CDS503:
Machine Learning

Topic 1
Machine Learning Basic

Assoc. Prof. DR UMI KALSOM YUSOF
SCHOOL OF COMPUTER SCIENCES
UNIVERSITI SAINS MALAYSIA (USM)

Contents

- Data Mining
- What is Machine Learning
- Supervised Learning
- Unsupervised Learning
- Learning Associations
- Reinforcement Learning
- Machine Learning Applications

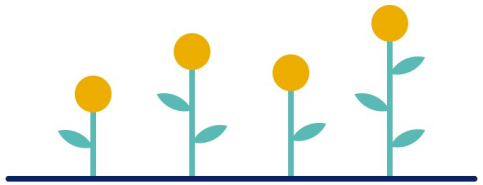
What is Data Mining?

Data mining is considered the process of extracting useful information from a vast amount of data.

To discover new, accurate, and useful patterns in data

Discover past patterns to predict future behaviour.

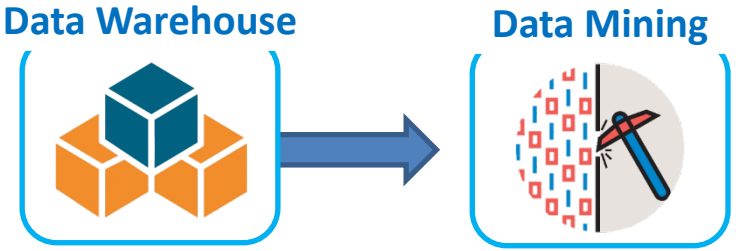
Dealing with **large data**, perhaps Gigabytes, perhaps in Terabytes.



Bigger data = **Higher confidence**

Higher confidence in the **discovered pattern**

Data mining in **small data** is still possible



Why Data Mining Now?

Evolution of Technology

- **Improved** database
- **Faster** & **cheaper** data collection
- **Ability** to analyze & synthesize information

Affordable Computing Power

- Data mining is computationally & resource extensive
- **Reduction** in the price of computer systems

Statistical & Learning Algorithm

- **Adapted** algorithms from fields like statistics & artificial intelligence
- **Enable** development of new algorithm

Large Amount of Data

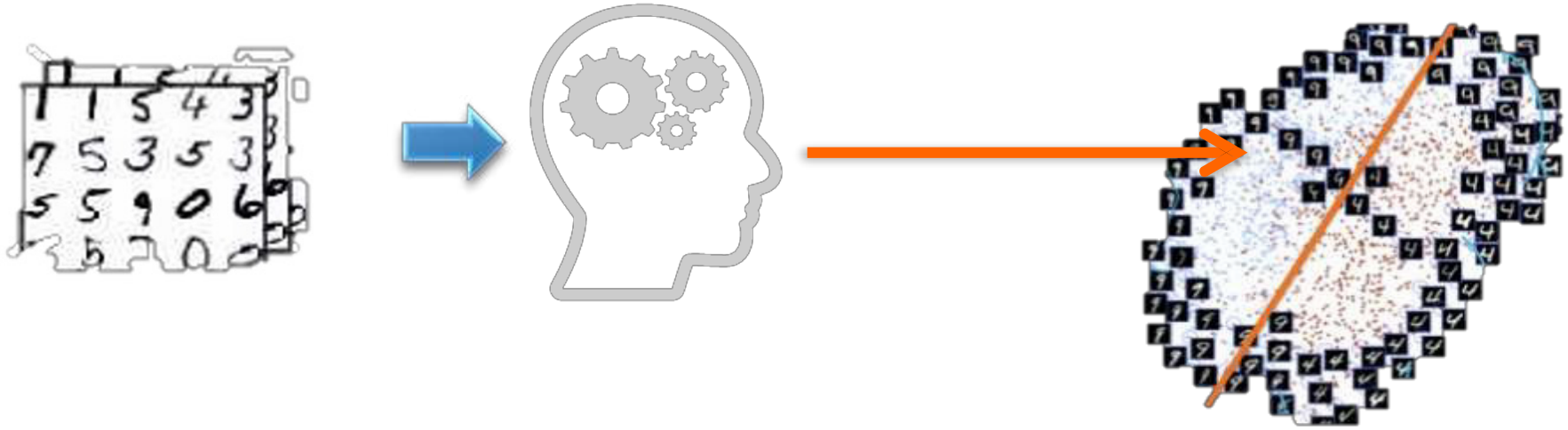
- **Accumulated** data (30+ years)
- Collect information for analysis
- **Example**: *loyalty cards*

Strong Business Competition

- **Growth** in service economies
- **Service economies** are information rich and very competitive.
- **Example:** *Mobile market*

What is Machine Learning?

- Machine learning is the process of discovering algorithms that have improved courtesy of experience derived from data.
- It's the design, study, and development of algorithms that permit machines to learn without human intervention.
- It's a tool to make machines smarter, eliminating the human element (but not eliminating humans themselves; that would be wrong).



Machine Learning Introduction

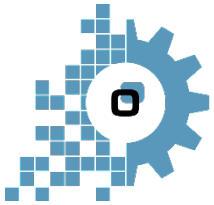
What is Machine Learning?

Learning

- Learning **general model** from **data**
- **Data** is **cheap** & **abundant**
- **Knowledge** is **expensive** & **scarce**
- **Example**: *Consumer behavior from buying transactions*

Machine Learning

Algorithms that **improves** its **performance** at some **tasks** with **experience**



Optimizes performance using **example data** or **past experience**

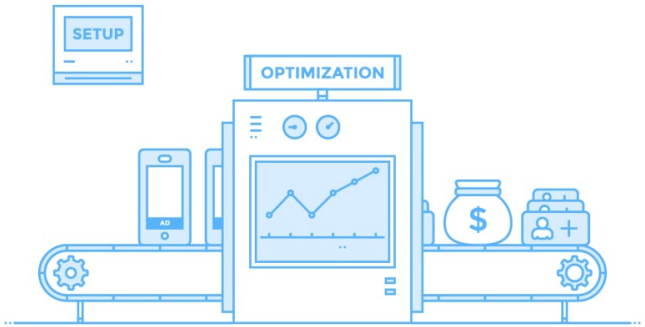
Represent and **evaluate** the **model** for **inference**

Machine Learning Works By

- **Pattern discovery**
- **Weighting important** input
- **Dismiss non-influential** factors

- Data **exploration**
- Clearly defined **business goals**
- Model **train/refine/validate**
- Reliant on **domain/data expertise**
- Model **deployment**

- Model **export** in **common formats** (.xml, .pmml)
- **Automated update**



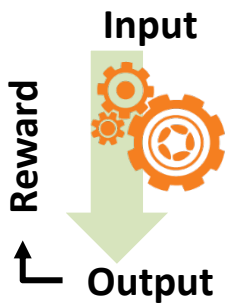
Machine Learning Categories & Classes

Machine Learning Category

Supervised (I/O-based)

Unsupervised (Input-based)

Reinforcement (Reward-based)



Machine Learning Classes

- **Predict "Relations"** — Exhibits specific behaviors in the future
- **Associate "Patterns"** — Event occurs together; Given a series of actions, determine the next likely action
- **Cluster "Differences"** — Group cases with similar characteristics

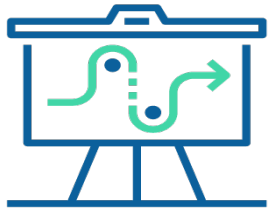
Supervised Learning: Profile and Predict

Build **predictive profile** of **historical outcome**

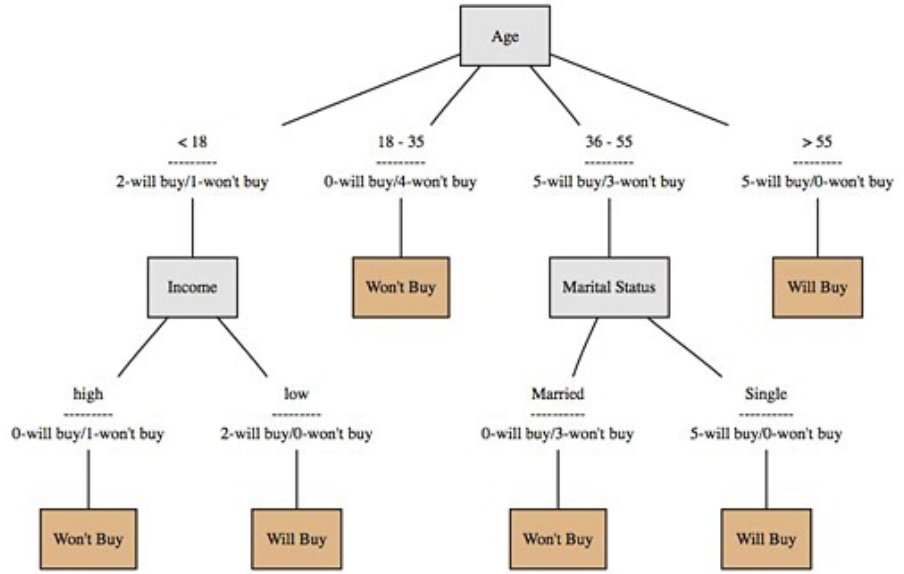
Utilize **collection** of **potential inputs**

Supervise the **process** as the **algorithm** attempts to **model** the **outcome** using **provided inputs**

Explore all **possible** combinations, interactions & contingencies



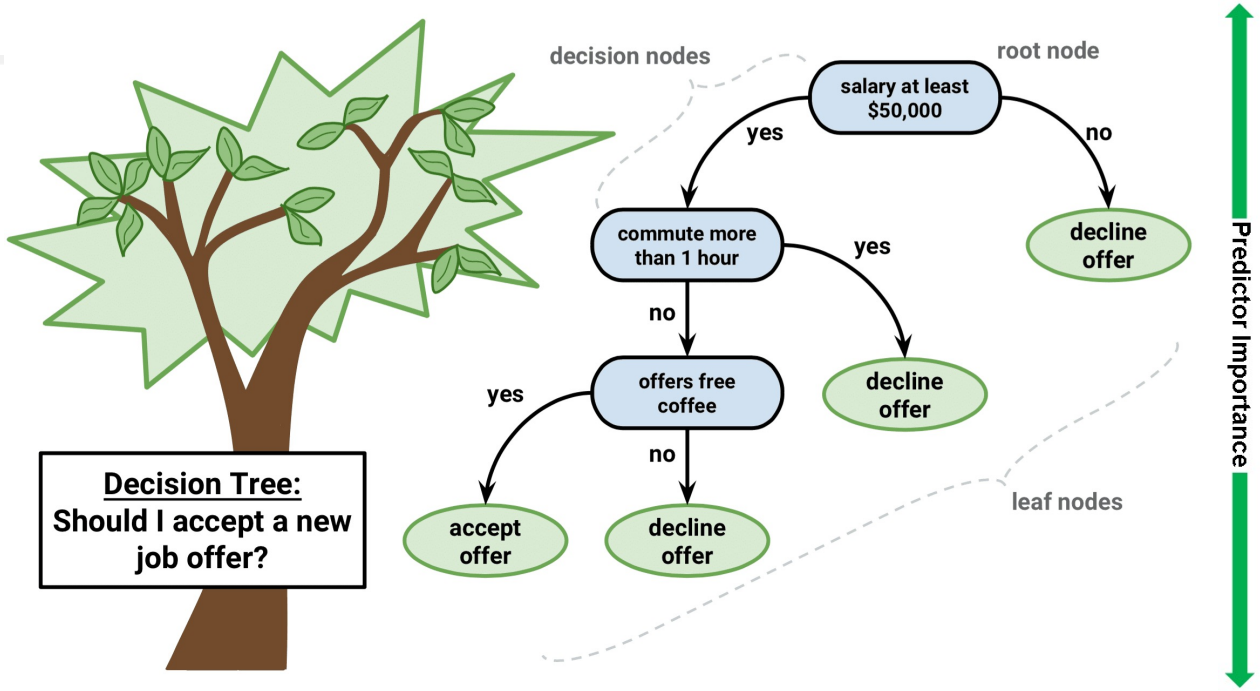
Use that **profile** to **predict future** cases



Supervised Learning:
Decision Trees and Rule Induction

Key Features

- **Predict** and **classify** systems
- **Technique** that shows '**reasoning**'
- Build on **easy** "if– then" rules
- **Eliminates unimportant** factors
- **Contrast** to Neural Network



Summary

- Excellent in **modeling complex** relationships
- Handle **nonlinearity** and **interactions** with **ease**
- Very easy to **understand** and **describe** to others
- Time to **insight** in **minutes**
- Very **accurate** on **small data** set to **inform decision making**

Supervised Learning: Uses

Predict future cases

Use **rules** to **predict** based on **future** input

Knowledge Extraction

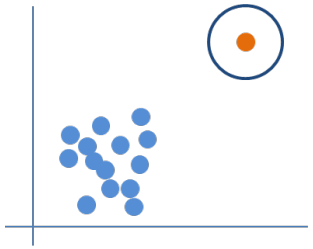
The rule is **easy** to **understand**

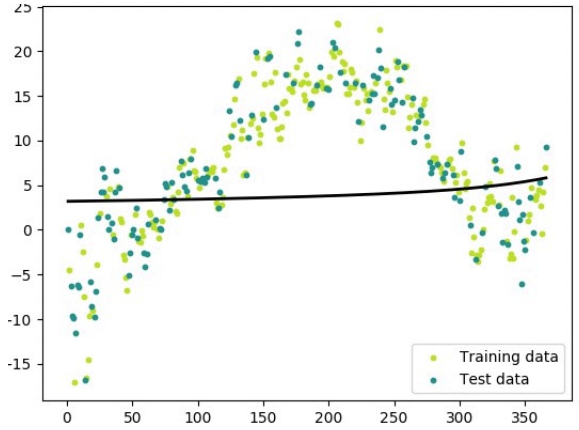
Compression

The rule is **simpler** than the **data** it explains

Outlier Detection

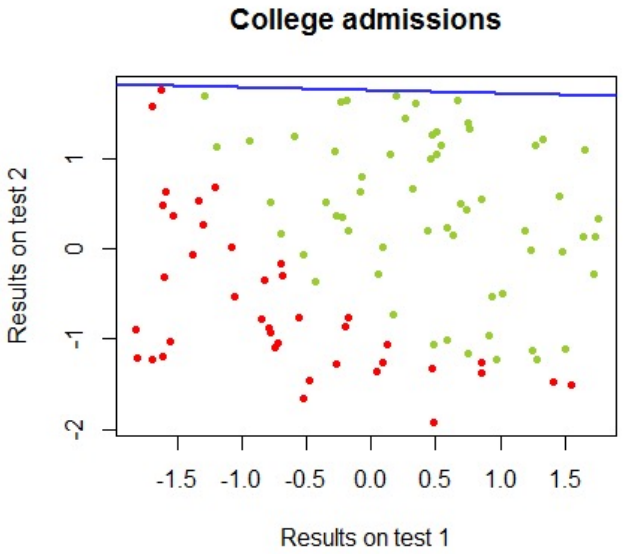
Exceptions not covered by the rule (i.e. **fraud**)





Regression

Example: "**Learn**" the **pattern** of housing prices in order to predict future house prices



Classification

Example: **Classifying** college admission of student based on their admission tests

Training Sets

Test Sets

Face Recognition

Example: "Learn" **different features** of a person **faces**

What is Unsupervised Learning?

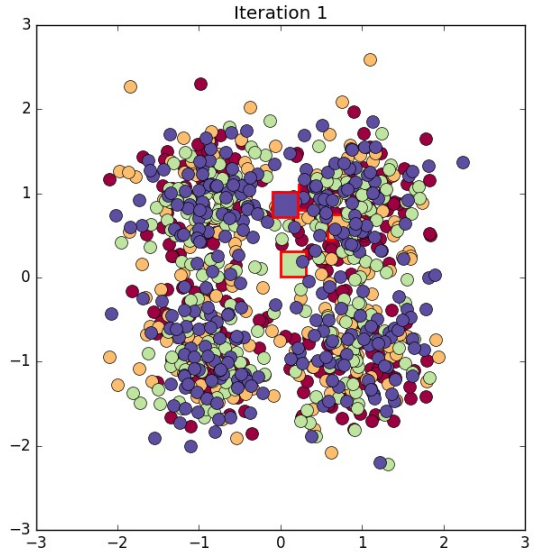
Machine learning is **useful** when we **don't** know the output

Finding '**useful**' **patterns** above & beyond noise

"**Fishing**" for **unknown** yet may be **insightful information**

Clustering

- **Exploratory** data analysis techniques
- Reveal **natural groups** within a **dataset**
- No **prior knowledge** about **groups** or **characteristics**.



Association

- Find things that **occurs together**
- Example: *events in a crime incident*
- Association **can exist** between **any** of the attributes (**no single outcome** like Decision Tree)

Sequential Association

- Discover association rules in **time-oriented data**
- Find the **sequence** or **order** of the **events**

Machine Learning
Learning Association

What one customer **purchased** at **one time**?

What **merchandise** customers are **buying** and **when**?

How to **uncover associations** of items that are **frequently bought together**?

Market Basket Analysis

Transaction 1	🍎 🍺 🍚 🍗
Transaction 2	🍎 🍺 🍚
Transaction 3	🍎 🍺
Transaction 4	🍎 🍐
Transaction 5	🍼 🍺 🍚 🍗
Transaction 6	🍼 🍺 🍚
Transaction 7	🍼 🍺
Transaction 8	🍼 🍐

- If **X** and **Y** are **products/services**
- **P(Y|X)** is the **probability** that somebody who **buys X** also **buys Y**
- **Example**: **P(*beer* | *apple*) = 0.7**

- **Apriori algorithm** & **association rules** can be used to **mine frequent item sets** to **discover association** among items.
- Three **measures**: *Support*, *Confidence*, *Lift*

$$\text{Support}\{🍎\} = \frac{4}{8}$$

$$\text{Confidence}\{🍎 \rightarrow 🍺\} = \frac{\text{Support}\{🍎, 🍺\}}{\text{Support}\{🍎\}}$$

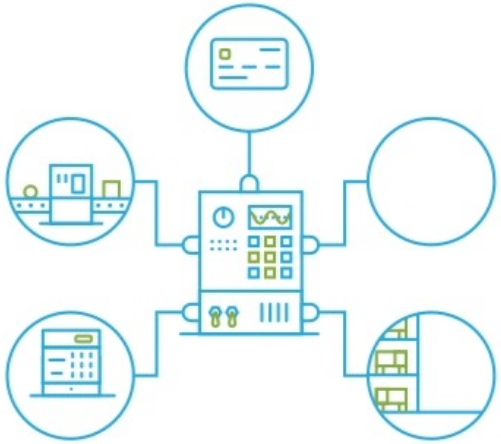
$$\text{Lift}\{🍎 \rightarrow 🍺\} = \frac{\text{Support}\{🍎, 🍺\}}{\text{Support}\{🍎\} \times \text{Support}\{🍺\}}$$

Learning Associations
Example Application

Amazon.com uses associations to recommend customers based on their past purchases and what other customers are purchasing.

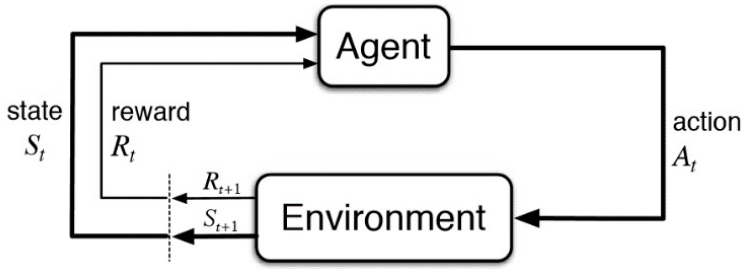
A store in USA "*Just for Feet*" has about 200 stores, each carrying up to 6000 shoe styles, each style in several sizes.

Data mining is used to find the right shoes to stock in the right store.



What is Reinforcement Learning?

- A **reward-based** learning
- Input/state-driven
- Learns to **react** to its **environment**
- **Example**: *teaching a child to memorize a word or a dog to fetch*



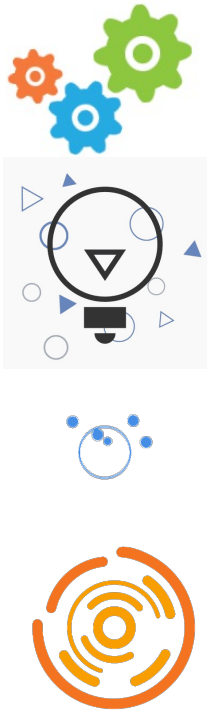
Main Elements

Policy (Core)
How to behave in a state; stimulus-response rules or action

Reward Signal
Short term reward, defines positive and negative ones

Value Function
Long term reward, a refined & farsighted judgement; desired goal

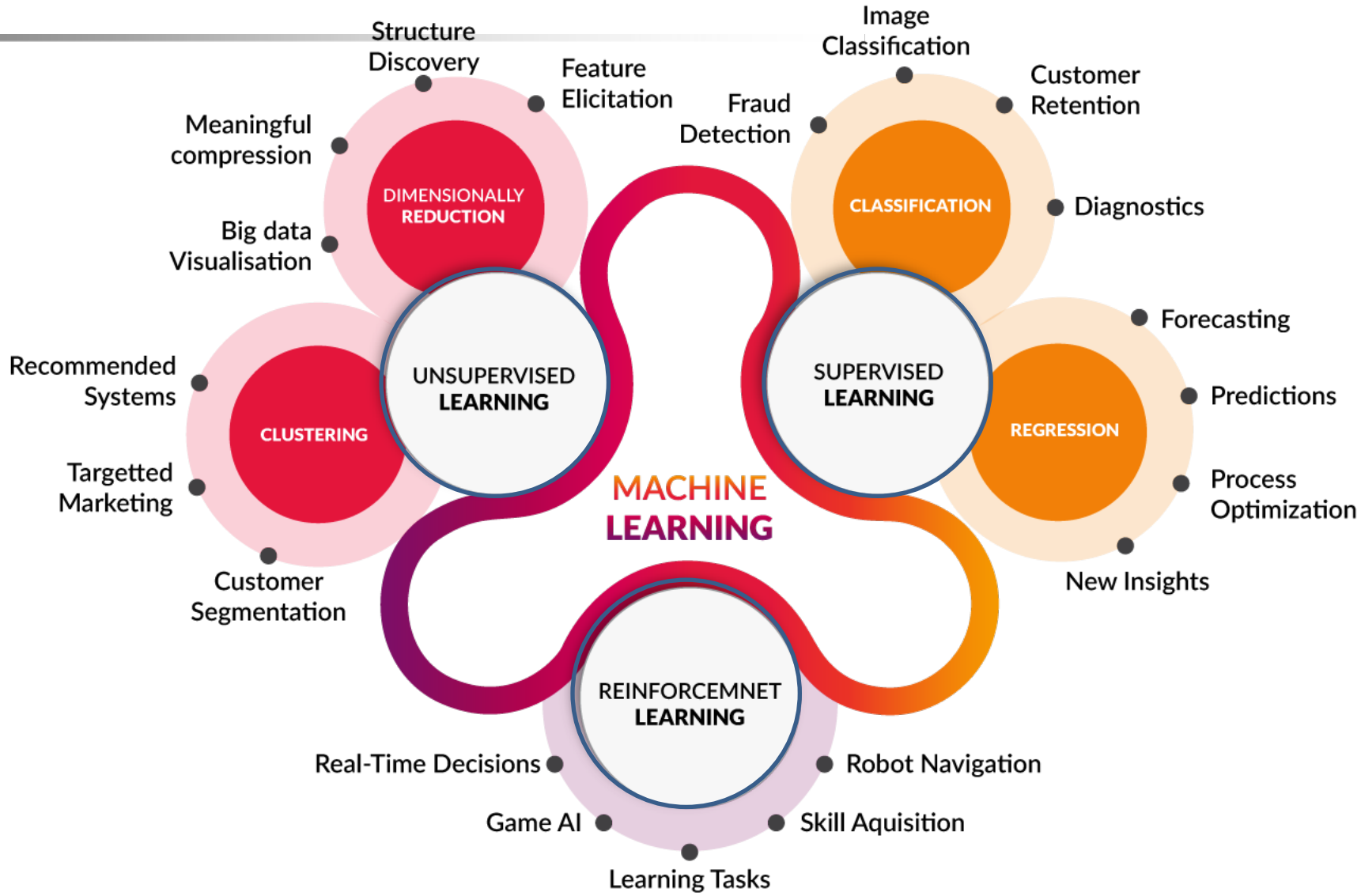
Model (*optional*)
Allow inference on how the environment will behave



Automate Ideal Behaviors	**Stochastically explore** and **exploit** to maximize **rewards**
Learn by Experience	**Interacts** and **learn** from its own **experience**
Generalized Learning	Considers the "**whole**" problem (**big picture**)
Robustness	**Robust** to react against **uncertainties**

Source: Sutton, Richard S., and Andrew G. Barto. "Reinforcement learning: An introduction." (2011).

Machine Learning Applications



Machine Learning Applications

Sophisticated applications of modern **enterprises** include:

- Sales forecasting and analysis
- Marketing and promotion planning
- Business modeling

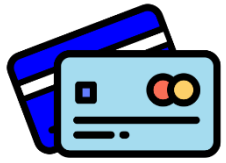
OLTP is **not designed** for such applications

Large enterprises operate a **number** of database **systems**.

It is necessary to **integrate information** for **decision making applications**.

Question: Why OLTP cannot be used for sales forecasting and analysis?

Machine Learning Applications in Finance, Retails, Telecom, Insurance, etc.



Loan & Credit Card Approval

Trend Analysis

Market Segmentation

Market Basket Analysis

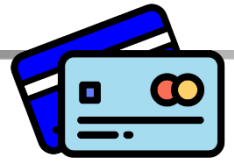
Fraud Detection

Customer Churn

Better Marketing

Website Design & Promotion

Machine Learning Applications



Loan & Credit Card Approval

- Bank **does not know** its **customers**
- Only have their **information**
- Collect customers' **behavioural data**
- Collect from **many sources**
- To **predict** the chances of a customer **paying back a loan**

Market Segmentation

- **Large** amount of data about customers
- May contain **valuable information**
- Segments them by **subgroups**
- Use **variables** that are **good discriminators**
- **Difficult** to find these **variables**

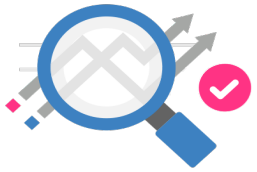
Fraud Detection

- Difficult to **define characteristics** of fraud
- Often **detected** based on **change in norm**
- In statistics, **common** to throw out **outliers**
- In machine learning, **useful** to identify them due to **errors** or **fraud**

Better Marketing

- Customer buy a **new** products, **suggest** other products when they ready
- Will the customer **respond**?
- Will the customer **purchase** and **how much**?
- Will the customer **return purchase**?
- Will the customer **pay** for the **purchase**?

Machine Learning Applications



Trend Analysis

- In a large company, not all **trends** are **visible**
- Useful to use **data mining** to **identify trends**
- Trends may be **long term** trends, **cyclic** trends or **seasonal** trends

Market Basket Analysis

- What customer **buy** & **buy together**
- Useful in **designing store layouts** or in **deciding** which items to **put on sale**
- Can also be **used** for **other applications**

Website Design & Promotion

- Web design that are **dynamic** and **adaptive** to **changing contents** & **platforms**
- Perceptive of **emotion** and **highly personalized** to the **individual** customer

Customer Churn

- Keep **good customers** and **persuade good customers** of their competitors to **switch**
- Find **those customers**, why they **switch** and what makes customers **loyal**.
- Cheaper to develop a **retention plan** (**retain old customers**) than to bring in **new customers**.
- Get to **know** the customers **better** (**keep them longer**)
- Remove **customers** that **cost more** (**than their worth**)

Machine Learning Applications

Machine Learning Applications (Classification)

Pattern Recognition
Classify important **feature**

Character Recognition
Different hand writing **styles**

Face Recognition
- Pose
- Lighting
- Occlusions
- Makeup
- Hair Styles

Medical Diagnosis
From **symptoms** to **illnesses**

Web Advertising
Predict user clicking an ads

Speech Recognition:
Temporal Dependency

- Utilizes **language syntax**/ **dictionary**
- Combine **multiple modalities**
- **Example**: **visual** of lip images & **acoustic** of speech

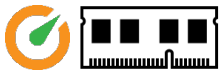
Machine Learning Applications (Reinforcement Learning)

Game Theory & Multi-Agent Interaction

Robotics

Personalized Web Services

Vehicular Navigation



Optimize Memory Control

Thank you