

Machine Learning

Introduction

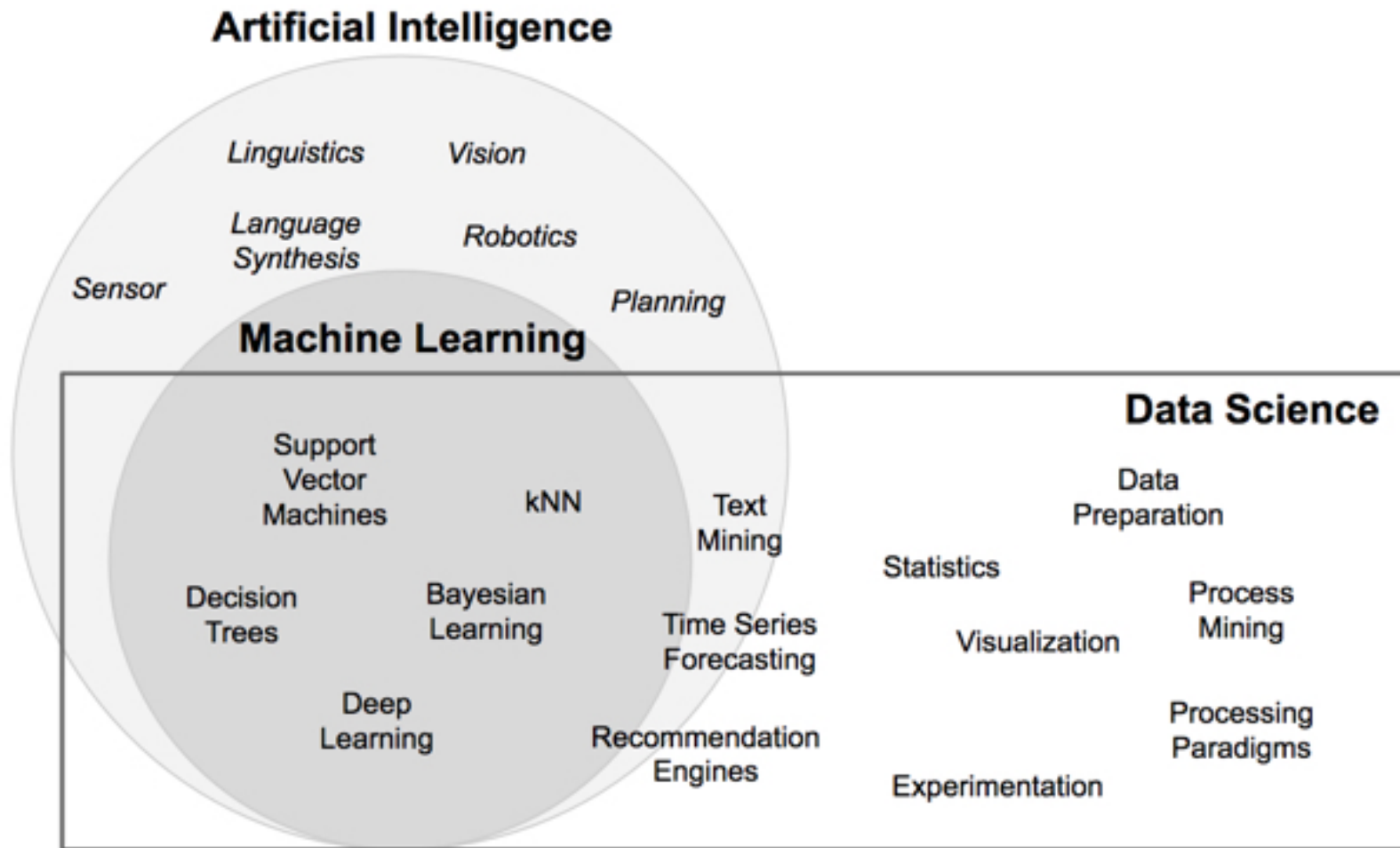
Resources

- The Hundred-Page Machine Learning Book
- Machine Learning – An Algorithmic Perspective
- Machine Learning – A Probabilistic Perspective
- Pattern Recognition and Machine Learning
- Introduction to ML with Python – A Guide for Data Scientists
- Data Science – Concepts and Practice
- Coursera - Machine Learning by Andrew Ng

Data Science

- Data science is also commonly referred to as
 - knowledge discovery
 - machine learning
 - predictive analytics
 - data mining.
- However, each term has a slightly different connotation depending on the context.

AI, Machine Learning, and Data Science



AI

- Artificial intelligence is about giving machines the capability of mimicking human behavior, particularly cognitive functions.
- Examples would be: facial recognition, automated driving, sorting mail based on postal code.
- There are quite a range of techniques that fall under artificial intelligence: linguistics, natural language processing, decision science, bias, vision, robotics, planning, etc.
- Learning is an important part of human capability.

Machine Learning

- Machine learning can either be considered a sub-field or one of the tools of artificial intelligence, is providing machines with the capability of learning from experience.
- Experience for machines comes in the form of data.

Data Science

- Data science is the business application of machine learning, artificial intelligence, and other quantitative fields like statistics, visualization, and mathematics.
- In the context of how data science is used today, it relies heavily on machine learning and is sometimes called data mining.
 - Examples of data science user cases are:
 - recommendation engines that can recommend movies for a particular user,
 - a fraud alert model that detects fraudulent credit card transactions,
 - find customers who will most likely churn next month,
 - predict revenue for the next quarter.

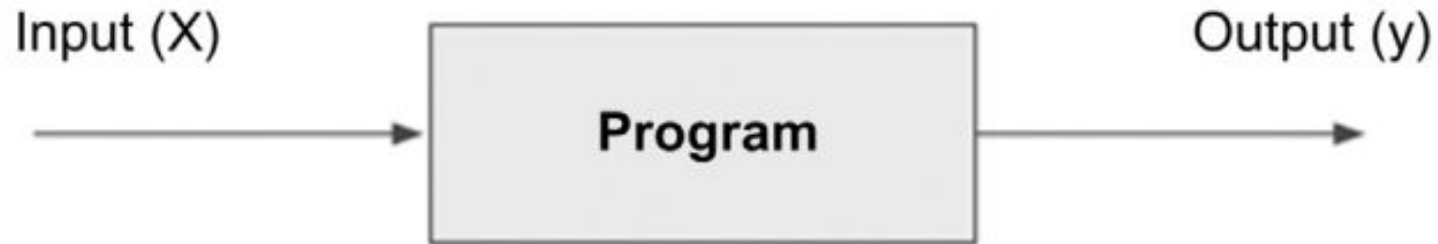
AI, Machine Learning, and Data Science

- Artificial intelligence is about giving machines the capability of mimicking human behavior, particularly cognitive functions. Learning is an important part of human capability.
- Machine learning can either be considered a sub-field or one of the tools of artificial intelligence, is providing machines with the capability of learning from experience. Experience for machines comes in the form of data.
- Data science is the business application of machine learning, artificial intelligence, and other quantitative fields like statistics, visualization, and mathematics.

Machine Learning

- Machine learning is providing machines with the capability of learning from experience. Experience for machines comes in the form of data.
- Data that is used to teach machines is called training data.
- Machine learning turns the traditional programming model upside down.

Machine Learning



Machine Learning

- A program, a set of instructions to a computer, transforms input signals into output signals using predetermined rules and relationships.
- Machine learning algorithms, also called “learners”, take both the known input and output (training data) to figure out a model for the program which converts input to output.
- For example, many organizations like social media platforms, review sites, or forums are required to moderate posts and remove abusive content.
- How can machines be taught to automate the removal of abusive content?

Machine Learning

- The machines need to be shown examples of both abusive and non-abusive posts with a clear indication of which one is abusive.
- The learners will generalize a pattern based on certain words or sequences of words in order to conclude whether the overall post is abusive or not. The model can take the form of a set of “if-then” rules.
- Once the data science rules or model is developed, machines can start categorizing the disposition of any new posts.

Machine Learning Types

- Machine learning is usually divided into two main types:
 - **Supervised learning**
 - **Unsupervised learning**
- There is a third type of machine learning, known as **Reinforcement learning**

Supervised Learning

- A *training set* of examples with the correct responses (*targets*) is provided and, based on this training set, the algorithm *generalises* to respond correctly to all possible inputs.
- This is also called learning from *exemplars*.

Unsupervised Learning

- Correct responses are not provided, but instead the algorithm tries to identify similarities between the inputs so that inputs that have something in common are *categorised* together.
- The statistical approach to unsupervised learning is known as *density estimation*.

Reinforcement Learning

- This is somewhere between supervised and unsupervised learning.
- The algorithm gets told when the answer is wrong, but does not get told how to correct it.
- It has to explore and try out different possibilities until it works out how to get the answer right.
- Reinforcement learning is sometime called learning with a *critic* because of this monitor that scores the answer, but does not suggest improvements.

Supervised Learning Applications

- **Classification**

- Document classification and email spam filtering
- Classifying flowers
- Image classification and handwriting recognition
- Face detection and recognition

Supervised Learning Applications

- **Regression**

- Predict tomorrow's stock market price given current market conditions and other possible side information.
- Predict the age of a viewer watching a given video on YouTube.
- Predict the location in 3d space of a robot arm end effector, given control signals (torques) sent to its various motors.
- Predict the amount of prostate specific antigen (PSA) in the body as a function of a number of different clinical measurements.
- Predict the temperature at any location inside a building using weather data, time, door sensors, etc.

Unsupervised Learning Applications

- Discovering clusters
- Discovering latent factors – dimensionality reduction
- Discovering graph structure
- Matrix completion
- Image inpainting
- Collaborative filtering
- Market basket analysis