MACHINE LEARNING TOPIC: INTRODUCTION
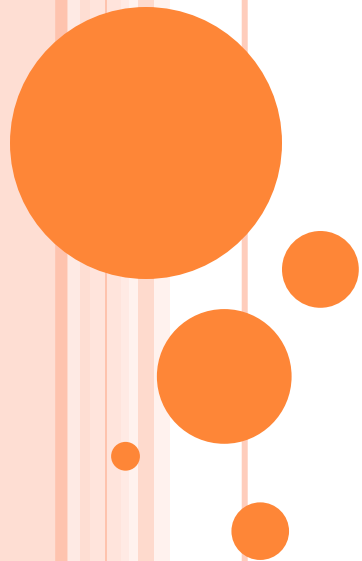
By

Prof. Dr. Sourav Saha

TOPICS COVERED

→ What is machine learning?

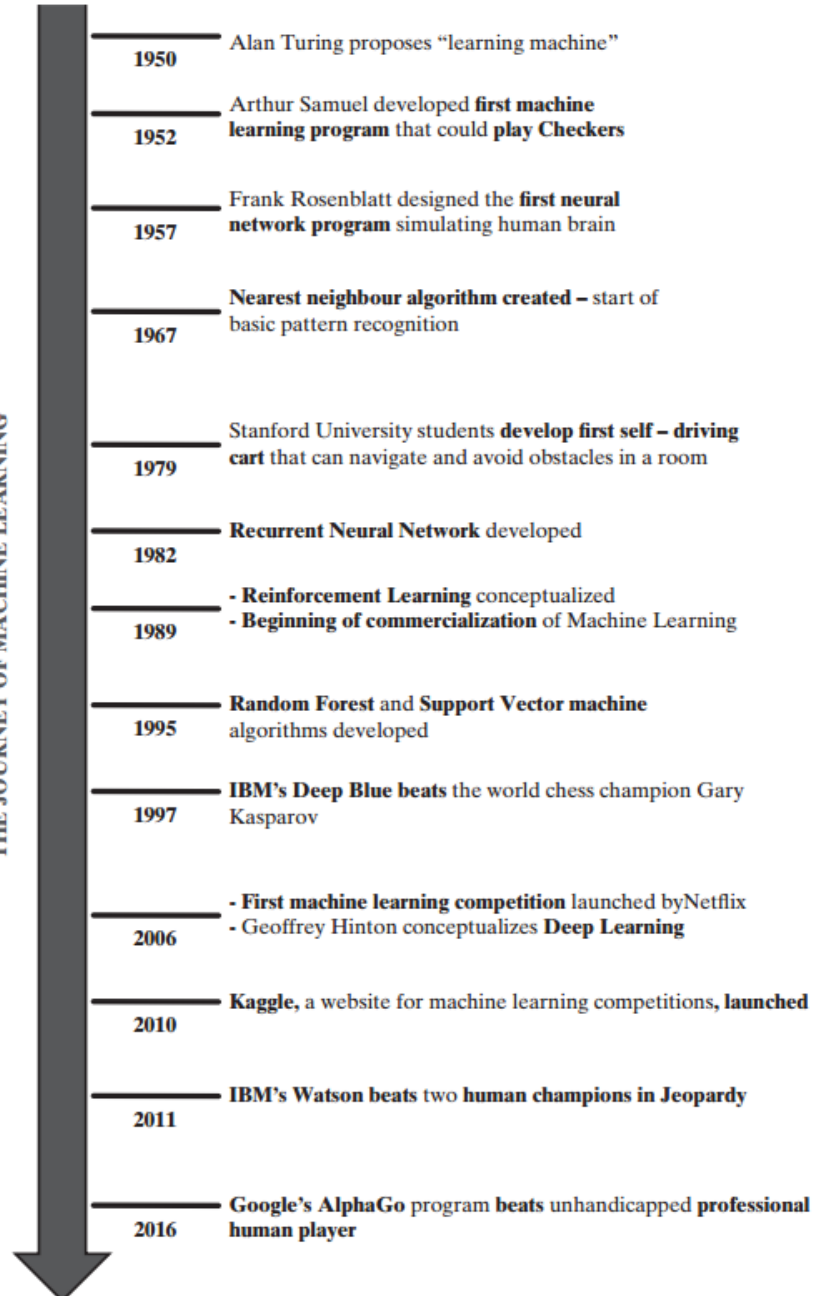→ Which problems can be solved by ML?

→ How industry leverages ML

SETTING THE CONTEXT …

- Do machines really learn?

- If so, can machines learn better than human?

- Most importantly, how do they learn?

SETTING THE CONTEXT …

THE JOURNEY OF MACHINE LEARNING

1950 — Alan Turing proposes "learning machine"

1952 — Arthur Samuel developed **first machine learning program** that could **play Checkers**

1957 — Frank Rosenblatt designed the **first neural network program** simulating human brain

1967 — **Nearest neighbour algorithm created –** start of basic pattern recognition

1979 — Stanford University students **develop first self – driving cart** that can navigate and avoid obstacles in a room

1982 — **Recurrent Neural Network** developed

1989 — - **Reinforcement Learning** conceptualized
- **Beginning of commercialization** of Machine Learning

1995 — **Random Forest** and **Support Vector machine** algorithms developed

1997 — **IBM's Deep Blue beats** the world chess champion Gary Kasparov

2006 — - **First machine learning competition** launched byNetflix
- Geoffrey Hinton conceptualizes **Deep Learning**

2010 — **Kaggle,** a website for machine learning competitions, **launched**

2011 — **IBM's Watson beats** two **human champions in Jeopardy**

2016 — **Google's AlphaGo** program **beats** unhandicapped **professional human player**

WHAT IS LEARNING?

It is the process of gaining information through expert guidance ...

Is that always true?

ANALOGY
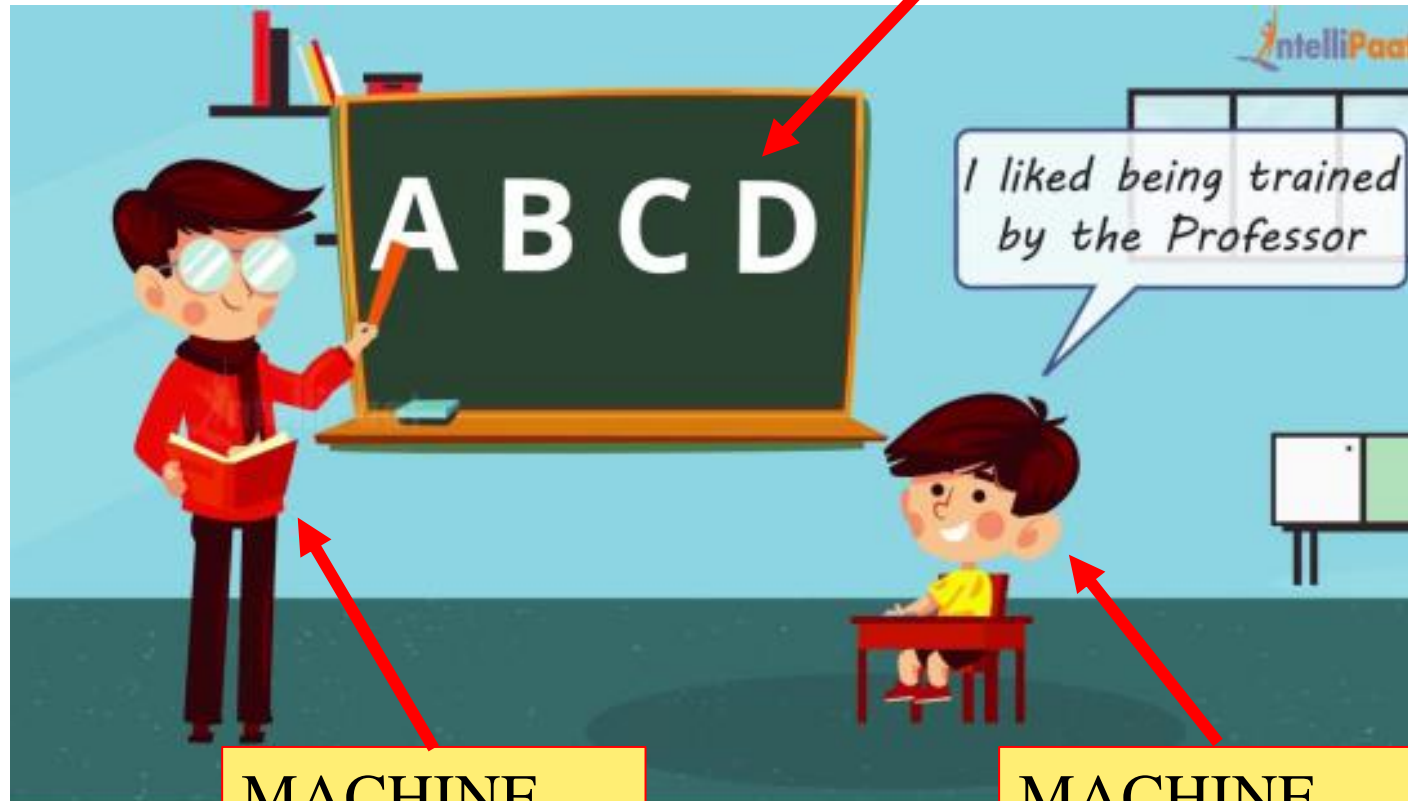
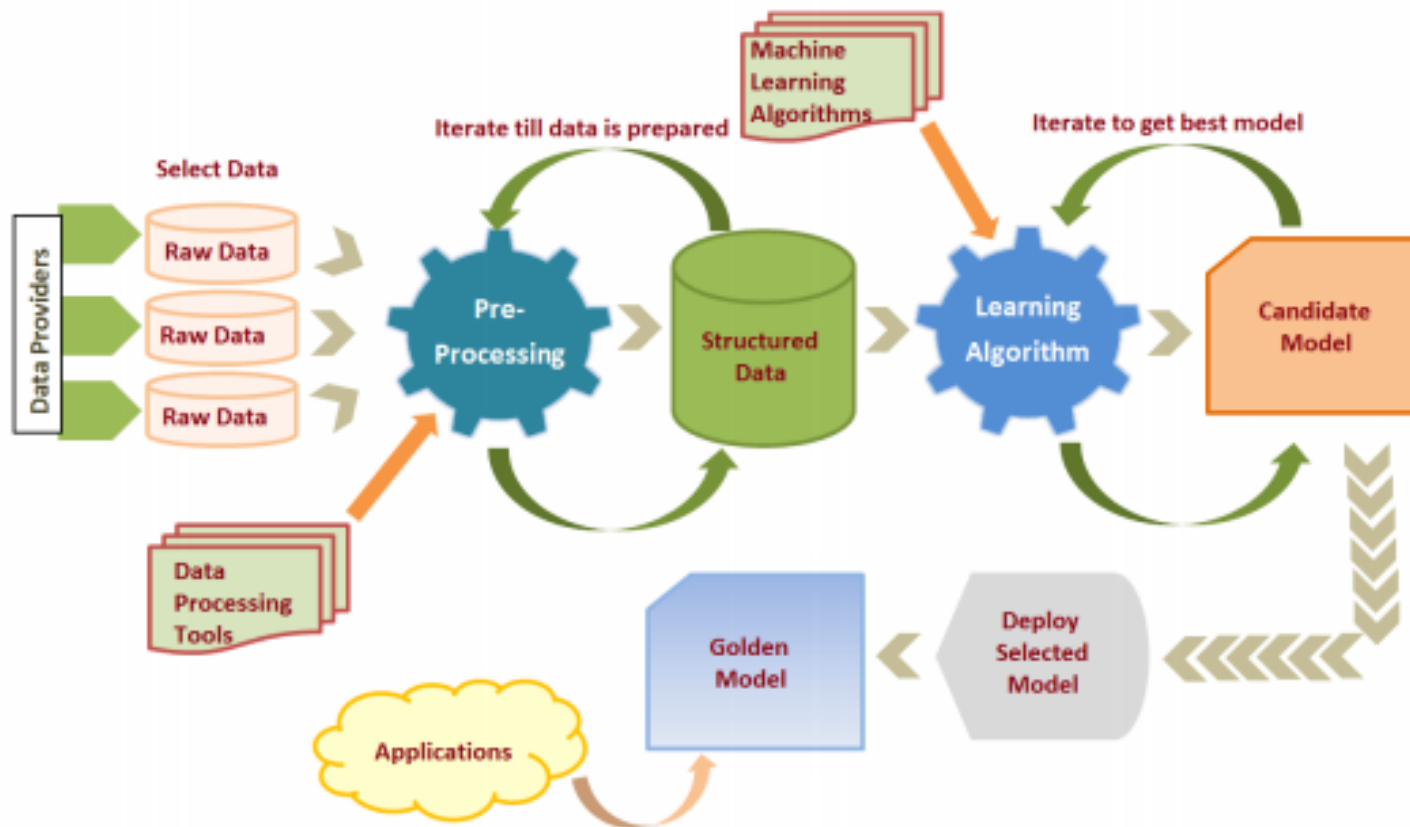# ANAOGY

TEST DATA

OVERALL IDEA

TYPES OF HUMAN LEARNING

- Learning through direct guidance from expert – is just one form …

- Learning through indirect guidance

- Learning by self

WHAT IS MACHINE LEARNING?

"A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."

Tom M. Mitchell,
Professor of Machine Learning Department,
School of Computer Science,
Carnegie Mellon University

What is machine learning?

*In context of the **learning** to **play chess**, **E** represents the experience of playing the game, **T** represents the task of playing chess and **P** is the performance measure indicated by the percentage of games won by the player.*

TRADITIONAL VS. ML PROGRAMMING
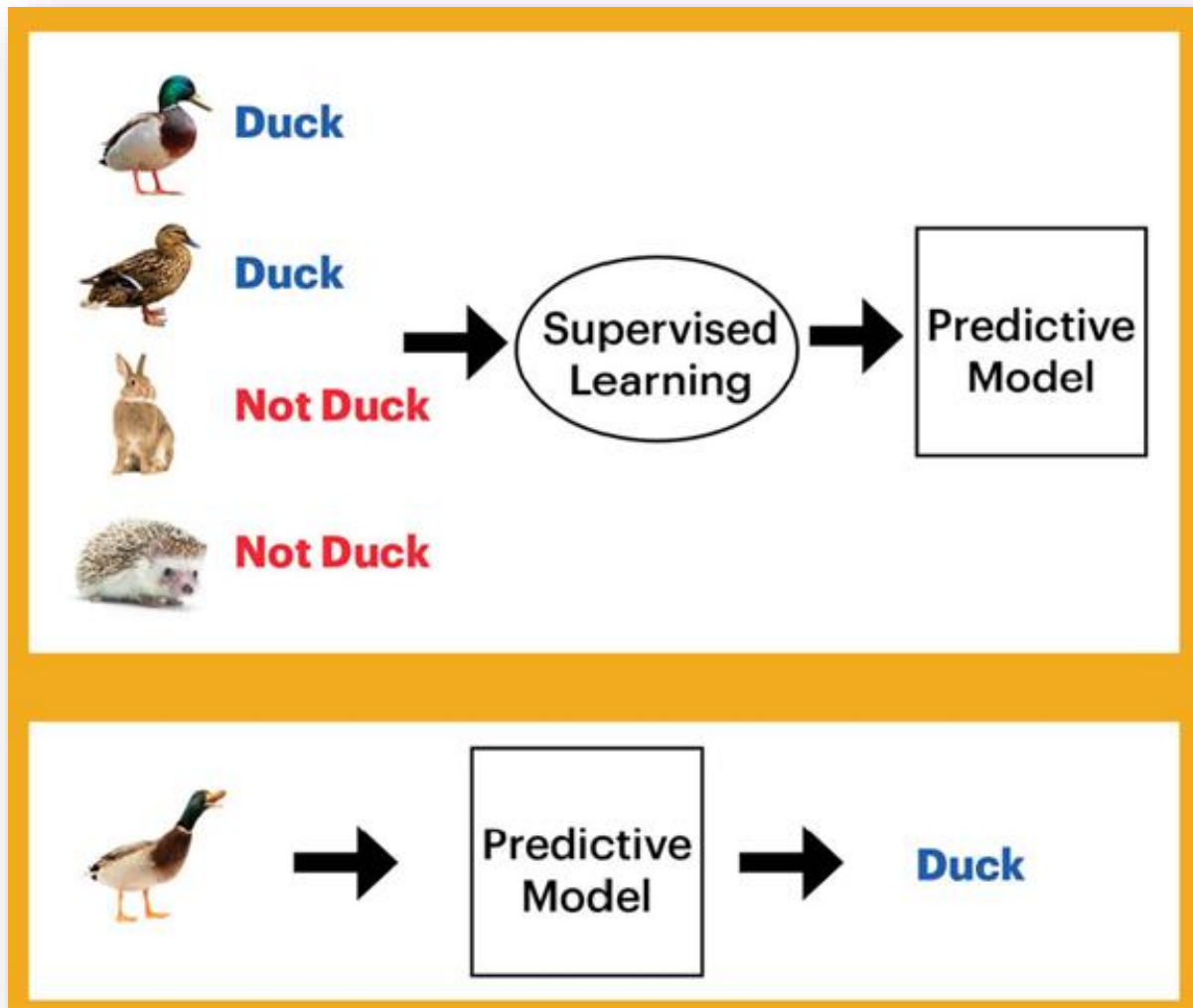
Traditional Programming
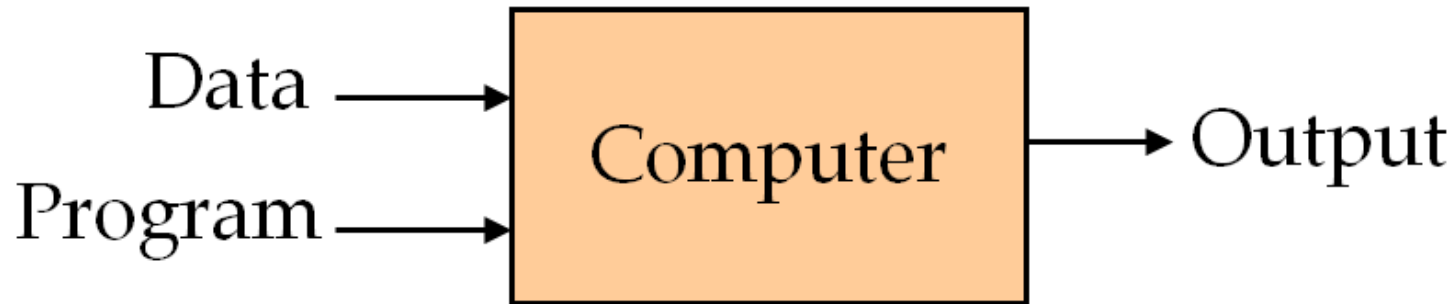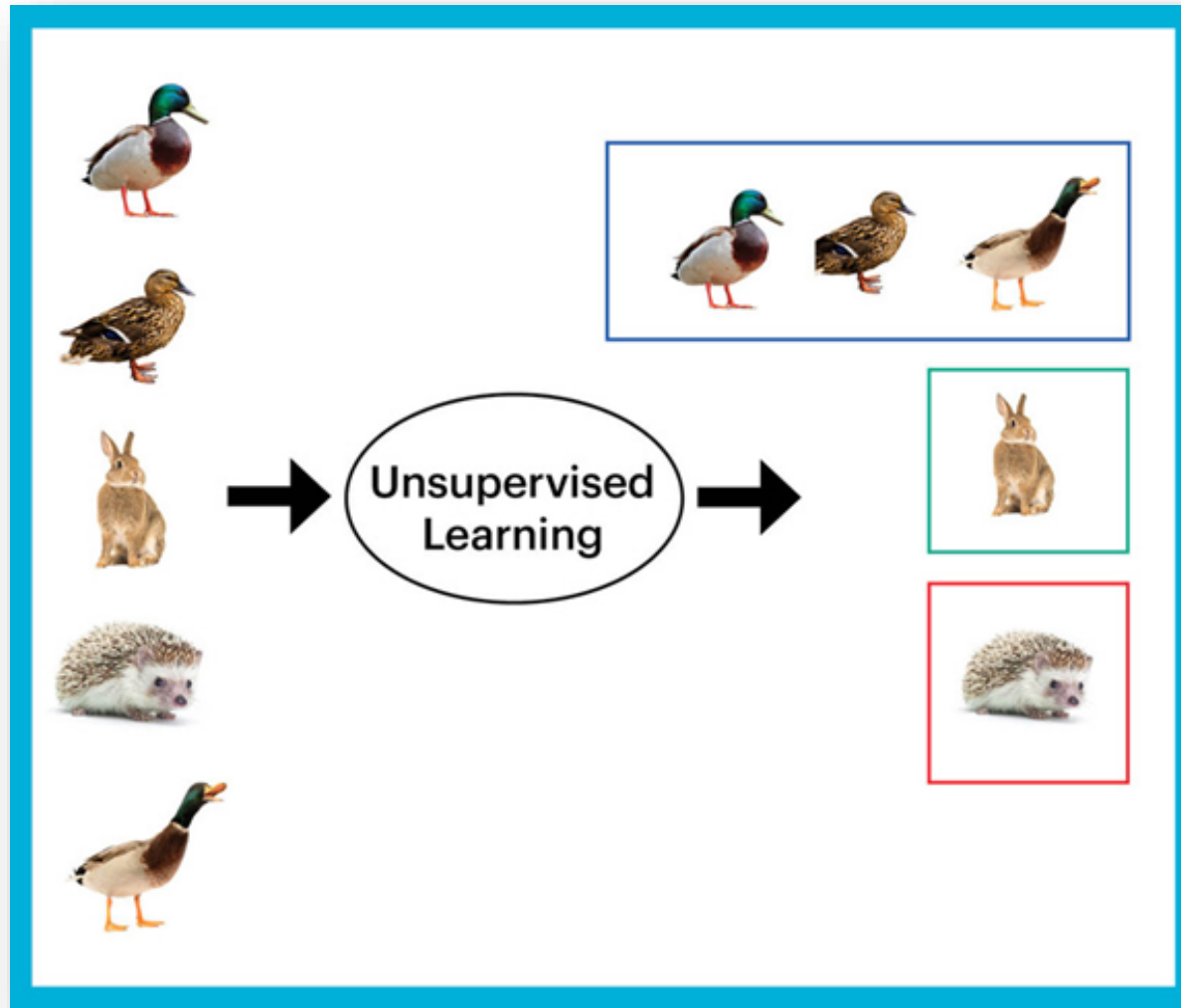

Machine Learning
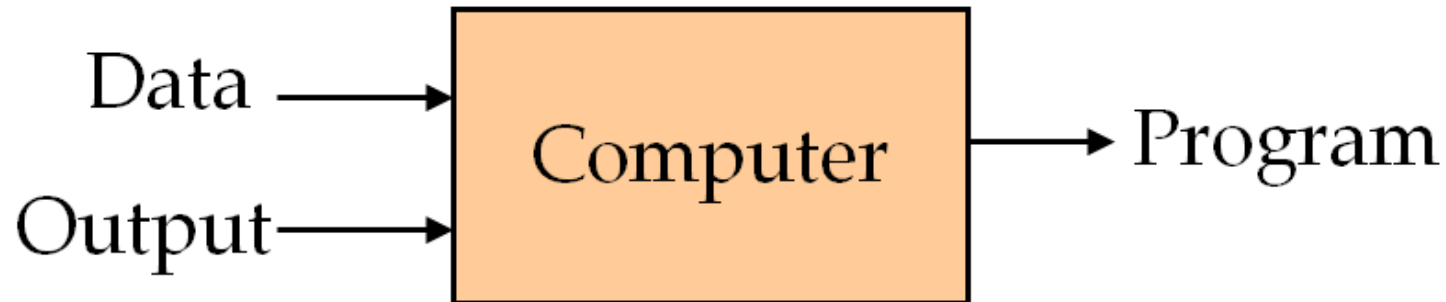
TYPES OF MACHINE LEARNING

- Supervised learning – also called predictive learning

- Unsupervised learning – also called descriptive learning

- Reinforcement learning

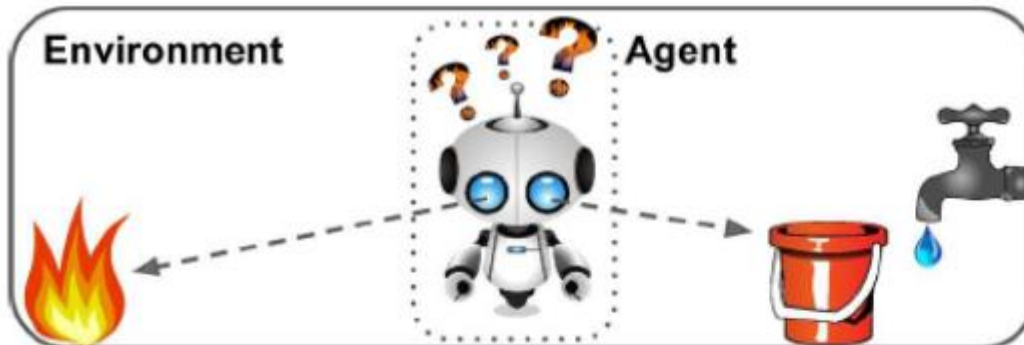# SUPERVISED LEARNING

UNSUPERVISED LEARNING

REINFORCEMENT LEARNING

MACHINE LEARNING PROCESS
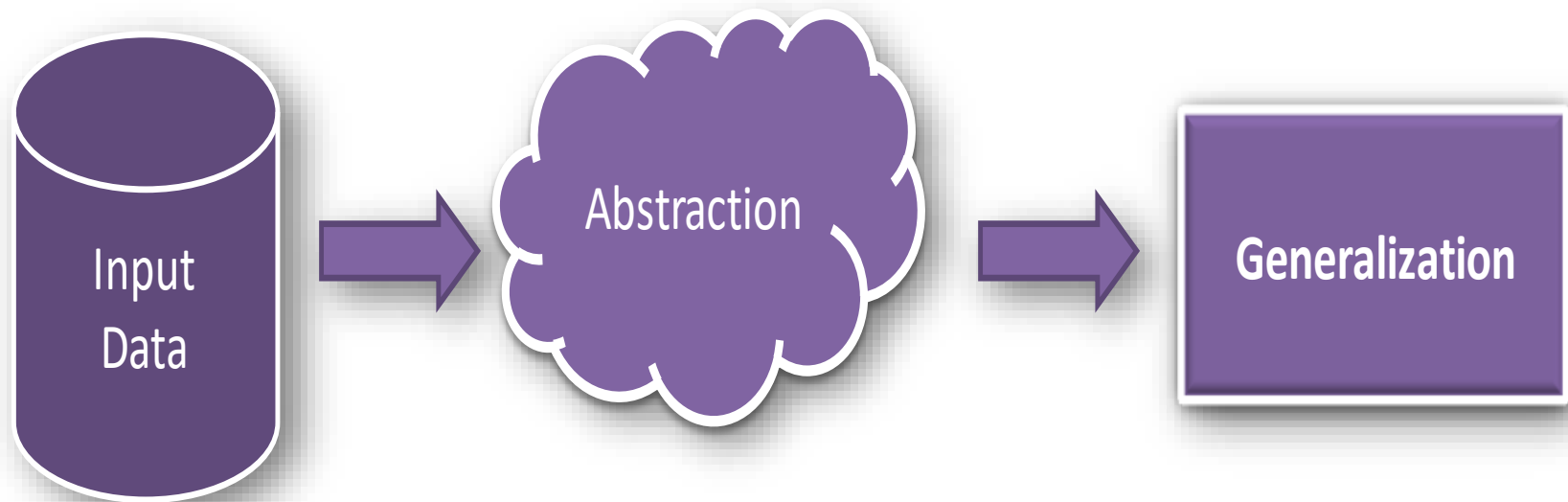
MACHINE LEARNING PROCESS

▸ **Data Input** – Past data or information is utilized as a basis for future decision-making

◉ **Abstraction** – The input data is represented in a broader way through the underlying algorithm

Input Data → Abstraction → Generalization

◉ **Generalization** – The abstracted representation is generalized to form a framework for making decisions

TYPICAL ML PROBLEMS

- Prediction of results of a game
- Predicting whether a tumor is malignant or benign
- Price prediction in domains like real estate, stocks, etc.
- Demand forecasting in retails
- Customer segmentation
- Self-driven cars

PROBLEMS NOT TO BE CONSIDERED FOR ML

- Bank interest calculation

- Inventory management (except the demand forecast module)

- Sorting, Searching

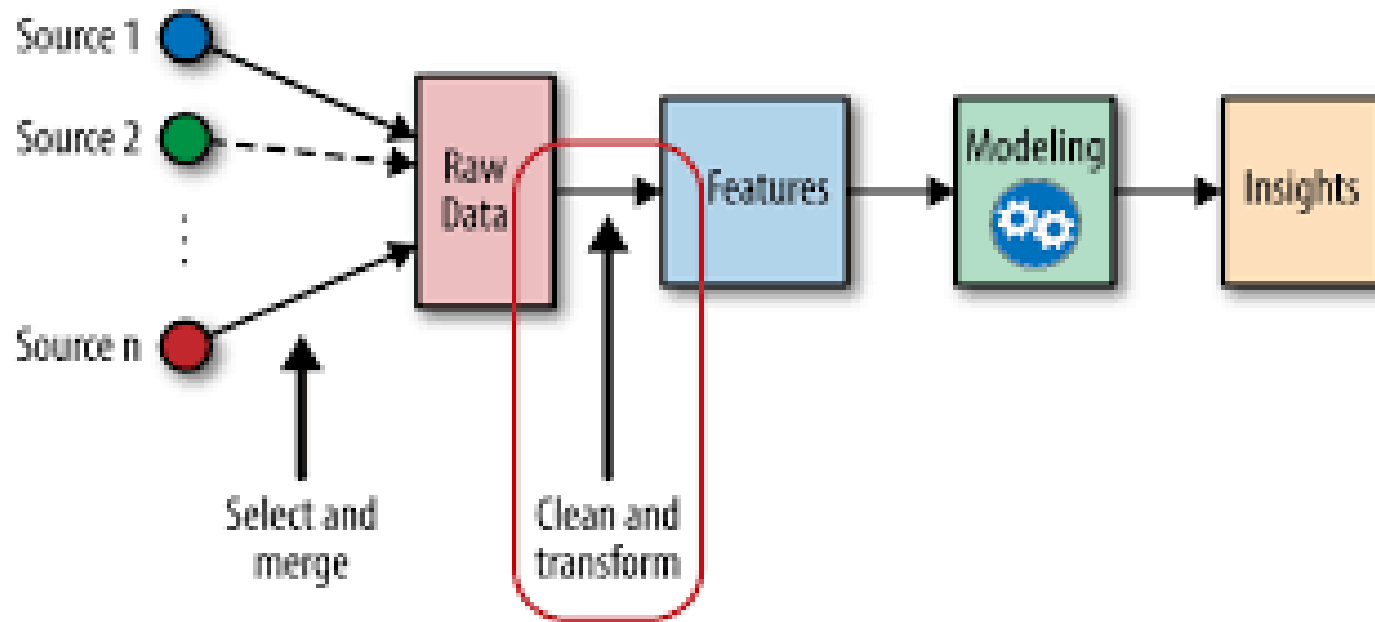- Tasks in which humans are very effective For example, painting

ML in Healthcare & Pharma

- **Proactive health monitoring and alerts**

- **Disease Identification/Diagnosis**

- **Personalized treatment**

- **Disease / epidemic outbreak prediction**

- **Clinical Trial Research**

DATA-FEATURES

DATA-FEATURES

CLASSIFICATION EXAMPLE

Is a Person Fit?

Age < 30 ?

Yes? / No?

Eat's a lot of pizzas?   Exercises in the morning?

Yes?   No?   Yes?   No?

Unfit!   Fit   Fit   Unfit!
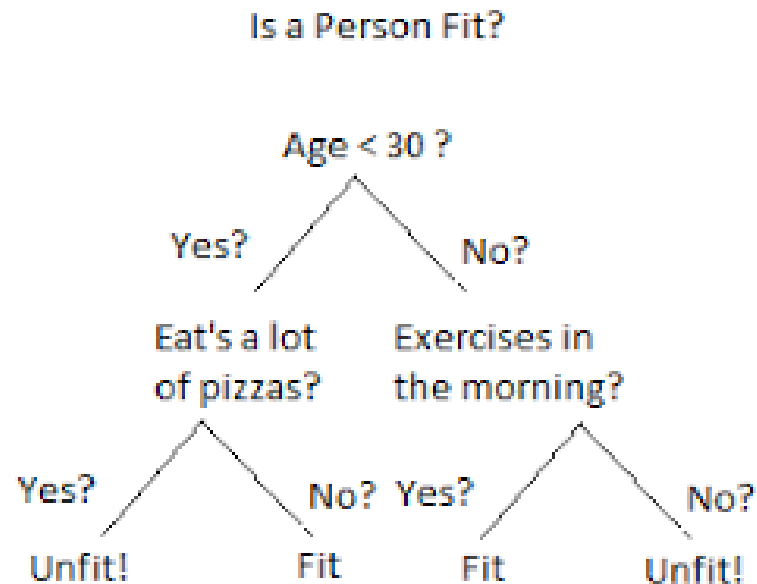
Classification example

Forma	Color	Size	Material	Class
square	red	big	metal	+
square	blue	small	plastic	+
triangle	yellow	medium	metal	+
triangle	pink	big	leather	−
square	pink	medium	leather	−
circle	red	small	plastic	−
circle	blue	small	metal	−
ellipse	yellow	small	plastic	−
ellipse	blue	big	leather	+
ellipse	pink	medium	wood	+
circle	blue	big	wood	+
triangle	blue	medium	plastic	+

CLASSIFICATION EXAMPLE

	case ID	predictors				target
	CUST_ID	CUST_GENDER	EDUCATION	OCCUPATION	AGE	AFFINITY_CARD
	101501	F	Masters	Prof.	41	0
	101502	M	Bach.	Sales	27	0
	101503	F	HS-grad	Cleric.	20	0
	101504	M	Bach.	Exec.	45	1
	101505	M	Masters	Sales	34	1
	101506	M	HS-grad	Other	38	0
	101507	M	< Bach.	Sales	28	0
	101508	M	HS-grad	Sales	19	0
	101509	M	Bach.	Other	52	0
	101510	M	Bach.	Sales	27	1

CLASSIFICATION EXAMPLE
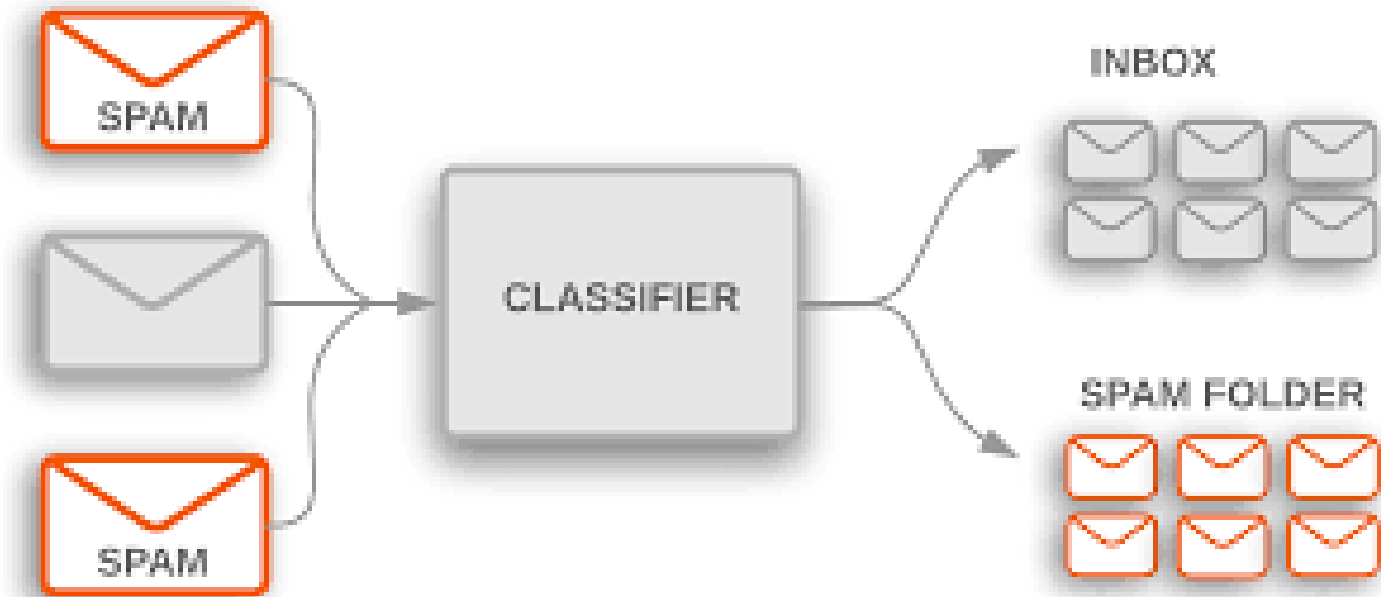
REGRESSION EXAMPLE

Example: House Prices

House Price in $1000s (y)	Square Feet (x)
245	1400
312	1600
279	1700
308	1875
199	1100
219	1550
405	2350
324	2450
319	1425
255	1700

Estimated Regression Equation:

$$\widehat{\text{house price}} = 98.25 + 0.1098\,(\text{sq.ft.})$$

Predict the price for a house with 2000 square feet
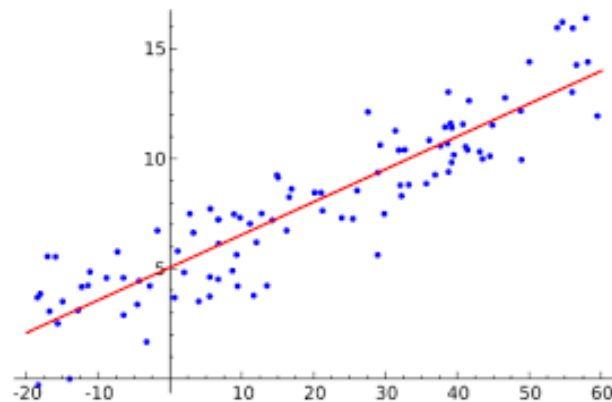
REGRESSION EXAMPLE

Multiple features (variables).

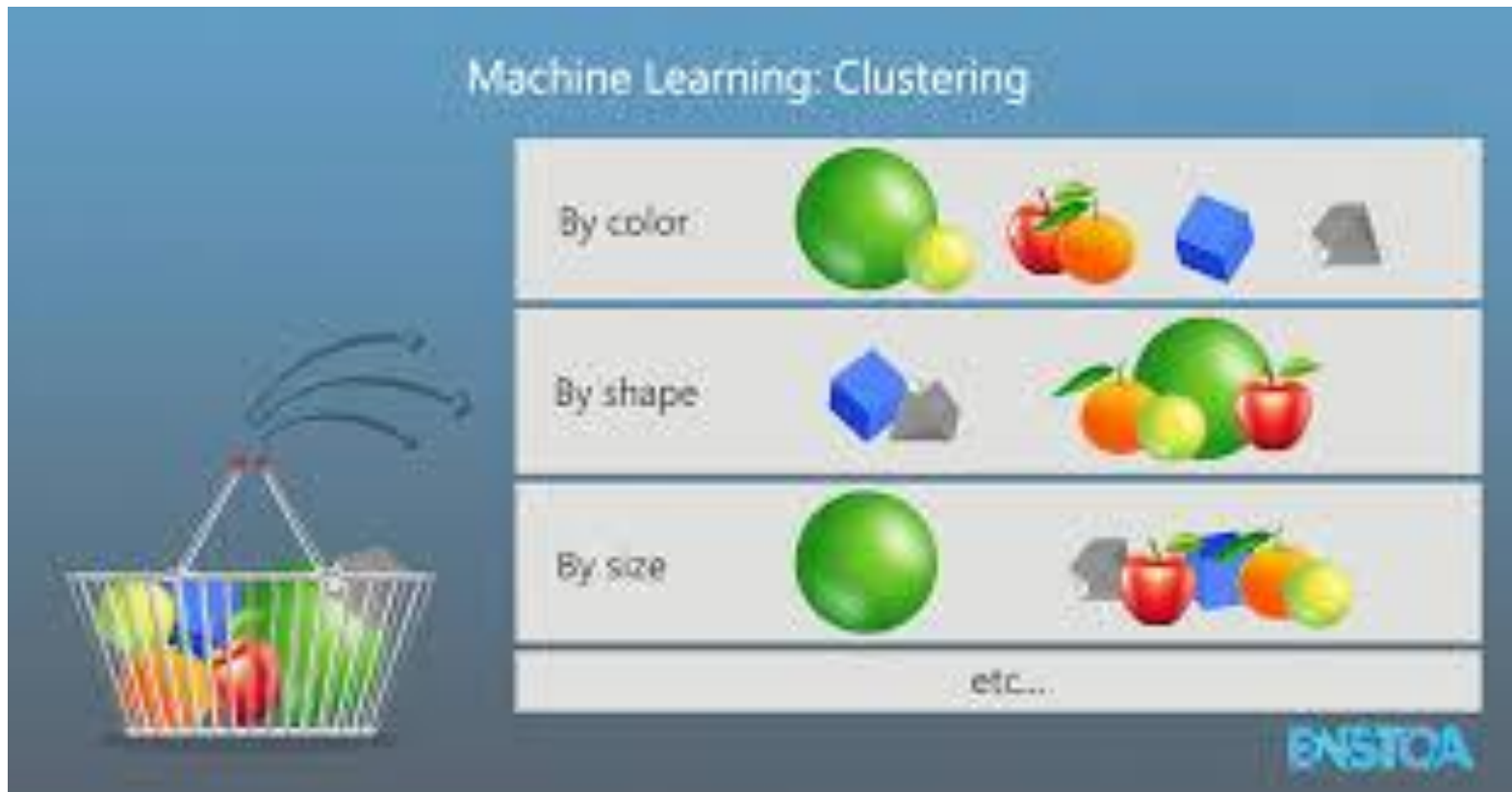| Size (feet²) x_1	Number of bedrooms x_2	Number of floors x_3	Age of home (years) x_4	Price ($1000) y
2104	5	1	45	460
1416	3	2	40	232
1534	3	2	30	315
852	2	1	36	178
...	...	...	...	...

Notation:
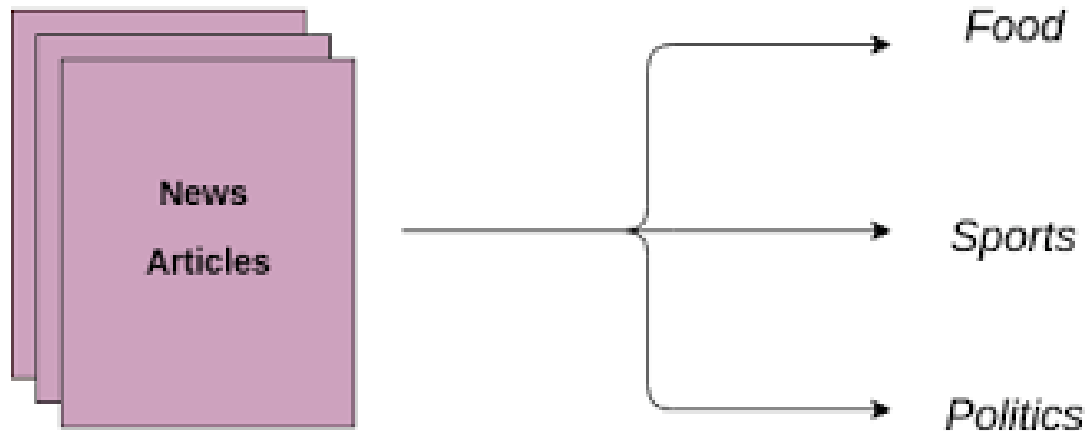
n = number of features $n = 4$

$x^{(i)}$ = input (features) of ith training example.

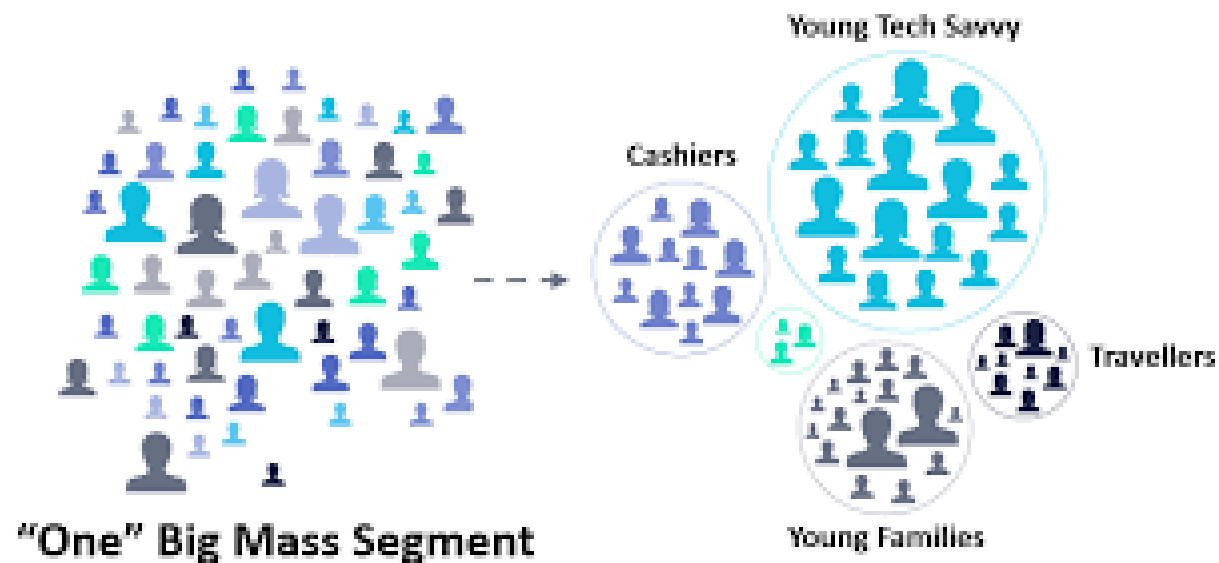$x_j^{(i)}$ = value of feature j in i^{th} training example.
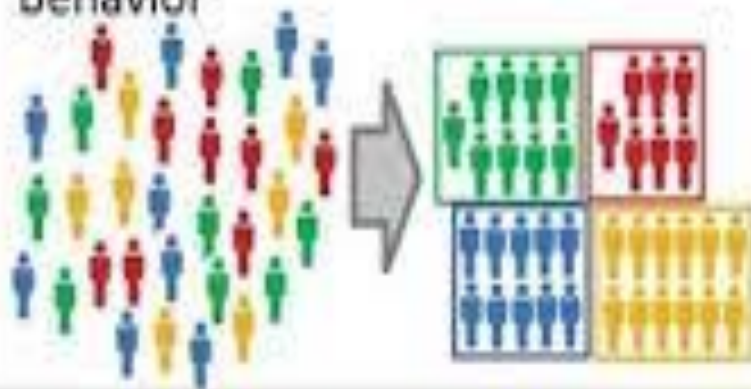
CLUSTERING EXAMPLE

CLUSTERING EXAMPLE

PROFILE SEGMENTATION

Young Tech Savvy

Cashiers

Travellers

"One" Big Mass Segment

Young Families

Unsupervised Learning

Clustering
Grouping customers by purchasing behavior

Association
People that buy X tend to buy Y

People that buy A+B tend to buy C

Introduction to AR

- **Ideas come from the market basket analysis (MBA)**

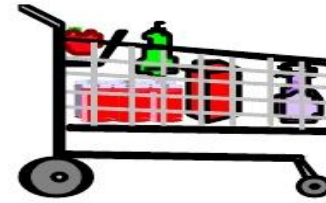 - Let's go shopping!

  Milk, eggs, sugar, bread

  Customer1

  Milk, eggs, cereal, bread

  Customer2

  Eggs, sugar

  Customer3

  - What do my customer buy? Which product are bought together?

  - **Aim:** Find associations and correlations between the different items that customers place in their shopping basket

SENTIMENT ANALYSIS

Customer Feedback Text	Sentiment
"This café is great, the staff are really friendly and the coffee is delicious"	Positive
"I would not recommend this café to anyone. Their coffee is terrible and is really expensive"	Negative

Types of Machine Learning

Supervised Learning

Unsupervised Learning

Reinforcement Learning

Classification
- Fraud detection
- Email Spam Detection
- Diagnostics
- Image Classification

Dimensionality Reduction
- Text Mining
- Face Recognition
- Big Data Visualization
- Image Recognition

Reinforcement Learning
- Gaming
- Finance Sector
- Manufacturing
- Inventory Management
- Robot Navigation

Regression
- Risk Assessment
- Score Prediction

Clustering
- Biology
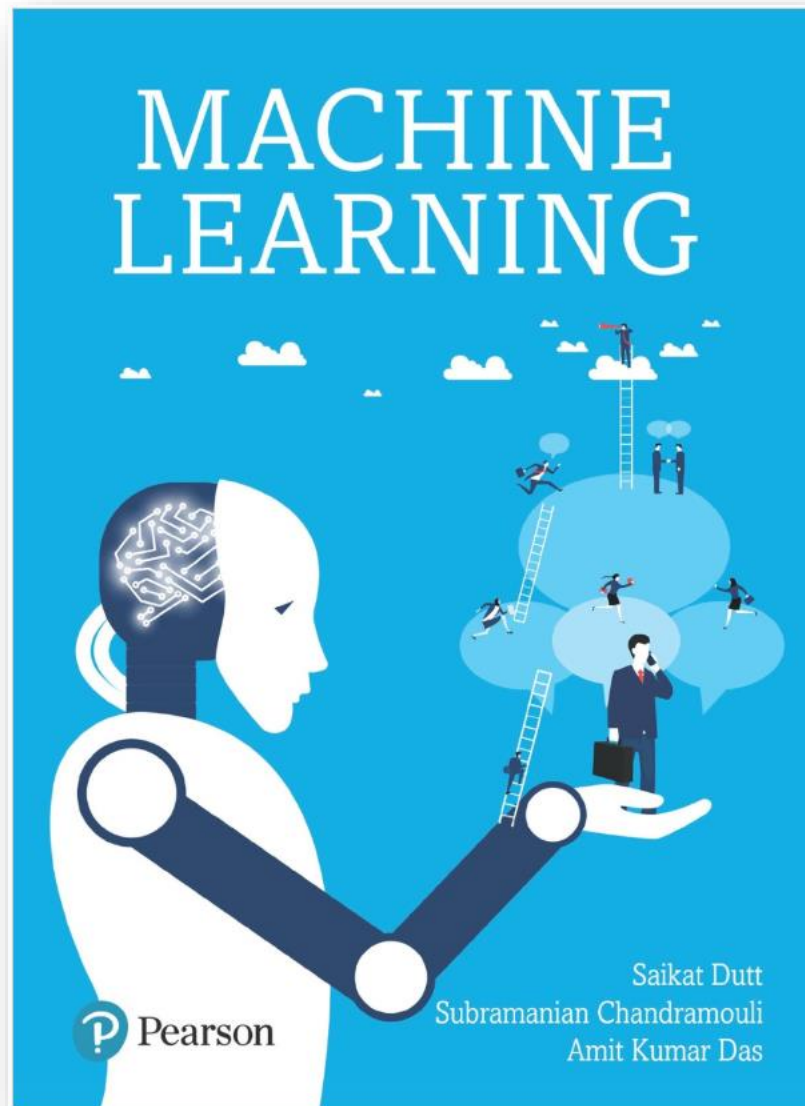- City Planning
- Targetted Marketing

State-of-the-art ML tools

- Python

- R

- MATLAB

FOR FURTHER READING ...

THANK YOU