Class 2- What is Machine Learning?

Econometrics

Mathematics and Statistics

WE ARE HERE

Machine Learning

Deep Learning

Artificial Intelligence

Prof. Pedram Jahangiry
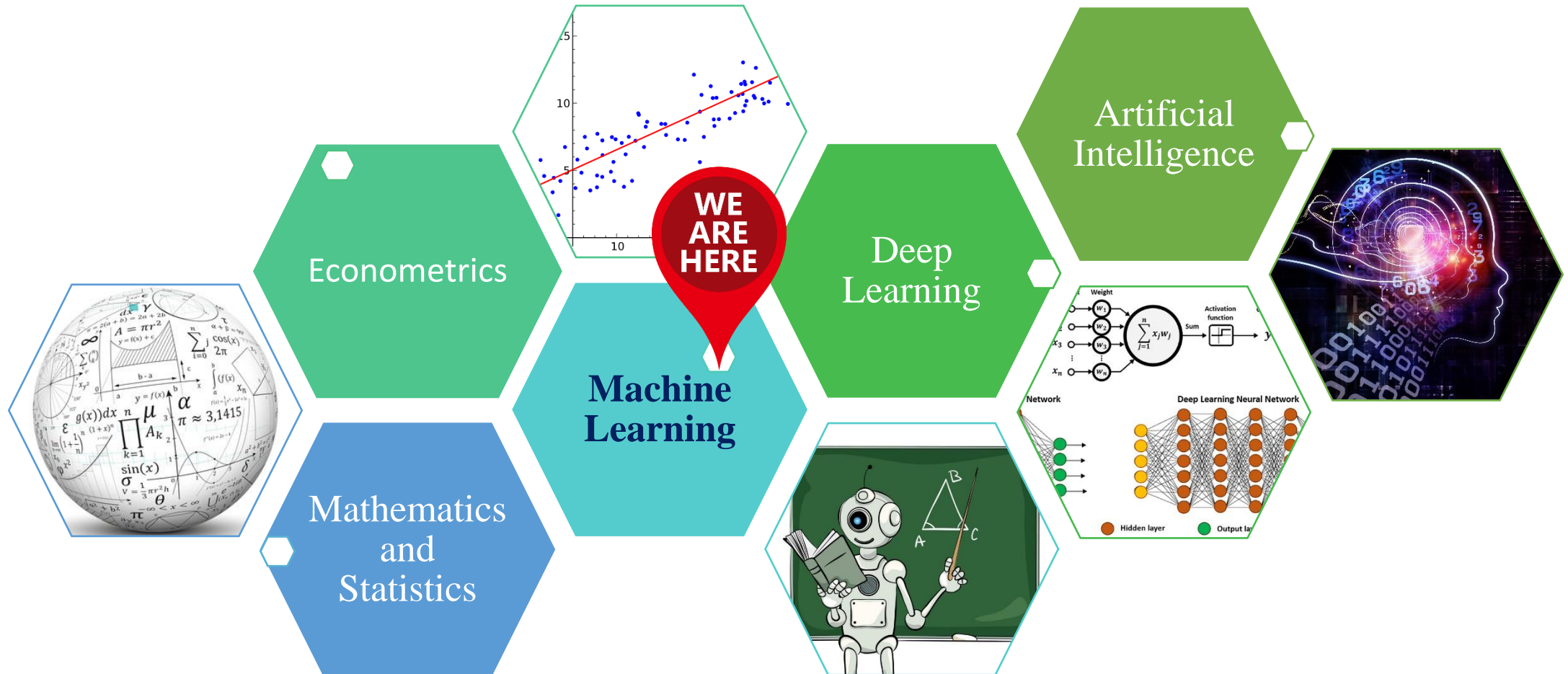
What is Machine Learning?

"A ML algorithm learns complex patterns in a high dimensional space without being specifically directed."

Advances in Financial Machine Learning (2018, p.15)

Let's break this statement into its components:

- ✓ **"learns … without being specifically directed"**: Unlike with other empirical tools, researchers do not impose a particular structure on the data. Instead, researchers let the data speak.

- ✓ **"learns complex patterns"**: The ML algorithm may find a pattern that cannot be easily represented with a finite set of equations.

- ✓ **"learns … in a high-dimensional space"**: Solutions often involve many variables and the interactions between them.

ML involves splitting the dataset into three distinct subsets: a training dataset, a validation dataset, and a test dataset.
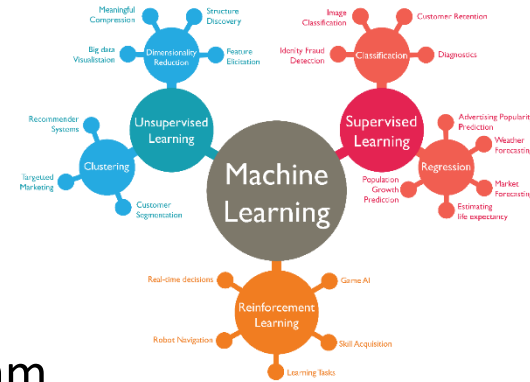
General Programming vs Machine Learning
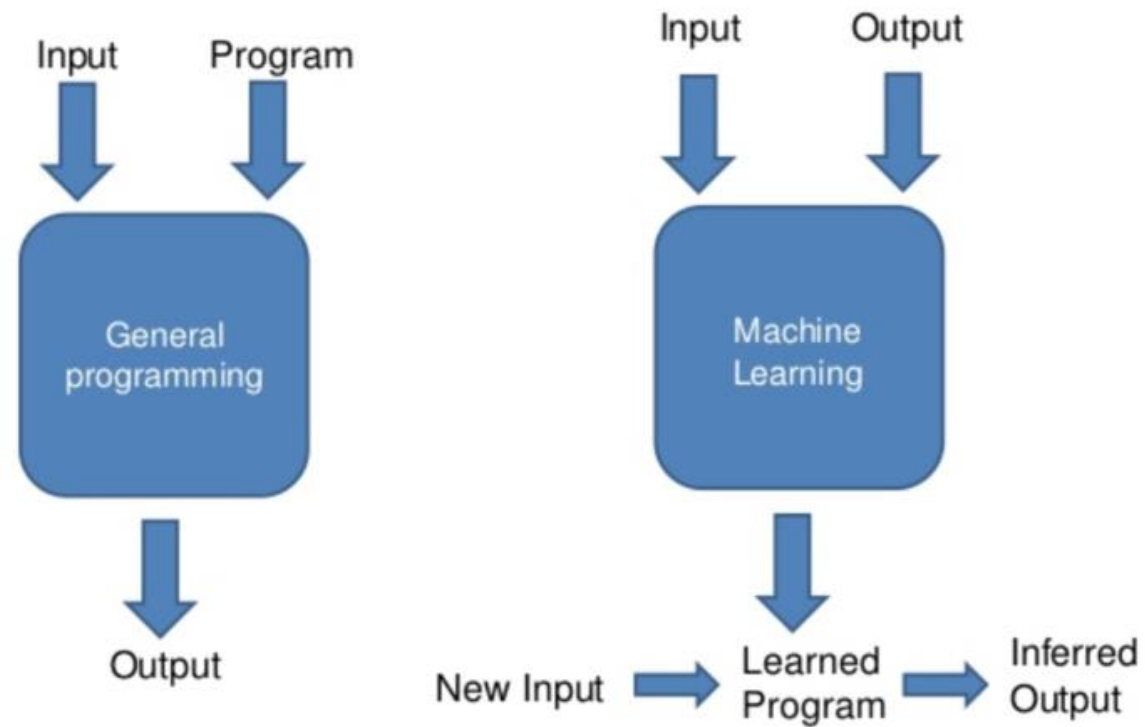
Types of Machine Learning
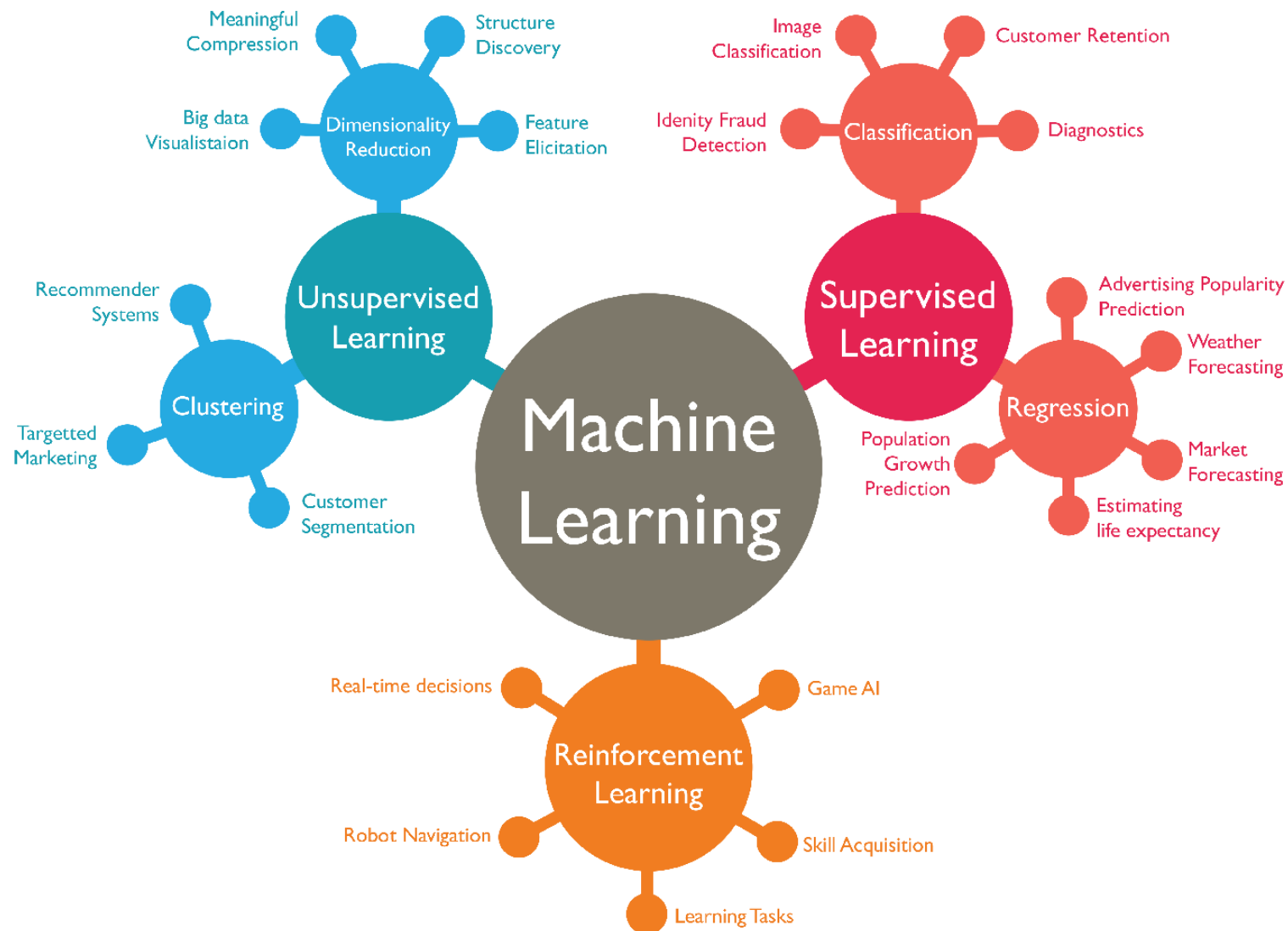
Supervised Learning

- In **supervised learning**, computers learn to model relationships based on labeled training data. In supervised learning, inputs and outputs are labeled for the algorithm. After learning how best to model relationships for the labeled data, the trained algorithms are used to model or predict outcomes for new datasets. Examples:

- **Regression:**
1. Predicting stock market returns
2. Predicting revenue growth
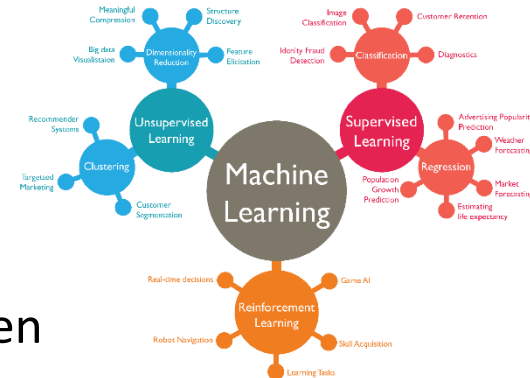3. ….

- **Classification:**
1. Generating buy, sell, hold signals.
2. Likelihood of a successful M&A or IPO
3. Enhancing detection of fraud in financial statements
4. Classification on winning and losing funds or ETFs
5. …
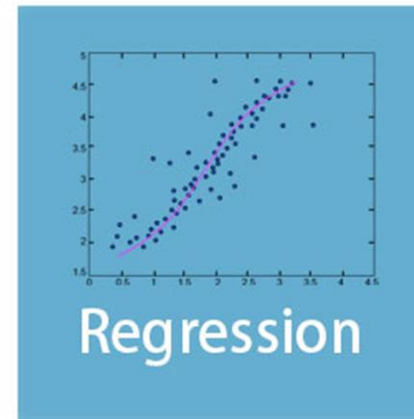
Unsupervised Learning

- In **unsupervised learning,** computers are not given labeled data but instead are given only data from which the algorithm seeks to describe the data and their structure. Examples:

- **Clustering:**
1. Grouping companies into peer groups based on some non-standard characteristics like financial statement data or corporate characteristics rather than sectors or countries.
2. Client profiling and asset allocation
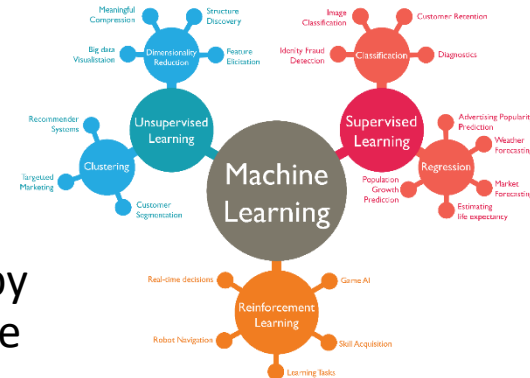3. Portfolio diversification and stock selection based on co-movements similarities
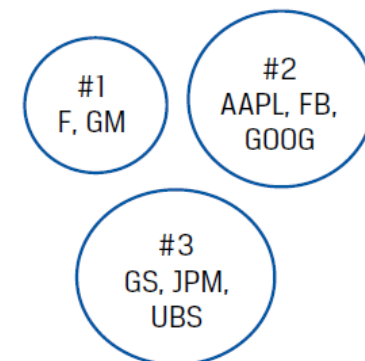
- **Dimensionality Reduction:**
1. Identify the most predictive factors underlying asset price movements (to avoid factor zoo)

Reinforcement Learning

- In **reinforcement learning**, a computer learns from interacting with itself or data generated by the same algorithm. You need to define the environment, actions and the reward system. The machine will then explore and exploit to maximize the reward. The new actions may not be immediately optimal. The learning subsequently occurs through millions of trials and errors.
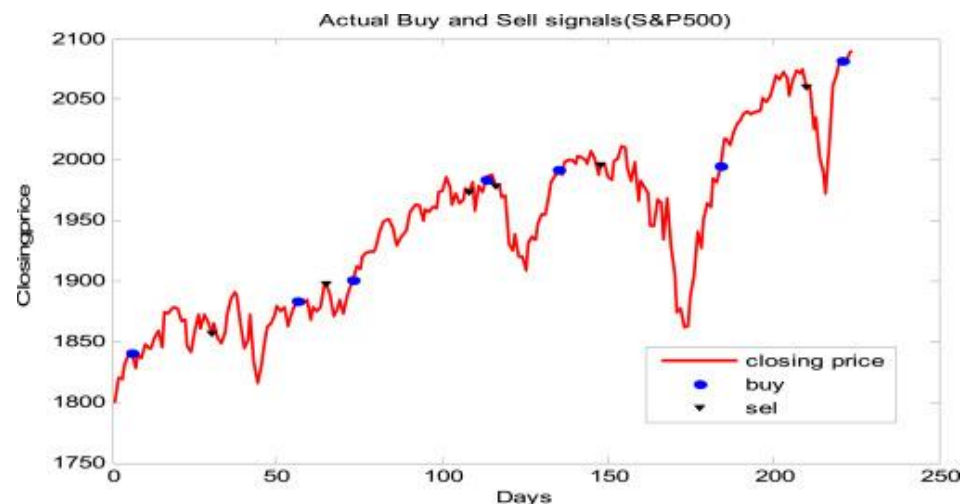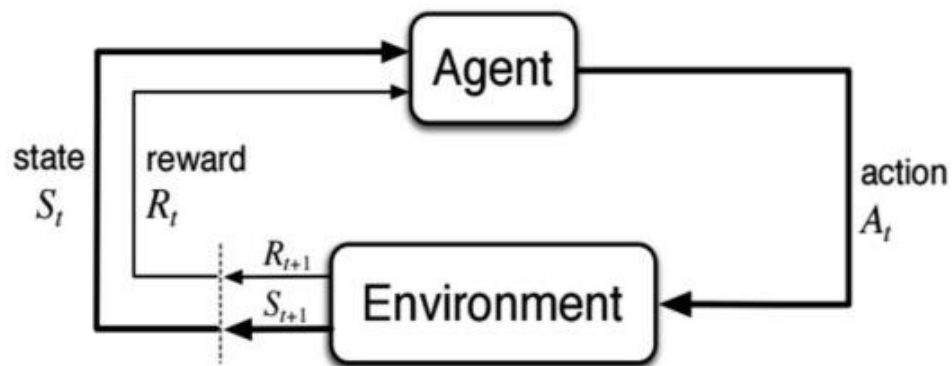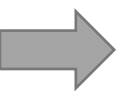

- Example: a virtual trader (agent) who follows certain trading rules (the actions) in a specific market (the environment) to maximize its profits (its reward).
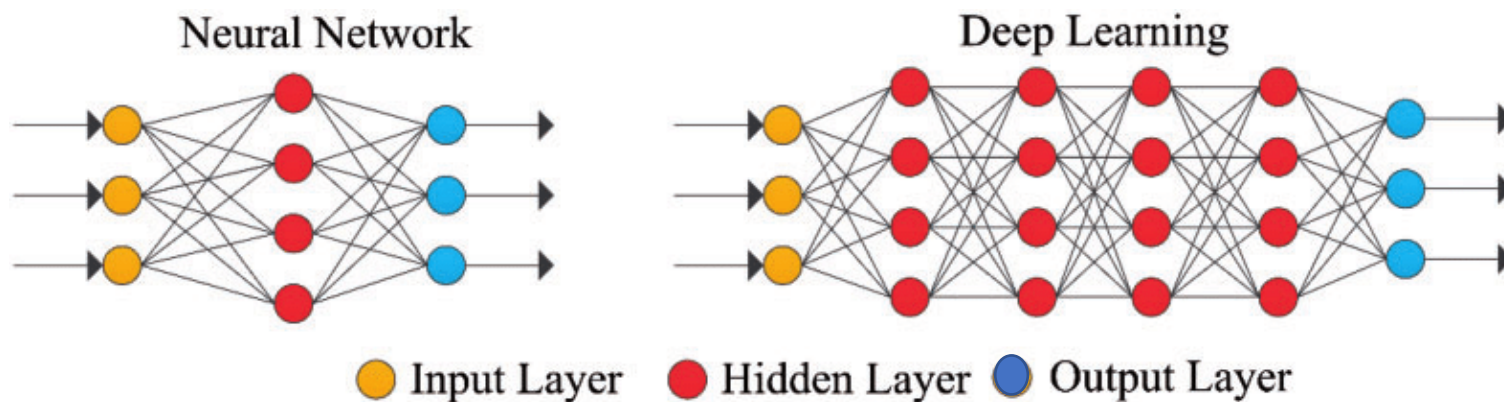
AI vs ML vs DL

Deep Learning

Neural Network   Deep Learning

○ Input Layer   ● Hidden Layer   ● Output Layer
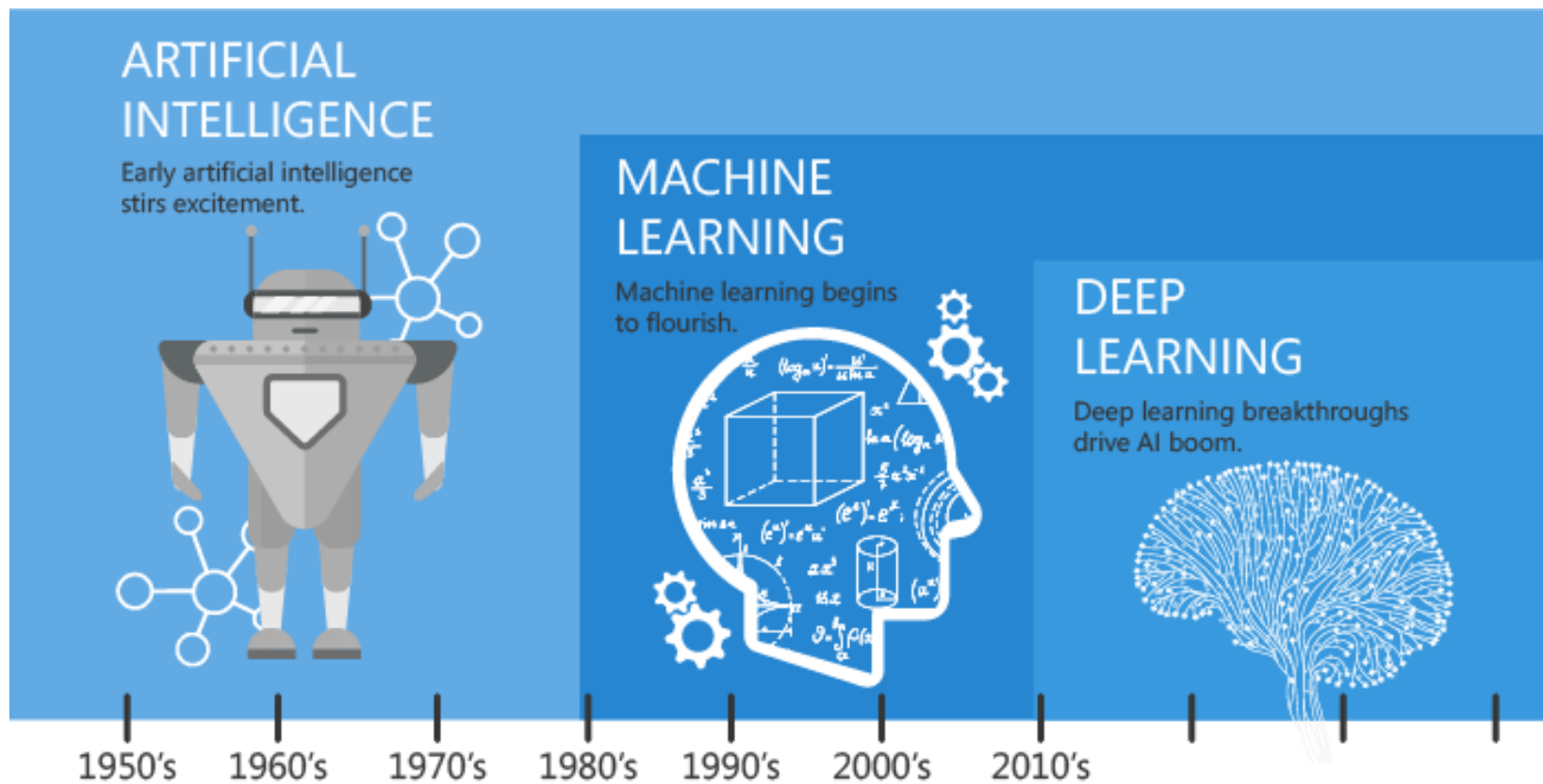
- Examples:
1. Image recognition algorithms can now analyze data from satellite-imaging systems to provide intelligence on the number of consumers in retail store parking lots,
2. Shipping activity and manufacturing facilities, and
3. Yields on agricultural crops

Natural Language Processing (NLP)

NLP is a field of research at the intersection of **computer science**, **artificial intelligence**, and **linguistics** that focuses on developing computer programs to analyze and interpret **human language**.

- Automated tasks using NLP include translation, speech recognition, text mining, sentiment analysis, and topic analysis.

- Examples:
  - Reading millions of pages of annual reports, thousands of hours of earning calls, transcripts, news articles and social media posts to identify trend in shorter timespans!
  - Analyzing communications and transcripts from policymakers (FED, ECB, …) to provide insights around trending topics like interest rate policy, GDP, inflation expectation and etc
  - Chatbots answer basic retirement savings questions, learning from their interactions with investors.
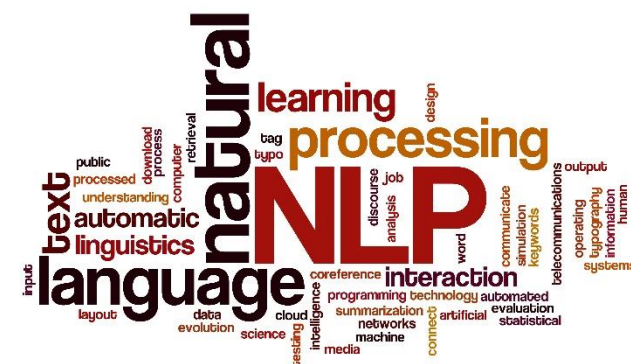
Cognitive computing

- **Cognitive Computing** focuses on mimicking human behavior and reasoning to solve complex problems.

- **Cognitive Computing** is not responsible for making the decision for humans. They simply supplement information for humans to make decisions.

  - **Algo trading**: Algorithmic trading is the computerized buying and selling of financial instruments, in accordance with pre-specified rules and guidelines.

  - **High-frequency trading (HFT)** is a form of algorithmic trading that makes use of vast quantities of granular financial data (tick data, for example) to automatically place trades when certain conditions are met. Trades are executed on ultra-high-speed, low-latency networks in fractions of a second. HFT algorithms decide what to buy or sell and where to execute based on real-time prices and market conditions, seeking to earn a profit from intraday market mispricing.
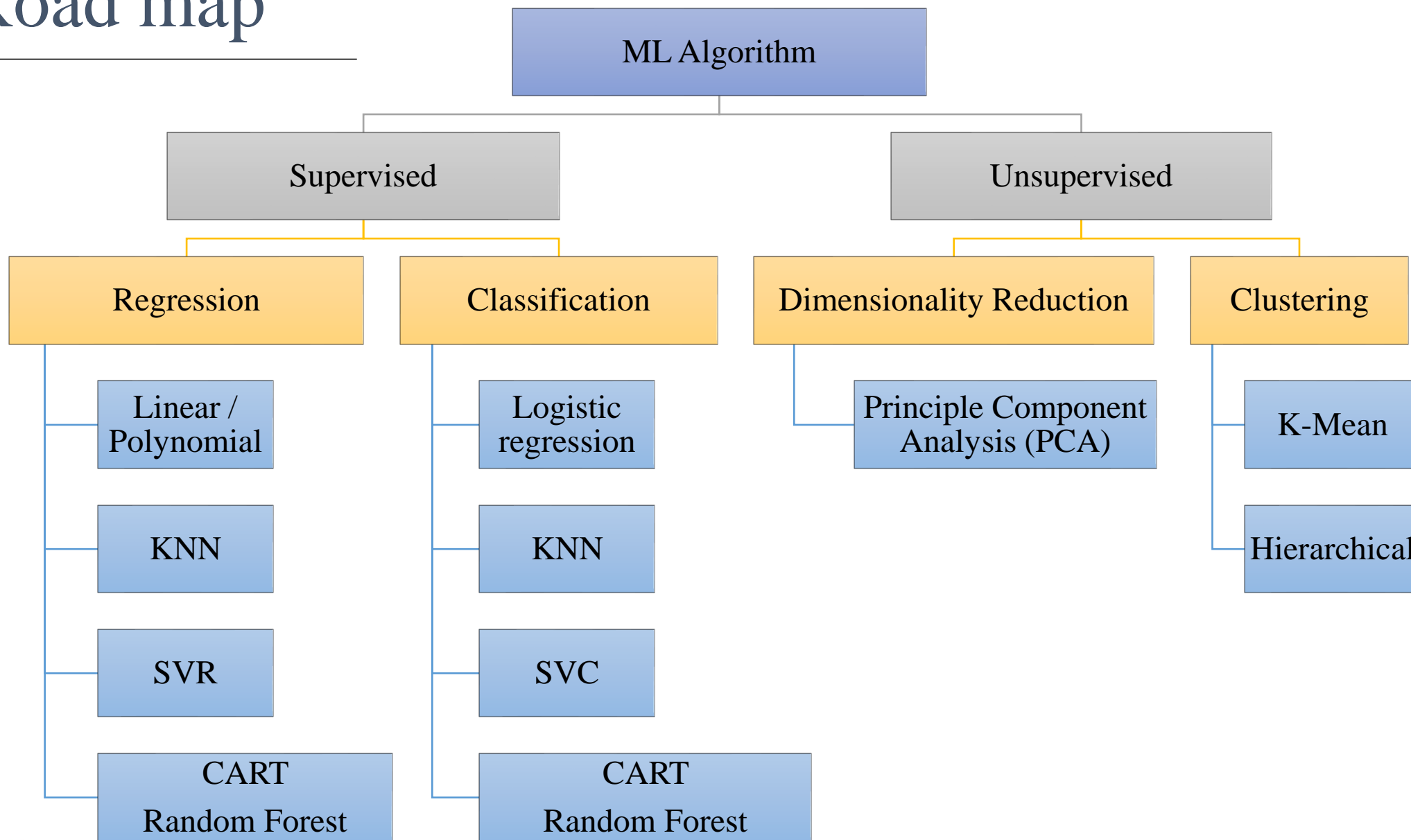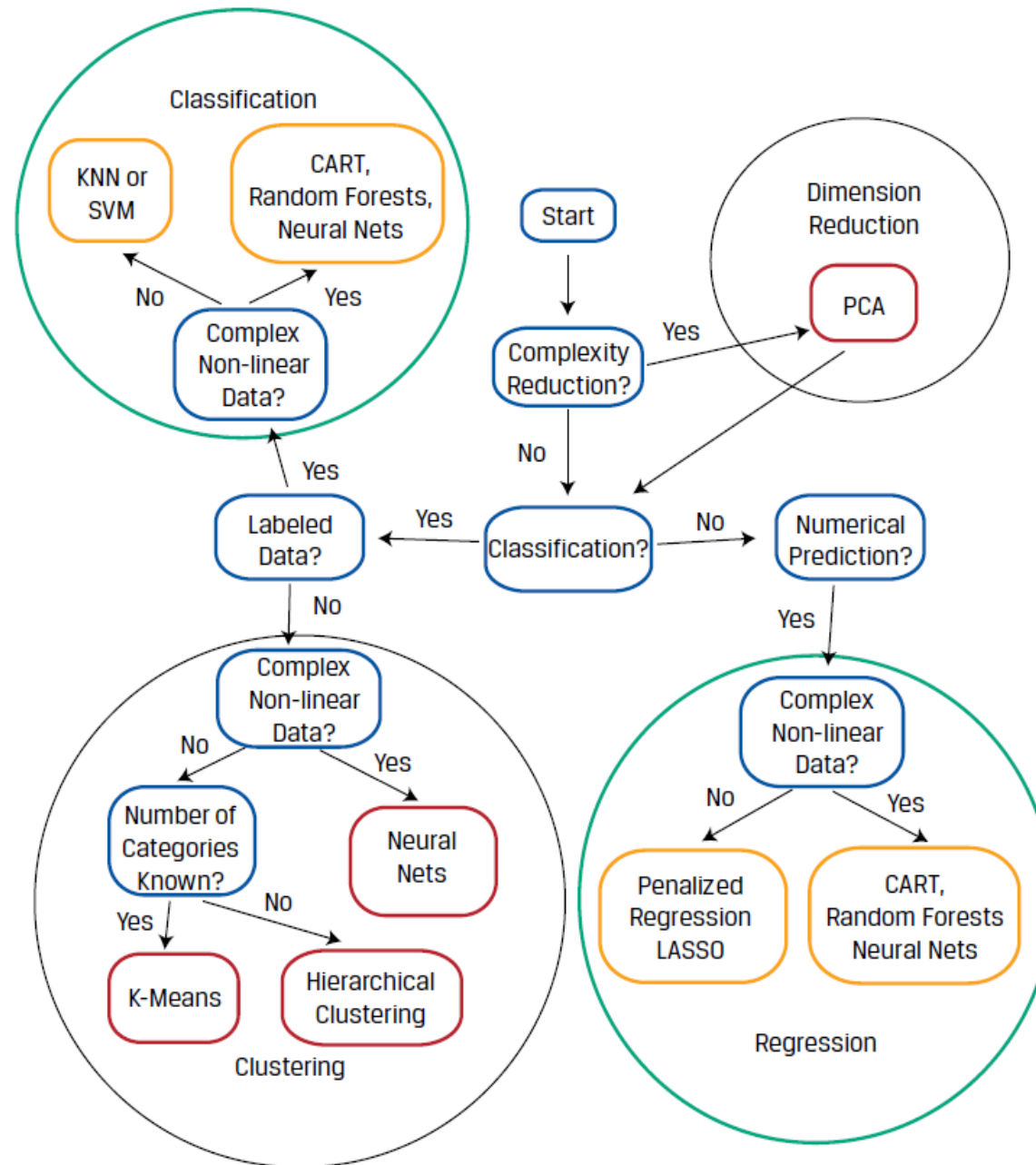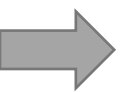
Road map

Exhibit 37   Stylized Decision Flowchart for Choosing ML Algorithms

Having said that…

- **Warning**: A ML algorithm will always find a pattern, even if there is none.

If you torture the data long enough, it will confess.

— Ronald Coase —