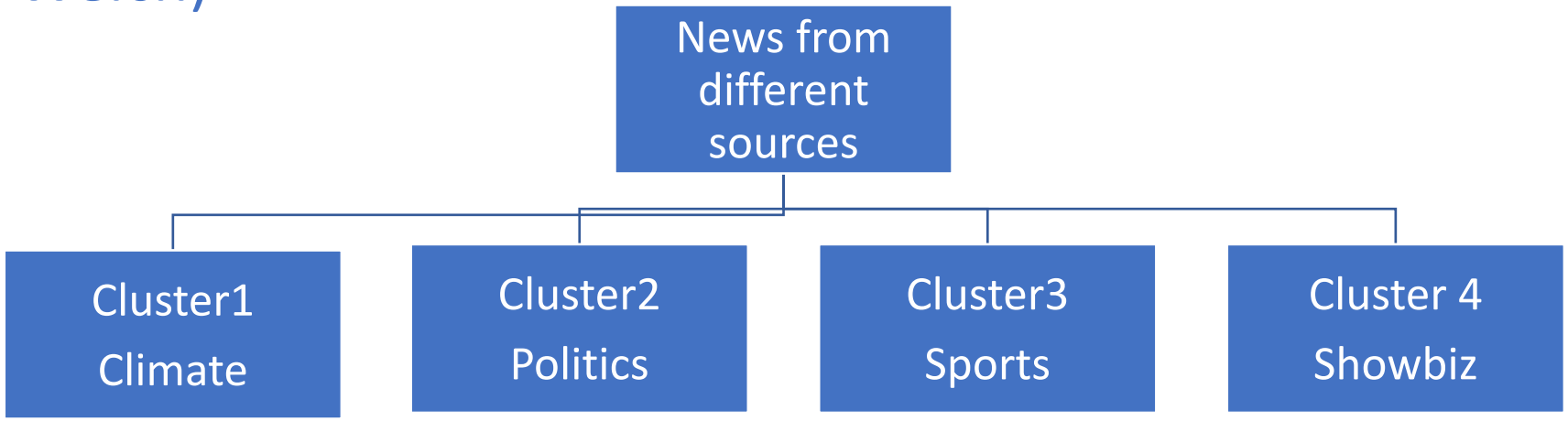
Price of mobile	Make of Mobile	Number of cameras	…….	Volume of sales
9K	Vivo	2		10K
12K	Redme	3		9K
16K	Oppo	4		12K
….	…	…		

Error reduction based learning

Unsupervised:

- Given dataset {features}, find the patterns to cluster the data

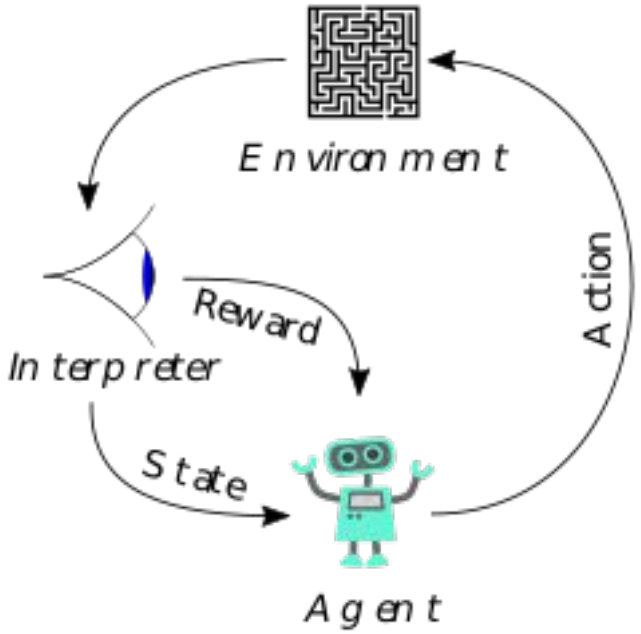
Clustering, Principal Component Analysis, Latent Semantic Analysis (LSA), Association Rules, Self-Organizing Maps (SOM), HMM Training (Baum Welch)



Reinforcement:

- **Semi-supervised**: Starting from few examples, progressively learn
- Learn from rewards and penalties from a series of interactions with the environment

Q-Learning, Monte Carlo Methods, Markov Decision Processes



Reinforcement Learning

- Short term v/s long term goals
- Games, Robotics
- Given Action A States S

$$P_a(S, S') = Pr(S_{t+1} = S' | S_t = S, A_t = a)$$

$R_a(S,S')$ = Immediate reward after transition

(i) Model of environment is there, but analytical solution not known

(ii) Simulation based optimization

(iii) To collect info, agent must interact with environment

(iv) Distributed Applications: a) Ad-hoc networks – Distributed communication b) Multi-robot action – Cooperative tasks with leader election

The process of undertaking ML based solutions

- Understand a domain
- Incorporate prior knowledge
- Know the goals
- **Data collection, integration, cleaning, completion, pre-processing**
- **Feature Selection**
- Choose and fine tune appropriate learning model
- Interpret results measure final performance
- Deploy knowledge gained
- Loop