Machine Learning for Finance

UCEMA 2020

Syllabus

- Lecture 1
  - Buzzwords: AI, ML, DL, BD, DM, DS, …
  - What is Artificial Intelligence?
  - What is Machine Learning?
  - Types of Machine Learning algorithms
  - Machine Learning in Finance
  - Machine Learning in Finances vs in Tech
  - Overview of an end-to-end Machine Learning process in Python demo

Syllabus

- Lecture 2
  - Bias-variance tradeoff
  - Training and test sets
  - Model capacity and overfitting
  - Learning in linear regressions
  - Regularizations
  - Hyperparameters and cross validation
  - Gradient descent-like algorithms
  - Probabilistic models, MLE, MAP
  - Probabilistic classifications, Logistic regressions
  - Scikit-learn + Finance Forecast Demos

Syllabus

- Lecture 3
  - TensorFlow
  - Data Flow Graph and Lazy execution
  - Automatic Differentiation
  - Automatic Differentiation in Quant Finance
  - First TensorFlow Demo
  - Artificial Neural Networks
  - Multi Layer Perceptron and Deep Learning
  - Neural Networks TensorFlow Demo

Syllabus

- Lecture 4
  - Stock Analysis: Fundamental, Technical, Quantitative, Alternative Data
  - Stock Forecast using Deep Learning Demo
  - Merton’s Corporate Default Model
  - Bank Failures
  - Bank Failures Forecast using Deep Learning Demos
  - Pricing American Options with Deep Learning (?)

Lectures Structure

- 1st block
  - ~1h: Presentation of subjects + code explanation.
  - ~15’: Discussion + Q&A.
- 2nd block
  - ~1h: Presentation of subjects + code explanation.
  - ~15’: Discussion + Q&A.

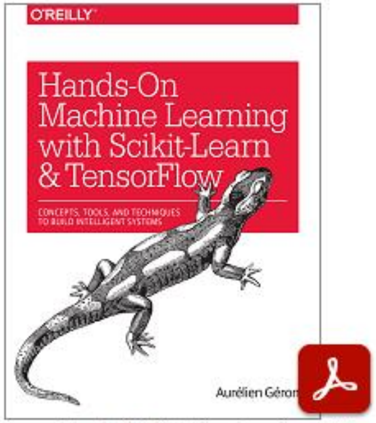
Objectives

- Regressions
  - Linear
  - Logit
  - Probit
  - Tobit
- Classifications
- Algorithmic Differentiation + Gradient Descent
- Deep Learning
- Applications in Finance using scikit-learn and TensorFlow

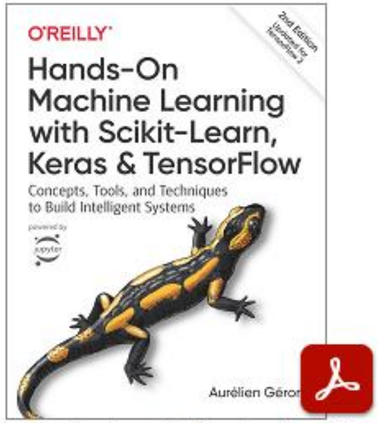
Prerequisites

- Python
- Probability
- Statistics
- Linear Algebra
- Real Calculus
- Matrix Calculus

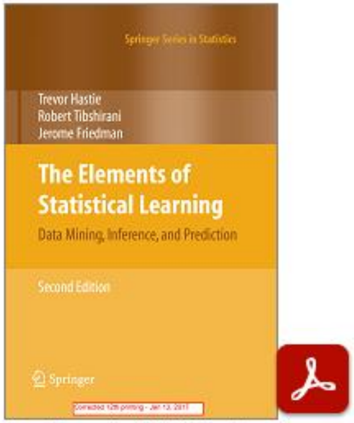
Bibliography



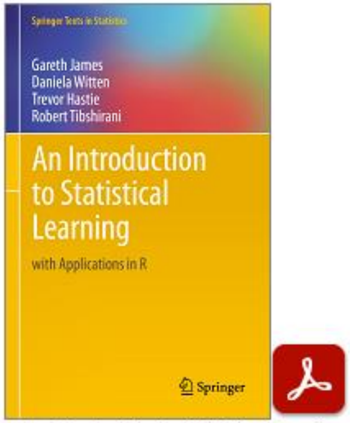
Géron - Hands-On Machine Learning with Scikit-Learn and TensorFlow - 2017.pdf



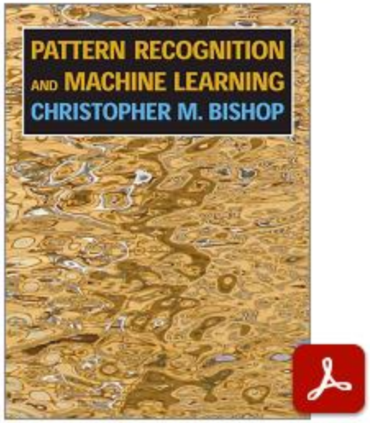
Geron - Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow - 2019.pdf



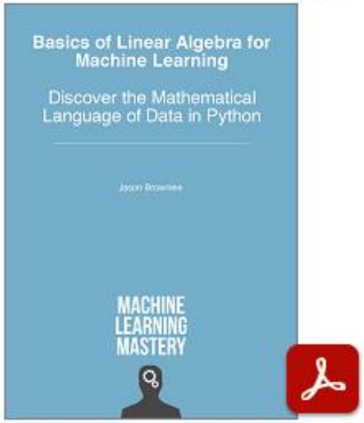
Hastie - Elements of Statistical Learning - 2017.pdf



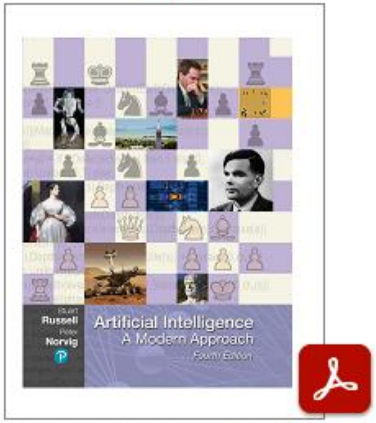
James - An Introduction to Statistical Learning with Applications in R.pdf



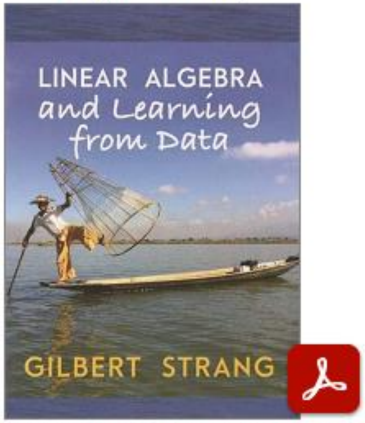
Bishop - Pattern Recognition and Machine Learning - 2007.pdf



Brownlee - Basics for Linear Algebra for Machine Learning - 2018.pdf



Russell - Artificial Intelligence -2020.pdf



Strang - Linear Algebra and Learning from Data - 2019.pdf

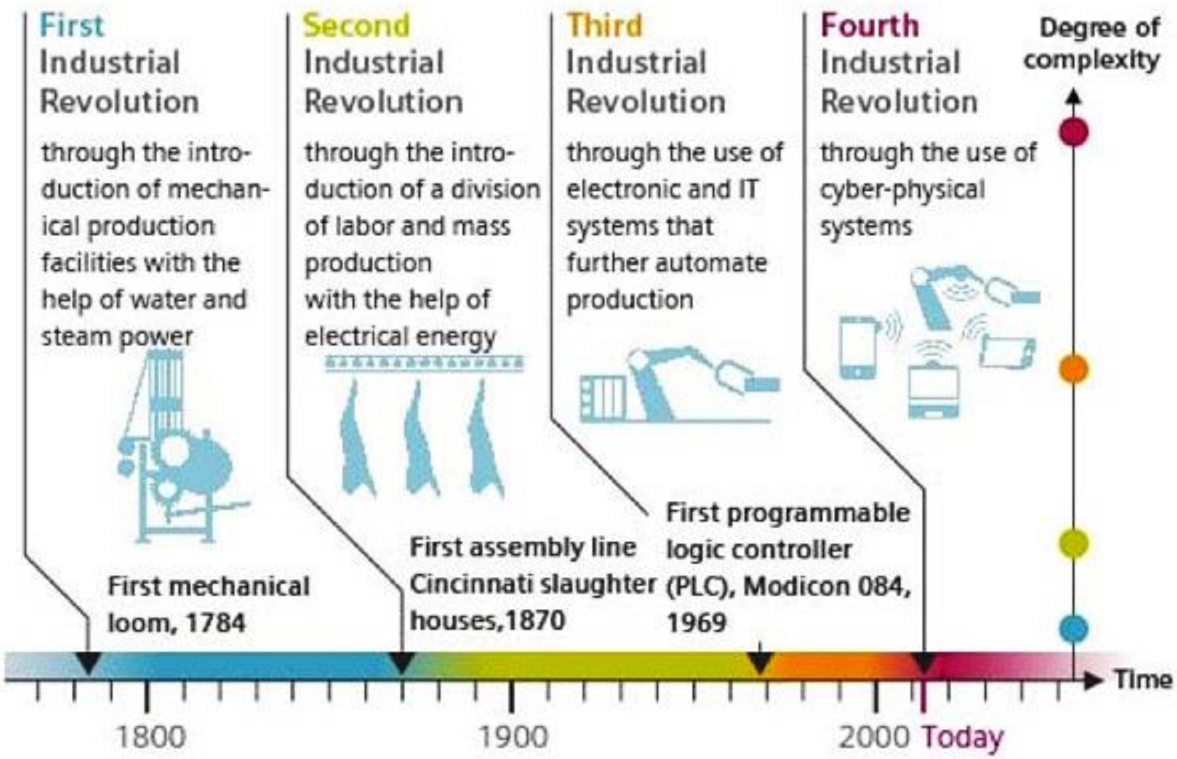
https://seeing-theory.brown.edu/index.html

Igor Halperin courses and book

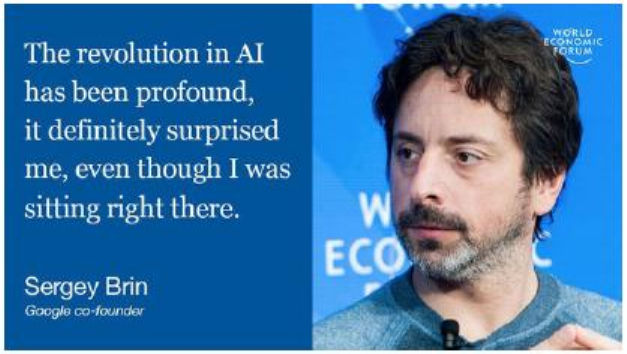
Machine Learning for Finance

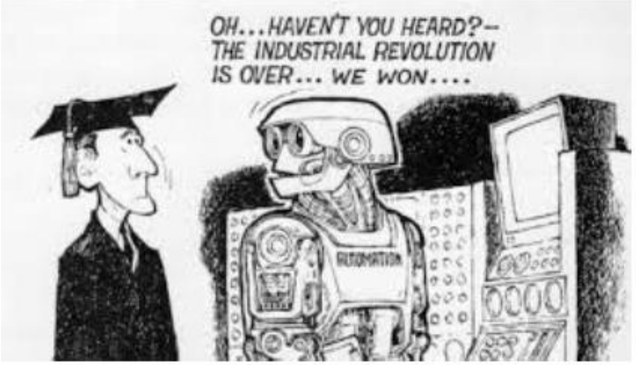
UCEMA 2020

Industrial Revolutions



Where will the AI revolution take us?



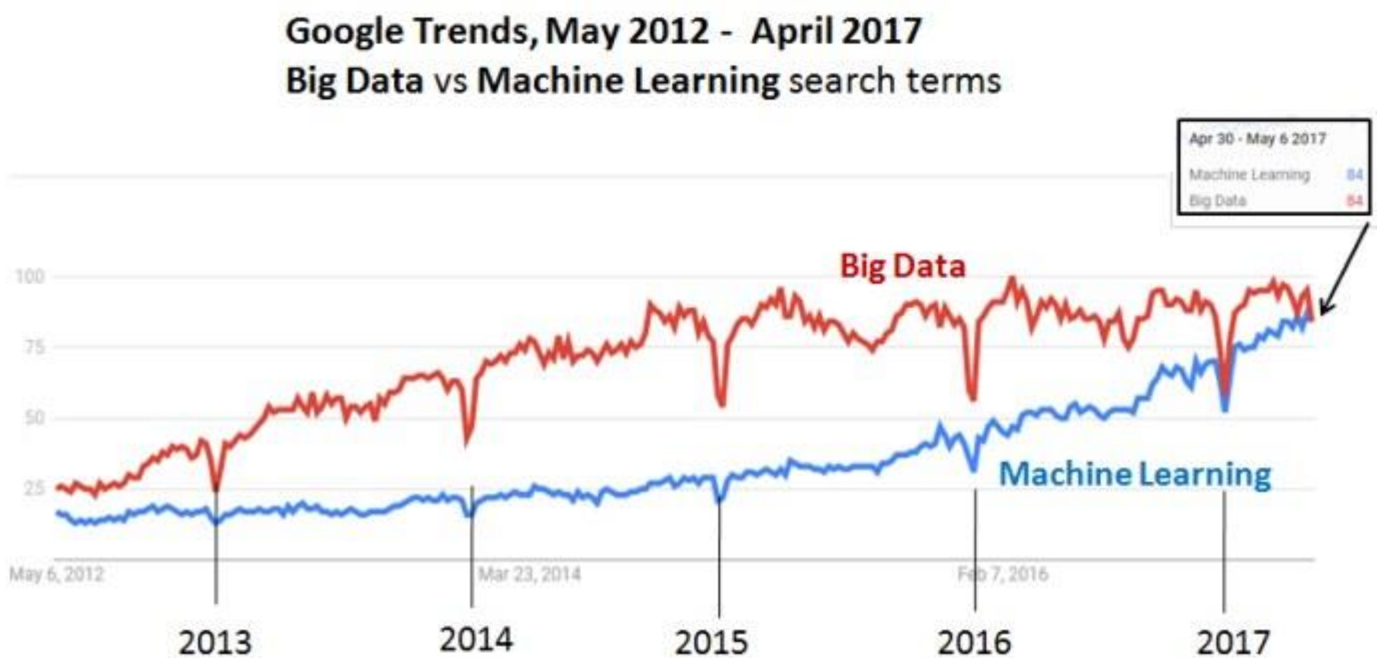


The future according to Stephen Hawking

The future according to Elon Musk

What is Machine Learning?

Buzzwords...


Google Trends, May 2012 - April 2017
Big Data vs Machine Learning search terms
Apr 30 - May 6 2017
Machine Learning 84
Big Data 84
Big Data
Machine Learning
100
75
50
25
May 6, 2012
Mar 23, 2014
Feb 7, 2016
2013
2014
2015
2016
2017

Buzzwords...

- **Artificial Intelligence (AI)** - machines achieving human-level performance at specific tasks (credit approval, face recognition, speech recognition etc.)
- **Machine Learning (ML)** - (a **heart of modern AI**) algorithms that teach a computer to perform a task from experience
- **Data Mining (DM)** - uses ML to find pattern in data in a quest for actionable data
- **Big Data (BD)** - DM on large sets of structured (numerical) and unstructured (text, speech) data
- **Data Science (DS)** - uses statistics and ML to monetize information in Big Data
- **ML** is a core element of all of the above fields
- **ML** is a way for a computer to learn about the world, much like physics and math for humans
- **Machine Intelligence** = ML/AI

What is AI?

What is AI?

Artificial Intelligence (AI) studies "intelligent agents" that perceive their environment and perform different actions to solve tasks that involve mimicking cognitive functions of humans (Russell, Norvig, "Artificial Intelligence: A Modern Approach", 2009)

$$\boxed{\text{Agent} = \text{architecture} + \text{algorithm}}$$

What is AI?

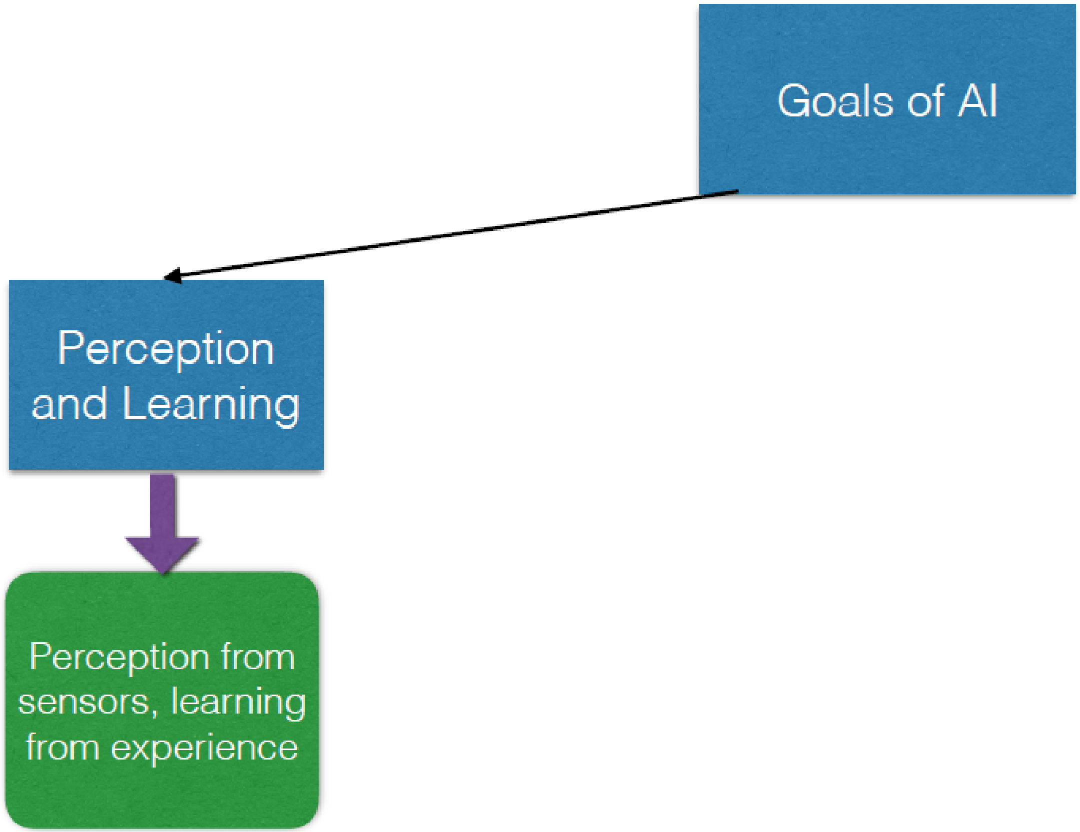
Artificial Intelligence (AI) studies "intelligent agents" that perceive their environment and perform different actions to solve tasks that involve mimicking cognitive functions of humans (Russell, Norvig, "Artificial Intelligence: A Modern Approach", 2009)

Goals of AI

Agent = architecture + algorithm

What is AI?

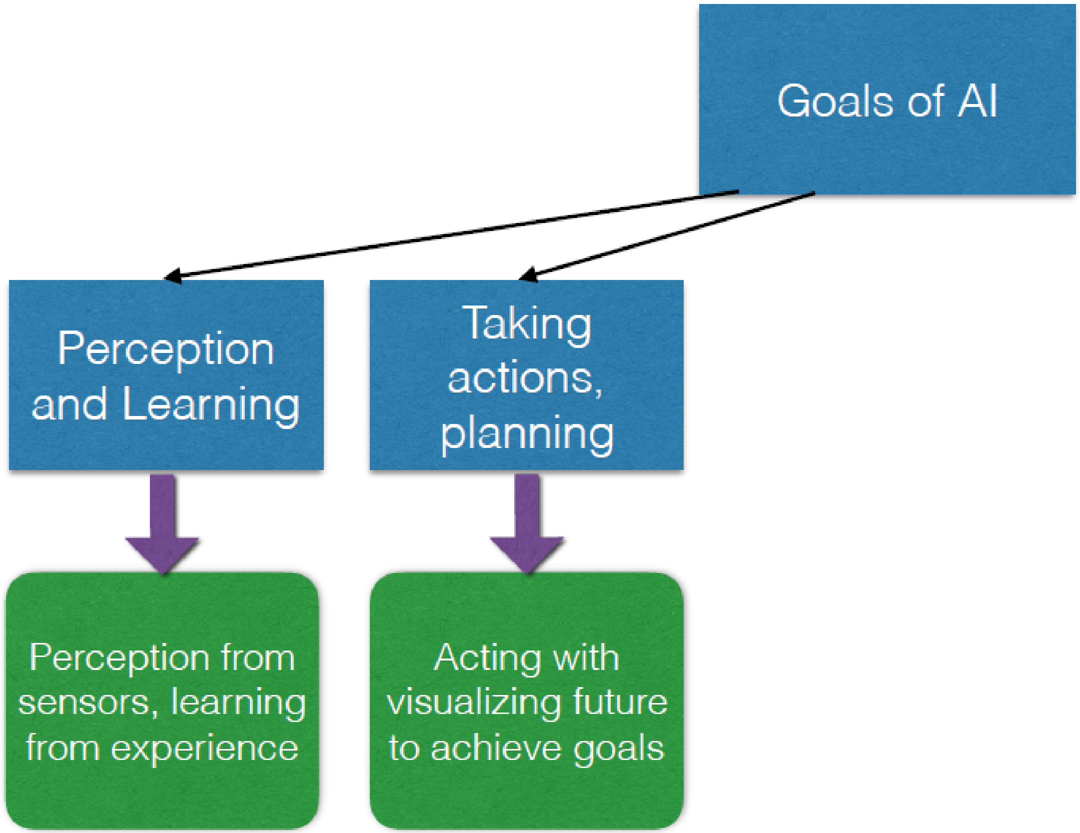
Artificial Intelligence (AI) studies "intelligent agents" that perceive their environment and perform different actions to solve tasks that involve mimicking cognitive functions of humans (Russell, Norvig, "Artificial Intelligence: A Modern Approach", 2009)



Agent = architecture + algorithm

What is AI?

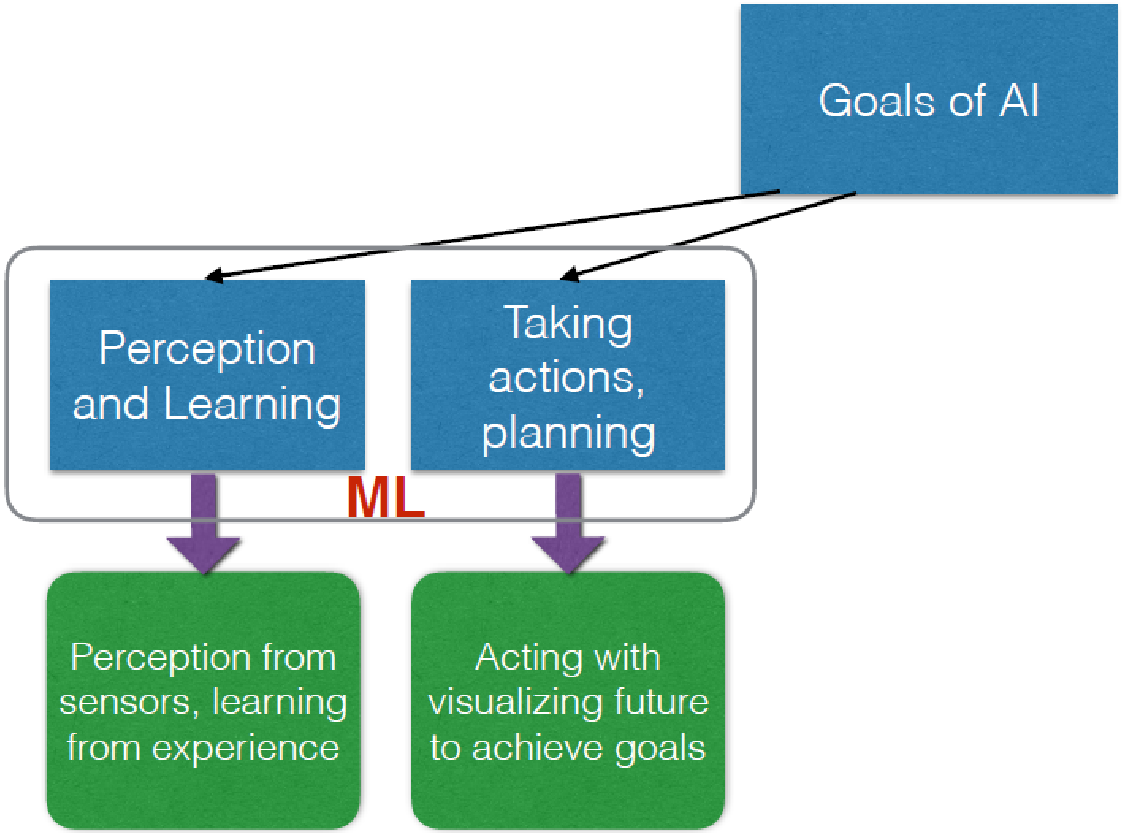
Artificial Intelligence (AI) studies "intelligent agents" that perceive their environment and perform different actions to solve tasks that involve mimicking cognitive functions of humans (Russell, Norvig, "Artificial Intelligence: A Modern Approach", 2009)



Agent = architecture + algorithm

What is AI?

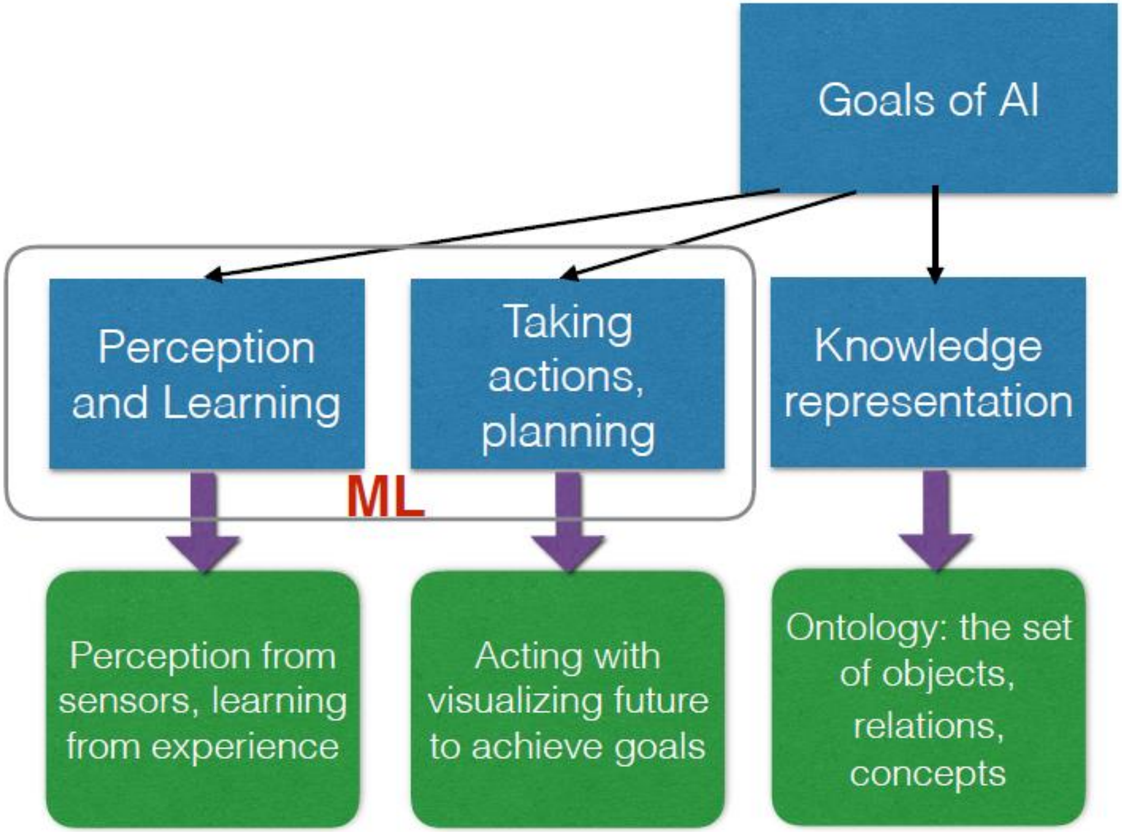
Artificial Intelligence (AI) studies "intelligent agents" that perceive their environment and perform different actions to solve tasks that involve mimicking cognitive functions of humans (Russell, Norvig, "Artificial Intelligence: A Modern Approach", 2009)



Agent = architecture + algorithm

What is AI?

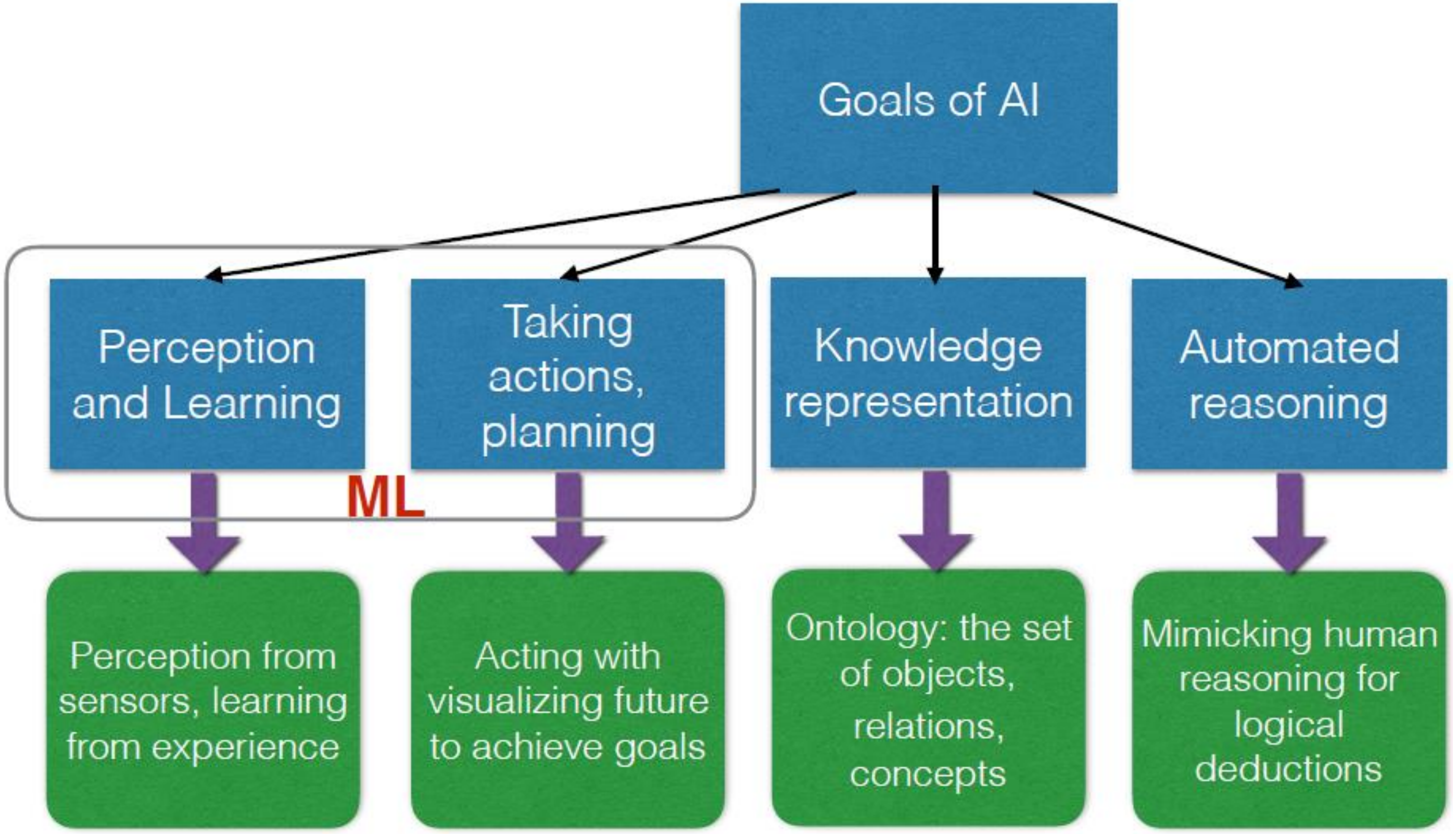
Artificial Intelligence (AI) studies “intelligent agents” that perceive their environment and perform different actions to solve tasks that involve mimicking cognitive functions of humans (Russell, Norvig, “Artificial Intelligence: A Modern Approach”, 2009)



Agent = architecture + algorithm

What is AI?

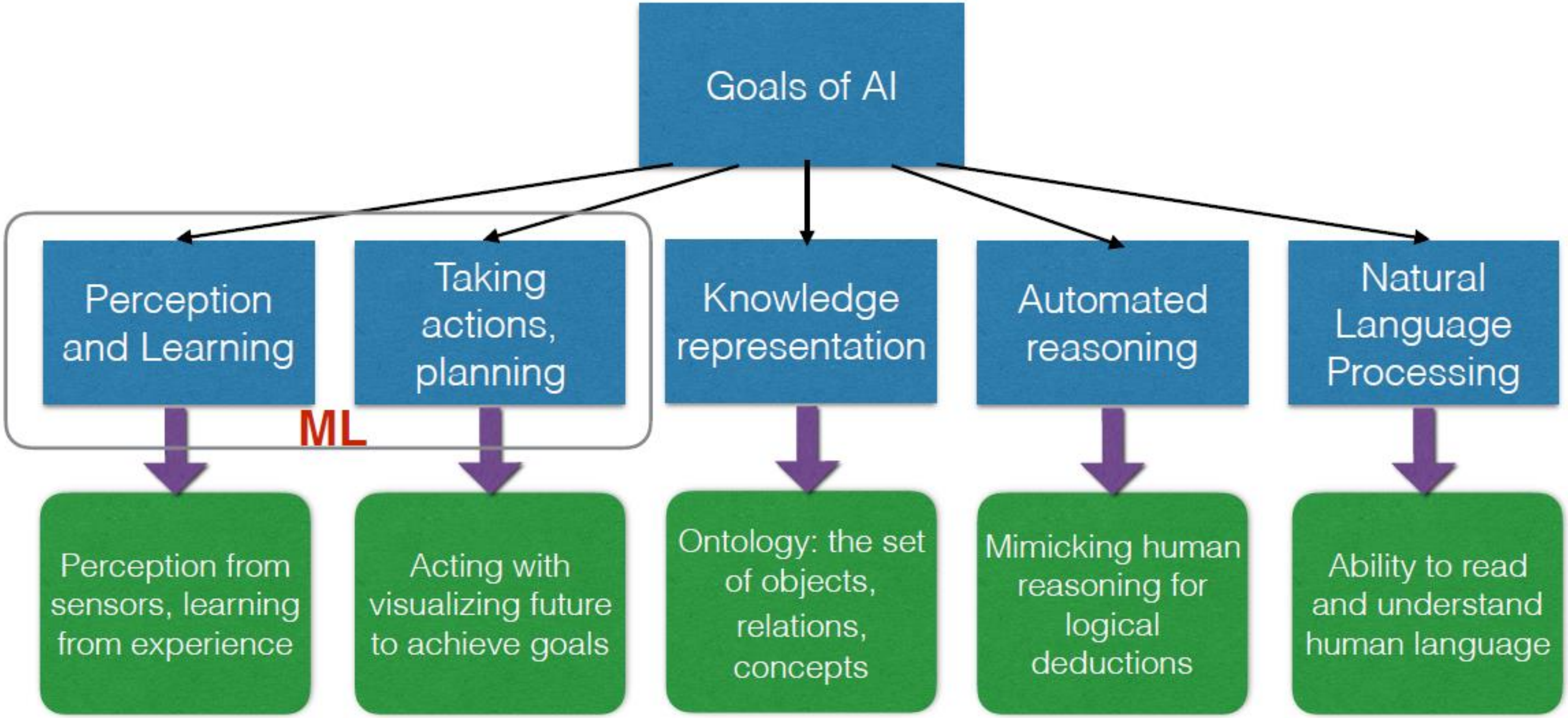
Artificial Intelligence (AI) studies "intelligent agents" that perceive their environment and perform different actions to solve tasks that involve mimicking cognitive functions of humans (Russell, Norvig, "Artificial Intelligence: A Modern Approach", 2009)



Agent = architecture + algorithm

What is AI?

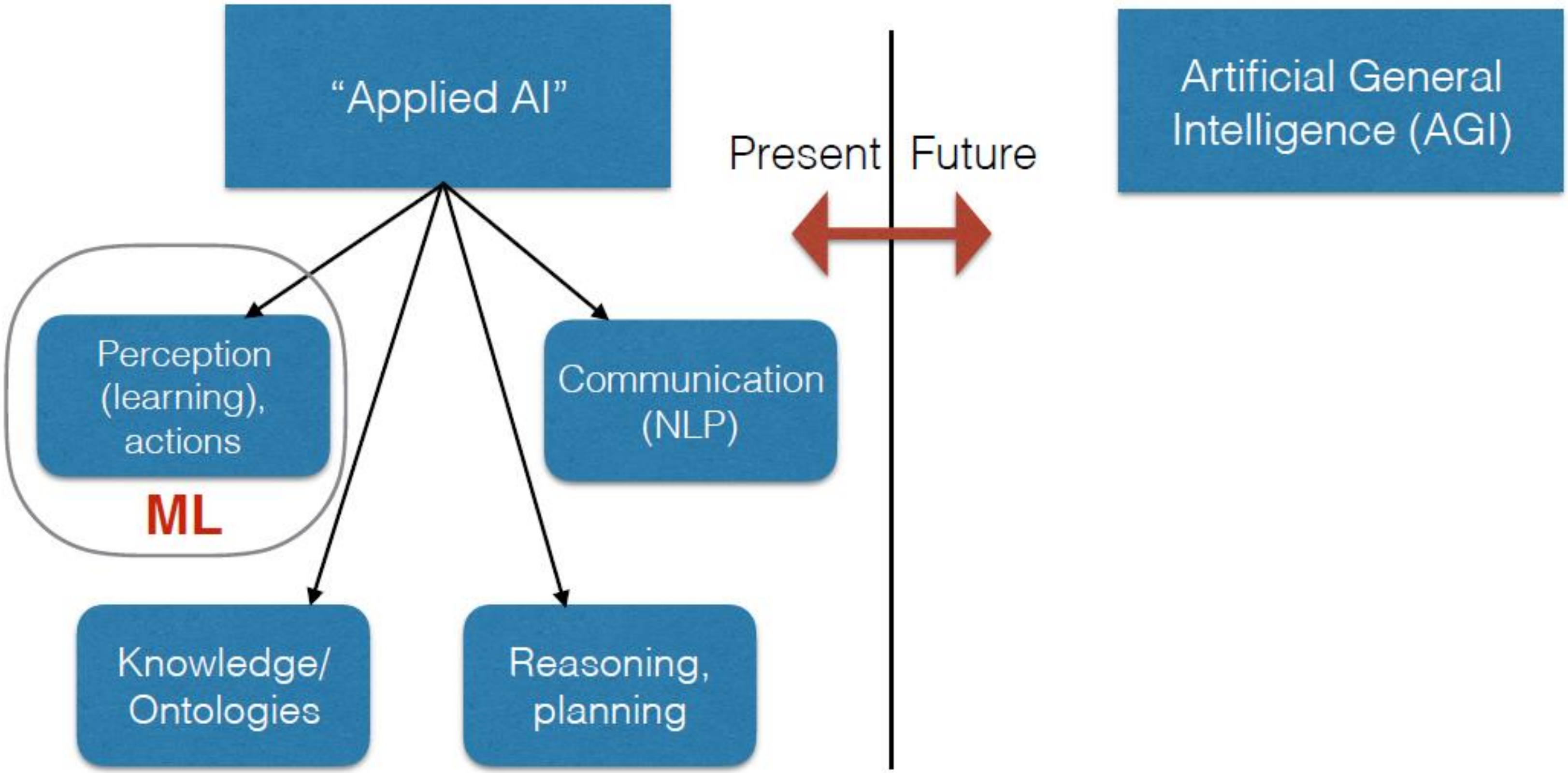
Artificial Intelligence (AI) studies "intelligent agents" that perceive their environment and perform different actions to solve tasks that involve mimicking cognitive functions of humans (Russell, Norvig, "Artificial Intelligence: A Modern Approach", 2009)



Agent = architecture + algorithm

Long Term Goals: Applied AI vs AGI

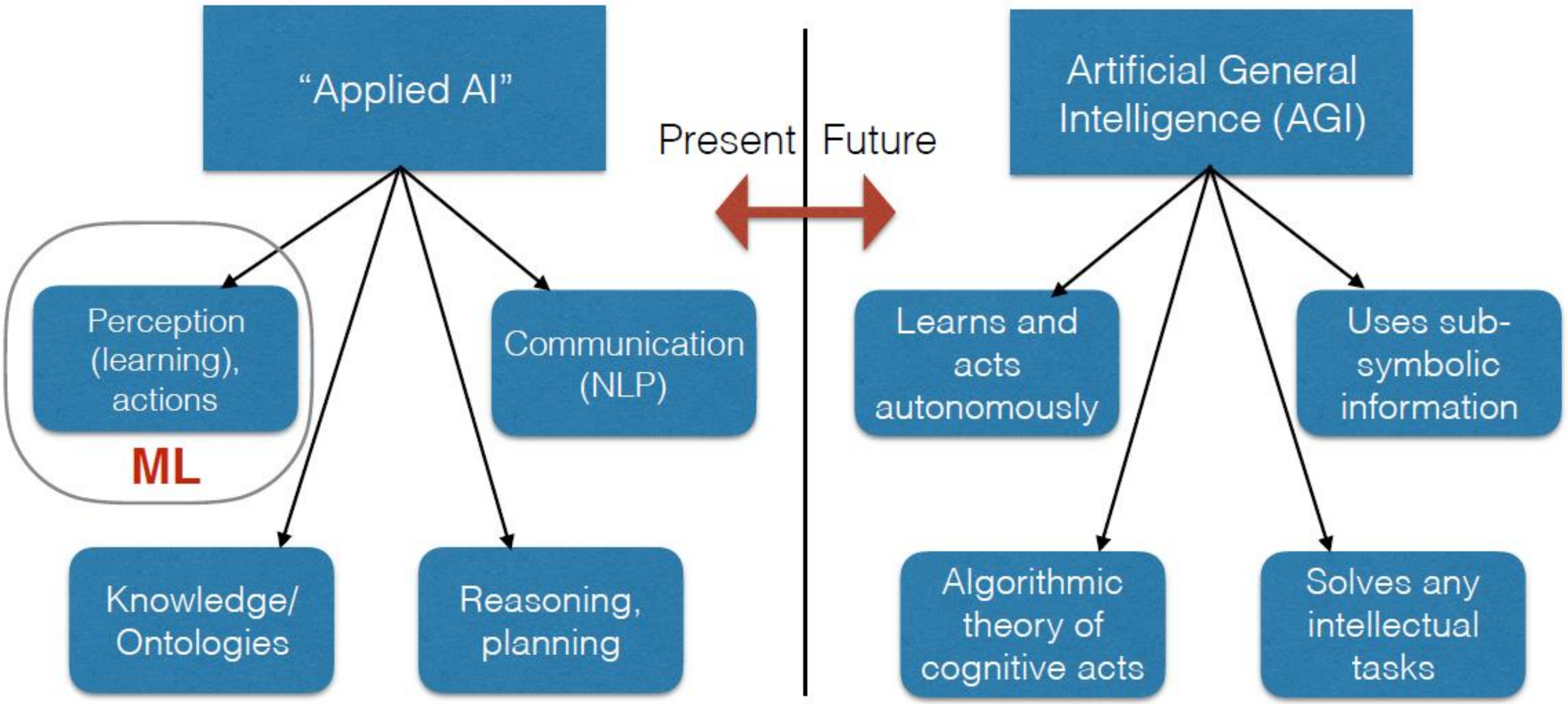
Long Term Goals: Applied AI vs AGI



"Applied AI" ("weak AI"):

- Performs one pre-specified task
- Operates with human-provided algorithms
- Works with numerical and symbolic information
- Remarkable technological and commercial success, strong impact on many industries

Long Term Goals: Applied AI vs AGI



"Applied AI" ("weak AI"):

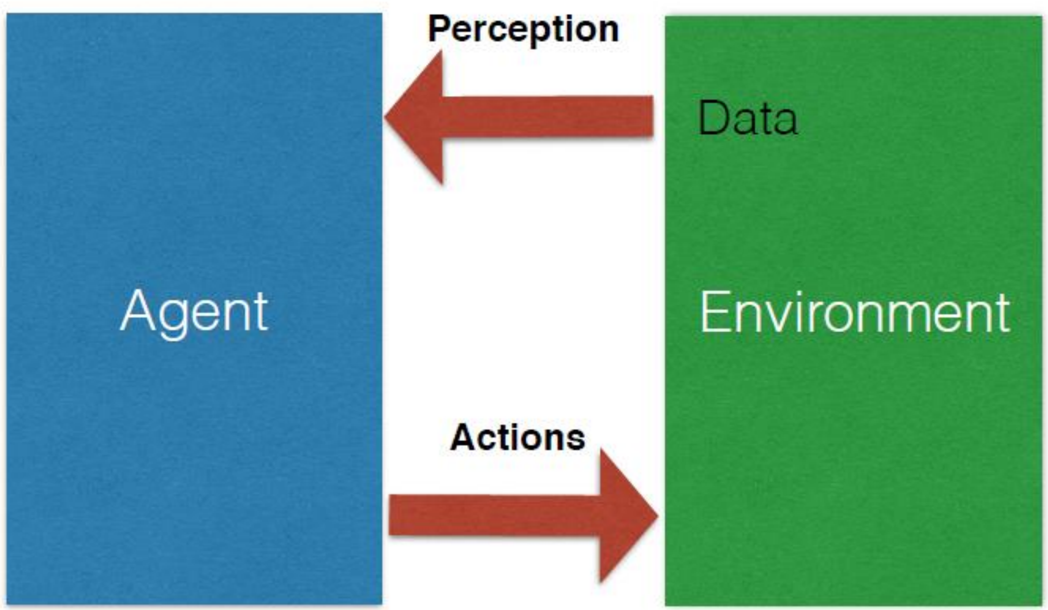
- Performs one pre-specified task
- Operates with human-provided algorithms
- Works with numerical and symbolic information
- Remarkable technological and commercial success, strong impact on many industries

AGI ("strong AI"):

- Autonomous learning and acting involves building its own models of the world
- Formalizes sub-symbolic information ("He took his bag and left the room")
- Algorithmic theory of novelty, surprise, creativity, curiosity etc. (Schmidhuber 2010)
- Solving any intellectual tasks: around 2045, per Ray Kurzweil

Agents and Environments

Artificial Intelligence (AI) studies "**intelligent agents**" that **perceive** their **environment** and perform different **actions** to solve tasks that involve mimicking cognitive functions of humans (Russell, Norvig, "Artificial Intelligence: A Modern Approach", 2009)



Perception: the physical world (through sensors), or digital data (read from a disk)

Actions: can be fixed, or can vary. May or may not change the environment

Rational AI Agents

Rational AI Agents

Rational agent: "For each possible percept sequence, a rational agent should select an action that is expected to maximize its performance measure, given the evidence provided by the percept sequence and whatever built-in knowledge the agent has." (RN 2009)

Rational AI Agents

Rational agent: "For each possible percept sequence, a rational agent should select an <u>action that is expected to maximize its performance measure</u>, given the evidence provided by the percept sequence and whatever built-in knowledge the agent has." (RN 2009)

Example: a vacuum-cleaner agent (RN 2009)

http://www.sharp-world.com/products/img/robotic_appliance/tech/voice-operation/img_voiceoperation.gif

ML as a Foundation of AI

ML as a Foundation of AI

- A rational agent should not only perceive its environment, but also **learn** as much as possible from what it perceives

ML as a Foundation of AI

- A rational agent should not only perceive its environment, but also **learn** as much as possible from what it perceives
- Learning = ability to **generalize** using raw data

ML as a Foundation of AI

- A rational agent should not only perceive its environment, but also **learn** as much as possible from what it perceives
- Learning != storing data
- Learning = ability to **generalize** using raw data

ML as a Foundation of AI

- A rational agent should not only perceive its environment, but also **learn** as much as possible from what it perceives
- Learning != storing data
- Learning = ability to **generalize** using raw data
- Enter **Machine Learning (ML):** a sub-field of AI that studies perception, learning and action tasks as **algorithms that learn from data**

ML as a Foundation of AI

- A rational agent should not only perceive its environment, but also **learn** as much as possible from what it perceives
- Learning != storing data
- Learning = ability to **generalize** using raw data
- Enter **Machine Learning (ML):** a sub-field of AI that studies perception, learning and action tasks as **algorithms that learn from data**
- ML modernized such fields as statistics and operational research by putting more <u>emphasis on computer algorithms to statistically estimate complex functions</u> that typically cannot be expressed in closed form

–

ML as a Foundation of AI

- A rational agent should not only perceive its environment, but also **learn** as much as possible from what it perceives
- Learning != storing data
- Learning = ability to **generalize** using raw data
- Enter **Machine Learning (ML):** a sub-field of AI that studies perception, learning and action tasks as **algorithms that learn from data**
- ML modernized such fields as statistics and operational research by putting more emphasis on computer algorithms to statistically estimate complex functions that typically cannot be expressed in closed form
- ML is an inter-disciplinary field that benefits from statistical modeling, information theory, physics, computer science, neuroscience, biology

ML as a Foundation of AI

- A rational agent should not only perceive its environment, but also **learn** as much as possible from what it perceives
- Learning != storing data
- Learning = ability to **generalize** using raw data
- Enter **Machine Learning (ML):** a sub-field of AI that studies perception, learning and action tasks as **algorithms that learn from data**
- ML modernized such fields as statistics and operational research by putting more <u>emphasis on computer algorithms to statistically estimate complex functions</u> that typically cannot be expressed in closed form
- ML is an inter-disciplinary field that benefits from statistical modeling, information theory, physics, computer science, neuroscience, biology
- ML is used in other AI tasks (NLP, knowledge base, AGI)

Machine Learning + AI = Machine Intelligence

ML vs Statistical Modeling

For more, read Leo Berriman's “Statistical Modeling: The Two Cultures ”

ML vs Statistical Modeling

Statistical Modeling	Machine Learning
Parametric models that try to "**explain**" the world. The focus is on modeling causality	**Non-parametric** models that try to "**mimic**" the world rather than "explain" it. Often uses correlations as proxies to causality

For more, read Leo Berriman's "Statistical Modeling: The Two Cultures "

ML vs Statistical Modeling

Statistical Modeling	Machine Learning
Parametric models that try to "**explain**" the world. The focus is on modeling causality	**Non-parametric** models that try to "**mimic**" the world rather than "explain" it. Often uses correlations as proxies to causality
Deduce relations for observed quantities by parameter estimation for a pre-specified model of the world	**Induce** relations between observable quantities, main goal is predictive power

For more, read Leo Berriman's "Statistical Modeling: The Two Cultures "

ML vs Statistical Modeling

Statistical Modeling	Machine Learning
Parametric models that try to "**explain**" the world. The focus is on modeling causality	**Non-parametric** models that try to "**mimic**" the world rather than "explain" it. Often uses correlations as proxies to causality
Deduce relations for observed quantities by parameter estimation for a pre-specified model of the world	**Induce** relations between observable quantities, main goal is predictive power
Small data (1-100 attributes,100-1000 examples)	Large data (10-100K attributes, 1K-100M examples)

For more, read Leo Berriman's "Statistical Modeling: The Two Cultures "

ML vs Statistical Modeling

Statistical Modeling	Machine Learning
Parametric models that try to "**explain**" the world. The focus is on modeling causality	**Non-parametric** models that try to "**mimic**" the world rather than "explain" it. Often uses correlations as proxies to causality
Deduce relations for observed quantities by parameter estimation for a pre-specified model of the world	**Induce** relations between observable quantities, main goal is predictive power
Small data (1-100 attributes,100-1000 examples)	Large data (10-100K attributes, 1K-100M examples)
Scalability is typically not the major concern	Scalability is often critical in applications

For more, read Leo Berriman's "Statistical Modeling: The Two Cultures "

ML vs Statistical Modeling

Statistical Modeling	Machine Learning
Parametric models that try to "**explain**" the world. The focus is on modeling causality	**Non-parametric** models that try to "**mimic**" the world rather than "explain" it. Often uses correlations as proxies to causality
Deduce relations for observed quantities by parameter estimation for a pre-specified model of the world	**Induce** relations between observable quantities, main goal is predictive power
Small data (1-100 attributes,100-1000 examples)	Large data (10-100K attributes, 1K-100M examples)
Scalability is typically not the major concern	Scalability is often critical in applications
Based on a probabilistic approach	Some ML methods are not probabilistic (SVM, neural networks, clustering, etc.)

For more, read Leo Berriman's "Statistical Modeling: The Two Cultures "

ML: Core Idea

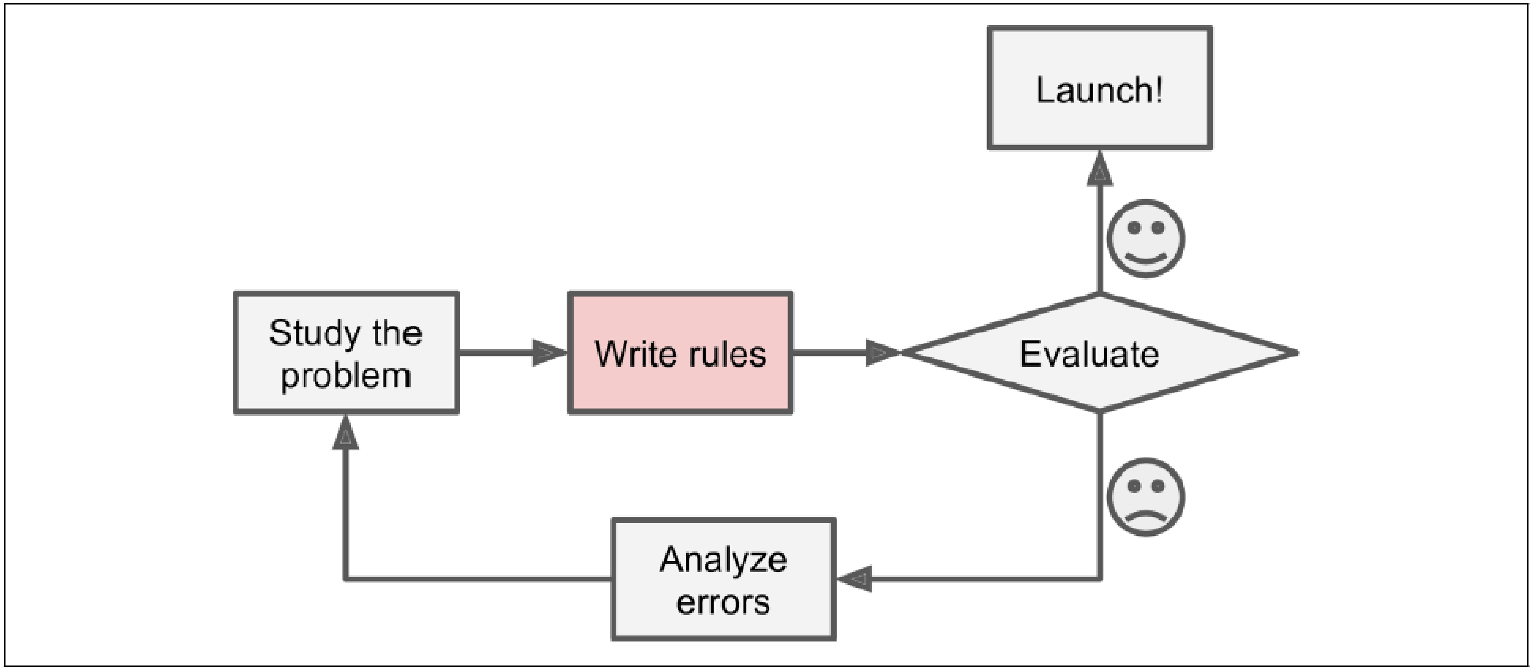
ML: Core Idea

Machine Learning is the science (and art) of programming computers so they can *learn from data*.

Here is a slightly more general definition:

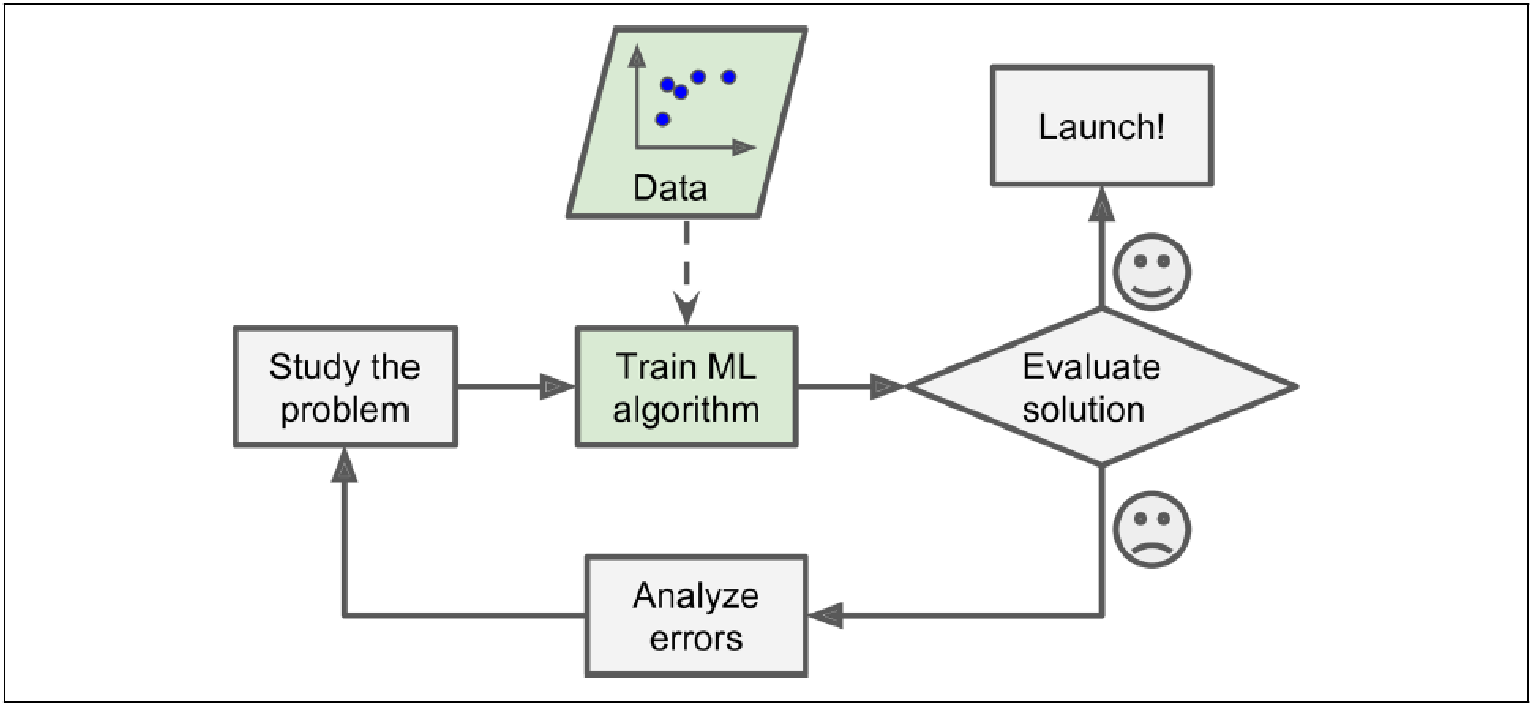
> [Machine Learning is the] field of study that gives computers the ability to learn without being explicitly programmed.
>
> —Arthur Samuel, *1959*

ML: Core Idea



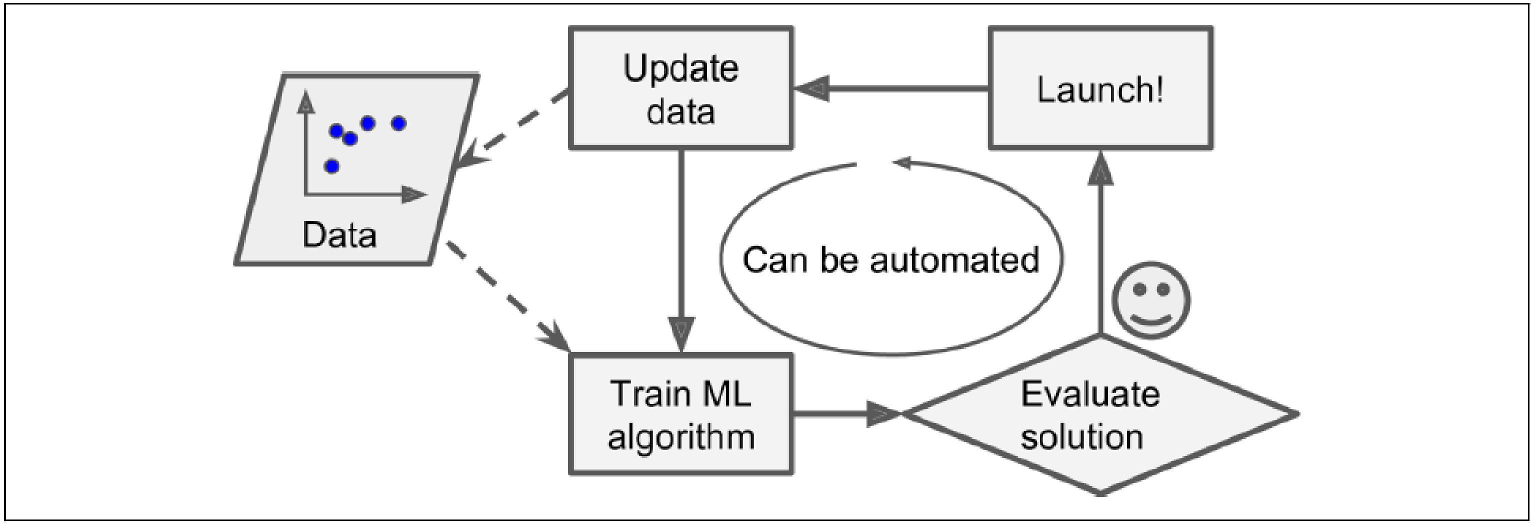
The traditional approach

ML: Core Idea



Machine Learning approach

ML: Core Idea

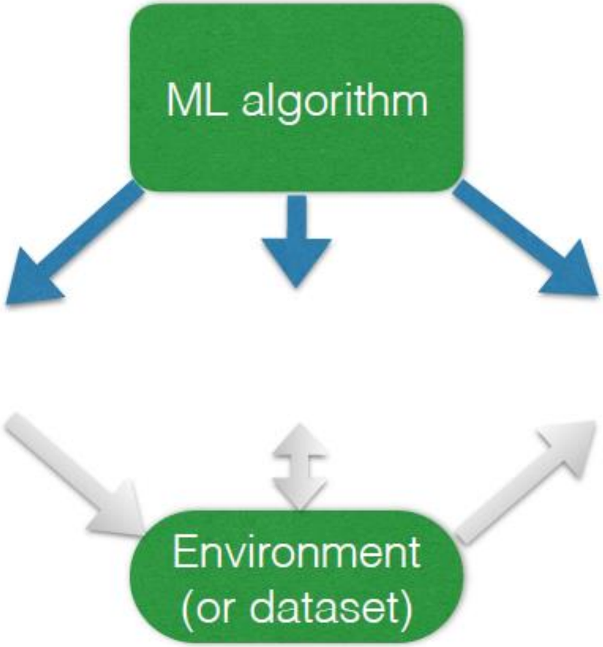


Automatically adapting to change

Pause (?)

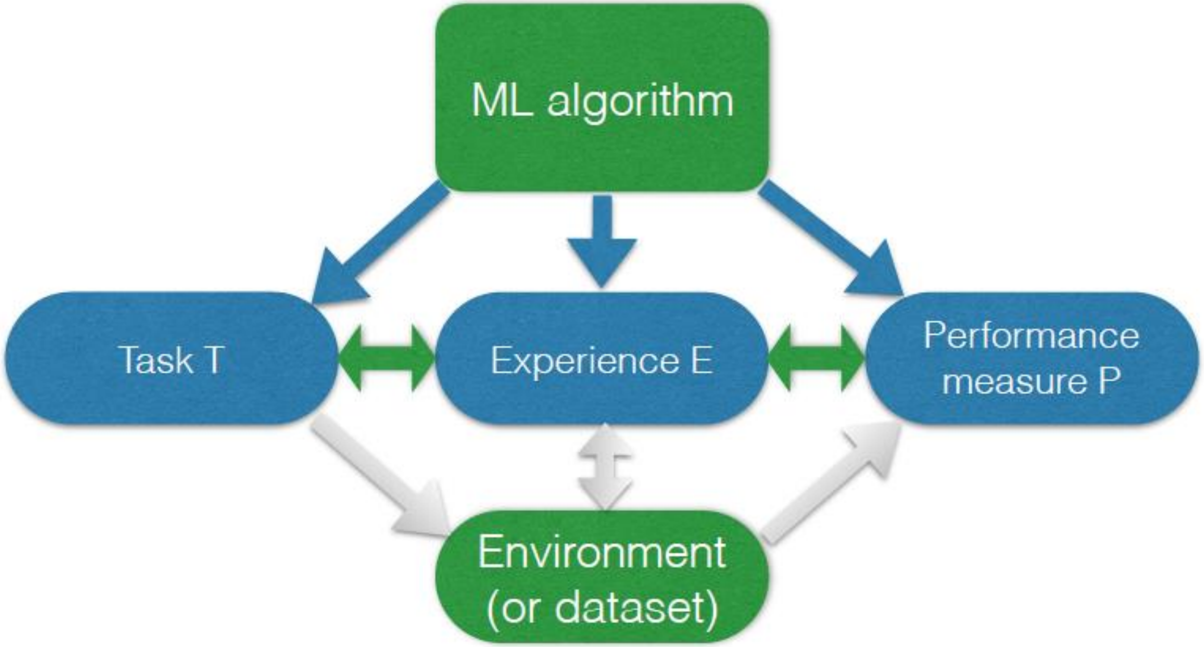
ML: Core Idea

ML: Core Idea



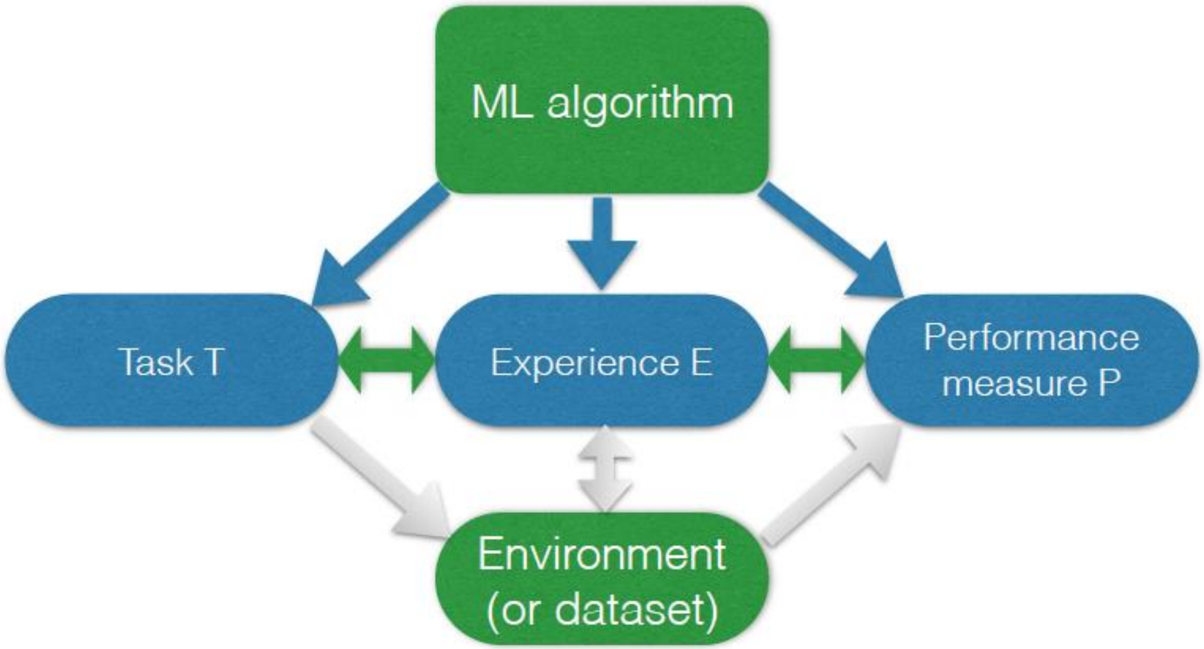
ML: Core Idea

"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**." (Mitchell, 1997)



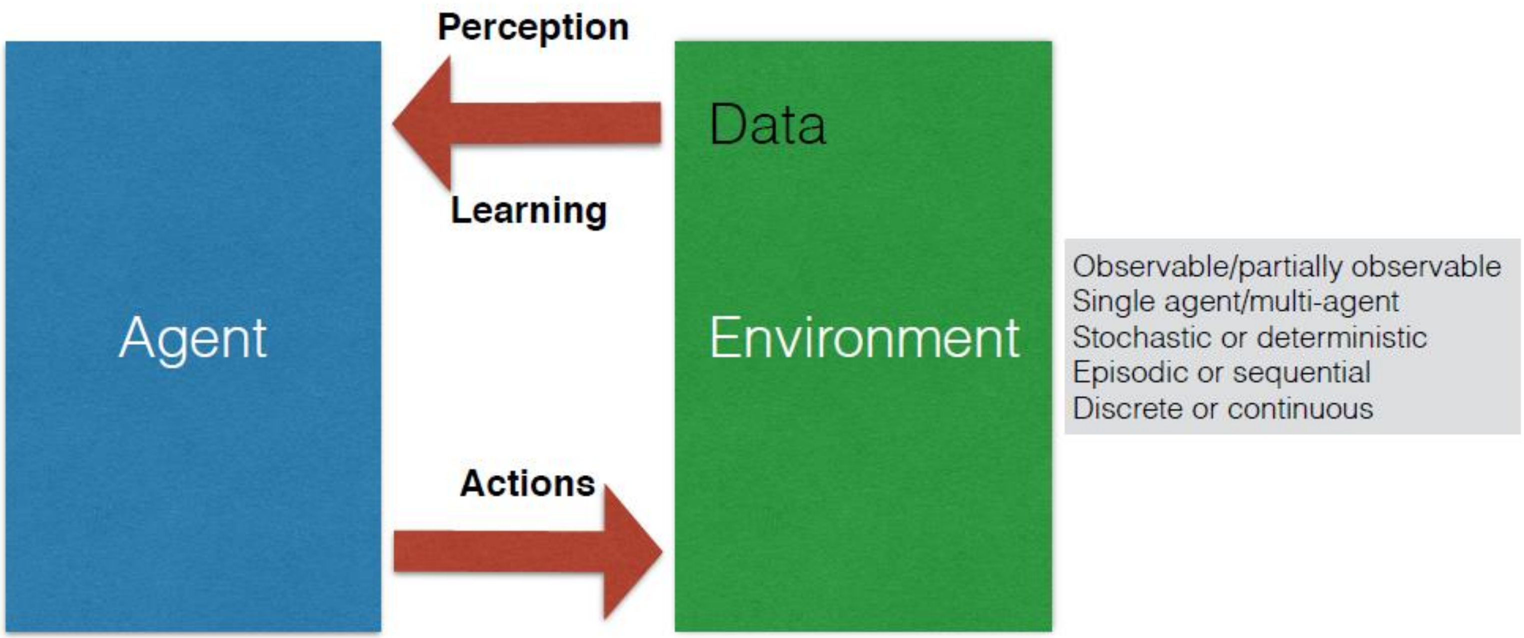
ML: Core Idea

"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**." (Mitchell, 1997)

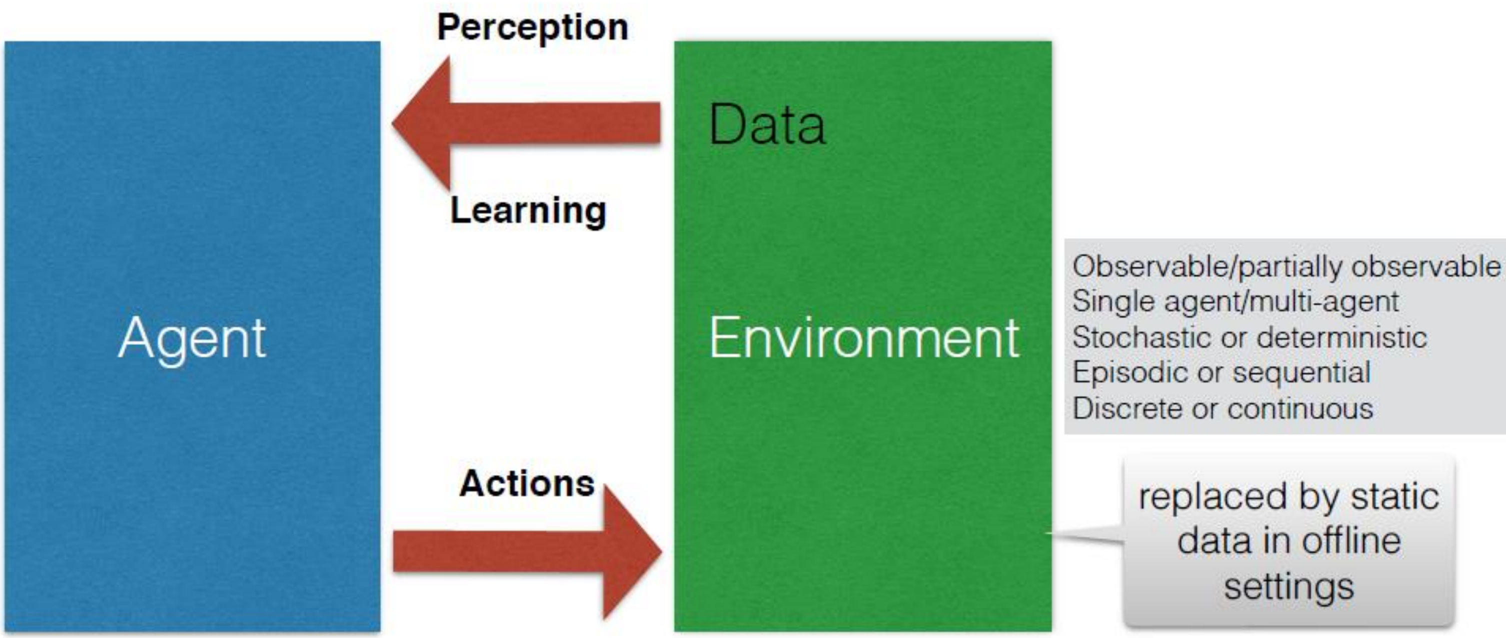


Next, we need to specify what we mean by **Tasks T**, **Performance measure P** and **Experience E**

Types of ML Tasks

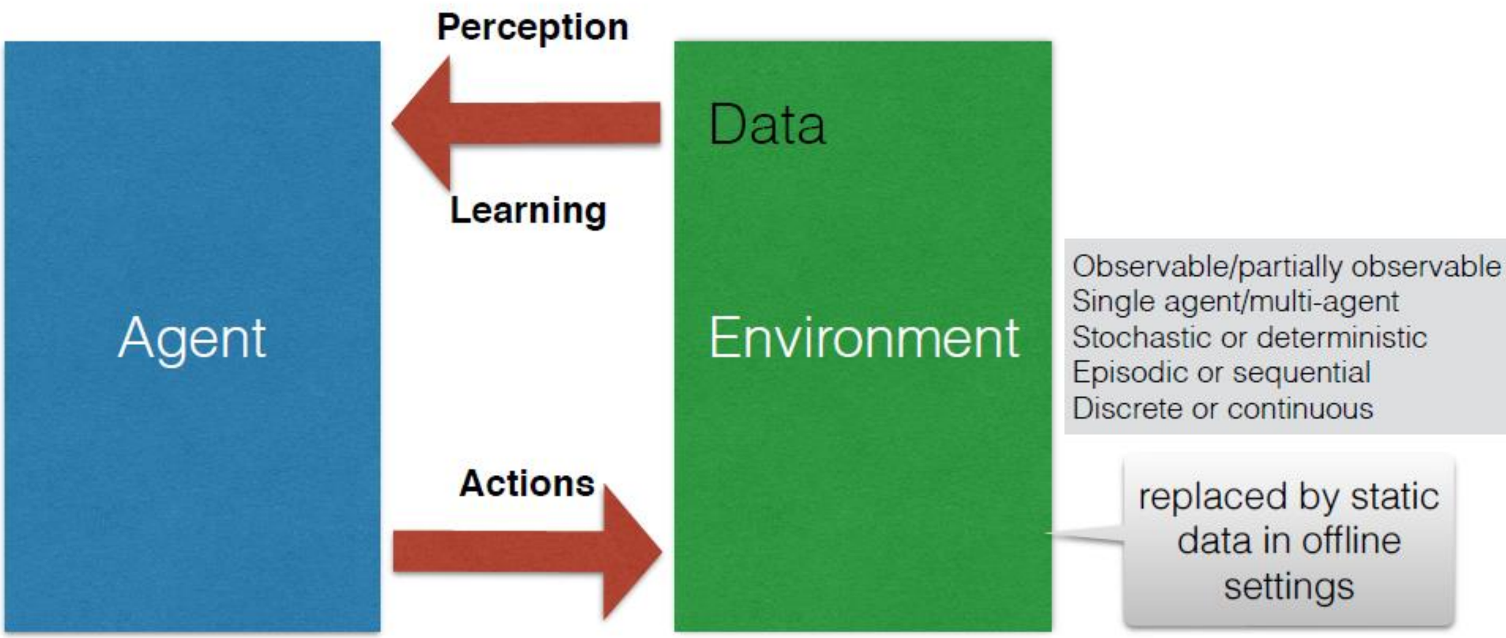


Types of ML Tasks



The agent may not have access to streaming data from the environment (on-line learning) and learn instead in a batch mode (off-line) from data obtained from this environment.

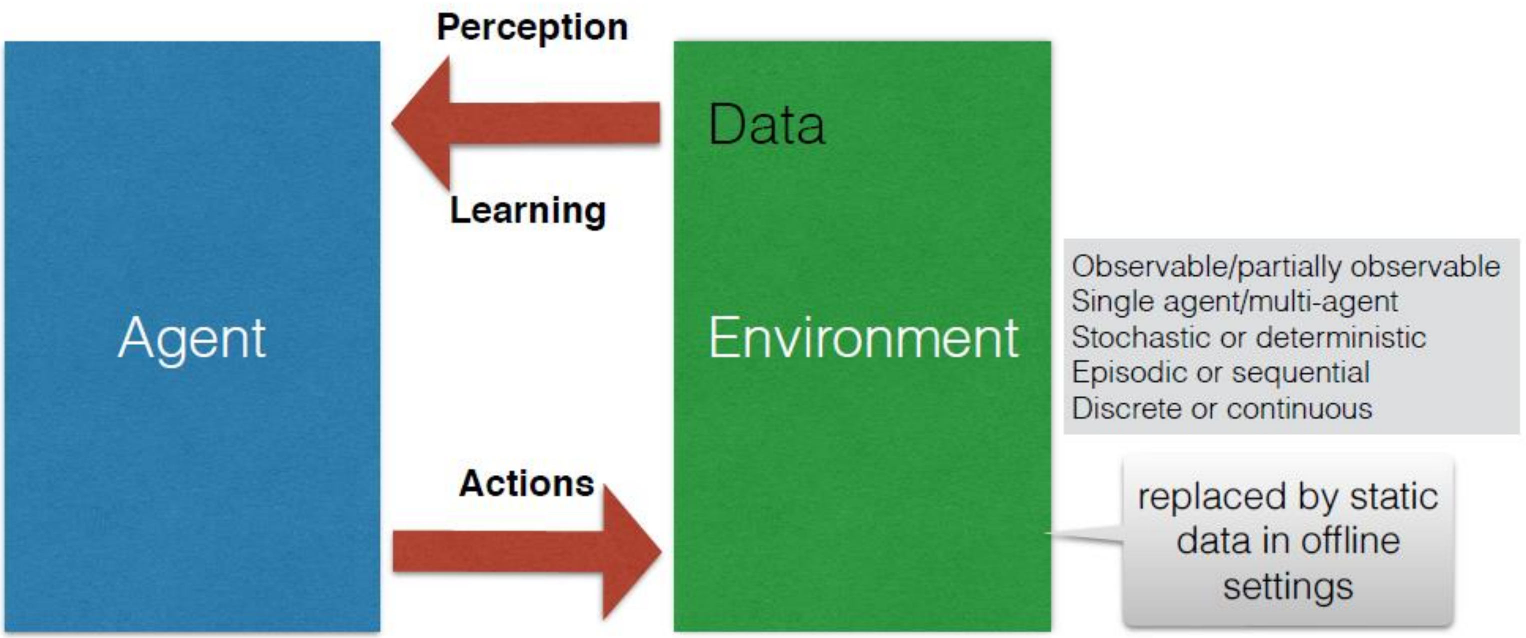
Types of ML Tasks



The agent may not have access to streaming data from the environment (on-line learning) and learn instead in a batch mode (off-line) from data obtained from this environment.

"**Perception tasks**": perception and learning from data. There is a fixed action, e.g. predict a loan default, classify an image, or translate a text. Regression and classifications are perception tasks. The output is a learned function of data $f(\mathbf{X})$

Types of ML Tasks



The agent may not have access to streaming data from the environment (on-line learning) and learn instead in a batch mode (off-line) from data obtained from this environment.

"**Perception tasks**"**:** perception and learning from data. There is a fixed action, e.g. predict a loan default, classify an image, or translate a text. Regression and classifications are perception tasks. The output is a learned function of data $f(\mathbf{X})$

"**Action tasks**": the same as perception tasks, but there are multiple possible actions. For sequential (multi-step) problems, action tasks involve planning and forecasting the future.

Performance Measure P

Typically, performance measure **P** is specific to the task **T**

Performance Measure P

Typically, performance measure **P** is specific to the task **T**

One possible choice for **classification tasks**

$$Error\ rate = \frac{N_{incorrectly\,classified}}{N_{total}} \Leftrightarrow Accuracy = \frac{N_{correctly\ classified}}{N_{total}} = 1 - Error\ rate$$

Error rate = expected 0-1 loss

Performance Measure P

Typically, performance measure **P** is specific to the task **T**

One possible choice for **classification tasks**

$$Error\ rate = \frac{N_{incorrectly\,classified}}{N_{total}} \Leftrightarrow Accuracy = \frac{N_{correctly\ classified}}{N_{total}} = 1 - Error\ rate$$

Error rate = expected 0-1 loss

The main problem with the the 0-1 loss is that it is not differentiable. A smooth version is available for probabilistic model: the **log-probability** given by the model to training examples

Performance Measure P

Typically, performance measure **P** is specific to the task **T**

One possible choice for **classification tasks**

$$Error\ rate = \frac{N_{incorrectly\ classified}}{N_{total}} \Leftrightarrow Accuracy = \frac{N_{correctly\ classified}}{N_{total}} = 1 - Error\ rate$$

Error rate = expected 0-1 loss

The main problem with the the 0-1 loss is that it is not differentiable. A smooth version is available for probabilistic model: the **log-probability** given by the model to training examples

Possible choices for **regression tasks**:

Performance Measure P

Typically, performance measure **P** is specific to the task **T**

One possible choice for **classification tasks**

$$Error\ rate = \frac{N_{incorrectly\,classified}}{N_{total}} \Leftrightarrow Accuracy = \frac{N_{correctly\ classified}}{N_{total}} = 1 - Error\ rate$$

Error rate = expected 0-1 loss

The main problem with the the 0-1 loss is that it is not differentiable. A smooth version is available for probabilistic model: the **log-probability** given by the model to training examples

Possible choices for **regression tasks**:

Mean square loss: $$L = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{Y_i - \hat{Y}_i}{Y_i}\right)^2$$

Performance Measure P

Typically, performance measure **P** is specific to the task **T**

One possible choice for **classification tasks**

$$Error\ rate = \frac{N_{incorrectly\,classified}}{N_{total}} \Leftrightarrow Accuracy = \frac{N_{correctly\ classified}}{N_{total}} = 1 - Error\ rate$$

Error rate = expected 0-1 loss

The main problem with the the 0-1 loss is that it is not differentiable. A smooth version is available for probabilistic model: the **log-probability** given by the model to training examples

Possible choices for **regression tasks**:

Mean square loss: $$L = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{Y_i - \hat{Y}_i}{Y_i}\right)^2$$

L1-loss: $$L = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right|$$

Learning from Experience E

Learning from Experience E

The performance measure **P** improves with Experience **E** as a result of learning

Learning from Experience E

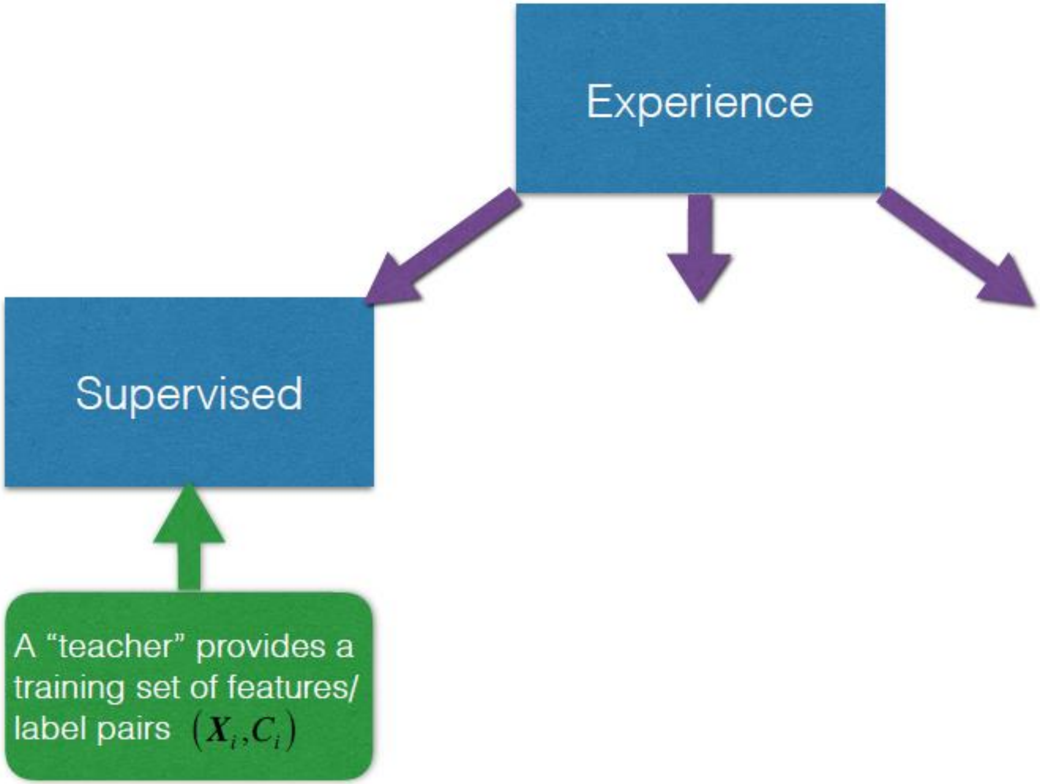
The performance measure **P** improves with Experience **E** as a result of learning

Types of learning from experience **E**

Learning from Experience E

The performance measure **P** improves with Experience **E** as a result of learning

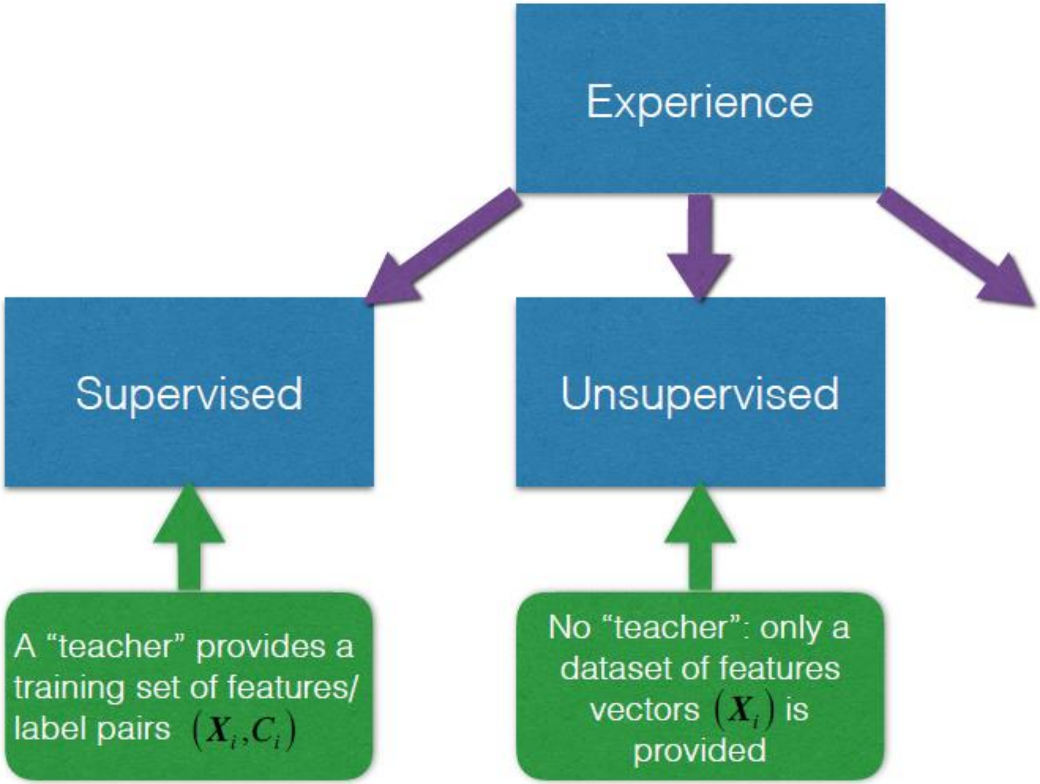
Types of learning from experience **E**



Learning from Experience E

The performance measure **P** improves with Experience **E** as a result of learning

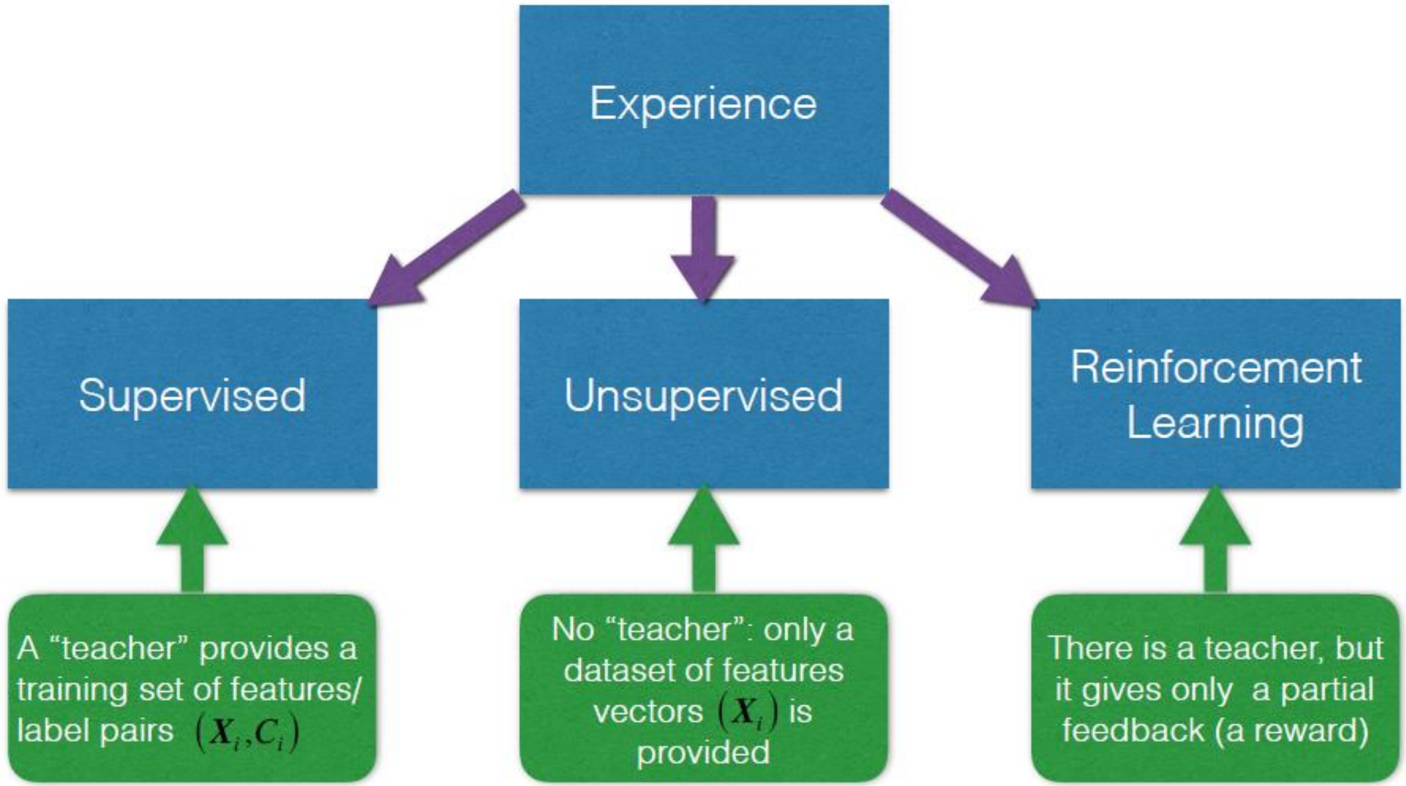
Types of learning from experience **E**



Learning from Experience E

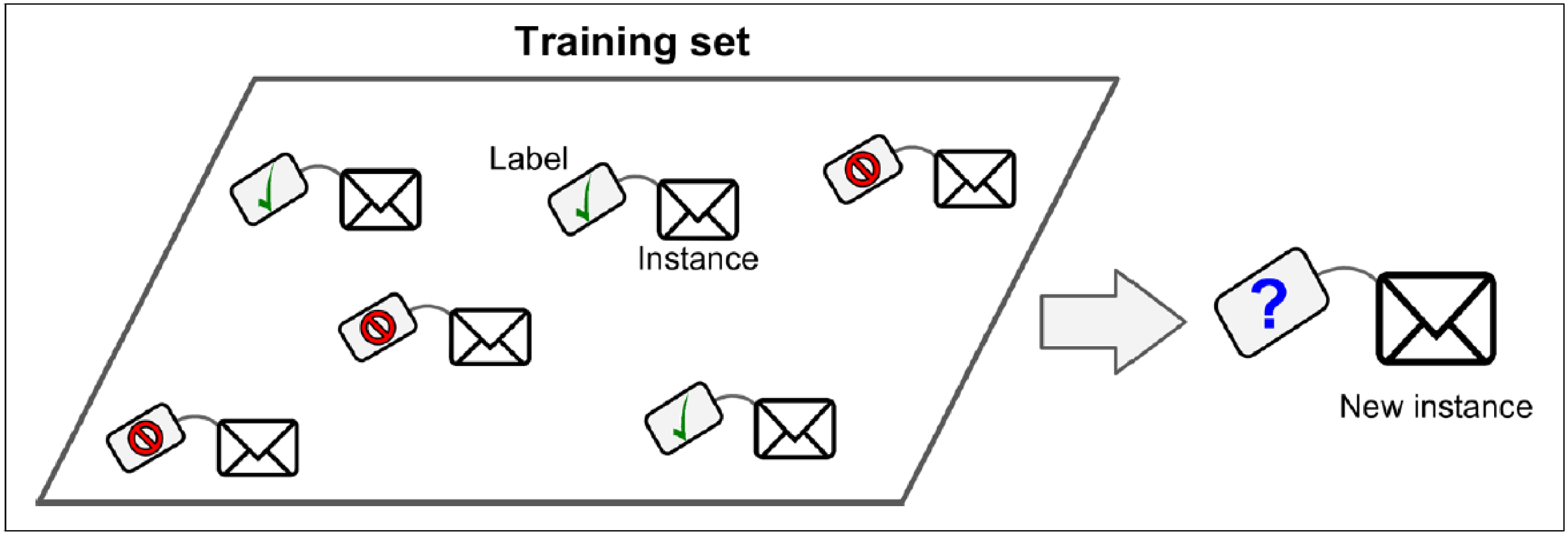
The performance measure **P** improves with Experience **E** as a result of learning

Types of learning from experience **E**



Types of ML: Supervised Learning

Types of ML: Supervised Learning



A labeled training set for supervised learning (e.g., spam classification)

Types of ML: Supervised Learning

Regression

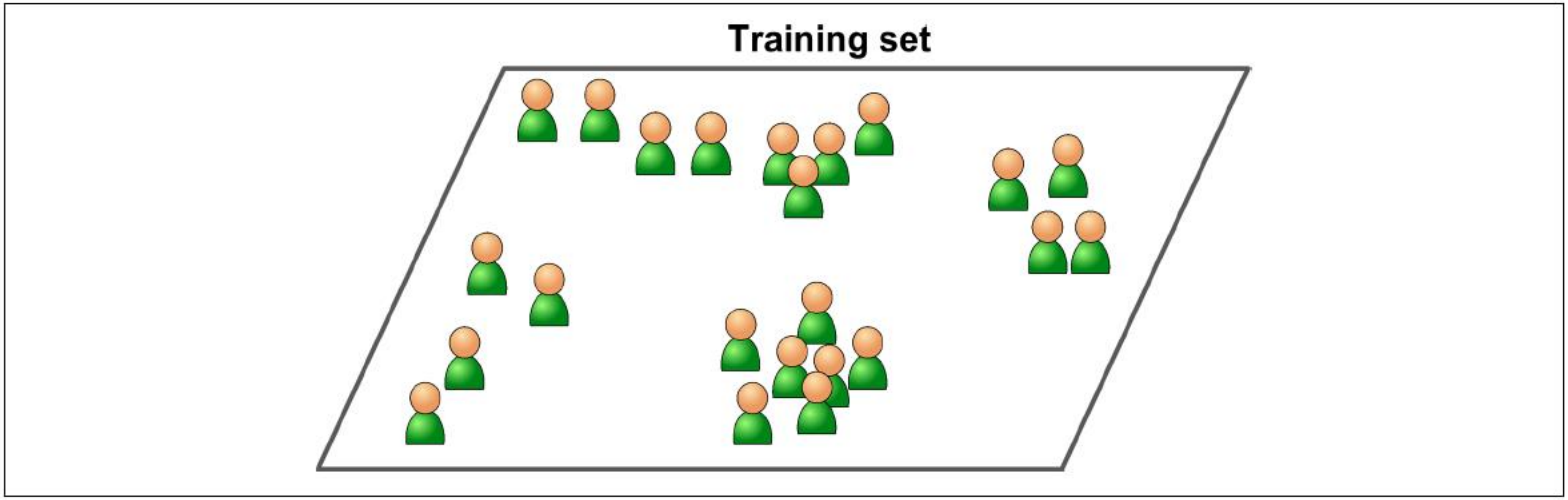
Types of ML: Supervised Leaning

- Most important supervised learning algorithms:
  - k-Nearest Neighbors ☒
  - Linear Regression ☑
  - Logistic Regression ☑
  - Support Vector Machines (SVMs) ☒
  - Decision Trees and Random Forests ☒
  - Neural networks ☑

☒ Not in this course

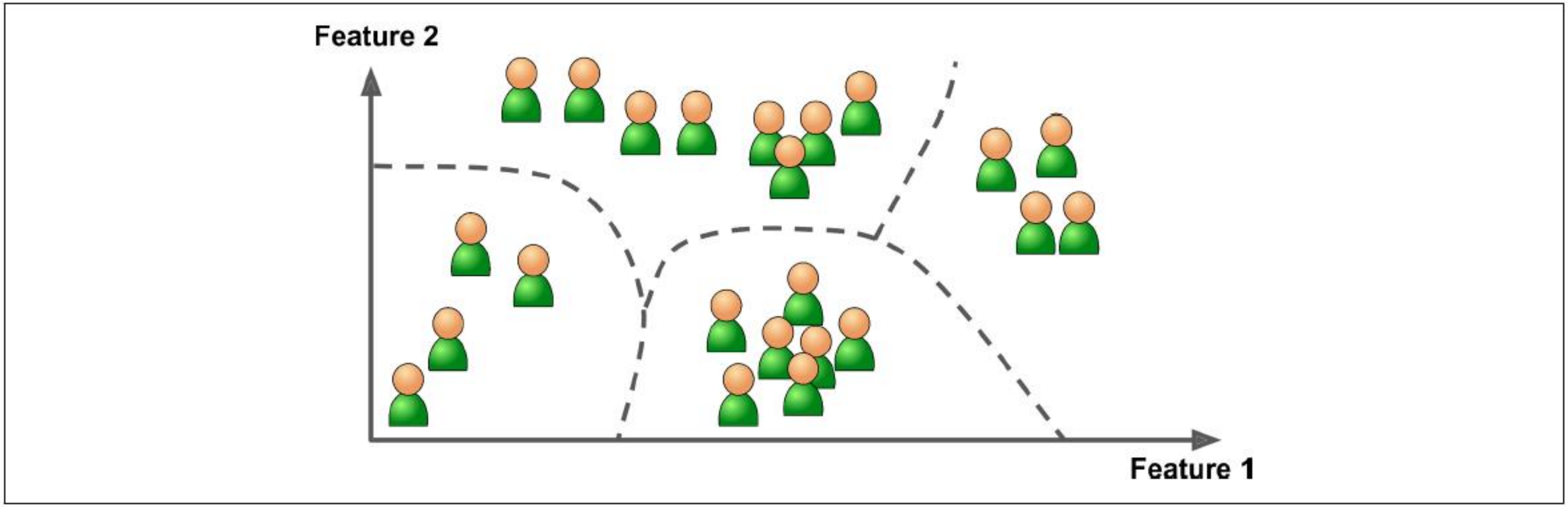
Types of ML: Unsupervised Learning

Types of ML: Unsupervised Learning



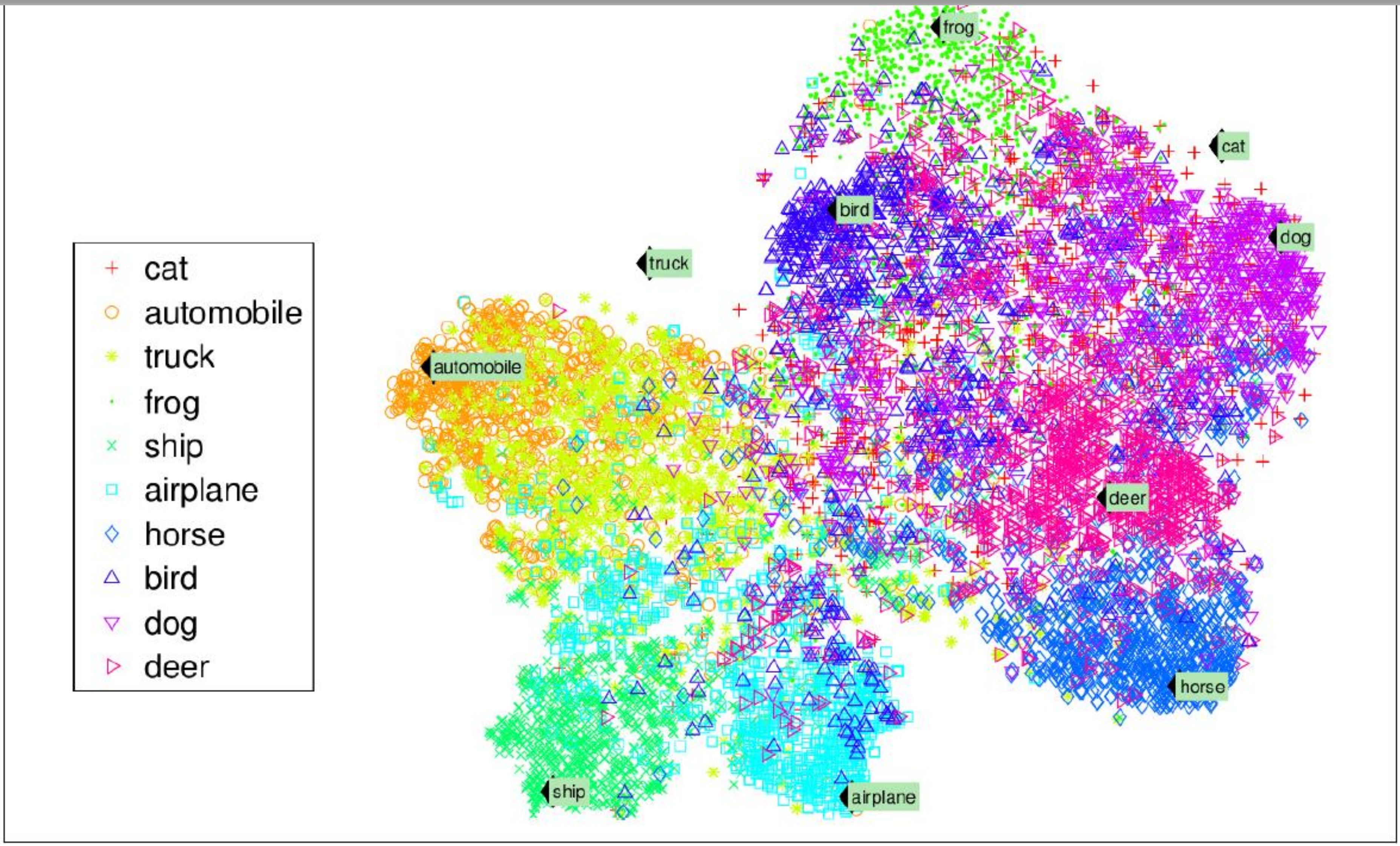
An unlabeled training set for unsupervised learning

Types of ML: Unsupervised Learning



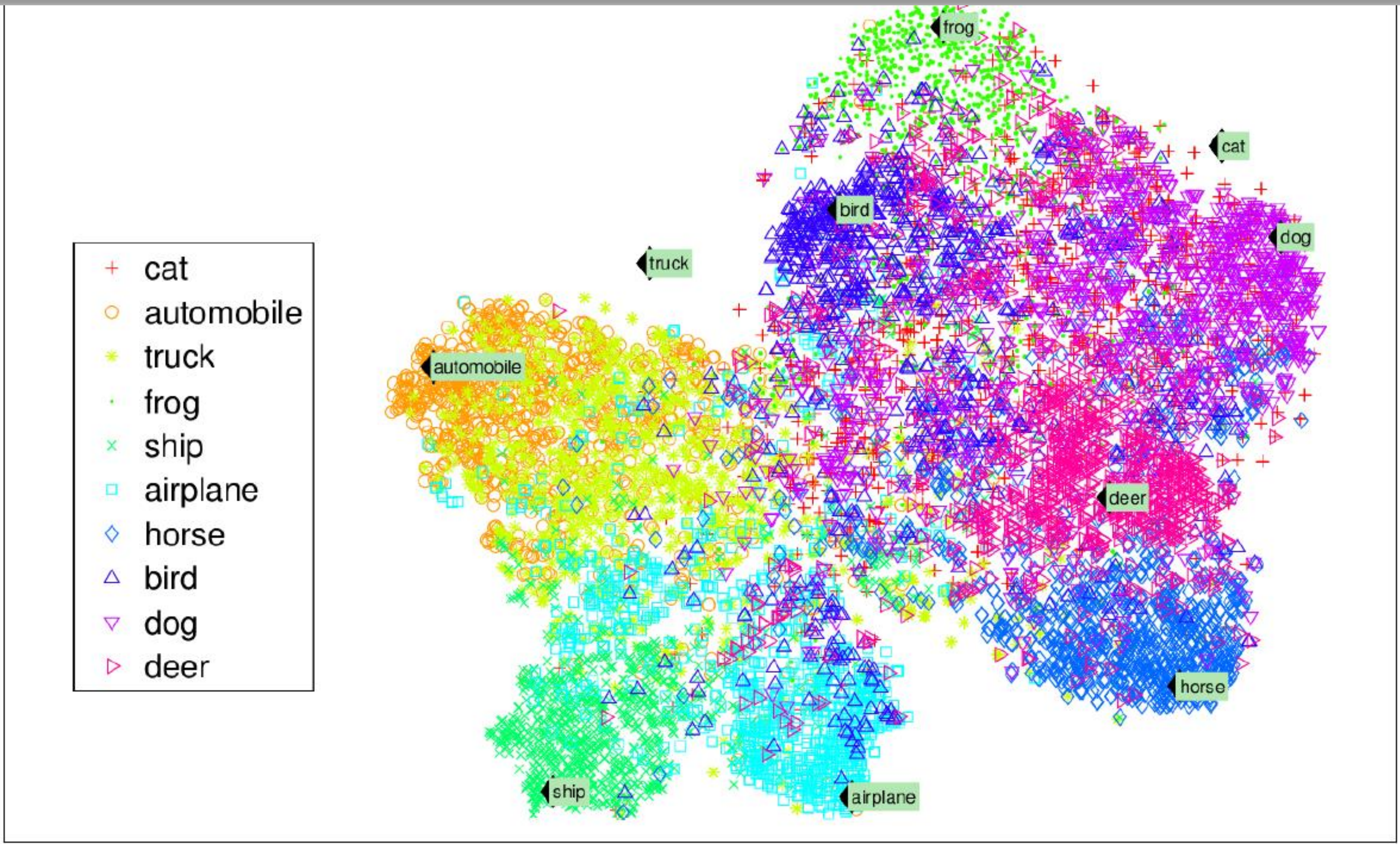
Clustering

Types of ML: Unsupervised Learning



Example of a t-SNE visualization highlighting semantic clusters

Types of ML: Unsupervised Learning

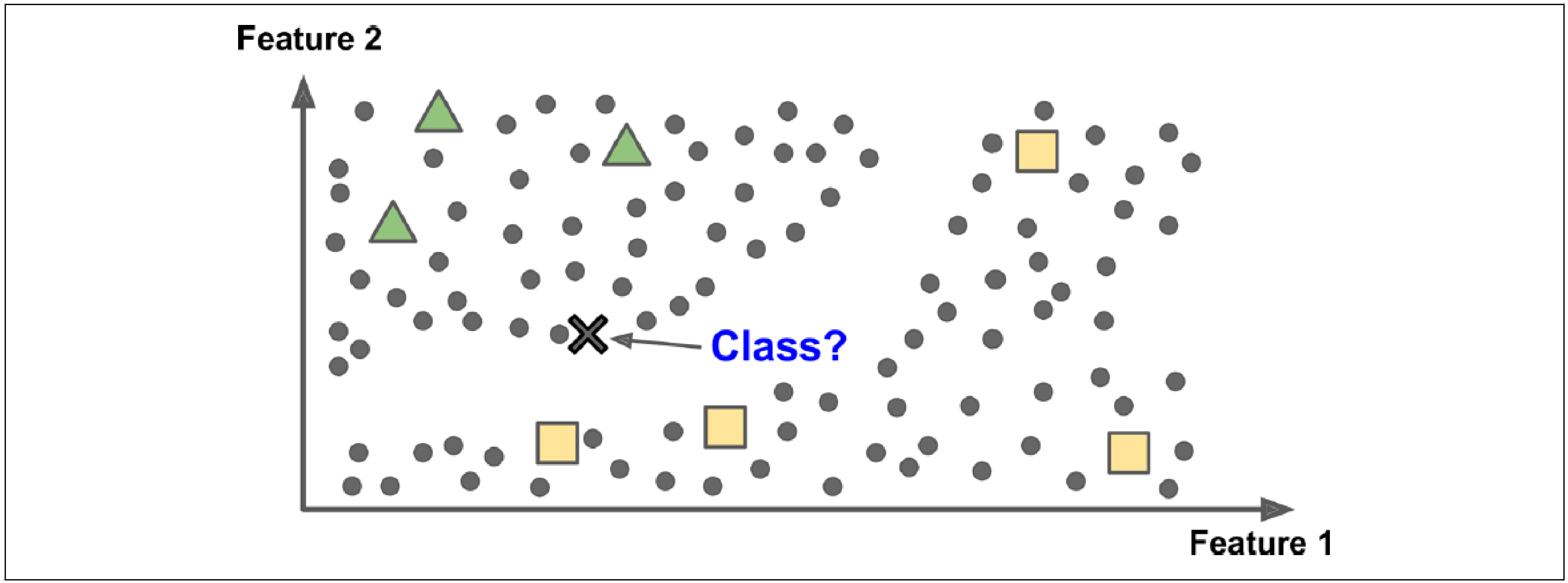


Example of a t-SNE visualization highlighting semantic clusters

Types of ML: Unsupervised Learning

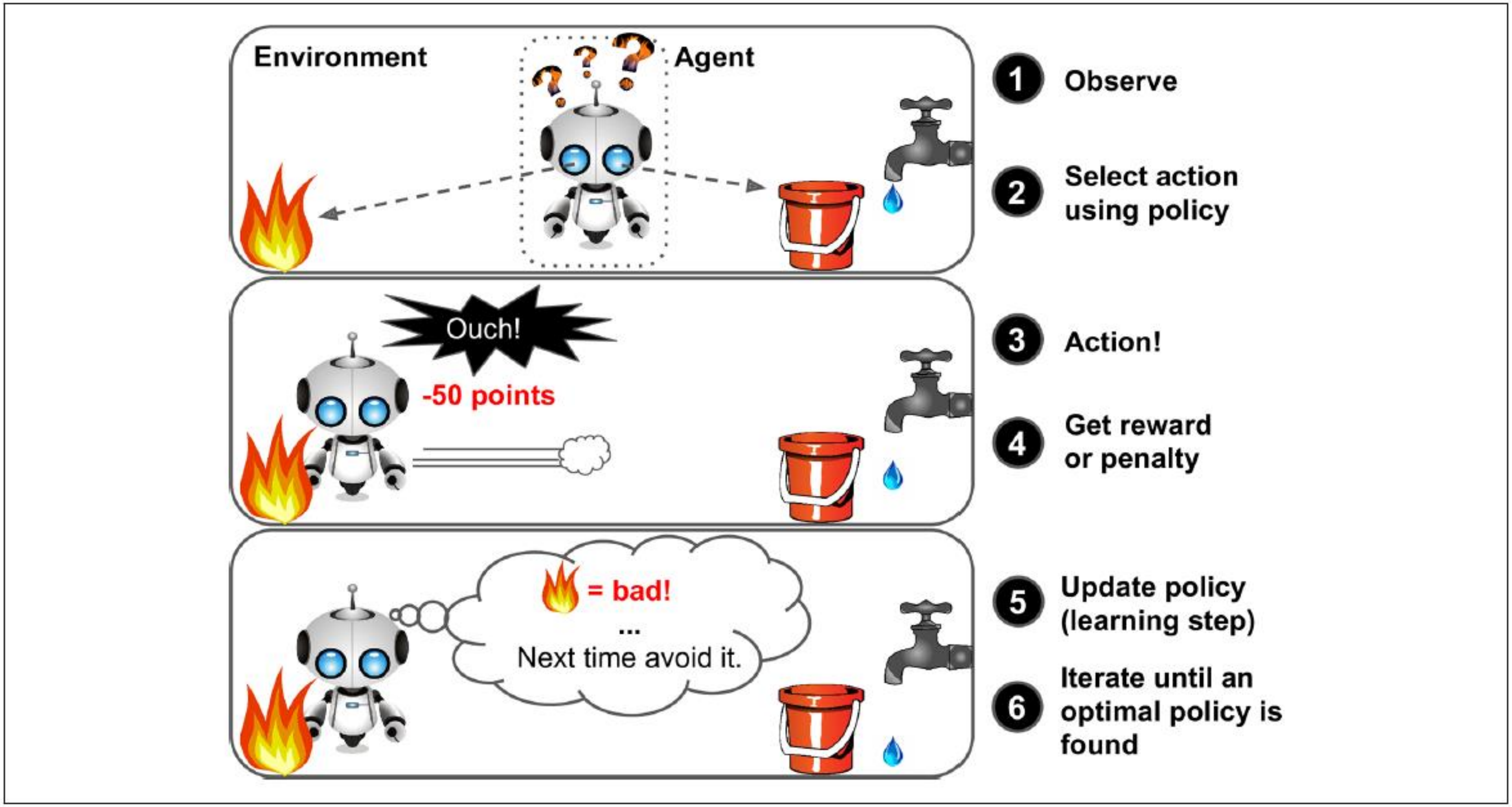
- Most important unsupervised learning algos:
  - Clustering
    - k-Means
    - Hierarchical Cluster Analysis (HCA)
    - Expectation Maximization
  - Visualization and dimensionality reduction
    - Principal Component Analysis (PCA)
    - Locally-Linear Embedding (LLE)
    - t-distributed Stochastic Neighbor Embedding (t-SNE)
  - Association rule learning
    - Apriori/Eclat

Types of ML: Semi Supervised



Semisupervised learning

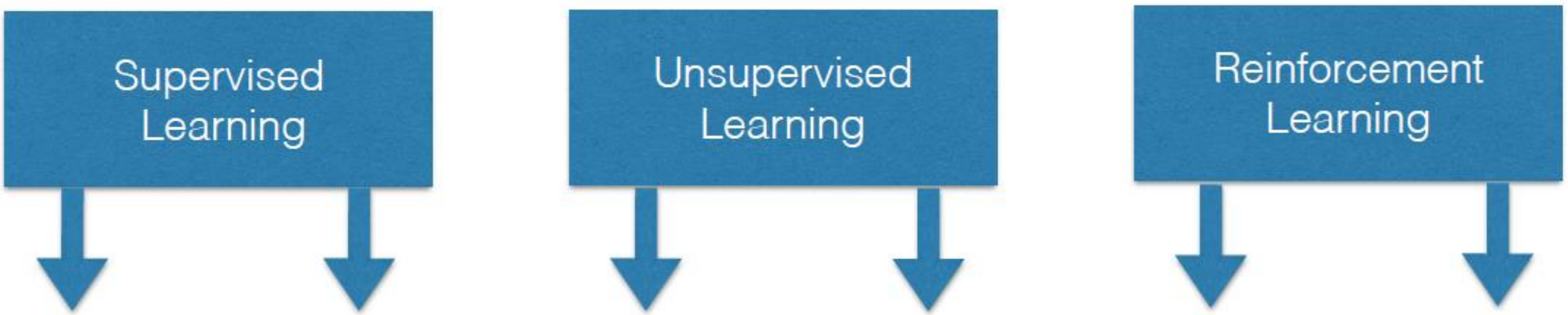
Types of ML: Reinforcement



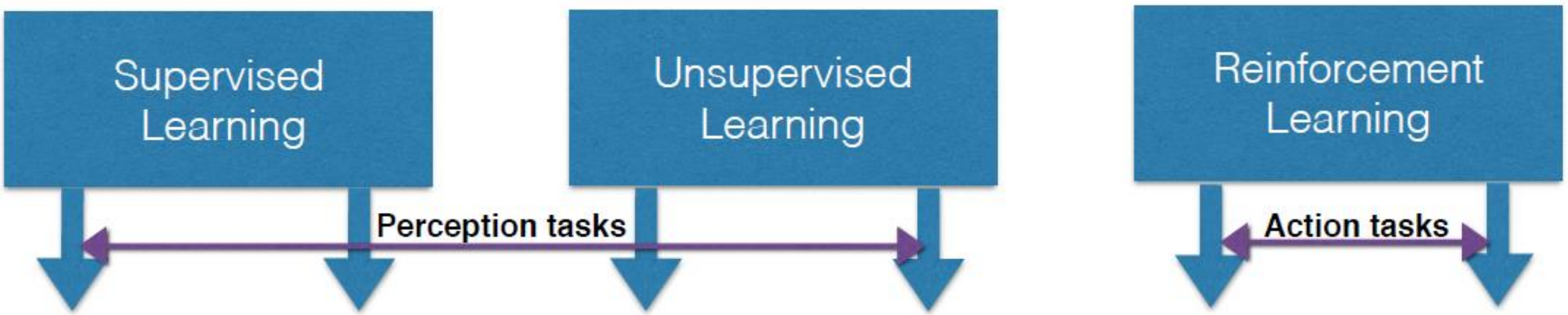
Reinforcement Learning

ML Landscape

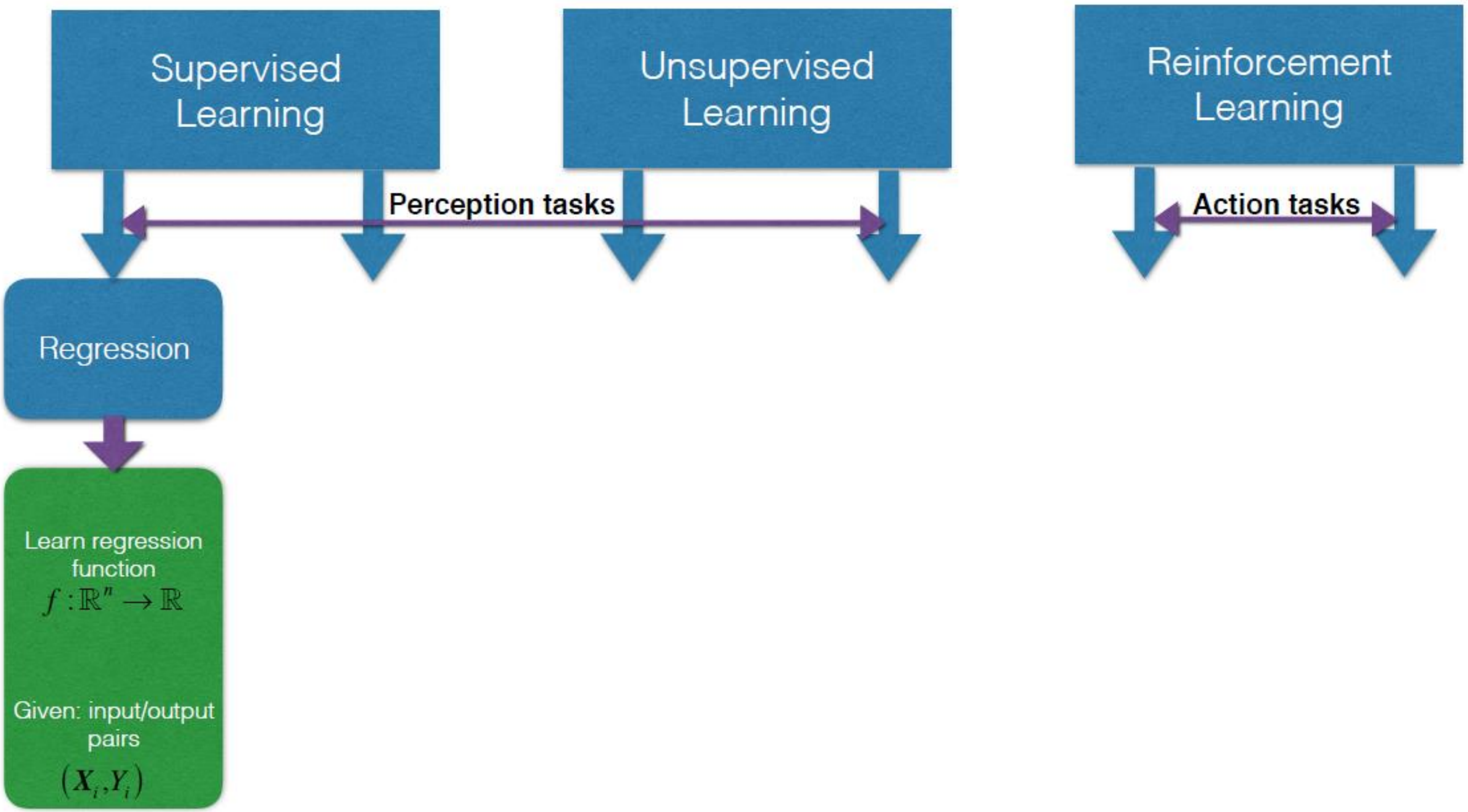
ML Landscape



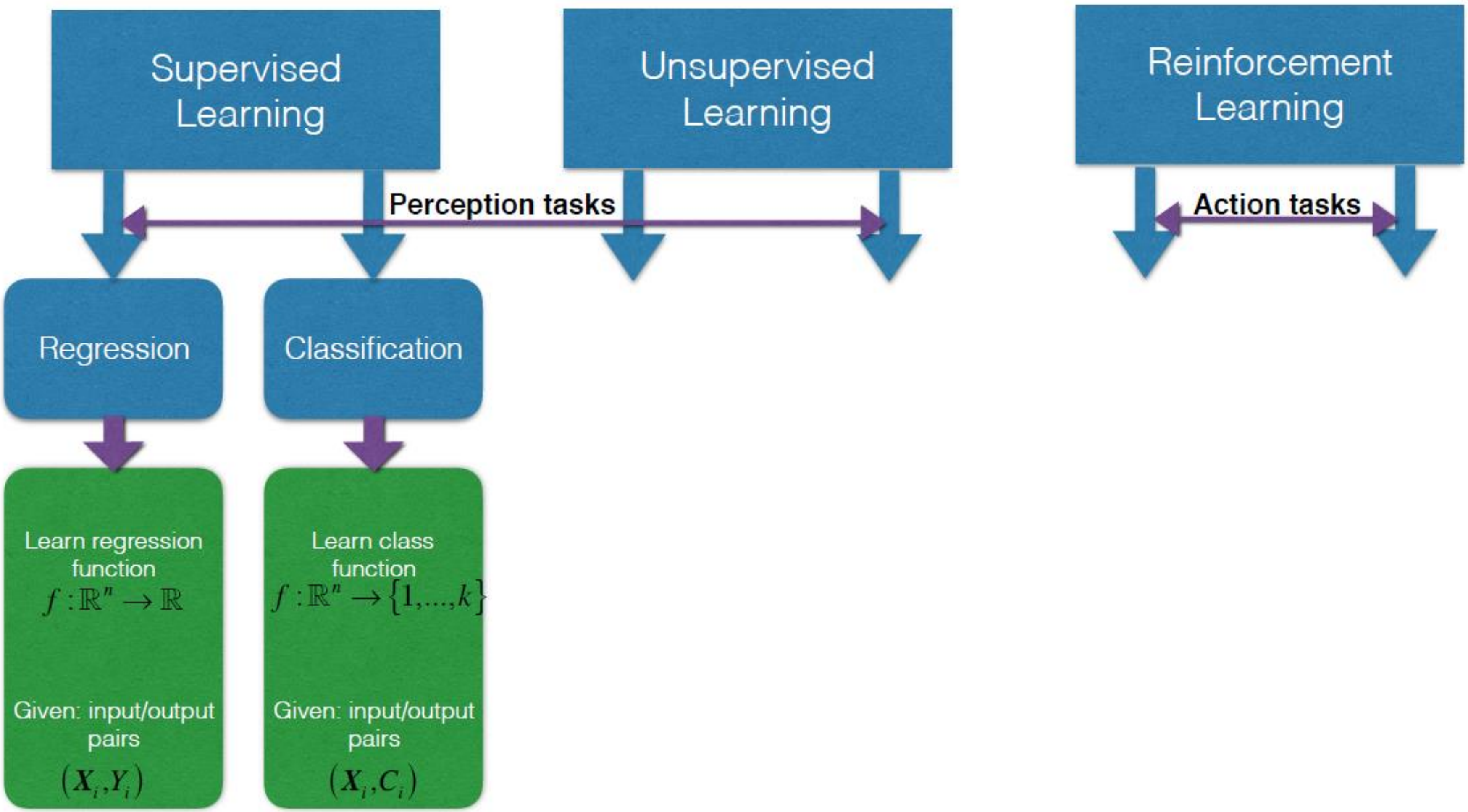
ML Landscape



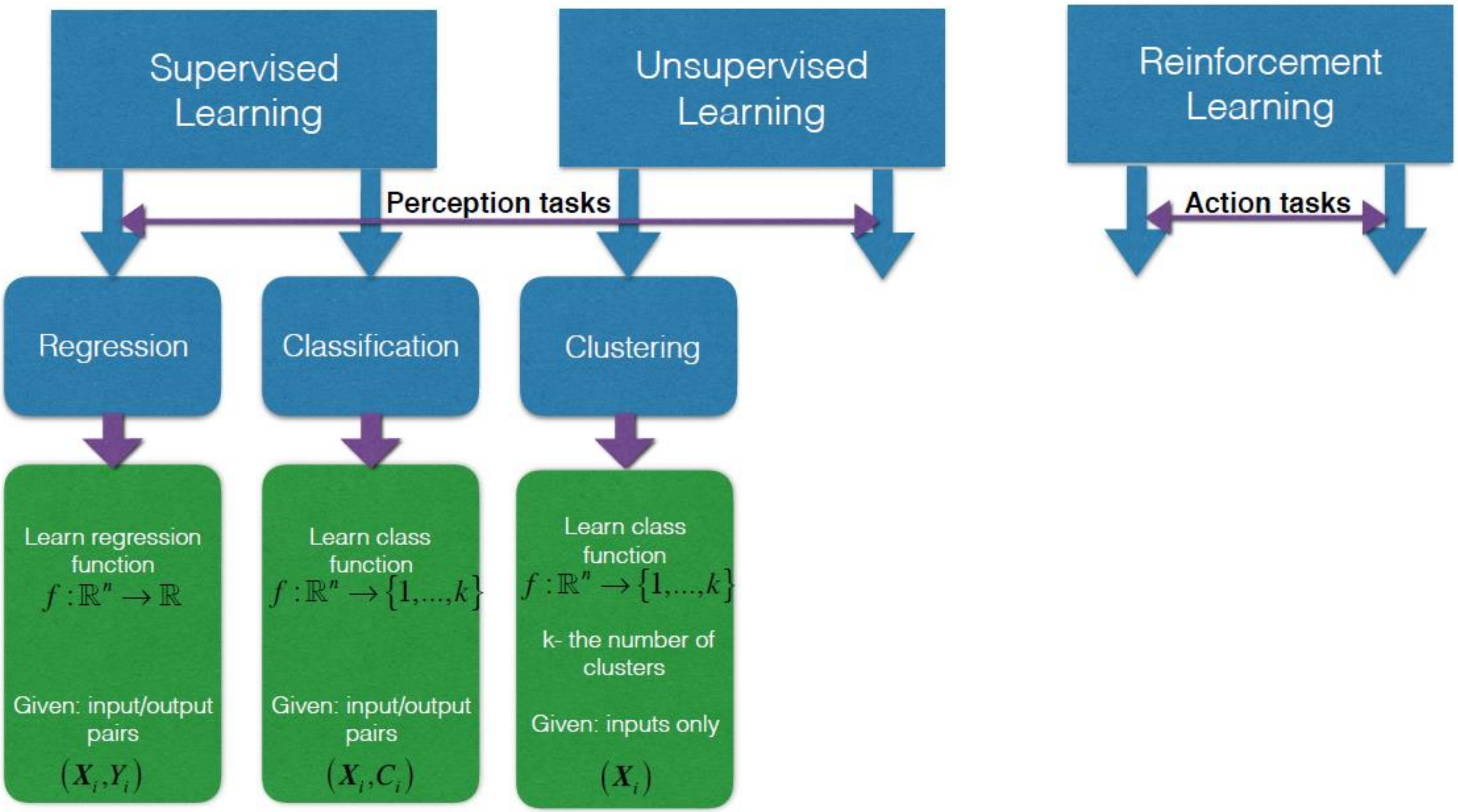
ML Landscape



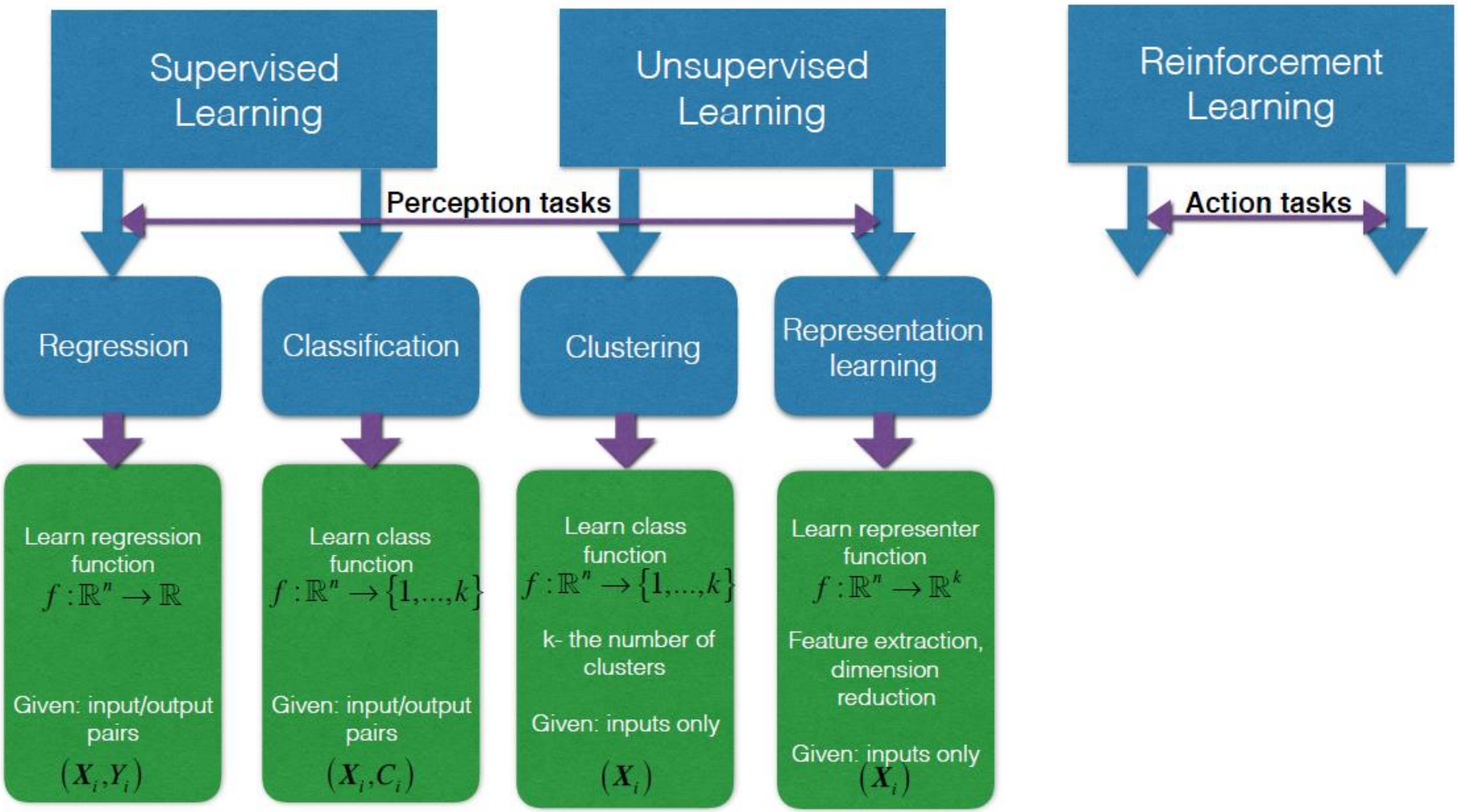
ML Landscape



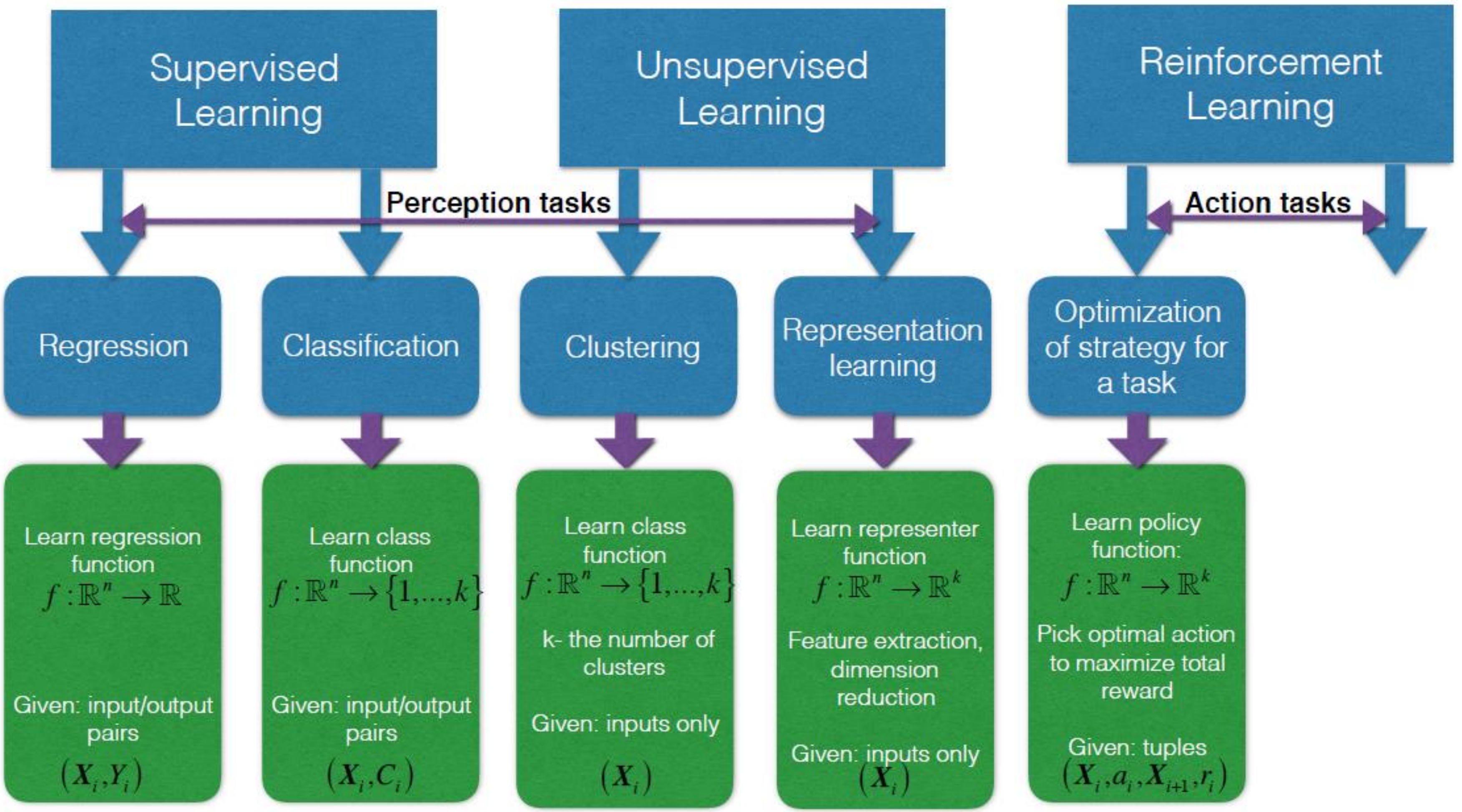
ML Landscape



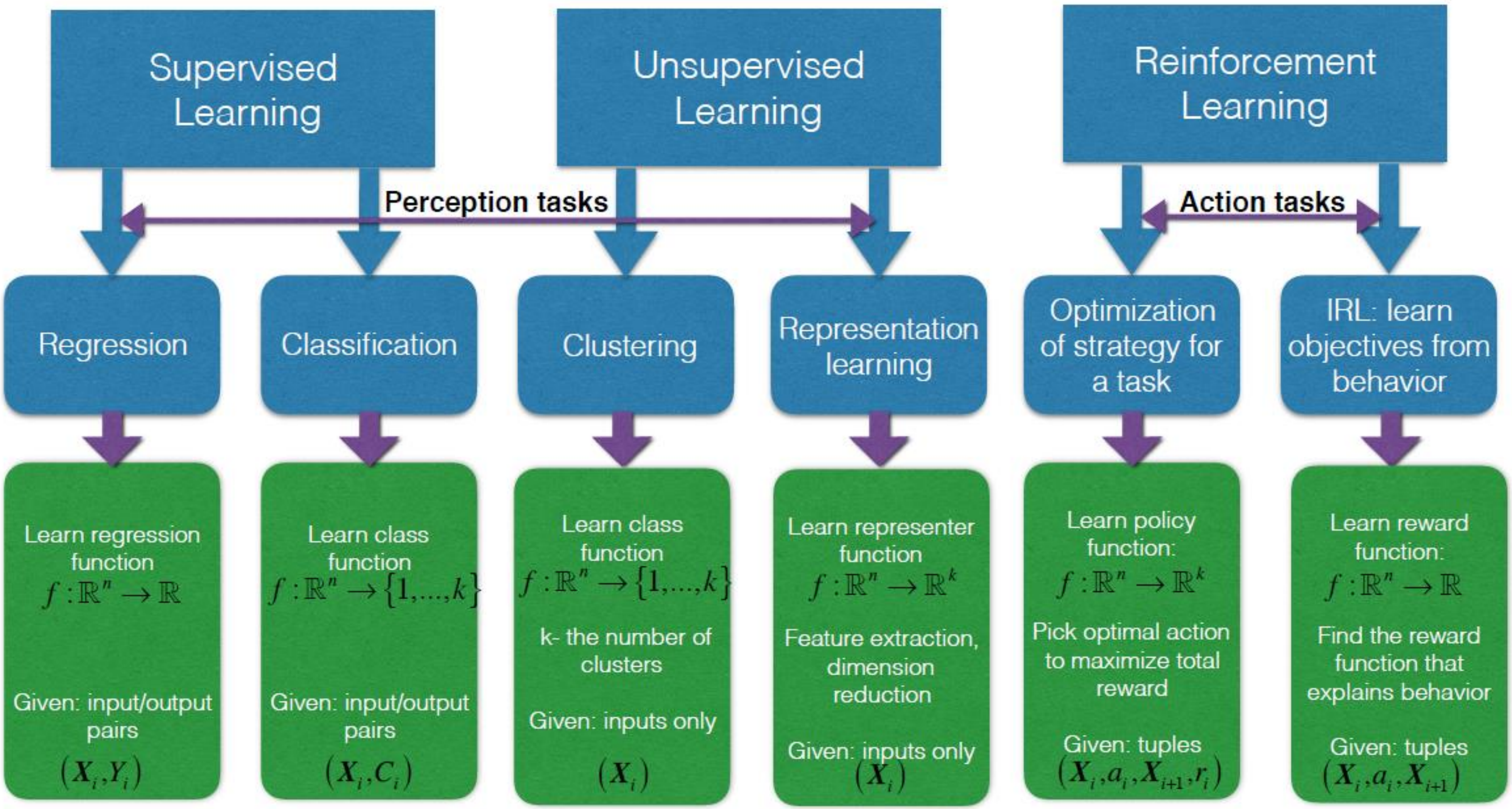
ML Landscape



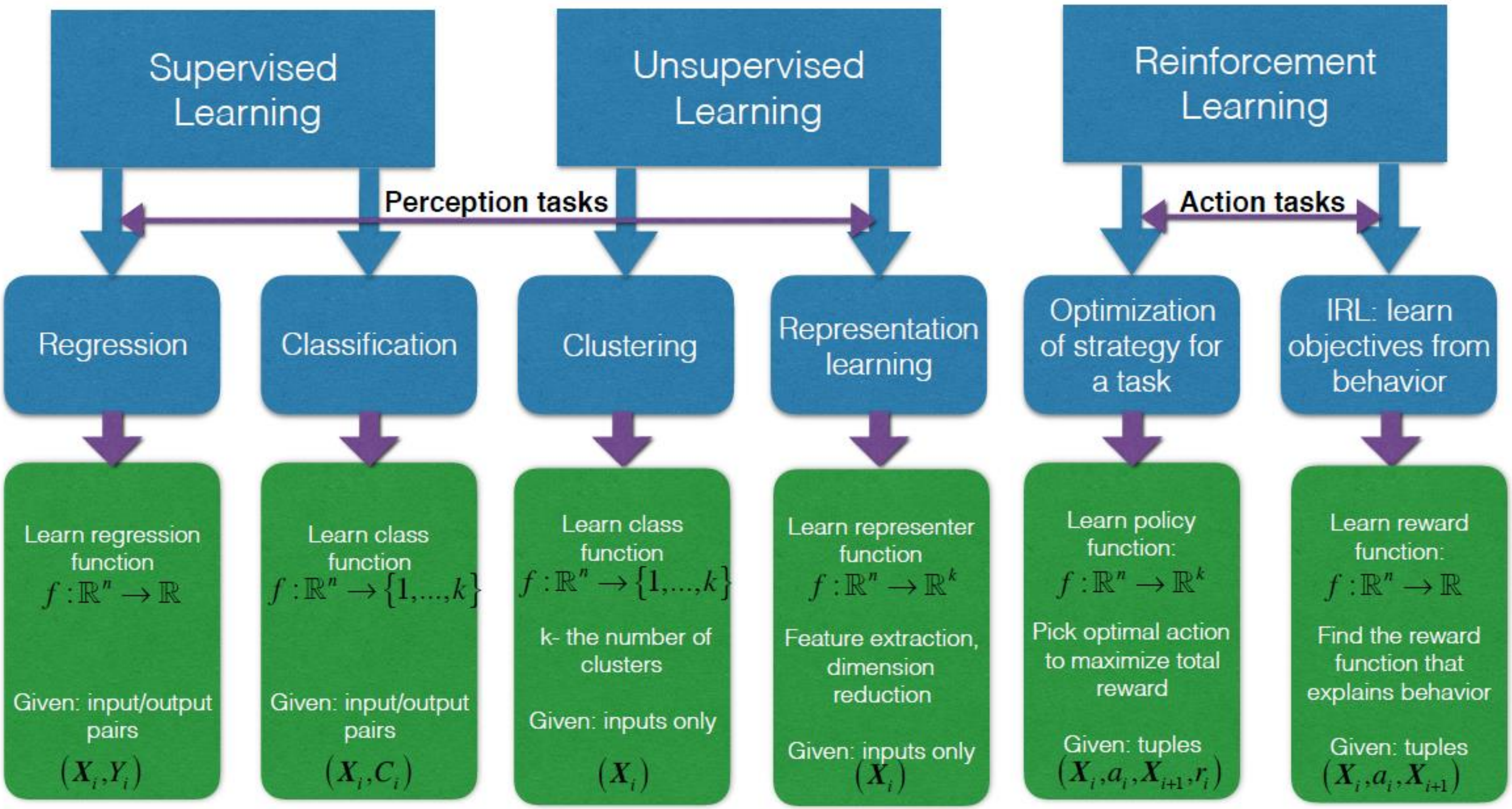
ML Landscape



ML Landscape



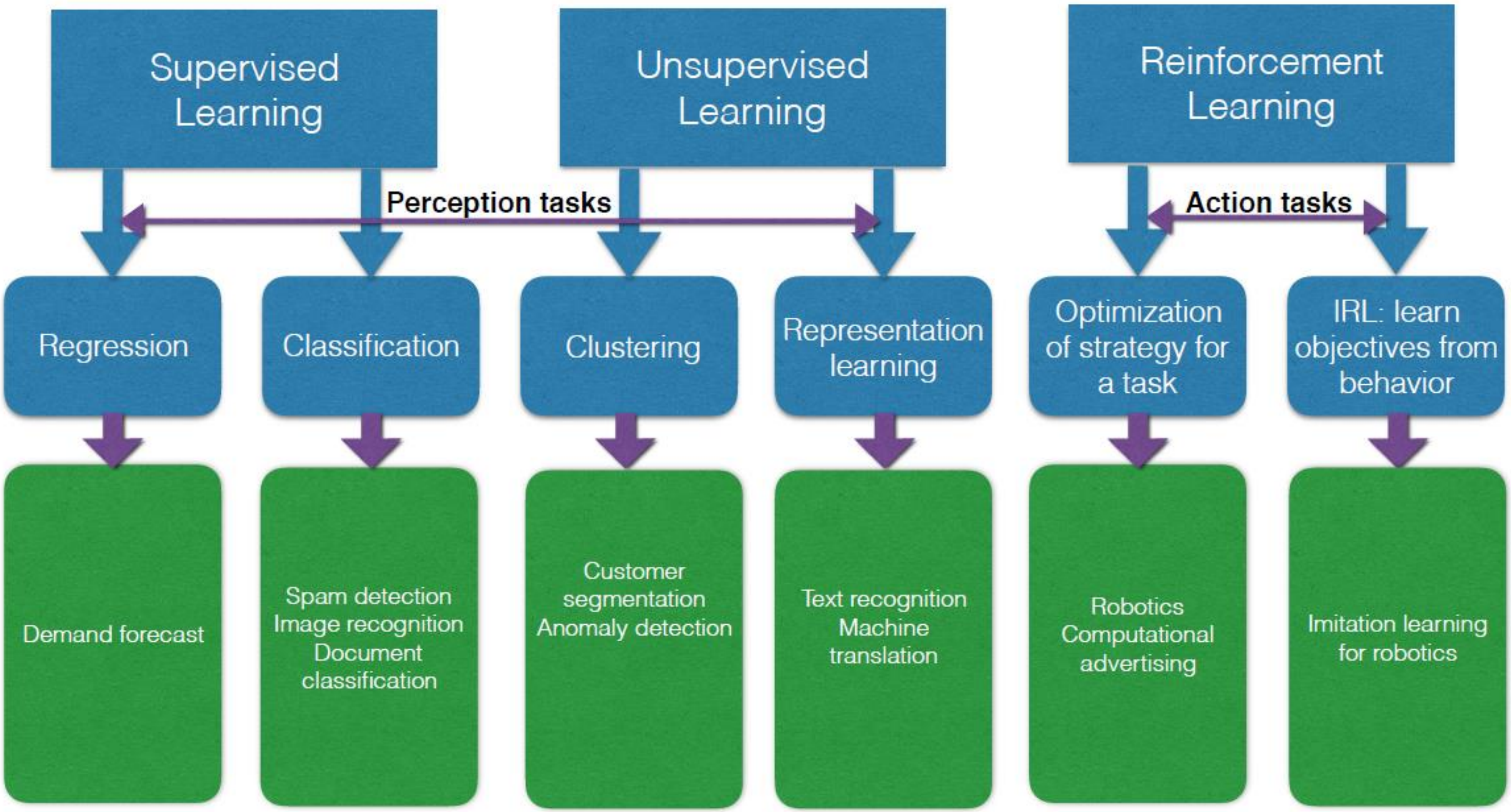
ML Landscape



Additional categories:

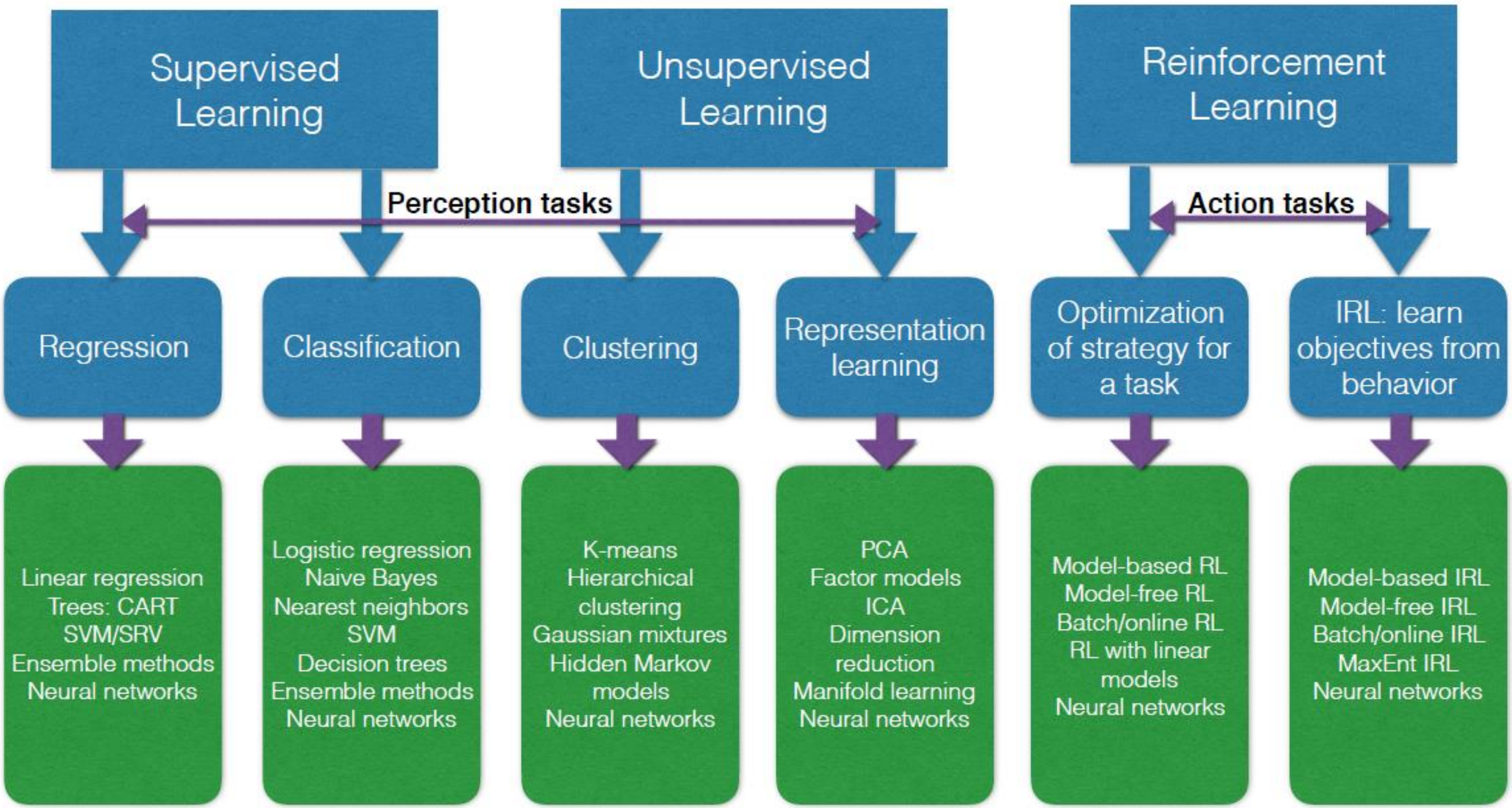
- Semi-supervised learning
- Active learning
- Sequence modeling
- RL methods for Supervised and Unsupervised Learning

ML: Examples in Tech



- These are general industrial applications
- Will be referred to as "ML in Tech" for short

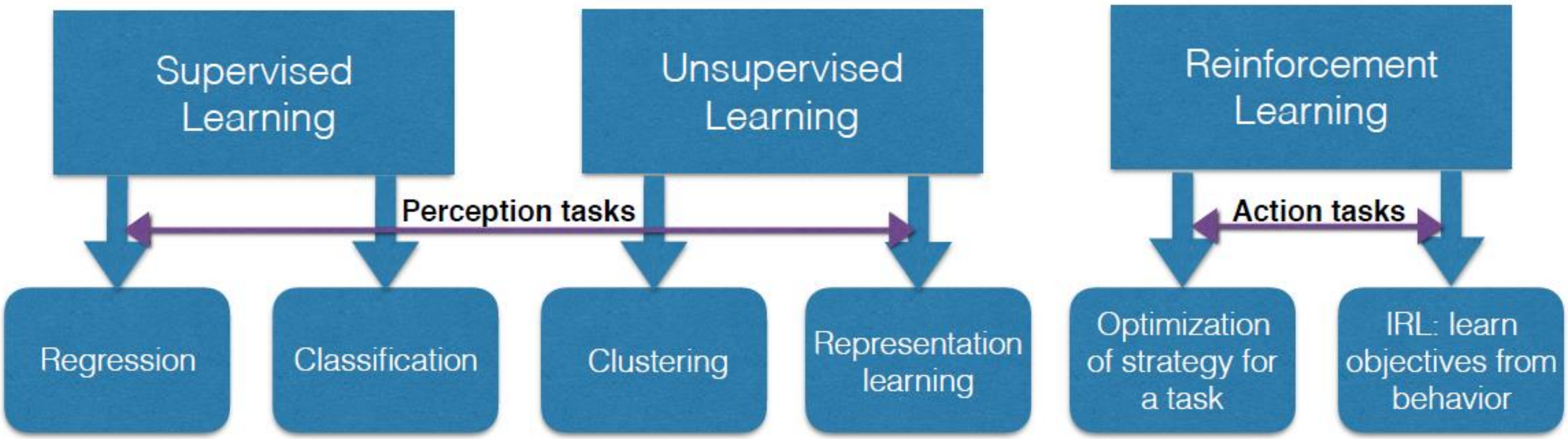
ML: Methods



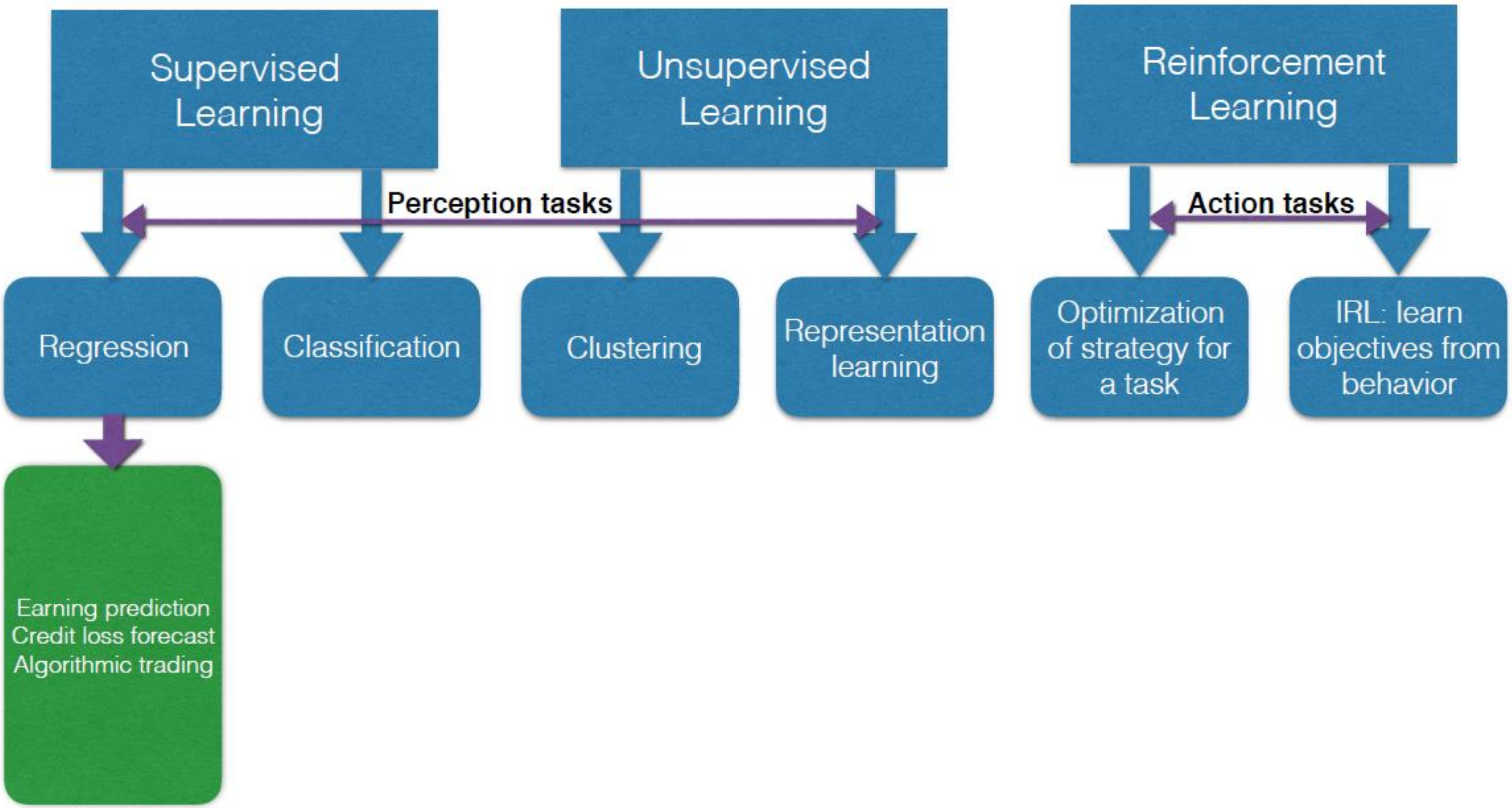
- **Neural networks** is the most universal (and scalable) approach
- **Deep Learning** revolution (2007-present)!

ML in Finance

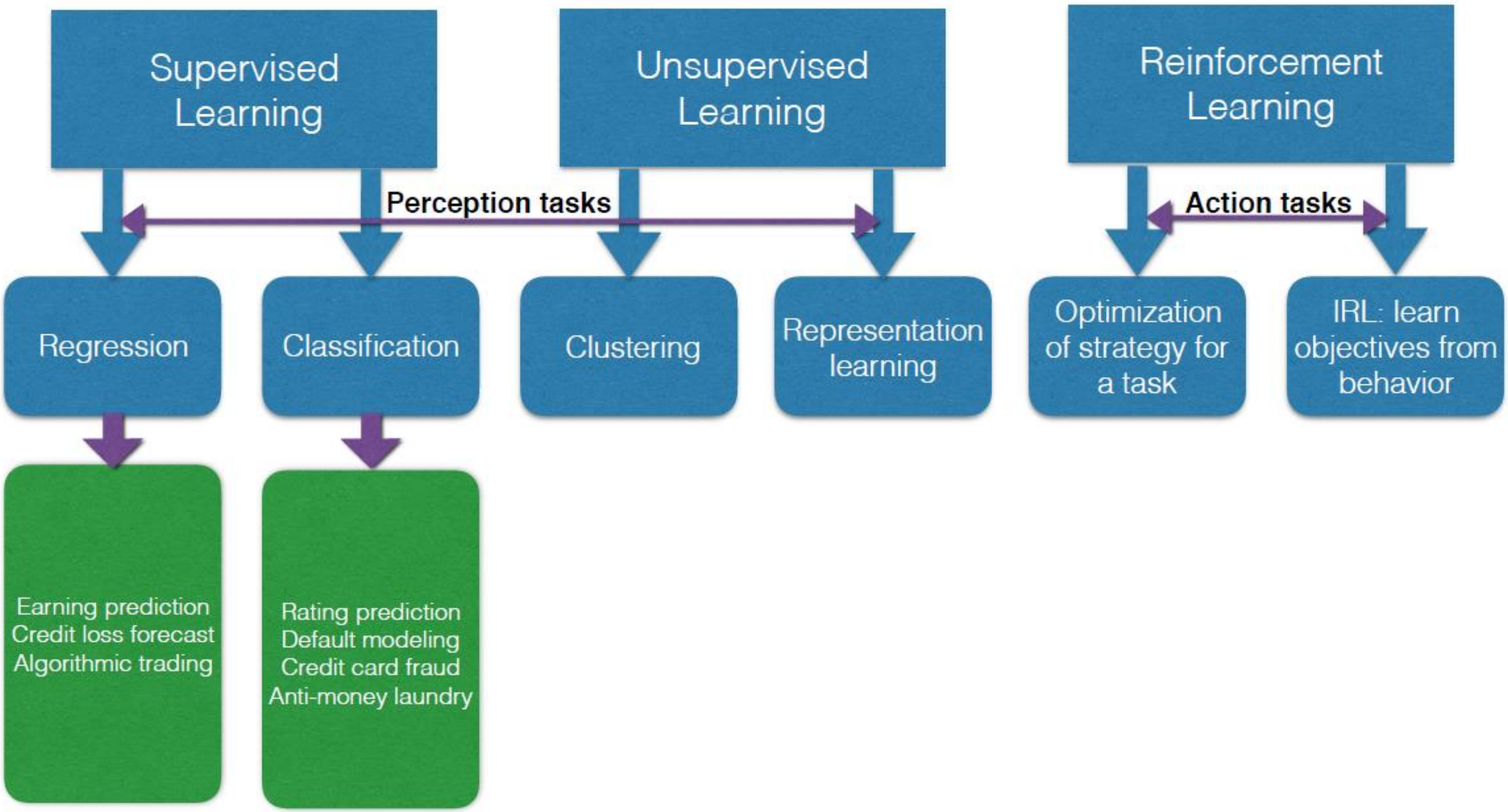
ML in Finance



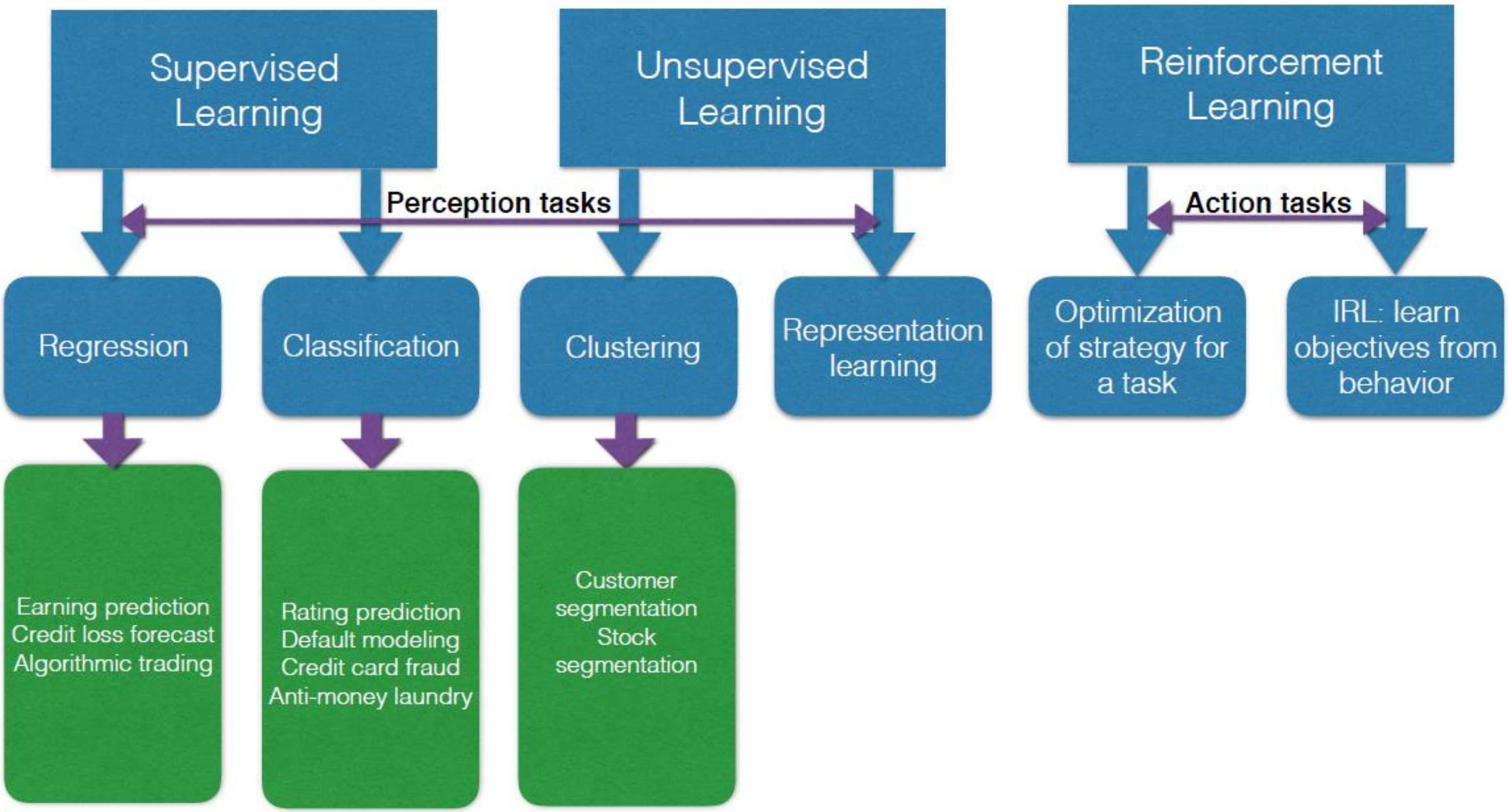
ML in Finance



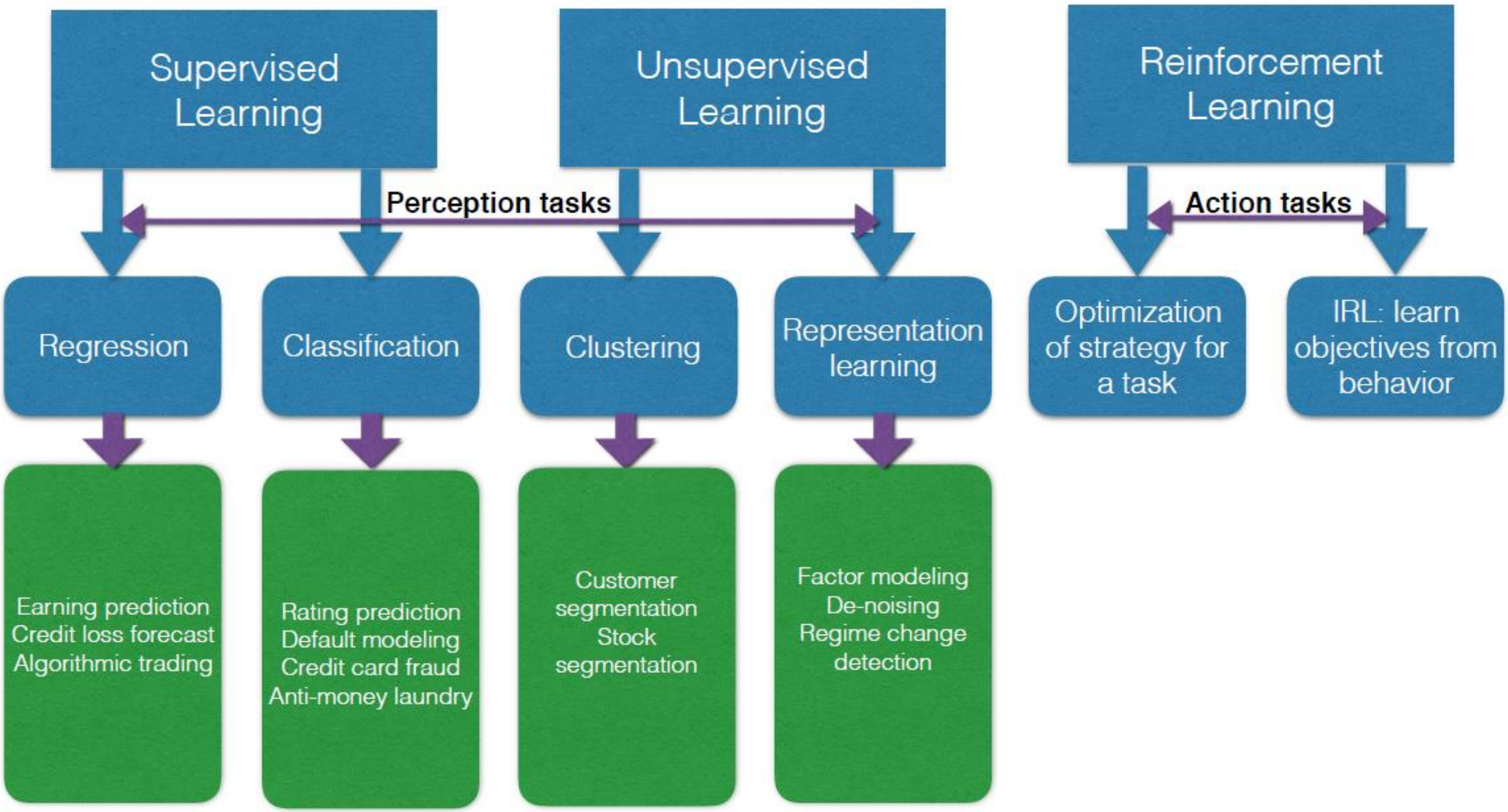
ML in Finance



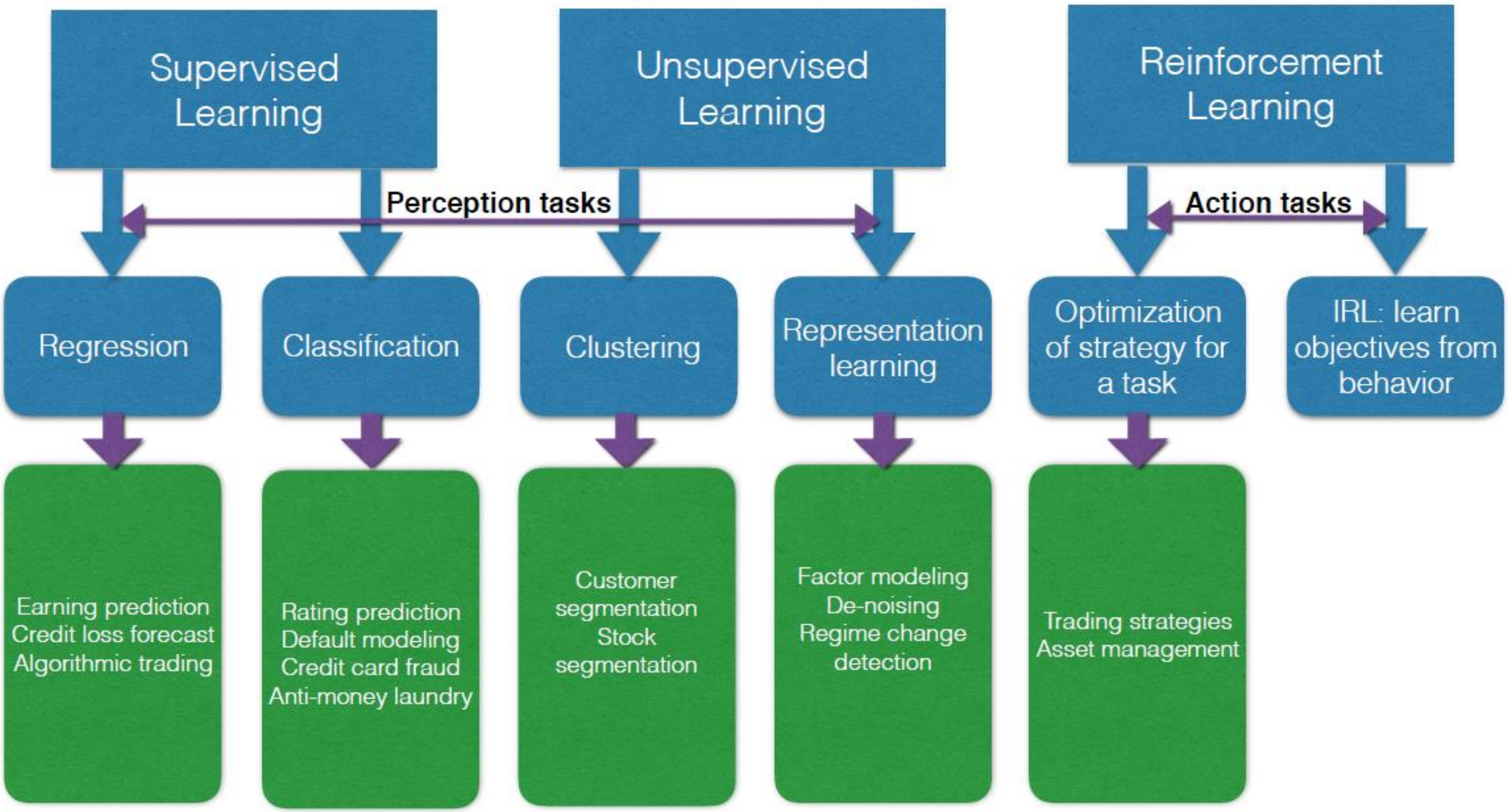
ML in Finance



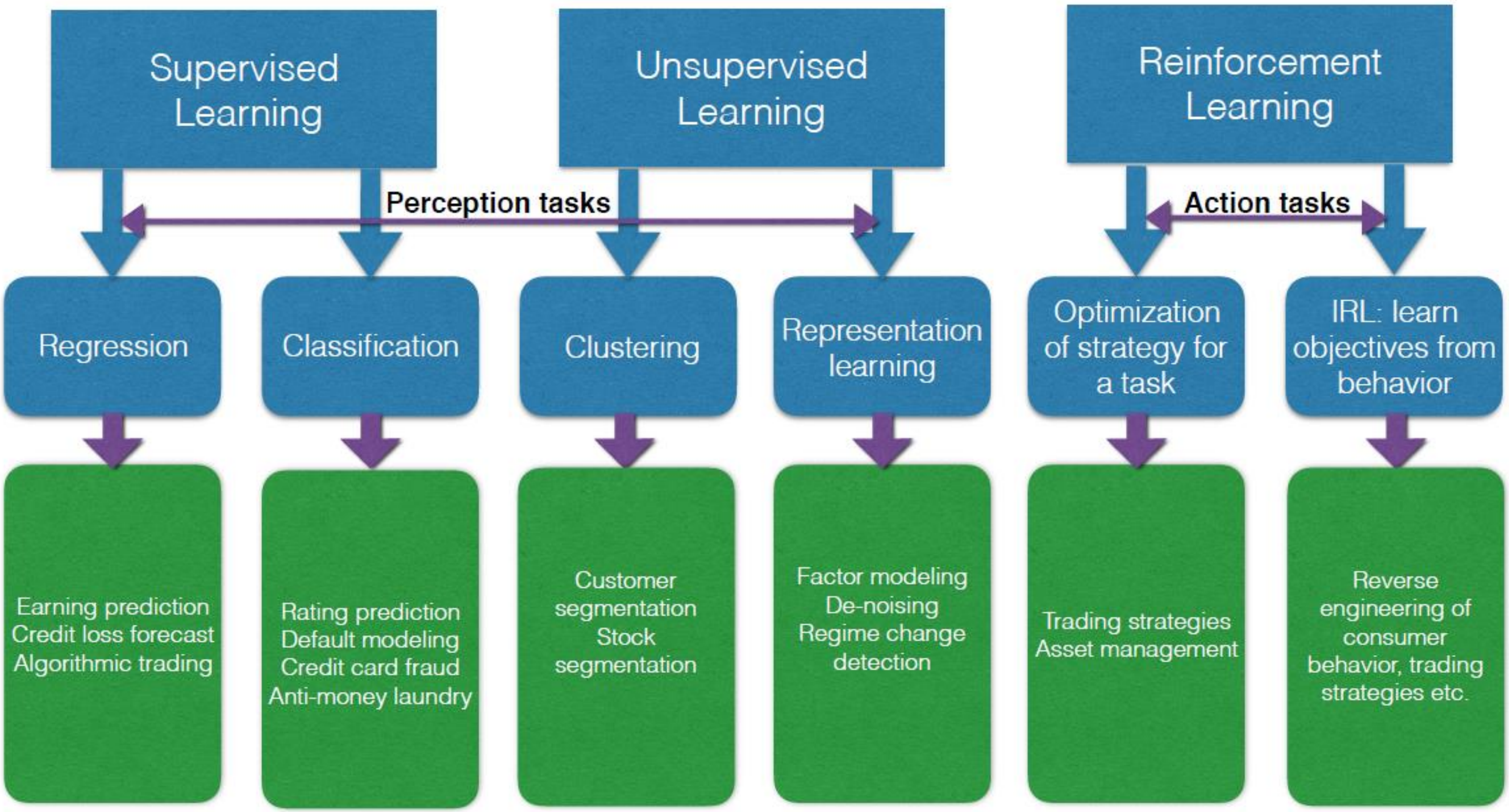
ML in Finance



ML in Finance



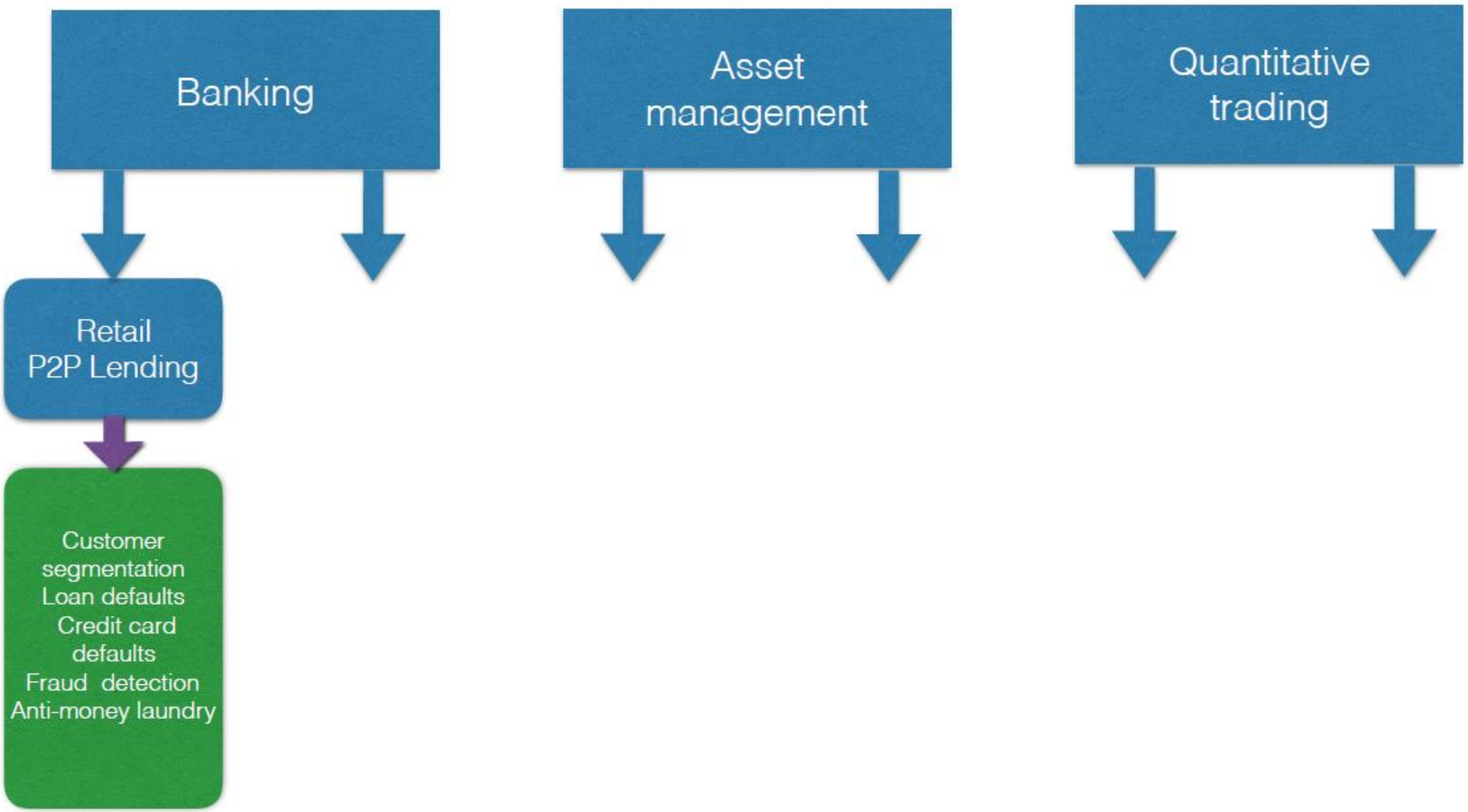
ML in Finance



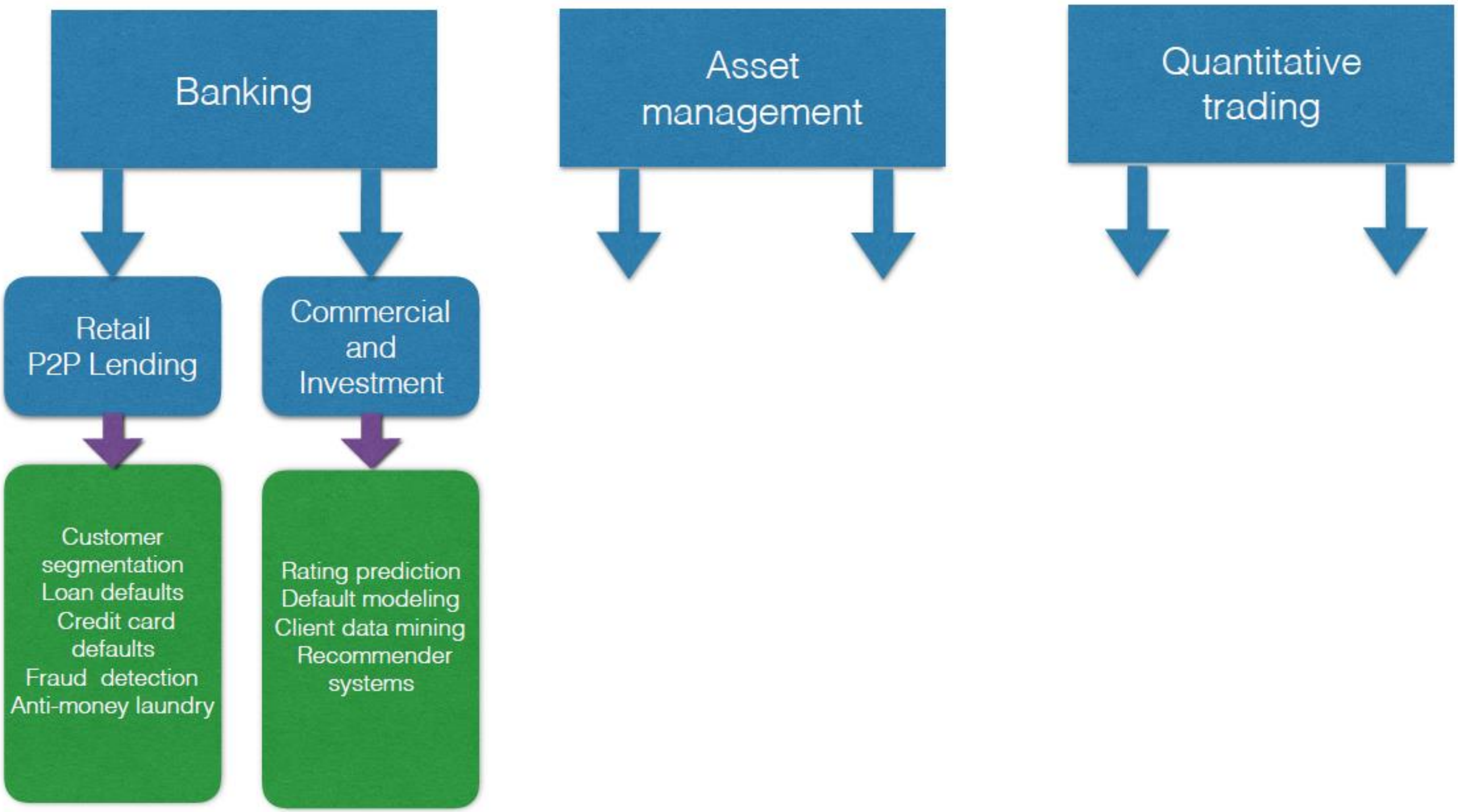
ML by Financial Application Areas

ML by Financial Application Areas

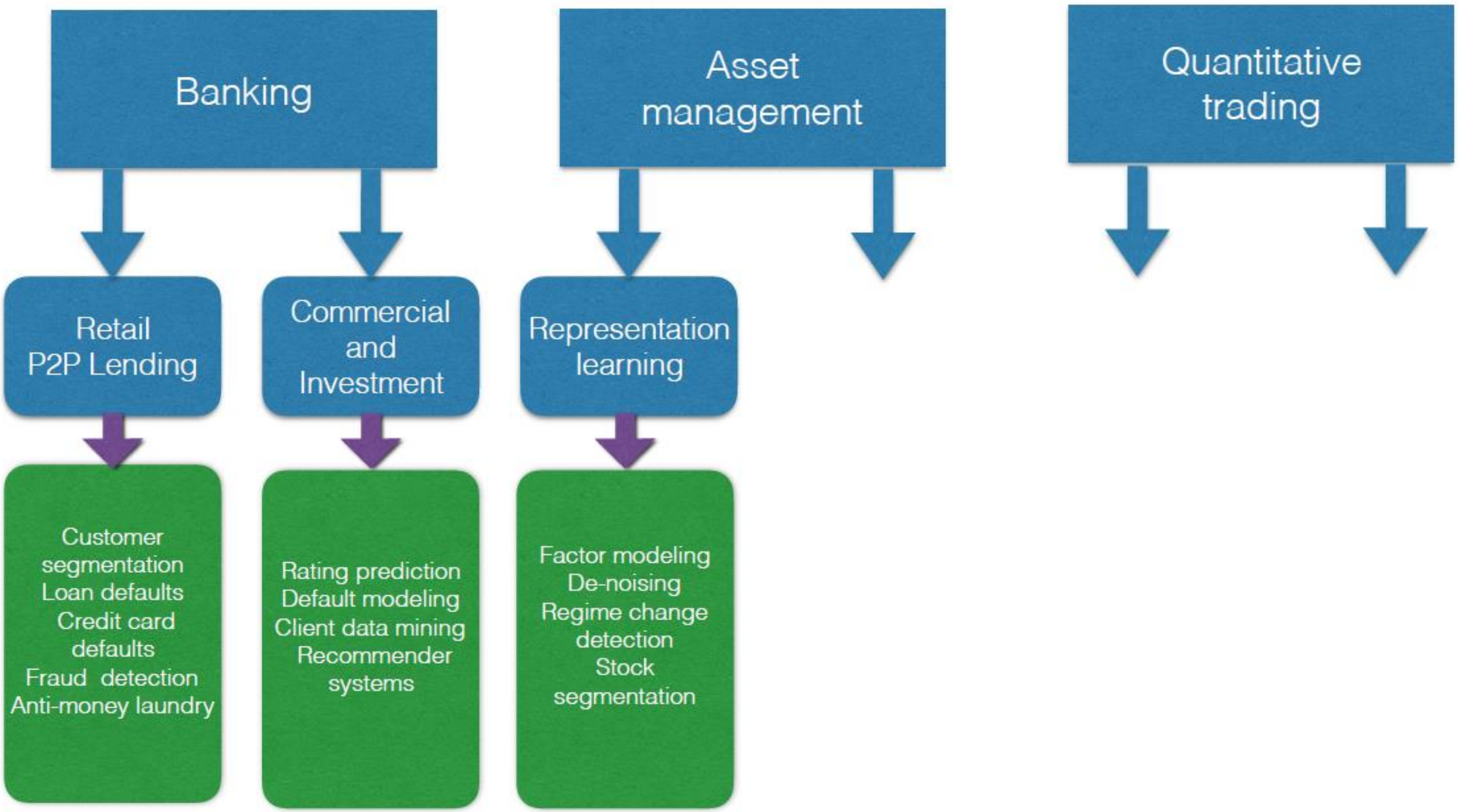
ML by Financial Application Areas



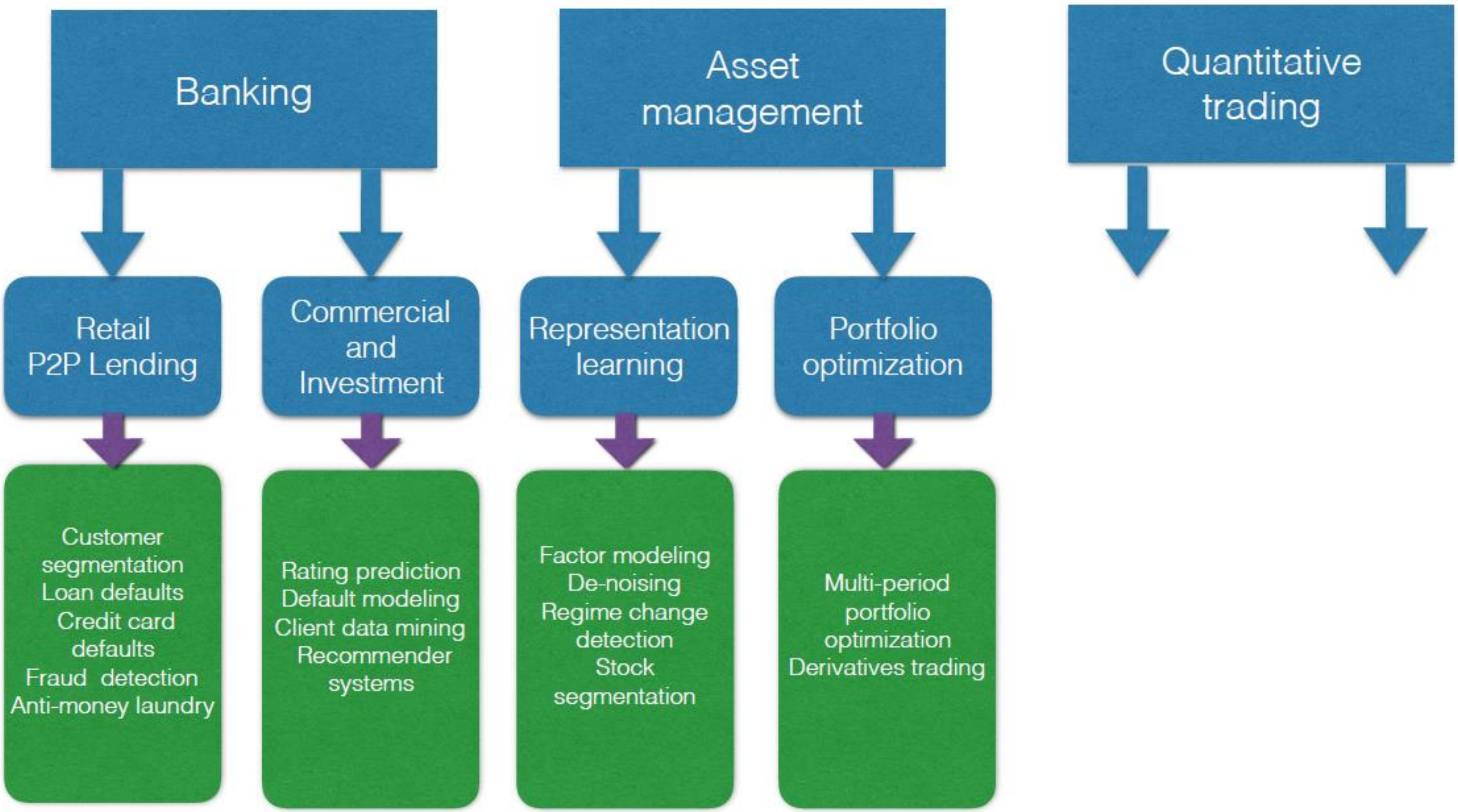
ML by Financial Application Areas



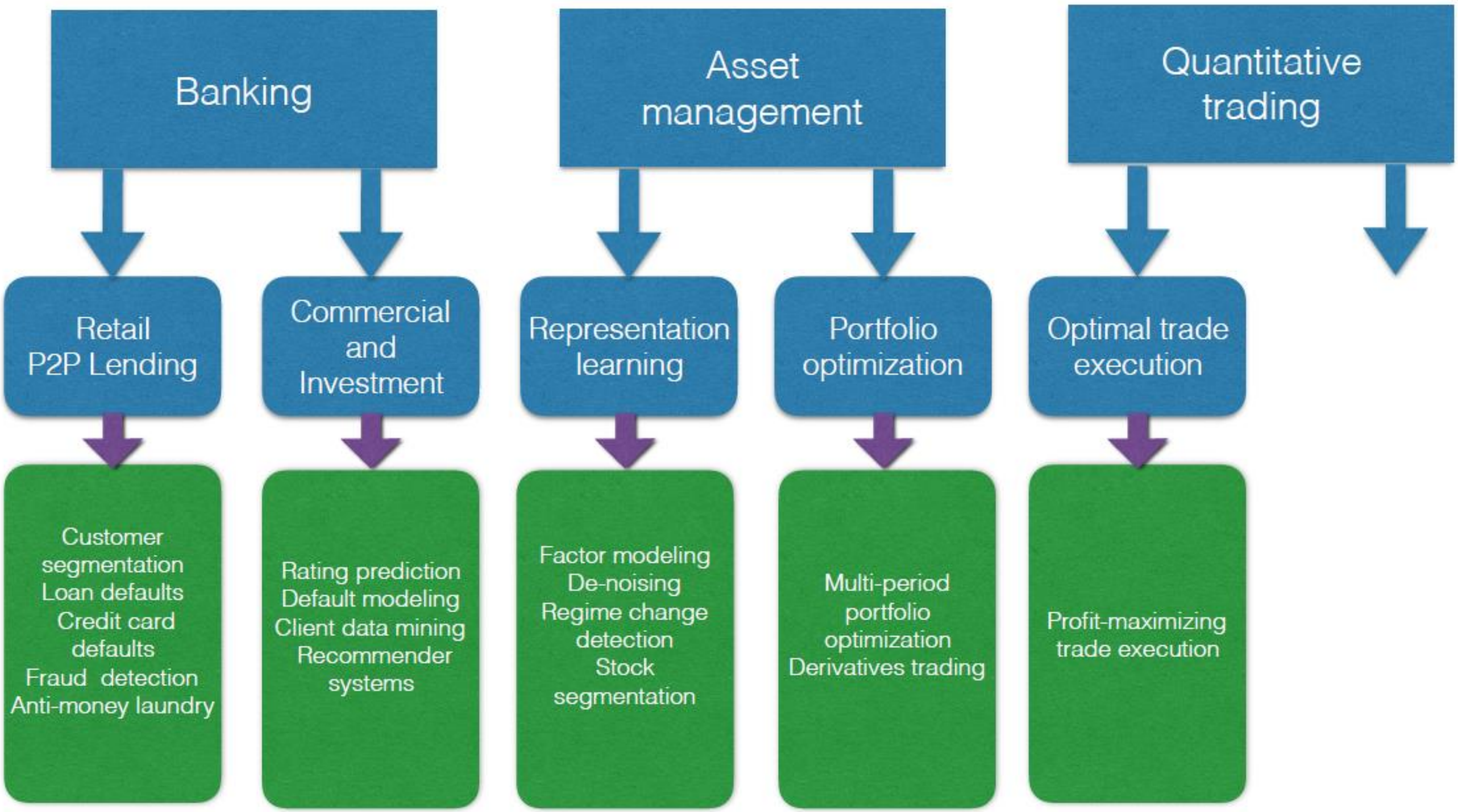
ML by Financial Application Areas



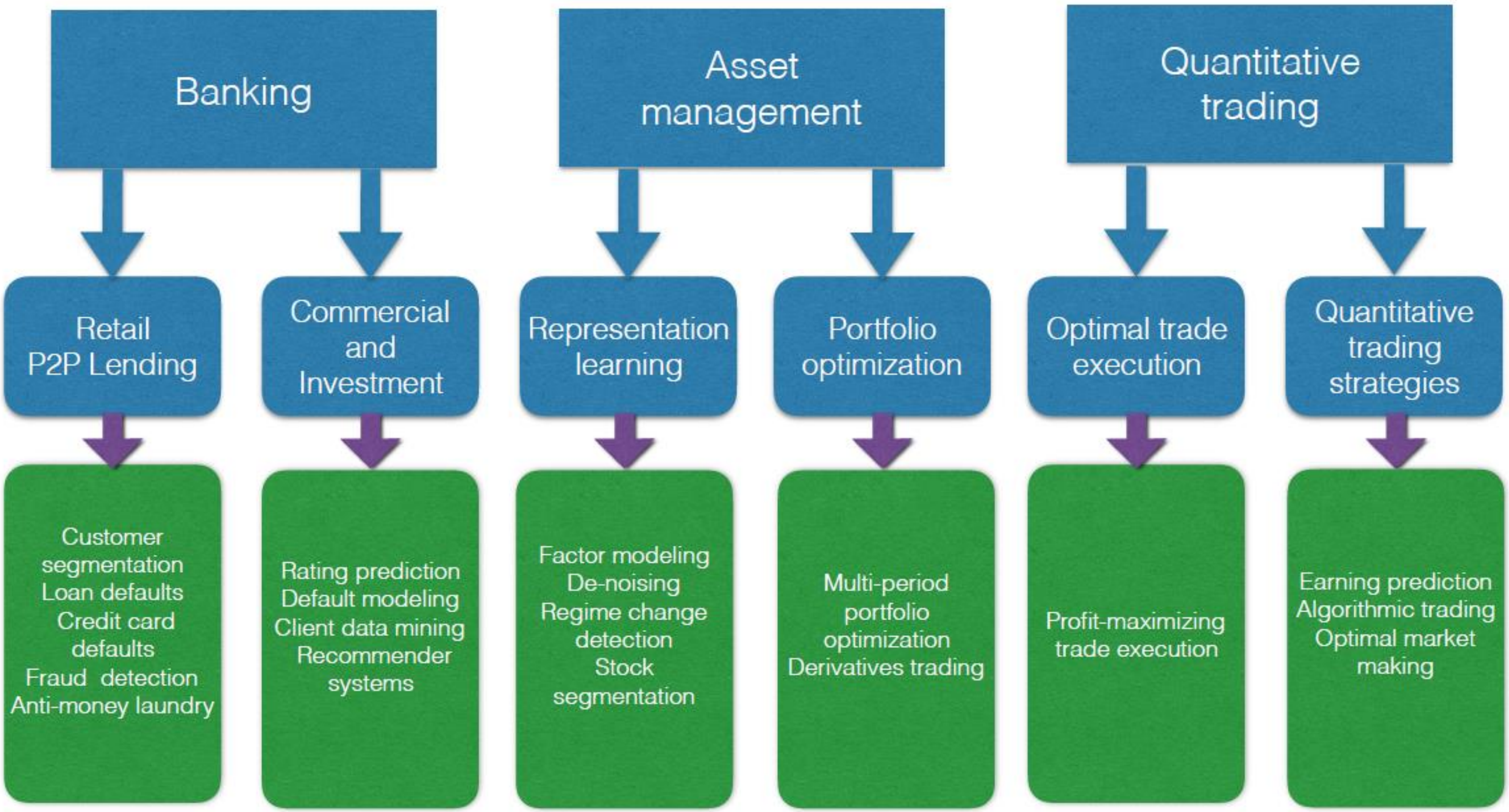
ML by Financial Application Areas



ML by Financial Application Areas

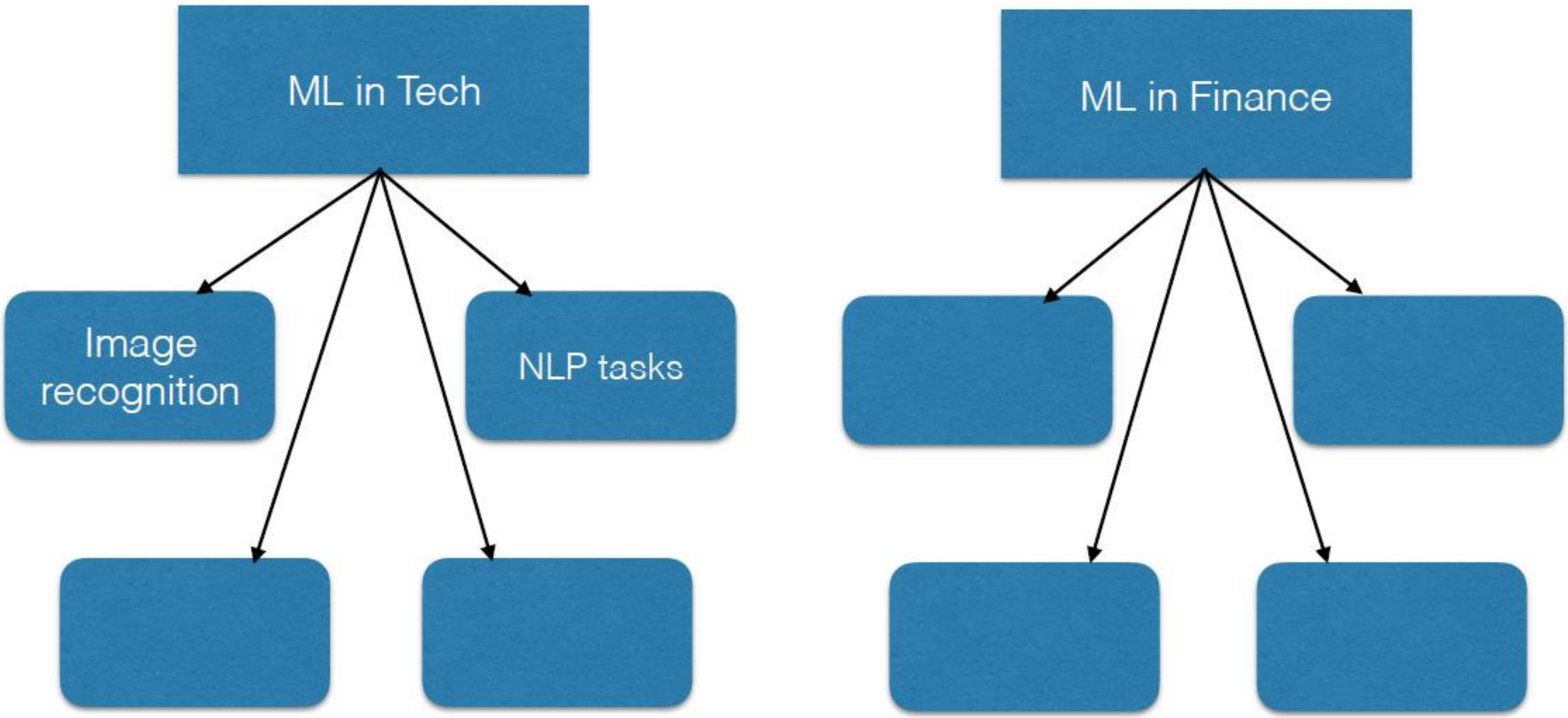


ML by Financial Application Areas



ML in Finance vs ML in Tech

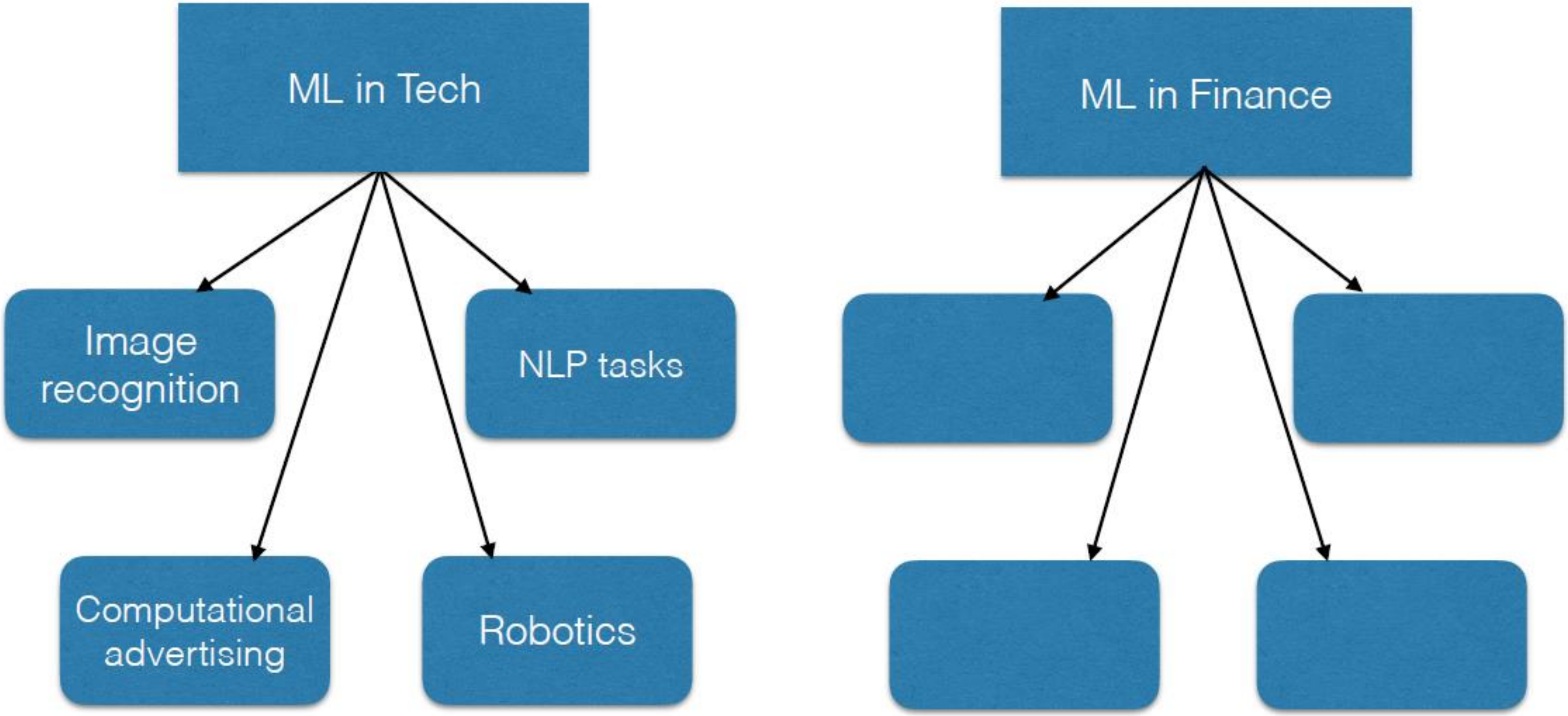
ML in Finance vs ML in Tech



ML in Tech:

- **Perception** (image recognition, NLP tasks, etc.). Methods: SL/UL

ML in Finance vs ML in Tech

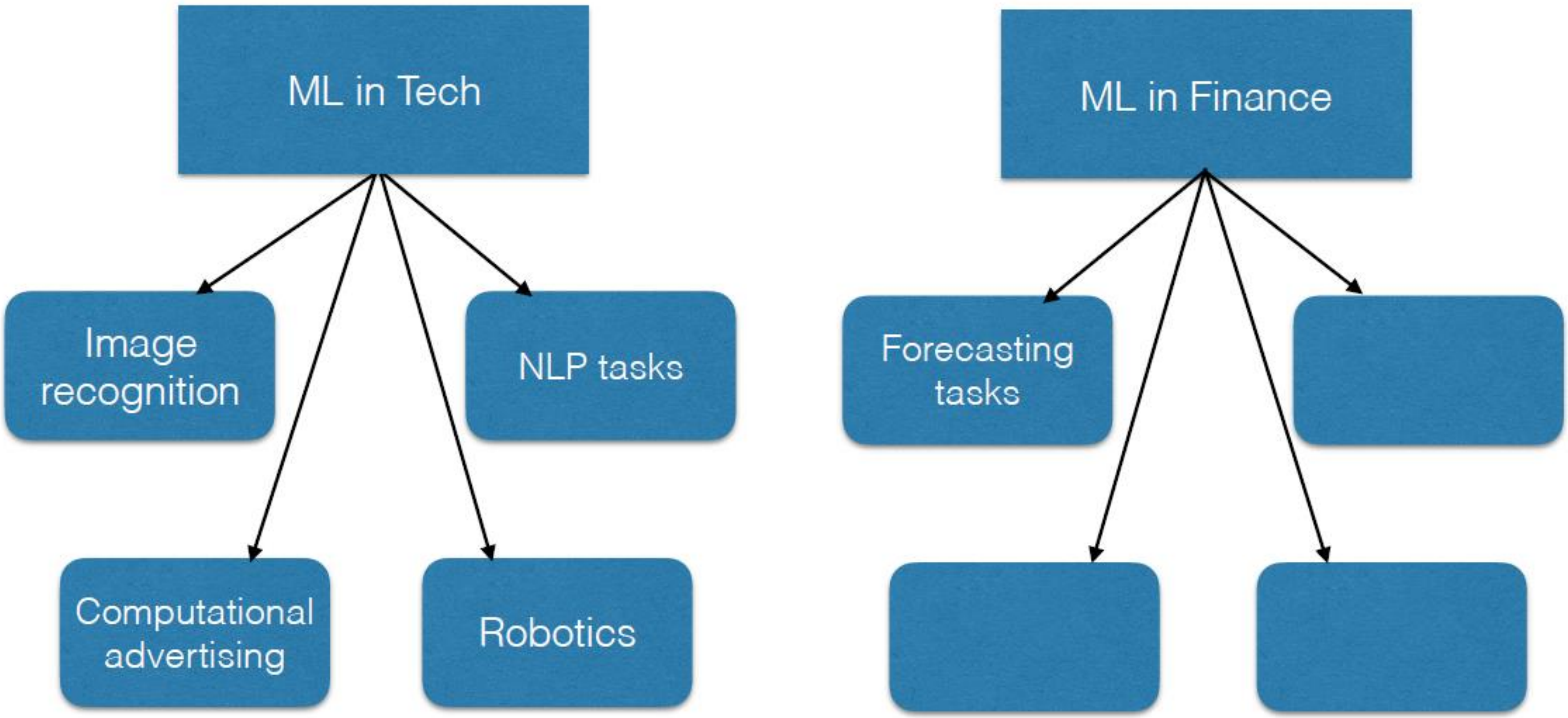


ML in Tech:

- **Perception** (image recognition, NLP tasks, etc.). Methods: SL/UL
- **Action** (computational advertising, robotics, self-driving cars, etc.). Methods: SL/UI/**RL**

What are typical ML tasks in Finance?

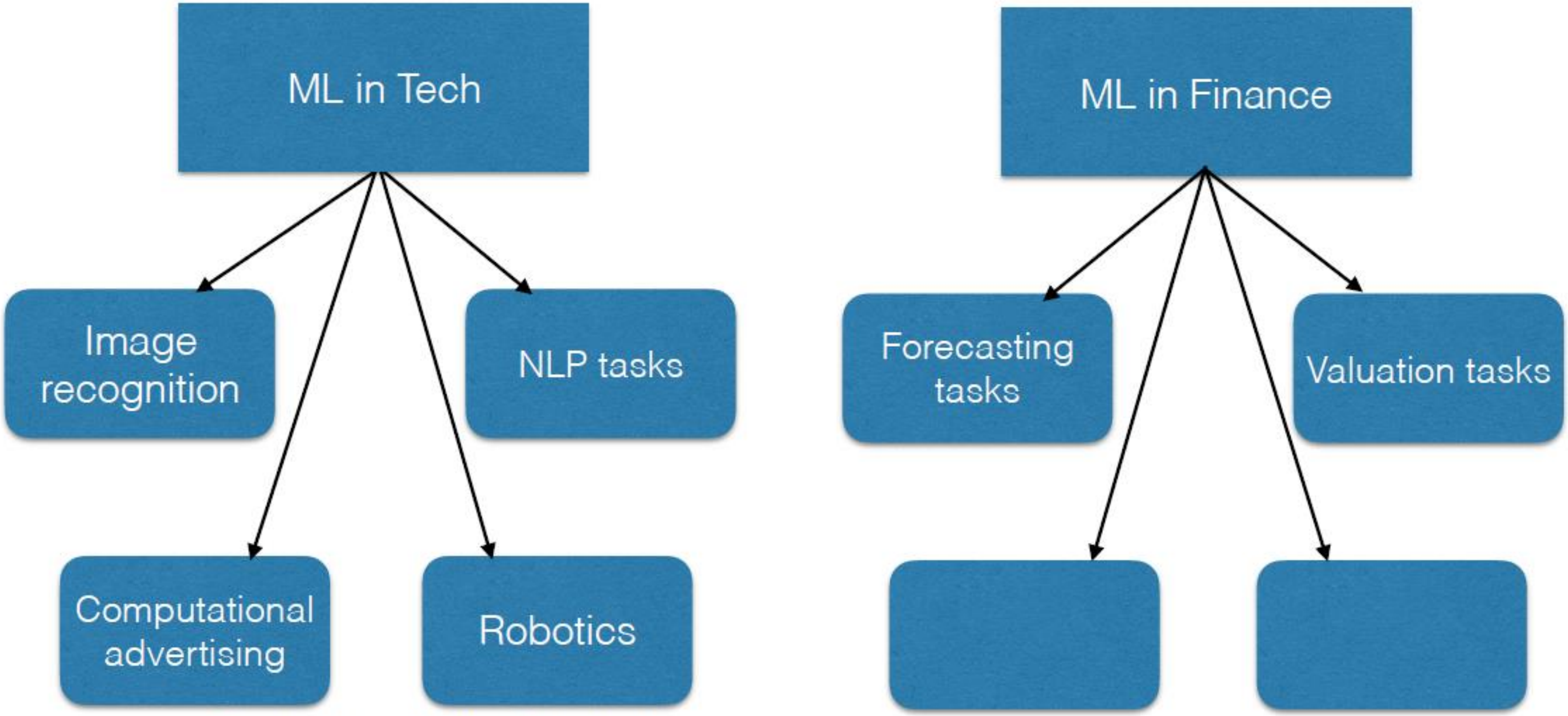
ML in Finance vs ML in Tech



Perception: Forecasting tasks

- Security price predictions (stocks, bonds, commodities etc.). Methods: SL/UL
- Corporate actors action prediction (dividends, mergers, defaults etc.). Methods: SL/UL/**RL**
- Individual actors action prediction (loan defaults, fraud, AML, etc.). Methods: SL/UL/**RL**

ML in Finance vs ML in Tech

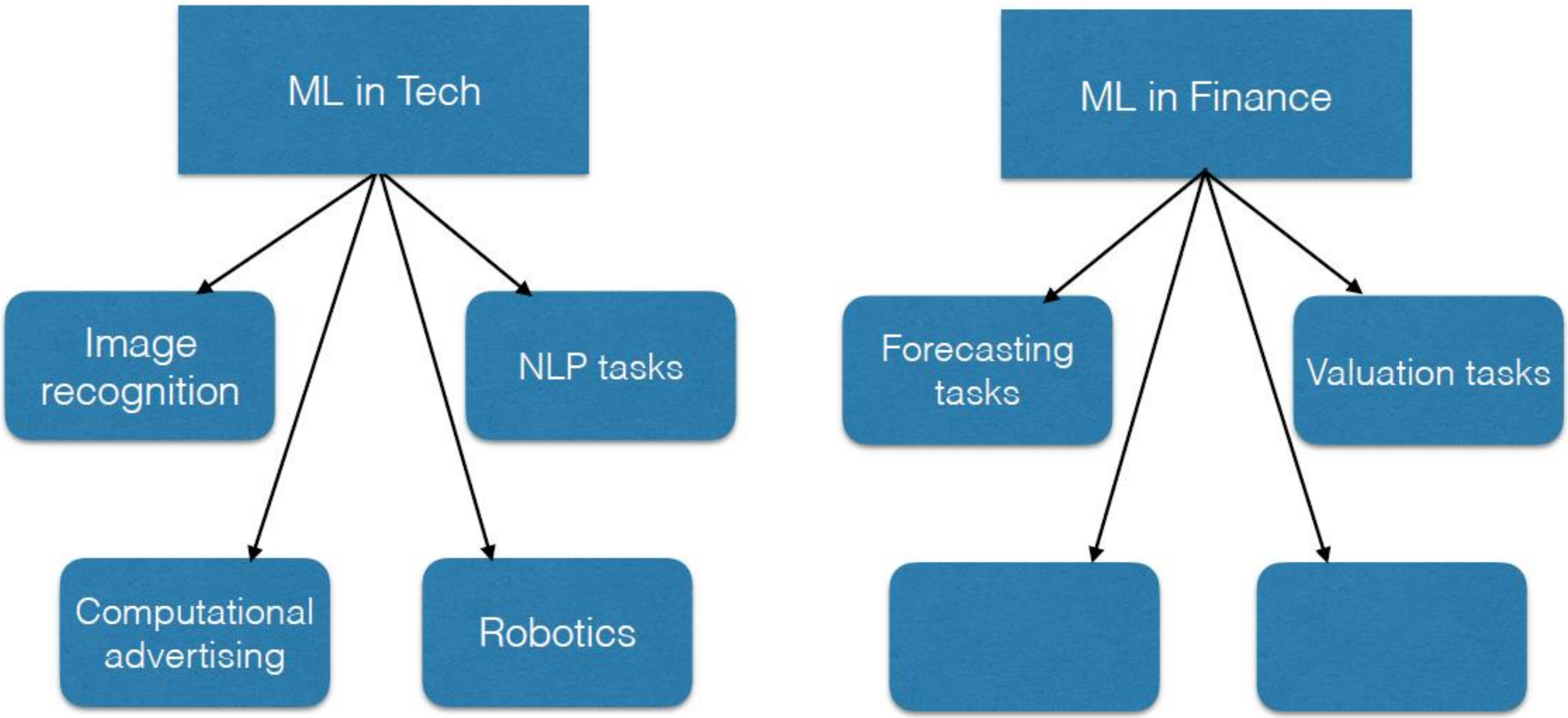


Perception: Valuation tasks

- Asset valuation (stocks, futures, commodities, bonds, etc.). Related to forecasting. Methods: SL/UL
- Derivatives valuation. Methods: SL/UL/**RL**

Question: why can perception tasks in Finance involve **RL**?

ML in Finance vs ML in Tech

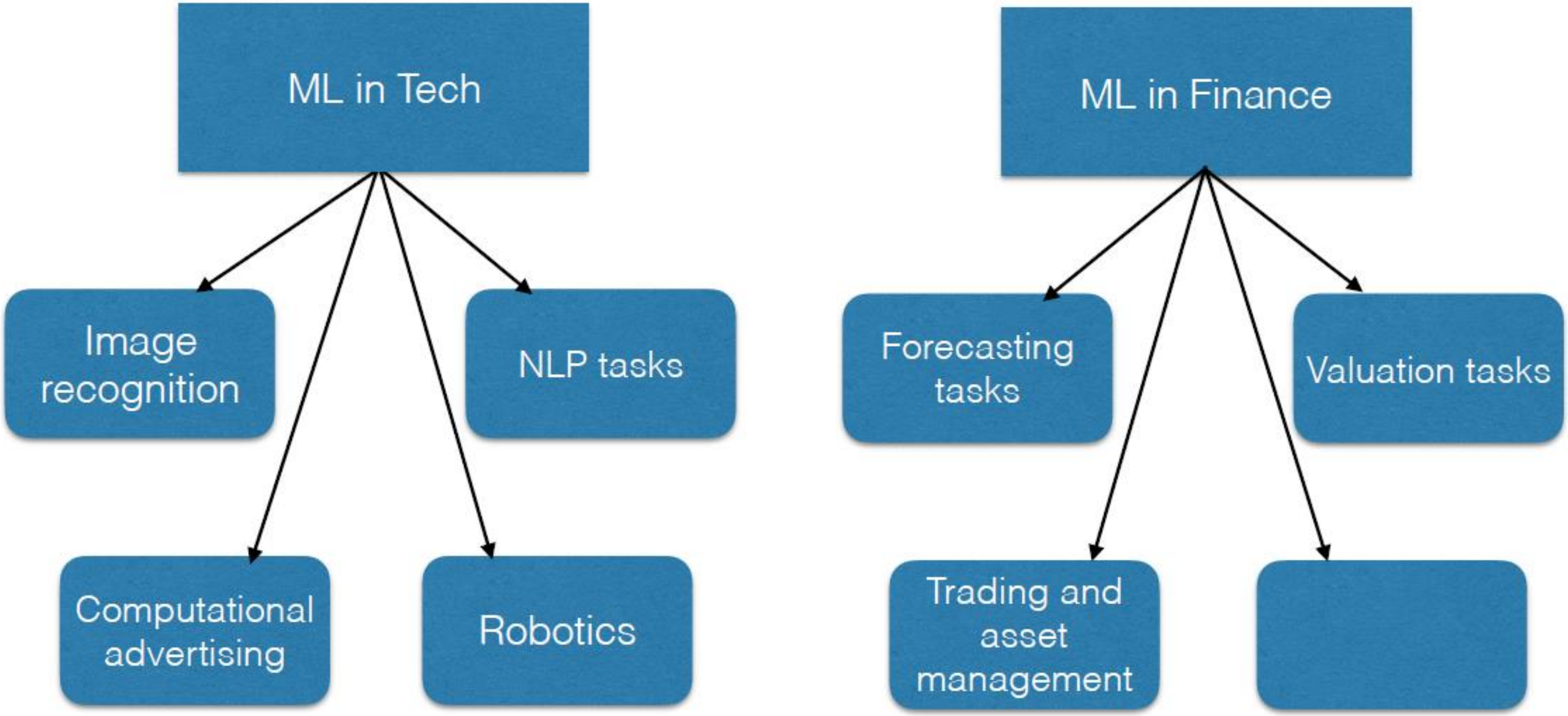


Question: why may perception tasks in Finance involve **RL**?

Answer:

- In Finance, perception tasks are often about the future
- The future is partly driven by future actions of decision makers
- This brings RL into the game even for "perception" tasks in Finance!

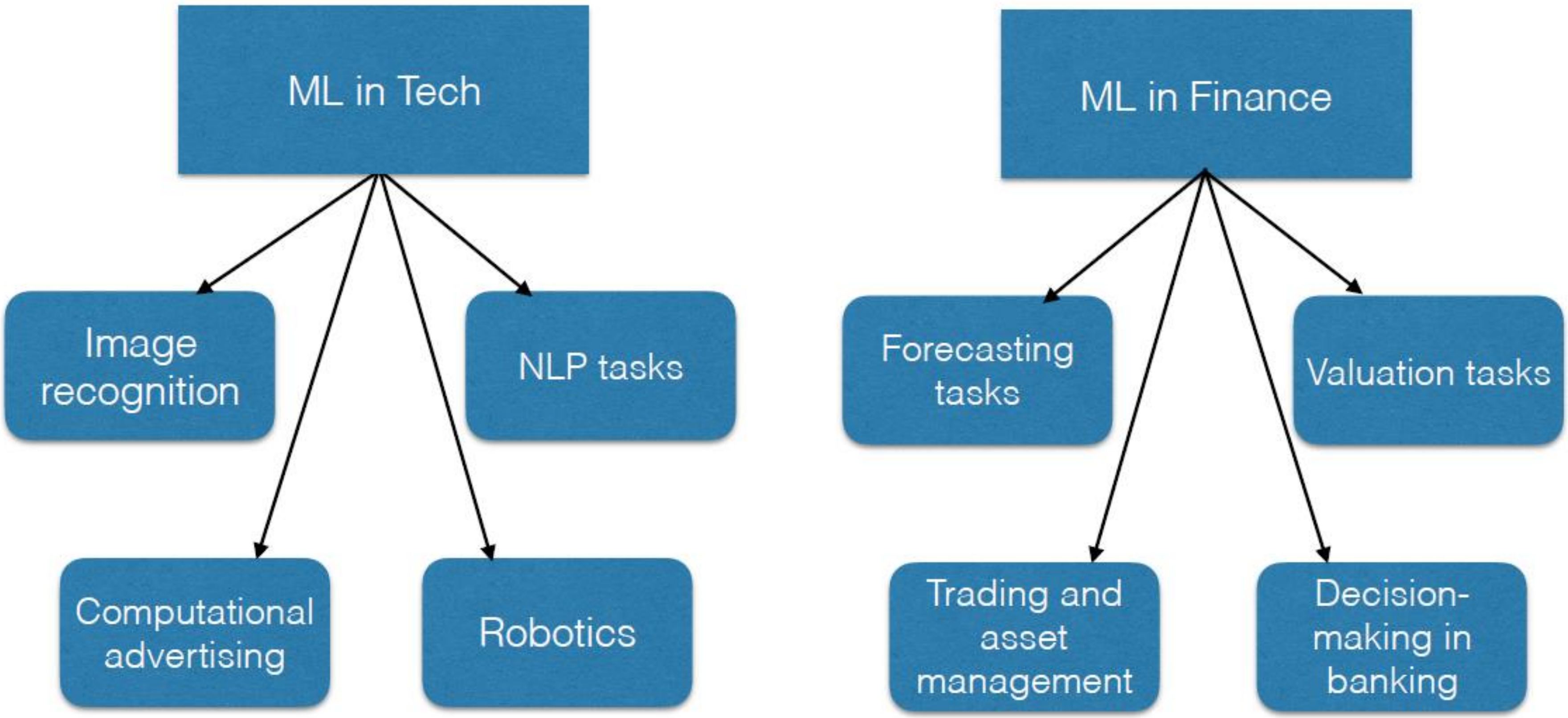
ML in Finance vs ML in Tech



Action: Trading and asset management

- Optimal execution for brokerage trading. Methods: SL/UL/**RL**
- Optimal strategies for day trading. Methods: SL/UL/**RL**
- Active portfolio management. Methods: SL/UL/**RL**

ML in Finance vs ML in Tech



Action: Decision-making in banking

- Loan approvals. Methods: SL/UL/**RL**, Bayesian networks
- Credit and operational risk management. Methods: SL/UL/**RL**, Bayesian networks
- Decision-making in compliance analytics (fraud, AML, etc.). Methods: SL/UL/**RL**, Bayesian networks

ML in Finance and ML in Tech

- We saw that in Finance, perception tasks might involve elements of RL. This is unlike typical perception tasks for ML in Tech

- This happens because some perception tasks in Finance involve **predicting future actions** of rational (or semi-rational) actors, or own actions (like e.g. with American options)

- What about other differences of ML in Finance from ML in Tech?

ML in Finance vs ML in Tech

ML in Finance vs ML in Tech

Tasks	ML in Tech	ML in Finance
Big Data?	typically **yes**	typically **no**

Most of data for ML in finance are medium-size, except HFT

ML in Finance vs ML in Tech

Tasks	ML in Tech	ML in Finance
Big Data?	typically yes	typically no
Stationary data?	typically **yes**	most often **no**

As most of financial data are non-stationary, collecting more data, even when possible, is not always helpful

ML in Finance vs ML in Tech

Tasks	ML in Tech	ML in Finance
Big Data?	typically yes	typically no
Stationary data?	typically yes	most often no
Signal-to-noise ratio	typically **low**	typically **high**

Financial data are typically quite noisy, “true” signals are unobservable!

ML in Finance vs ML in Tech

Tasks	ML in Tech	ML in Finance
Big Data?	typically yes	typically no
Stationary data?	typically yes	most often no
Signal-to-noise ratio	typically low	typically high
Action (RL) tasks	Low dimensional state-action space, low uncertainty	High-dimensional state-action space, high uncertainty

- **ML in Tech: dimensionality of the state-action space is usually in hundreds. The action space is often discrete (except in robotics). Uncertainty is low to moderate (think self-driving cars!)**
- **ML in Finance: dimensionality of the state-action space is often in thousands. The action space is usually continuous. Uncertainty is high (think Brexit!)**

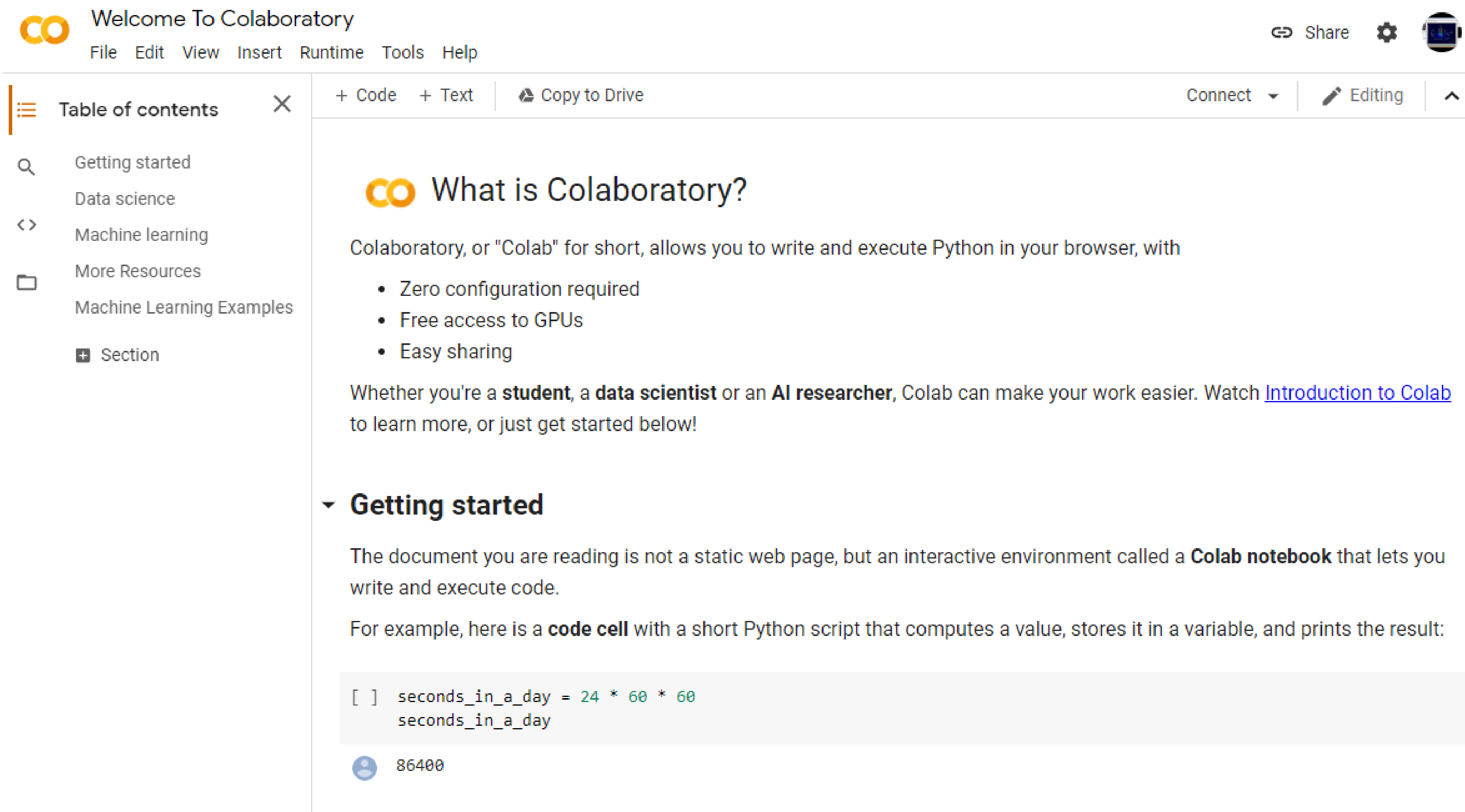
ML in Finance vs ML in Tech

Tasks	ML in Tech	ML in Finance
Big Data?	typically yes	typically no
Stationary data?	typically yes	most often no
Signal-to-noise ratio	typically low	typically high
Action (RL) tasks	Low dimensional state-action space, low uncertainty	High-dimensional state-action space, high uncertainty
Interpretability of results	**typically, not important, or not the main focus**	**Typically, either desired or required**

Interpretability of results is:

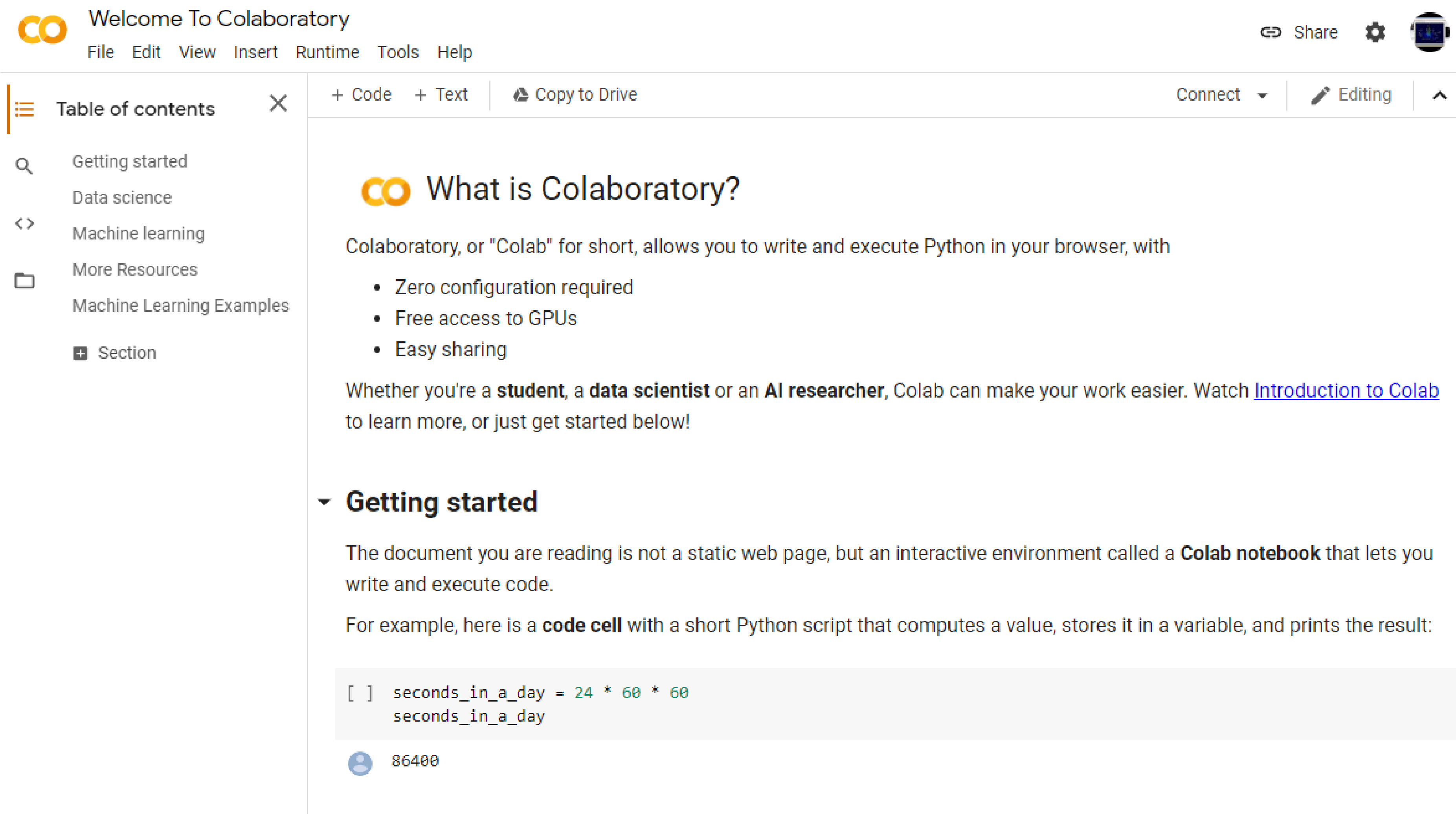
- **Desired for trading**
- **Required for regulation (General Data Protection Regulation, 2018)**

Google colab



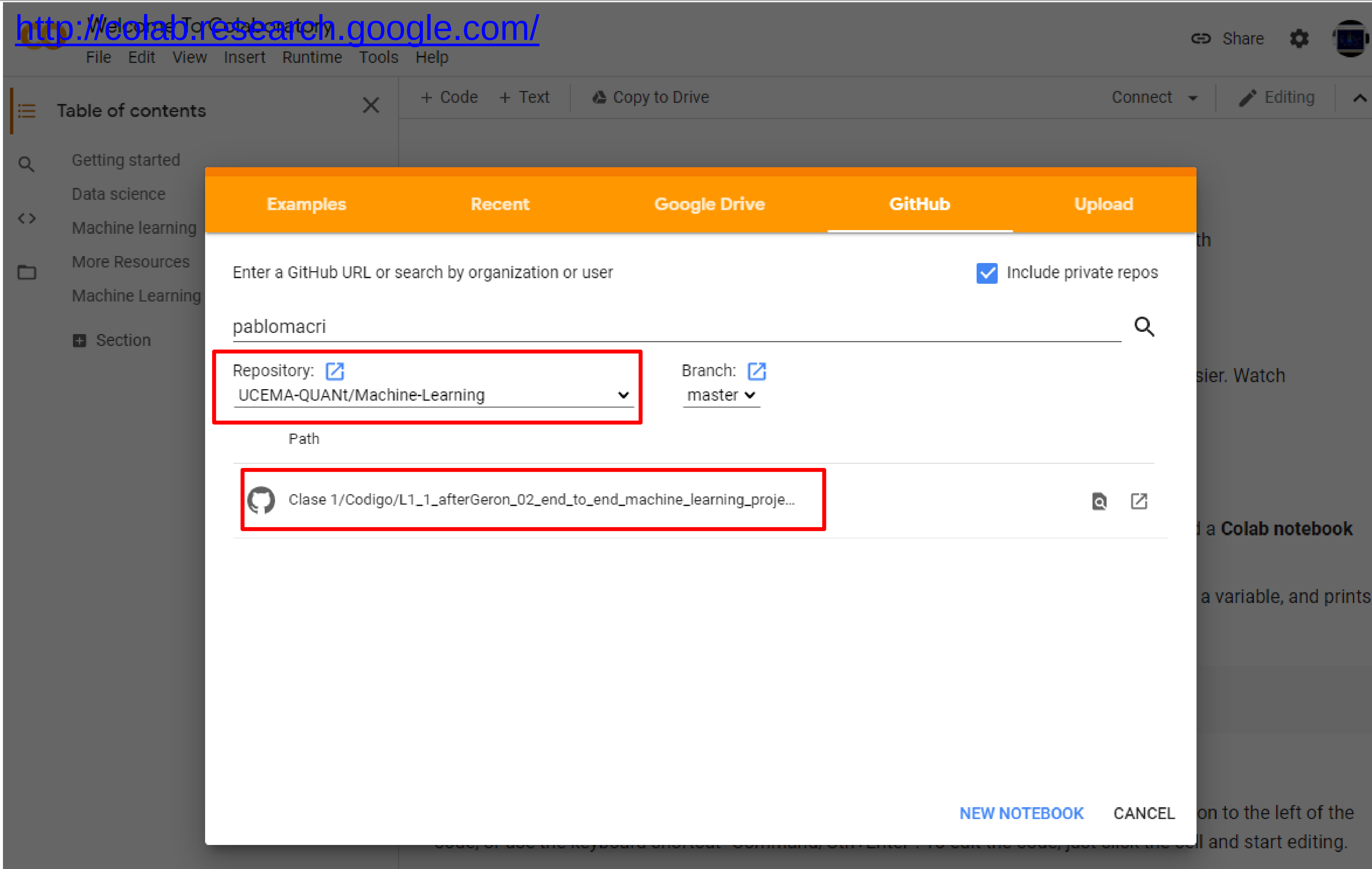
Google colab

http://colab.research.google.com/



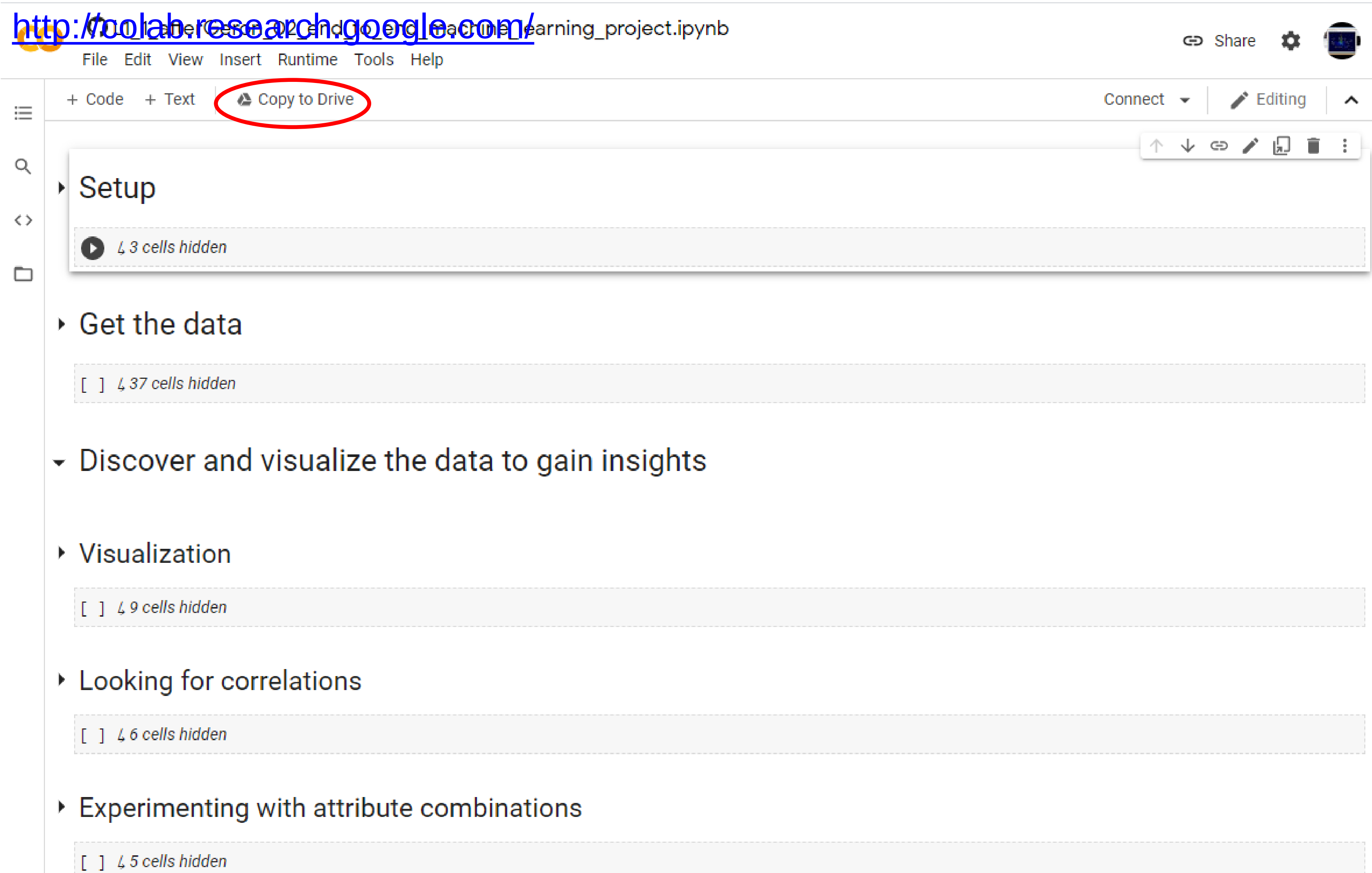
Google colab

http://colab.research.google.com/



Google colab

http://colab.research.google.com/



Demo: End-to-End ML Project

- Introduce scikit-learn as our first ML engine ☑
- Overview of a full ML project:
  - Get the data ☒
  - Look at the data structure ☑
  - Create training and testing data sets ☑
  - Visualize data to gain insights ☒
  - Data cleaning and handling ☒
  - Transformation pipelines ☒
  - Training and evaluating the model ☑
  - Cross validation and hyperparameters ☑
  - Model evaluation and comparison ☑

☒ for the data science course