Topic 1. What is a Learning System?

Reference:
[1] Tom Mitchell, Machine Learning, McGraw Hill, 1997

Tom Mitchell [1]:

A Computer learns a **Task T**

from **experience E**,

if its **Performance P** improves with **E**.

For Computers, Experience is a repository of historical data

Some observations on Mitchell's definition

- Focus on *Data (Experience) for carrying out Tasks*!

- *Any learning approach must perform the Task better, when given more Data*

- Continuous assessment of *Performance*

Applying Mitchell's definition to Recruitment task

T: Predict **+ve (Hired)** or **–ve (Rejected)** for candidates
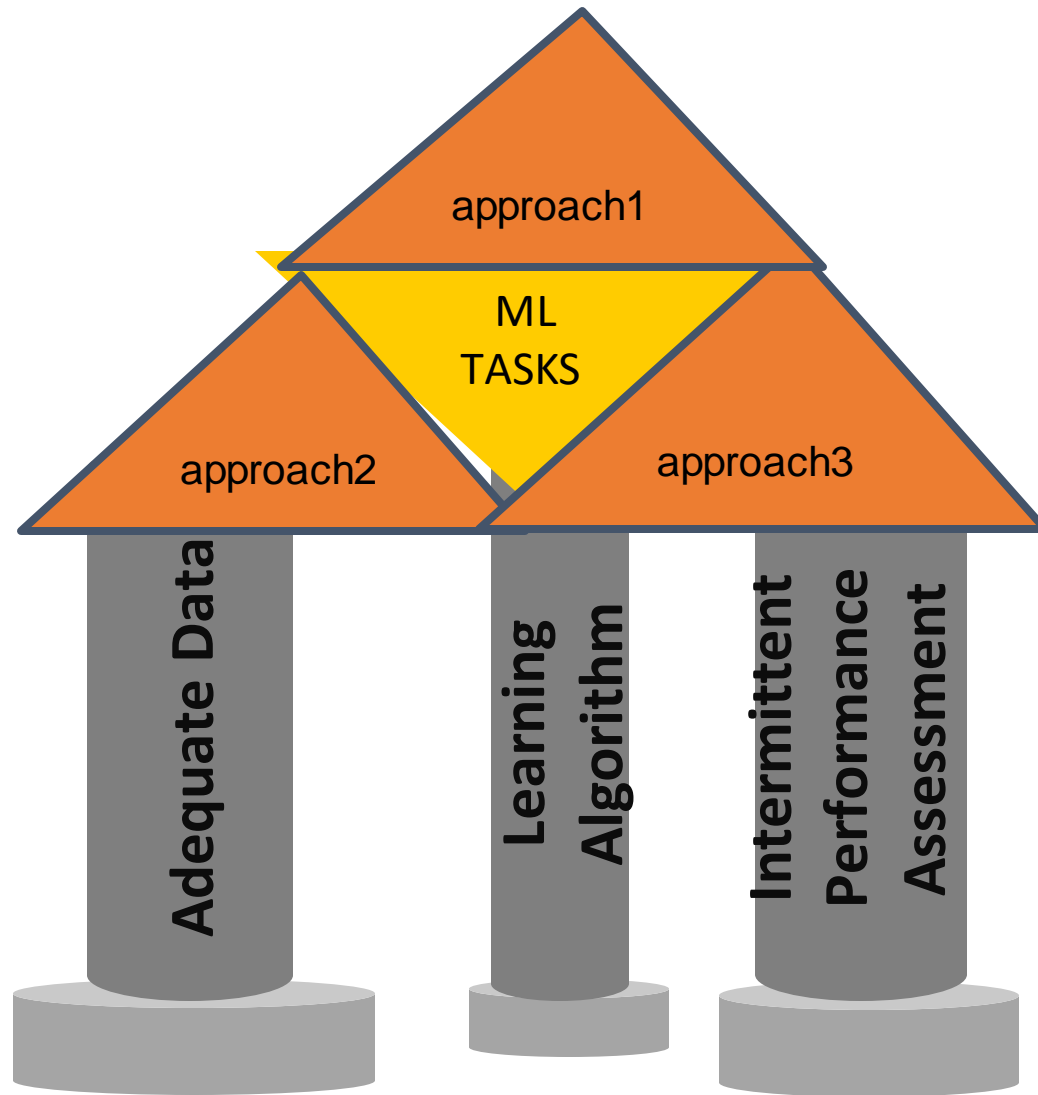
E: Experience: data repository of past hirings.

P: Performance metrics ▢ Tallies

True +ve (TP): Number predicted **+ve** and actually **Hired**

True –ve (TN): Number predicted **–ve** and actually *Rejected*

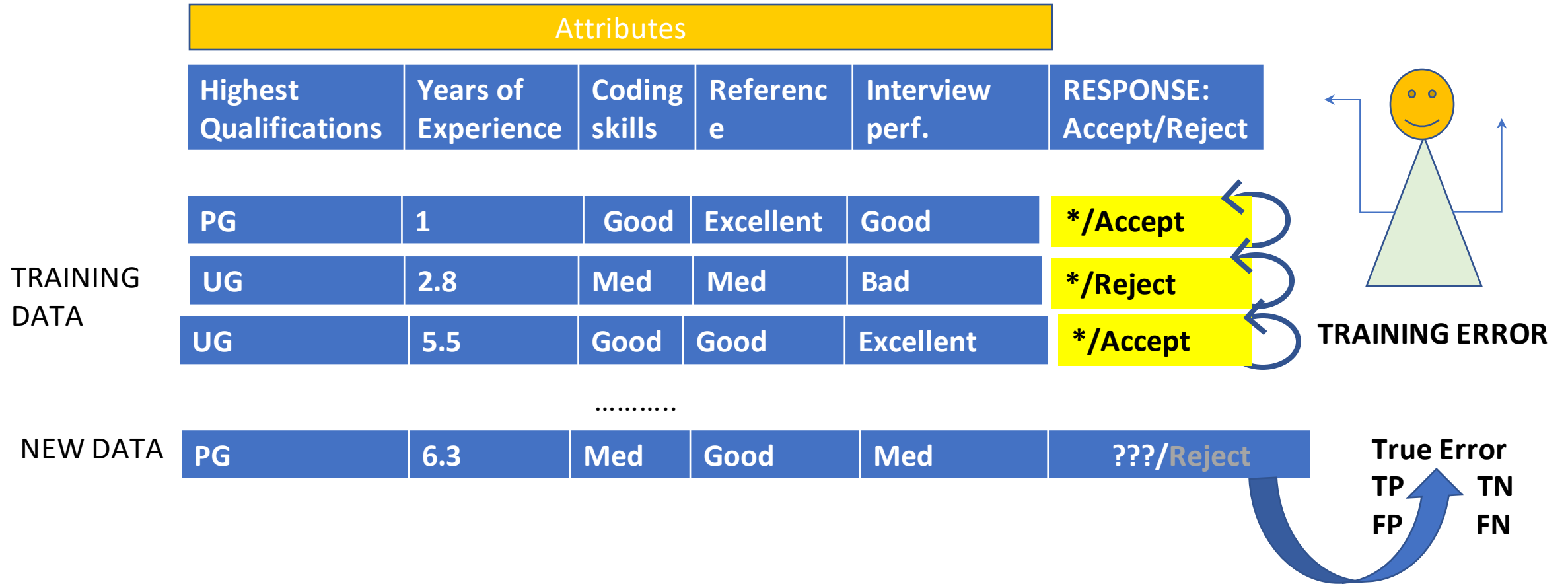**False +ve (FP):** Number predicted **+ve**, but actually **Rejected**

False –ve (FN): Number predicted **–ve,** But actually *Hired*

Topic 2. Approaches for Machine Learning

Approach I. Supervised Learning – Discreet Output

Task: Predict results of Recruitment

Attributes					
Highest Qualifications	Years of Experience	Coding skills	Reference	Interview perf.	RESPONSE: Accept/Reject
PG	1	Good	Excellent	Good	*/Accept
UG	2.8	Med	Med	Bad	*/Reject
UG	5.5	Good	Good	Excellent	*/Accept

TRAINING DATA

TRAINING ERROR

....

NEW DATA

PG	6.3	Med	Good	Med	???/Reject

True Error

TP   TN

FP   FN

Approach I. Supervised Learning – Continuous output

Task: Recruitment – predict scores

Highest Qualifications	Years of Experience	Coding skills	Reference	Interview marks	Score
PG	1	Good	Excellent	79	*/80
UG	2.8	Med	Med	55	*/50
UG	5.5	Good	Good	87	*/75

TRAINING DATA

....

NEW DATA

| PG | 6.3 | Med | Good | 67 | ???/70 |

TRAINING ERRORS

Errors
(Predicted score - Actual score)^2
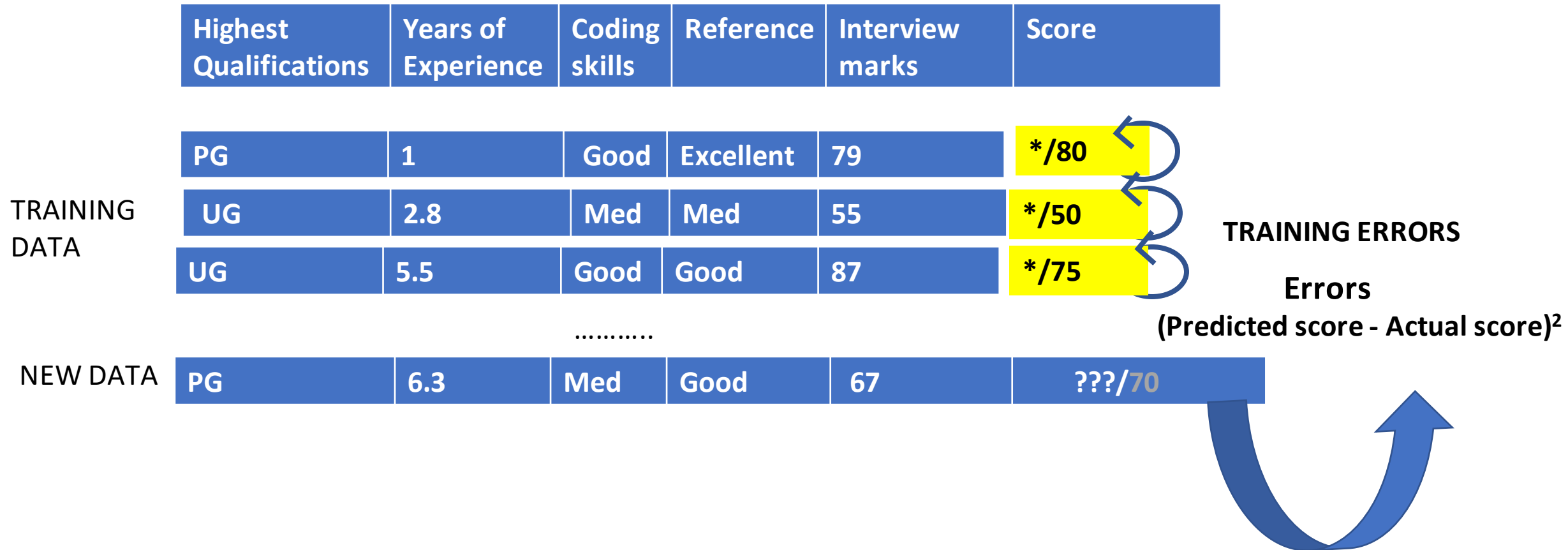
Types of supervised tasks

Supervised
Learning

Score

Regression
(Continuous
Response)

Accept/Reject

Classification
(Discreet
Response)
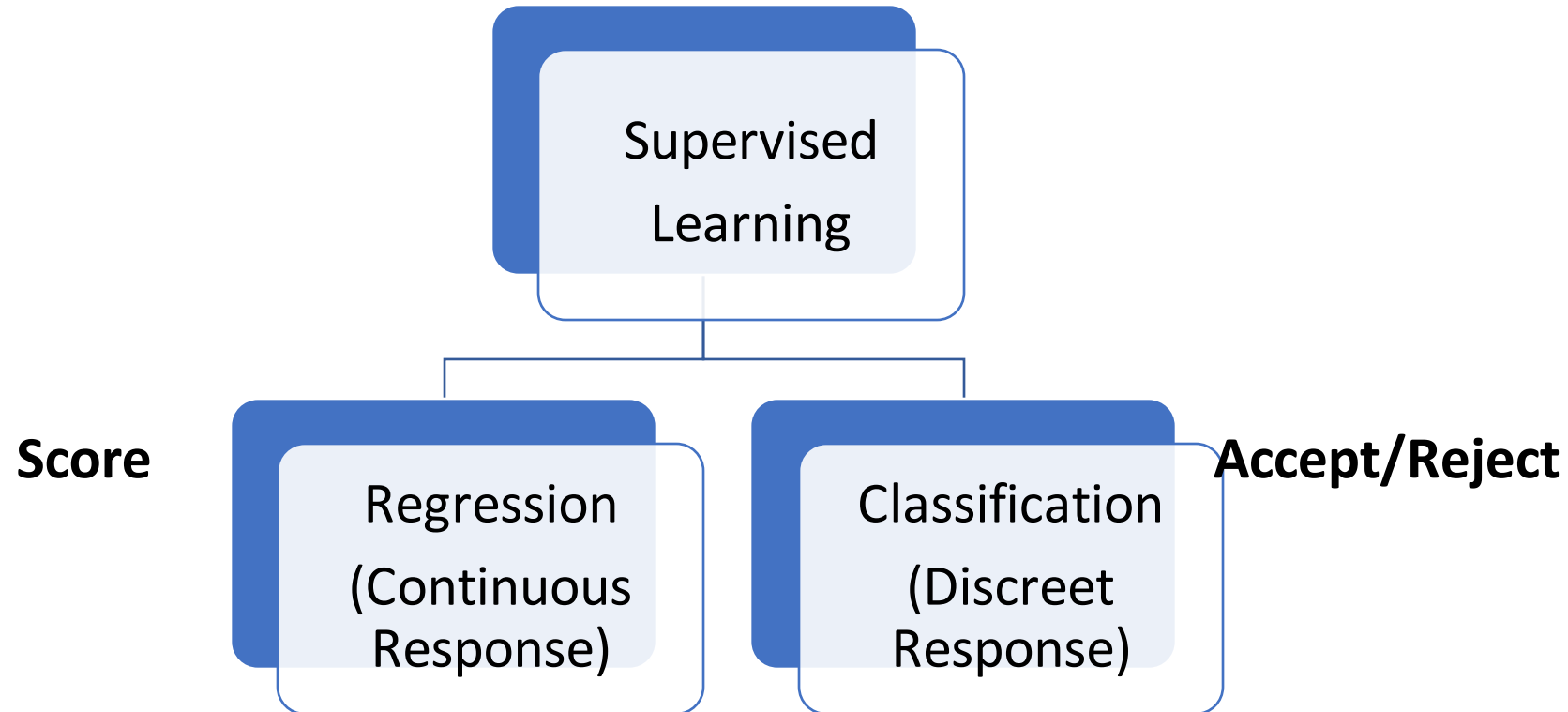
The Process of Supervised ML
TASK: Predict Recruitment Results

Learning Algorithm

TRAINING DATA

Highest Qualification	Years of Experience	Level of Coding Skill	Reference	Interview Performance	Result
UG	5	Good	Excellent	Good	*/Accept
PG	3	Med	Good	Bad	*/Reject
UG	6	Poor	Medium	Excellent	*/Accept
UG	2	Good	Excellent	Good	*/Accept
...	...	...	...	...	...

VALIDATION/TEST

Highest Qualification	Years of Experience	Level of Coding Skill	Reference	Interview Perf	Result
PG	5	Good	Good	Good	???/Accept
UG	3	Med	Good	Med	???/Reject
UG	6	Med	Medium	Good	???/Accept
UG	2	Good	Excellent	Bad	???Accept
...	...	...	...	...	...

1. Develop a Model

Model

TRAINING CYCLES:
TESTING CYCLES:

2. Optimize the Model

FIELD

Highest Qualification	Years of Experience	Level of Coding Skill	Reference	Interview Perf	Result
PG	5	Good	Good	Excellent	???
UG	3	Med	Good	Good	???
UG	6	Med	Medium	Med	???

3. Apply the Model

ALTERNATING TRAINING CYCLE

ASSESS PERFORMANCE ON
VALIDATION/TEST DATA
PROVIDE FEEDBACK / PAUSE AND LEARN

Algorithms for Supervised Learning

- **Linear Regression**

- **Logistic Regression**

- **Bayesian Classification**

- **Decision Trees**

- **Support Vector Machines**

- **K-Nearest Neighbours**

- **Artificial Neural Networks**

Column index / Row Index	Years of Experience	Level of Coding Skill	Reference	Interview Performance	Result
anil	5	Good	Excellent	Good	*/Accept
Sita	3	Med	Good	Bad	*/Reject
Gita	6	Poor	Medium	Excellent	*/Accept
Riva	2	Good	Excellent	Good	*/Accept
…	…	…	…		…

Summary

- Supervised learning is **Error reduction based learning**.

- Goal of ML agent: Learn a mapping function from input (feature values) to output (response) ▢ *Target Hypothesis.*

- Learning gained from examples is *induced to new cases*

Approach II. Unsupervised – Learning without a teacher

Extract Concepts in biodata of candidates

CONCEPT 1: Candidates who:
- UG
- Experience 1 to 2 years,
- Good/Excellent reference,
- Good/Med coding skills,
- Excellent or Good performance in interview

CONCEPT 2:
- UG/PG,
- Experience any,
- Bad reference,
- Med/Low coding skills,
- Med/Bad perf in interview

Types of unsupervised tasks
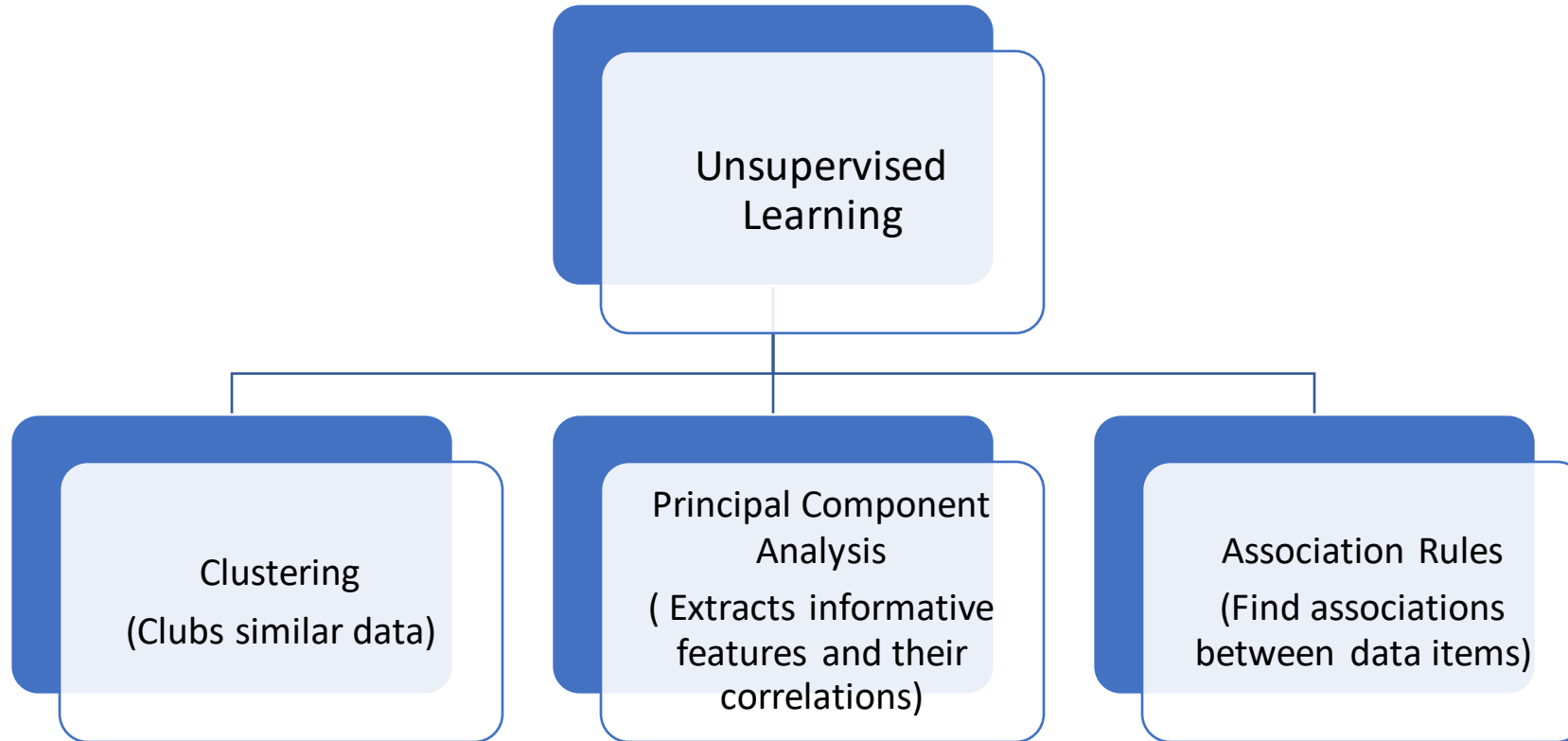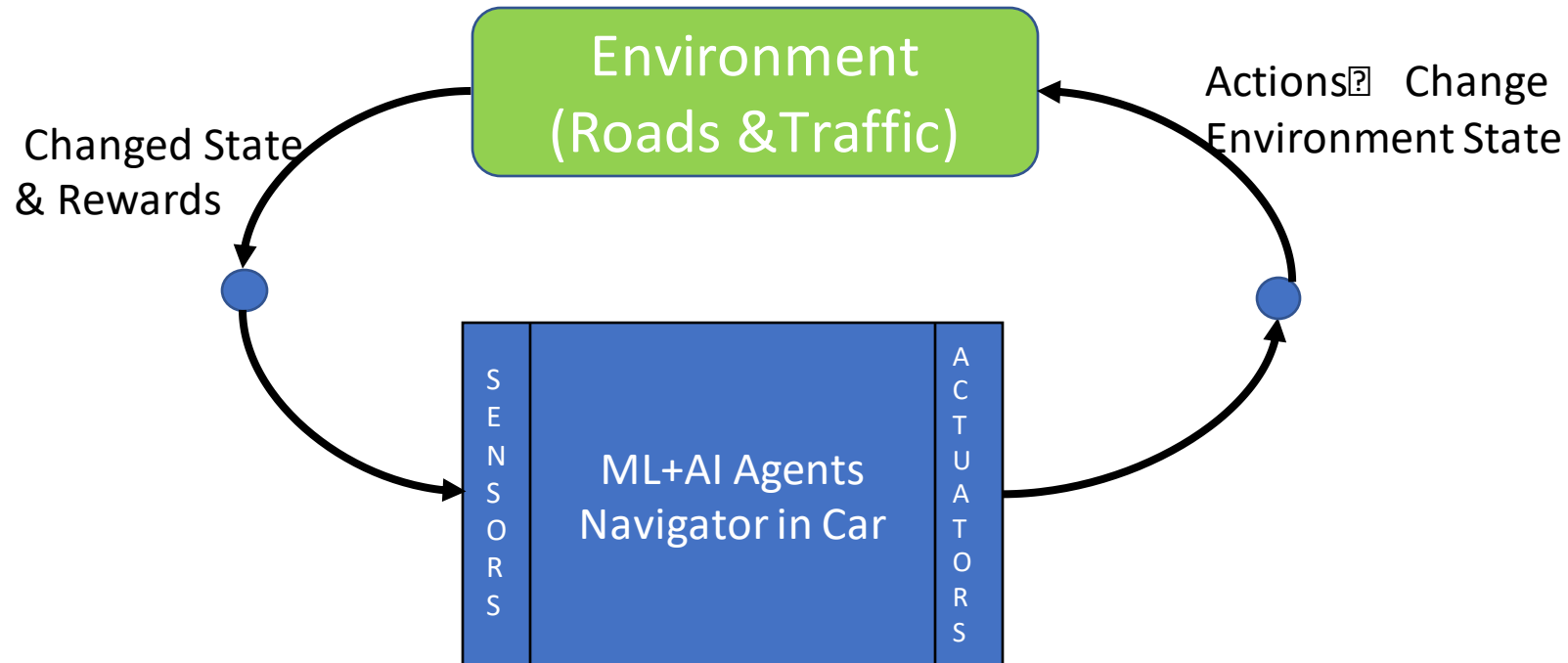
Algorithms for Unsupervised Learning

- Clustering:
  - K-means Clustering
  - Hierarchical Clustering
- Principal Component Analysis
- Association rules
- Self Organizing Maps (SOM)
- Singular Value Decomposition SVD
- Anomaly Detection

Approach III. Semi-Supervised Learning

- Labelled pool D_1 & Unlabelled pool D_2

- Repeat till acceptable performance:
  - Learn with labelled examples D_1
  - Classify D_2
  - Label confidently classified instances in D_2
  - Add them to D_1

- Can surpass performance with either D_1 or D_2

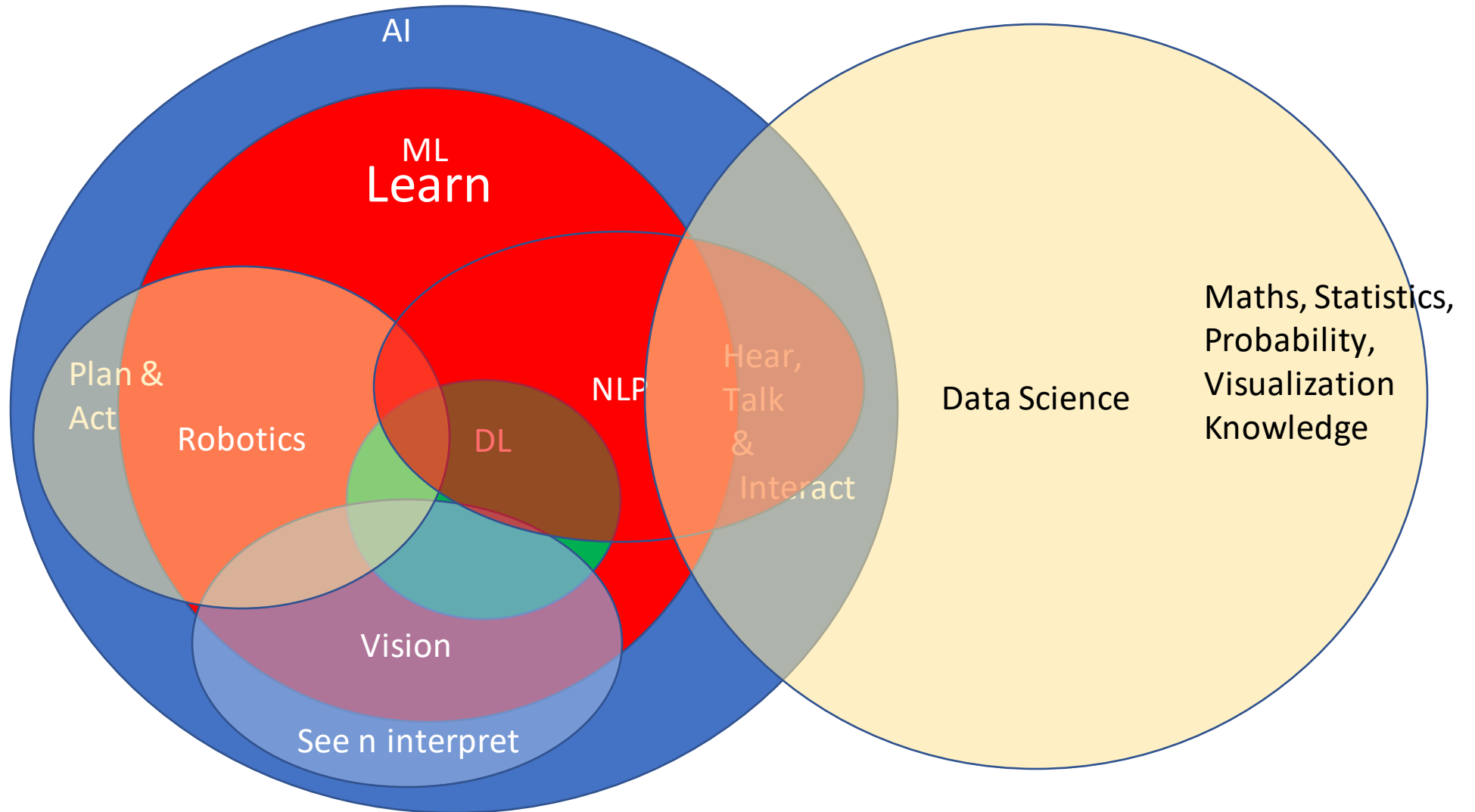- No burden of collecting lots of labelled data

Approach IV. Reinforcement Learning

- No examples to begin with!

- Some inkling of good or bad steps

- Gradually learns from rewards (or penalties) from a series of interactions with the environment

Environment (Roads &Traffic)

Changed State & Rewards

Actions   Change Environment State

ML+AI Agents Navigator in Car

SENSORS

ACTUATORS

Topic 3: Convergence of Technologies

Relationship between ML & other domains of AI

Great Applications!

- Robot-based manufacturing

- Human assistance systems

- Autonomous driving

- Understanding & speaking human language

- Translating one language to another

- Deep Sea/earth Space exploration

- ……..

Learning Points

➢Machine Learning has these components –

➢ A Task to be learnt

➢ Data on past experience

➢ Performance to be assessed intermittently

➢A learning approach to find mappings (hypothesis)

Learning Points

➢ The main approaches for ML are:

  ➢ Supervised learning – Error reduction based hypothesis learning

  ➢ Unsupervised – Finds latent similarities, correlations and associations in data

  ➢ Semi-supervised learning – Boosts unlabelled data starting with few labelled examples

  ➢ Reinforcement Learning – interactive & reward based learning

Learning Points

➢ The supervised ML Process includes:

  ➢ Training with labelled data ▢ Model

  ➢ Validating and testing performance of model

  ➢ Applying field data on trained model

Learning Points

➢ Machine Learning is:

   ➢ A sub-branch of AI

   ➢ Includes Deep Learning

   ➢ Interacts with interdisciplinary fields like Robotics, NLP and Vision

➢ Several sophisticated applications use convergence of technologies

► Link for Self Assessment:

https://forms.gle/m6kTSVd2NayVC44x9

The Secret to going ahead is getting started – Mark Twain