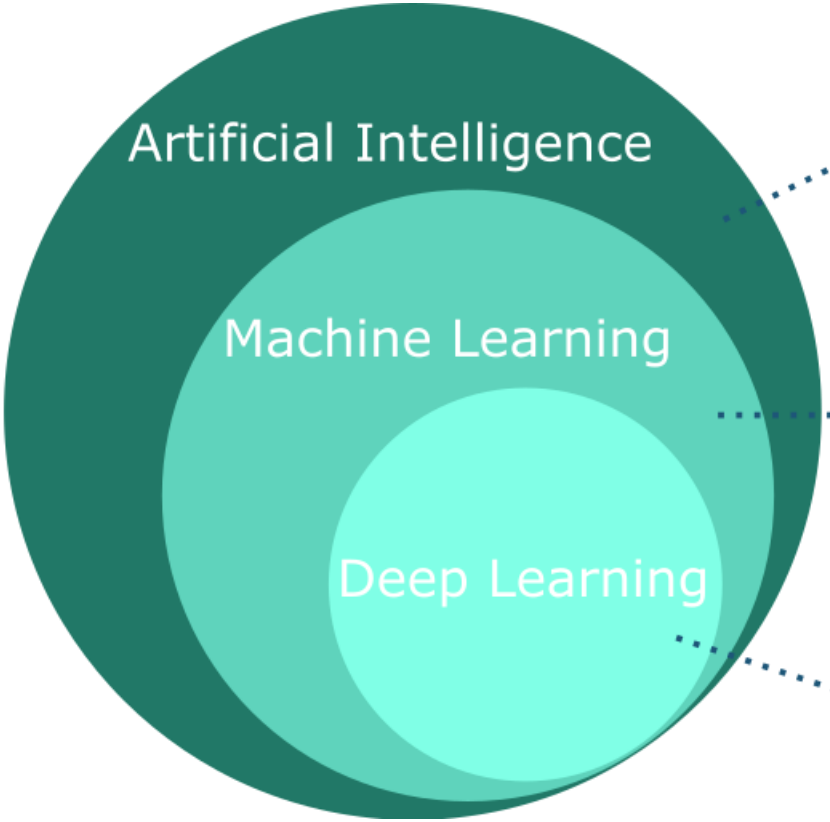
Introduction to Machine Learning

UNIVERSITY OF ARKANSAS

Machine Learning vs Deep Learning



A technique which enables machines to mimic human behaviour

Subset of AI technique which use statistical modeling to enable machines on real problems

Subset of ML which make the computation of multi-layer neural network feasible (GPU)

Fundamental Concepts of Machine Learning

- Task: make your favorite cheeseburger
- Method 1: Search for the ingredients online, or go to McDonald's
- Method 2: Eat enough burgers until you know how to do it

Traditional algorithm vs ML

Traditional programming

Inputs

Rules

* Make bread
* Ground beef
* Cut lettuce
* Cut tomato
* ...

Outputs

Machine learning algorithm

Inputs

Outputs

Rules

* Make bread
* Ground beef
* Cut lettuce
* Cut tomato
* ...

UNIVERSITY OF ARKANSAS

Why use ML/DL?

- For a complex problem, can you think of all the rules?
- ML is used to find patterns of anything as long as you can convert it into numbers.
- Question: Should you always use ML?
- “Don’t be afraid to launch a product without machine learning” – rule #1 of Google’s Machine Learning Handbook

UNIVERSITY OF ARKANSAS

What ML/DL is good for?

- Tackling complex problem with elusive rules
- Adapting to continuously changing environments
- Analyzing large data collections

What ML/DL is not good for?

- Requirement of explainability
- Superior performance of traditional methods
- Low tolerance for errors
- Scarcity of data

UNIVERSITY OF ARKANSAS

Machine Learning vs Deep Learning

- Effective with smaller datasets
- Can often be run on low-end systems
- Interpretable
- Structured data

- Requires large datasets
- More powerful hardware
- Less interpretable
- Unstructured data

Machine Learning vs Deep Learning

- Regression
- Decision Tree
- Naïve Bayes
- Nearest Neighbor
- Support Vector Machine
- …

- Neural Networks
- Feed Forward NN
- Recurrent NN
- Convolutional NN
- LSTM
- Transformer
- Graph NN
- …

Structured data vs Unstructured Data

- Structured Data: This type of data is highly organized and formatted in a way that makes it easily searchable and identifiable in datasets. Examples include names, ages, dates, credit score, stock information, etc.
- Unstructured Data: Unstructured data lacks a predefined format or organization, making it more complex to process and analyze using conventional database techniques. Examples include images, videos, text, genome sequence, etc.

Structured data vs Unstructured data

	sepal length (cm)	sepal width (cm)
2	4.8	3.1
3	5.2	3.5
4	4.6	3.2
5	4.5	2.3
6	5.5	2.6
7	5.5	4.2
8	6.7	3
9	5.7	2.6
10	5	3.5
11	5.4	3.4
12	4.9	3.1
13	4.8	3.4
14	6.3	3.4

	mpg	cylinders	displacemen	horsepower	weight	acceleration	year	origin	name
2	18	8	307	130	3504	12	70	1	chevrolet chevelle malibu
3	15	8	350	165	3693	11.5	70	1	buick skylark 320
4	18	8	318	150	3436	11	70	1	plymouth satellite
5	16	8	304	150	3433	12	70	1	amc rebel sst
6	17	8	302	140	3449	10.5	70	1	ford torino
7	15	8	429	198	4341	10	70	1	ford galaxie 500
8	14	8	454	220	4354	9	70	1	chevrolet impala
9	14	8	440	215	4312	8.5	70	1	plymouth fury iii
10	14	8	455	225	4425	10	70	1	pontiac catalina
11	15	8	390	190	3850	8.5	70	1	amc ambassador dpl
12	15	8	383	170	3563	10	70	1	dodge challenger se
13	14	8	340	160	3609	8	70	1	plymouth 'cuda 340
14	15	8	400	150	3761	9.5	70	1	chevrolet monte carlo

MNIS

Similarly, we supported the Corporate Social Responsibility strategy, which is being aligned with applicable UN Sustainable Development Goals. Our environmental and social goals include the philanthropic work of the Airbus Foundation and efforts to promote diversity at all levels.

The Board proposed a 2016 dividend of €1.35 per share. We intend to honour our commitment of increasing dividend per share on a sustainable basis by proposing this payment, which is about four percent higher than in 2015. The value is outside the range of the dividend policy exceptionally. It is based on our 2016 underlying performance and it demonstrates our confidence in our future operational cash generation.
In terms of governance, we introduced 'staggered' Board terms with one third of the Directors being reappointed or replaced every year. The extension of the mandates of three Directors and the nomination of one new Director at the 2017 Annual General Meeting follow this principle.

A-Z

UNIVERSITY OF ARKANSAS

Types of Learning

- Supervised Learning
- Unsupervised & Self-supervised Learning
- Reinforcement Learning
- Transfer Learning
- Semi-supervised, Deep, Ensemble, Federal, etc.

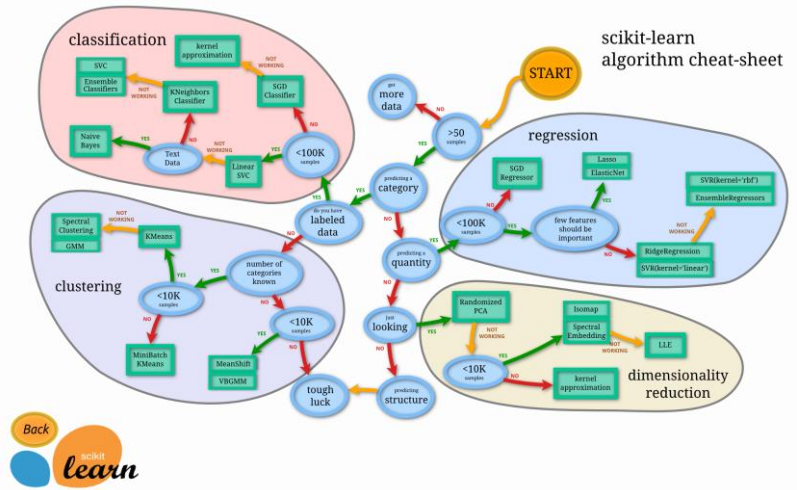
UNIVERSITY OF ARKANSAS

Types of Learning Tasks

- Classification
- Regression
- Clustering
- Dimensionality reduction
- Anomaly rule learning
- Reinforcement learning
- Natural language processing
- Recommendation systems
- Time series forecasting

Software

- Python, Matlab, c++
- Scikit-learn
- Tensorflow, Keras
- Pytorch

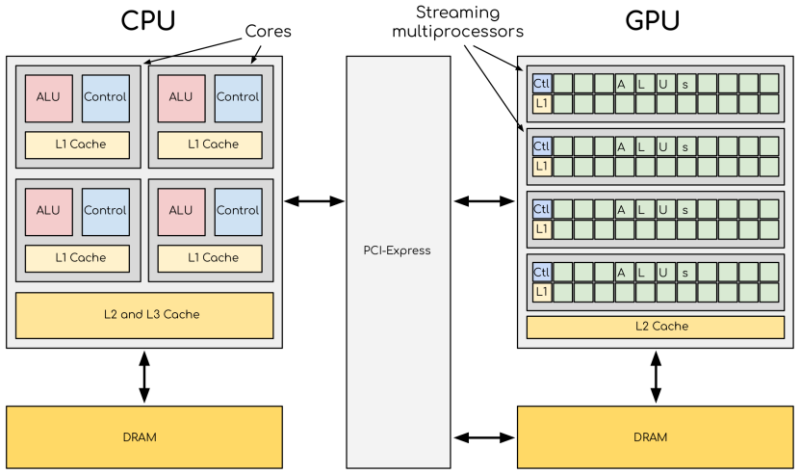


PyTorch

TensorFlow

Hardware

- CPU, GPU, TPU
- Cloud server: Azure, AWS, Colab
- AHPCC



Outline of this course

- Data preprocessing
- Gradient descent
- Linear methods for regression
- Linear methods for classification
- Dimensionality reduction
- Clustering
- Support Vector Machine
- Decision tree

(Beyond)

- Neural Network
- Convolutional Neural Network
- Recurrent Neural Network
- Hopfield, Boltzmann Machine
- Long Short-Term Memory
- Transformer
- Variational autoencoder
- Stable Diffusion
- GAN, ResNet, …

UNIVERSITY OF ARKANSAS

Transformer, GPT, Prompt Engineering

- ChatGPT
- Prompt
- Longchain
- BERT
- Hugging Face
- Auto agent
- Github copilot
- NLP & LLM
- Zero-shot learning, one-shot learning, etc.