Introduction to Machine Learning

Hands-On Crash Course!

Shahrokh Shahi
PhD in Computer Science, CSE

Georgia Institute of Technology
2023

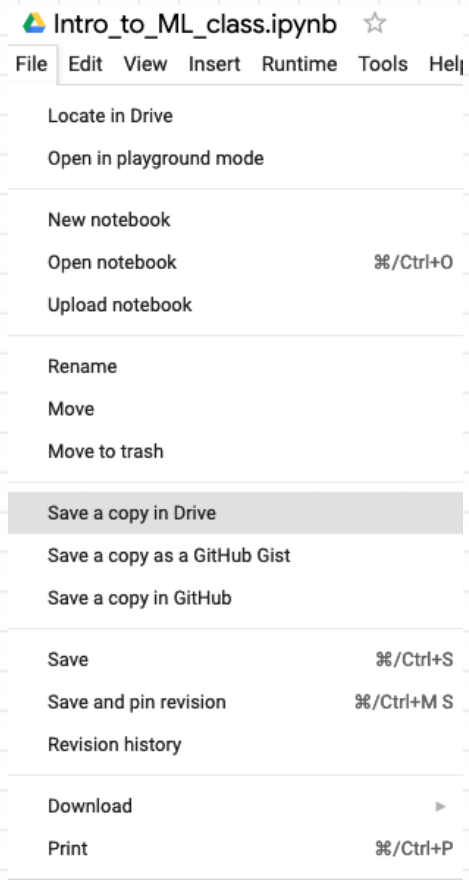
Overview

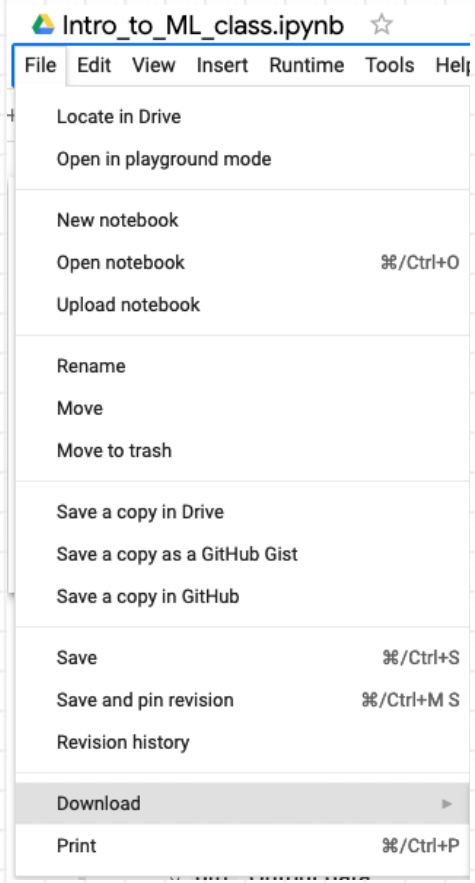
- Introduction to machine-learning techniques
  - Definitions
  - Applications
- Hands-on experience
  - Google colaboratory notebook (Colab)
  - Examples of “supervised” learning
  - Python packages
    - NumPy, matplotlib
    - Pandas
    - Scikit-learn
    - TensorFlow, keras
- Further reading/practice recommendations

Class Colab Notebook

- Link to the class notebook

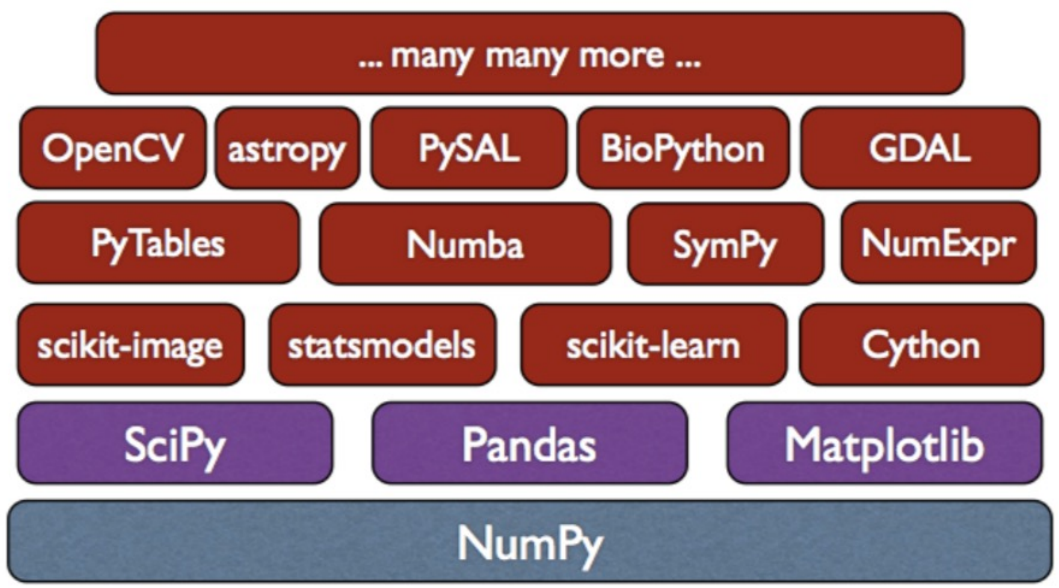
- You can clone this notebook by
  1. Save a copy on your Google Drive or,
  2. Download as a local Jupyter notebook





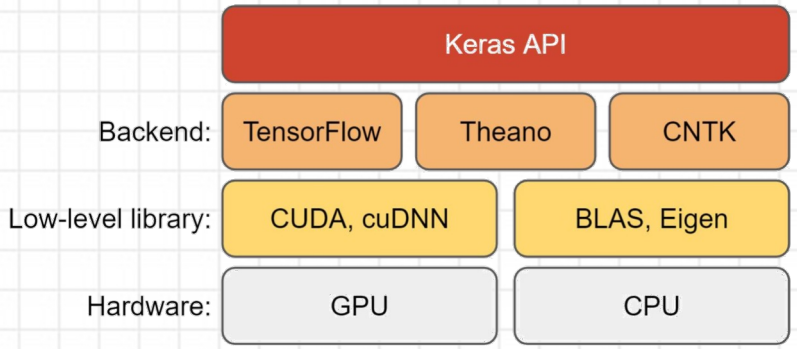
Machine Learning Packages

Well-known Python numerical libraries



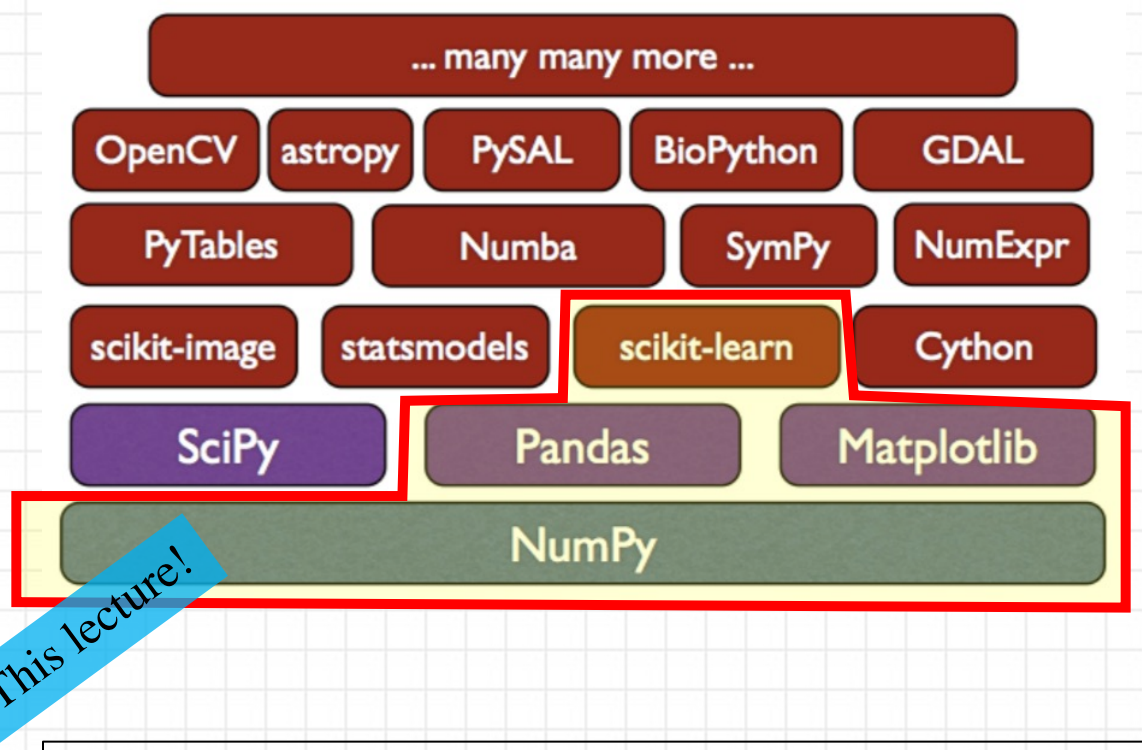
http://quasiben.github.io/PyDataEMC/#/9

Building deep-learning models

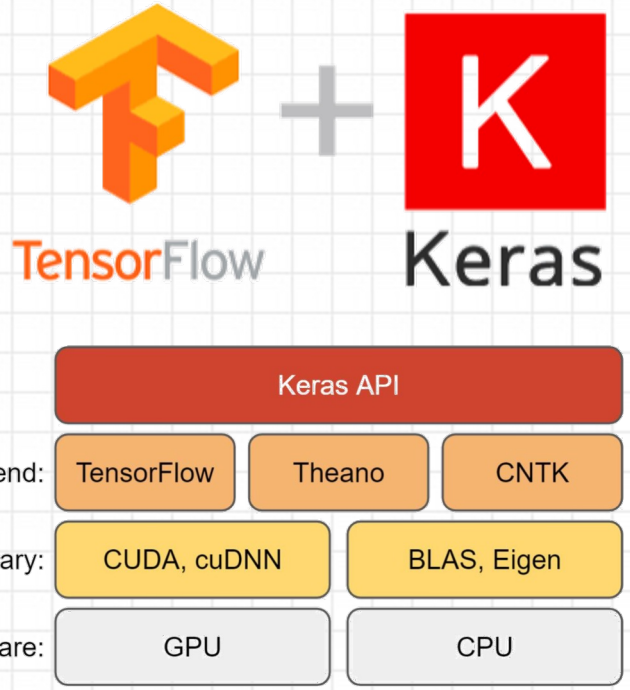


Machine Learning Packages

What we use in this lecture for general machine learning and data science tasks



Building deep-learning models

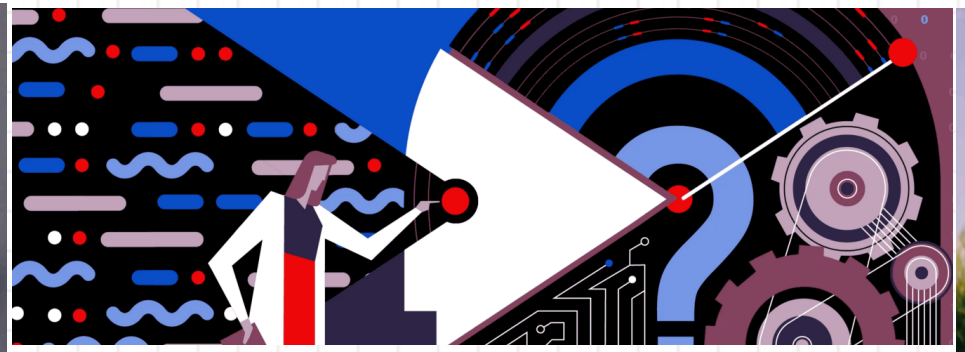


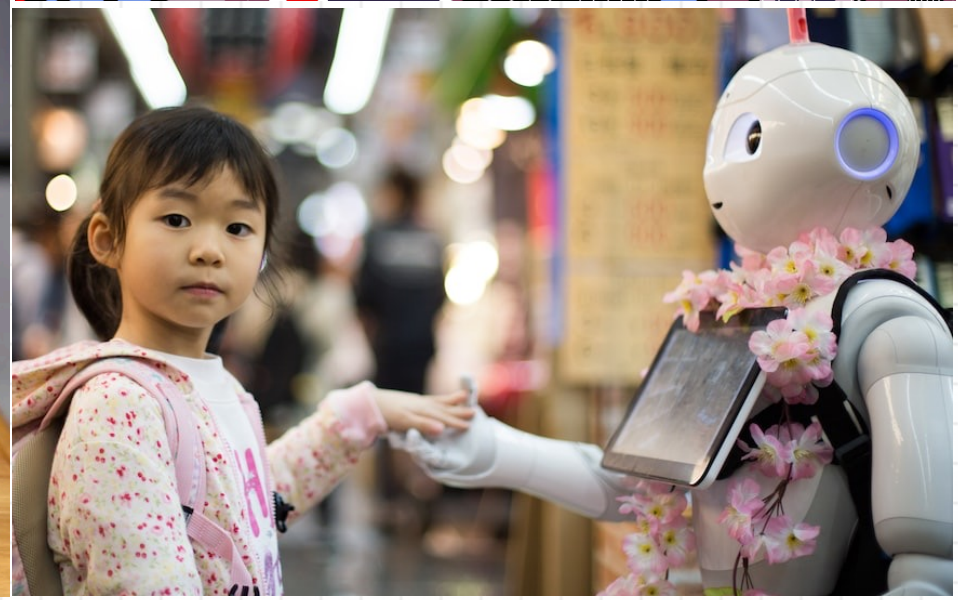
Prerequisites

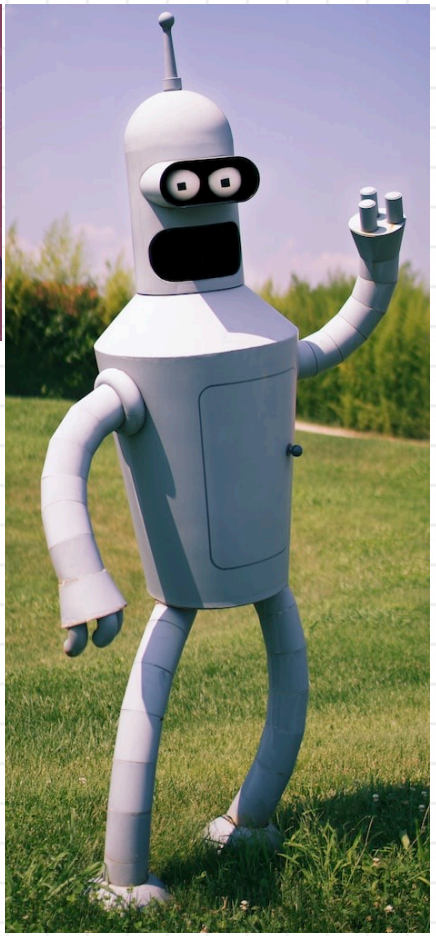
- Basic Python
  - Variables, control (if, for, while), basic data structures (list, dictionary), …
  - Some knowledge of NumPy and matplotlib would be helpful

- Enthusiasm for fun!

What is Machine Learning?







Terminology and Buzzwords

- **Artificial Intelligence (AI):** Simulation of human intelligence in machines to perform tasks that would naturally need human cognitive function, e.g., learning and perception, decision-making.

- **Machine Learning (ML):** A branch of AI that focuses on allowing computer systems to learn and improve over time.

- **Deep Learning (DL):** A subset of ML approaches based on artificial neural networks (ANN)

Terminology and Buzzwords

- **Artificial Intelligence (AI):** Simulation of human intelligence in machines to perform tasks that would naturally need human cognitive function, e.g., learning and perception, decision-making.

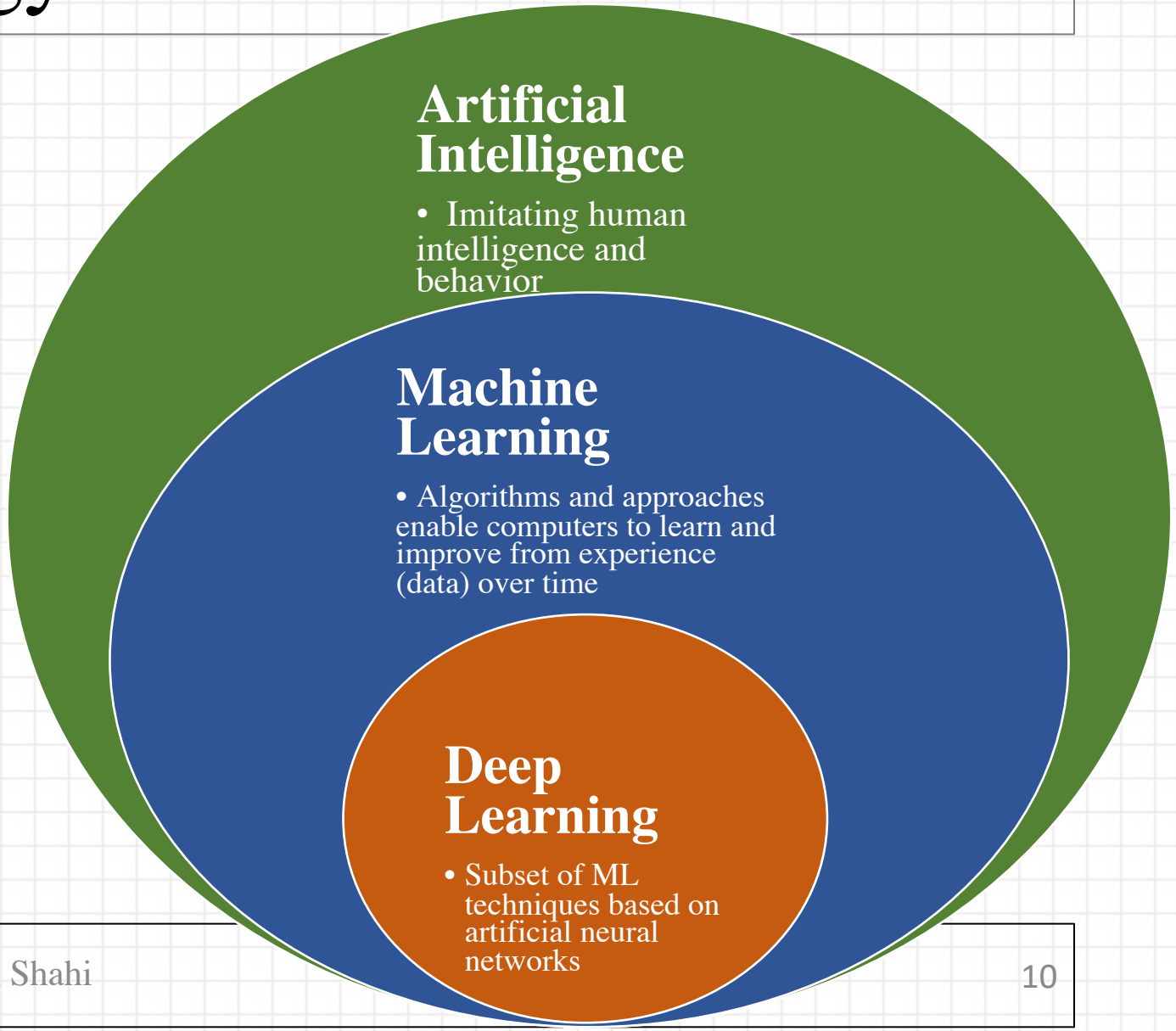
- **Machine Learning (ML):** A branch of AI that focuses on allowing computer systems to learn and improve over time.

- **Deep Learning (DL):** A subset of ML approaches based on artificial neural networks (ANN)

- How these topics are related?

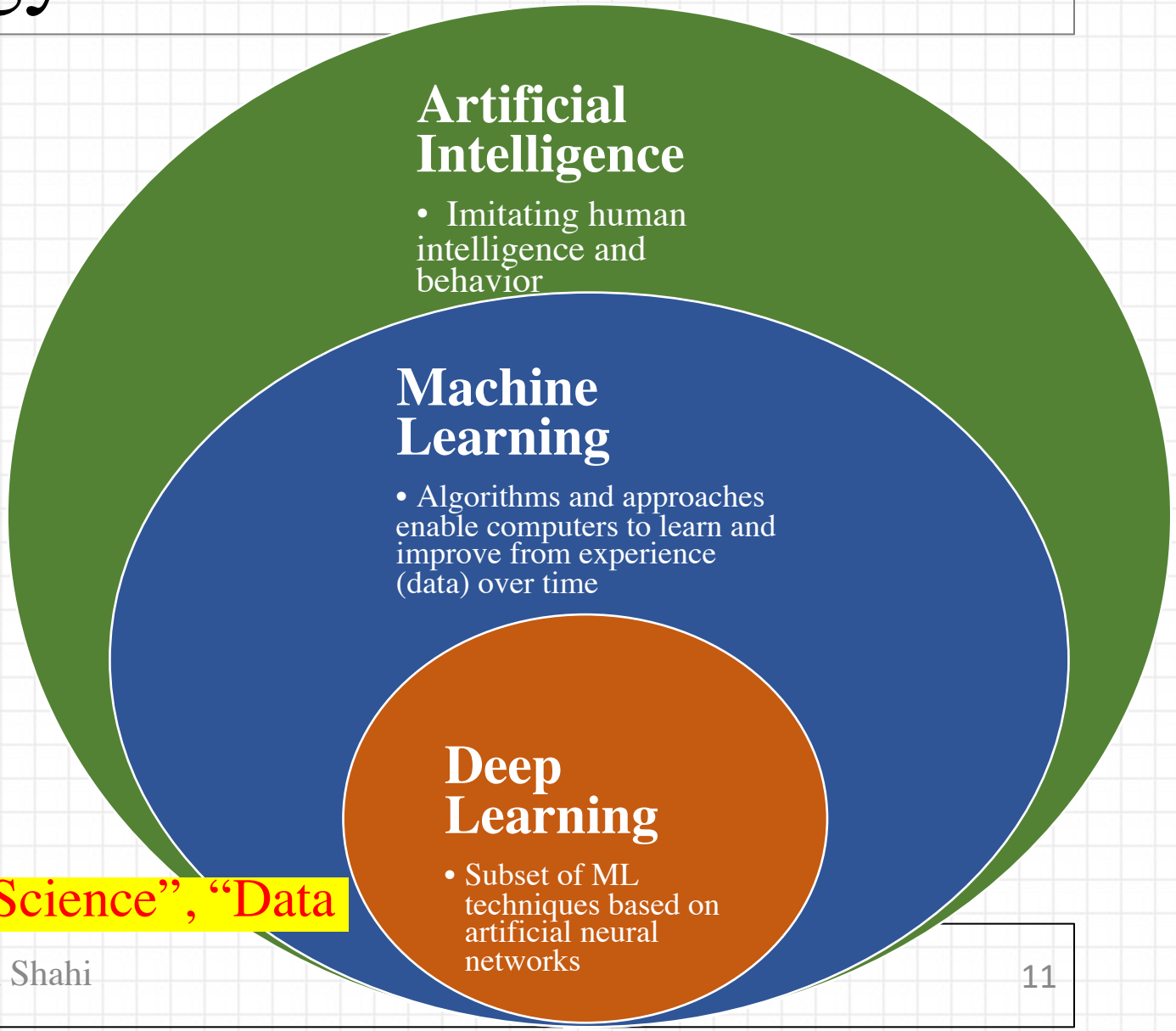
Terminology and Buzzwords

- **Artificial Intelligence (AI):** Simulation of human intelligence in machines to perform tasks that would naturally need human cognitive function, e.g., learning and perception, decision-making.
- **Machine Learning (ML):** A branch of AI that focuses on allowing computer systems to learn and improve over time.
- **Deep Learning (DL):** A subset of ML approaches based on artificial neural networks (ANN)



Terminology and Buzzwords

- **Artificial Intelligence (AI):** Simulation of human intelligence in machines to perform tasks that would naturally need human cognitive function, e.g., learning and perception, decision-making.
- **Machine Learning (ML):** A branch of AI that focuses on allowing computer systems to learn and improve over time.
- **Deep Learning (DL):** A subset of ML approaches based on artificial neural networks (ANN)



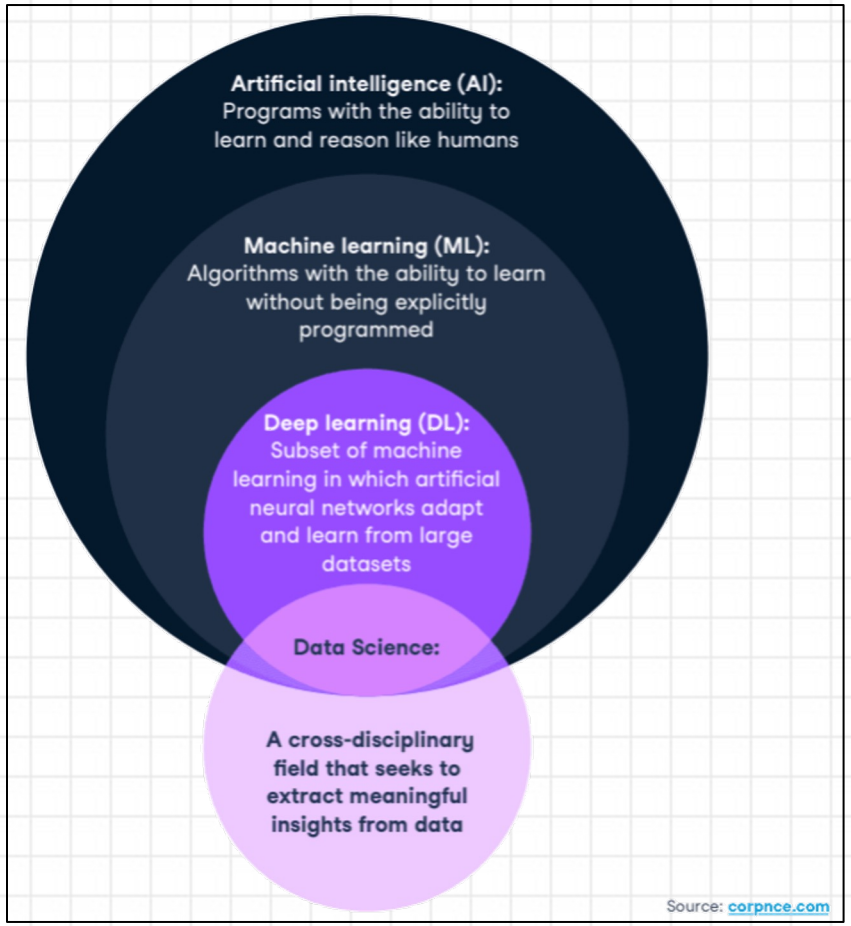
More terms: "Data Science", "Data Mining", ….

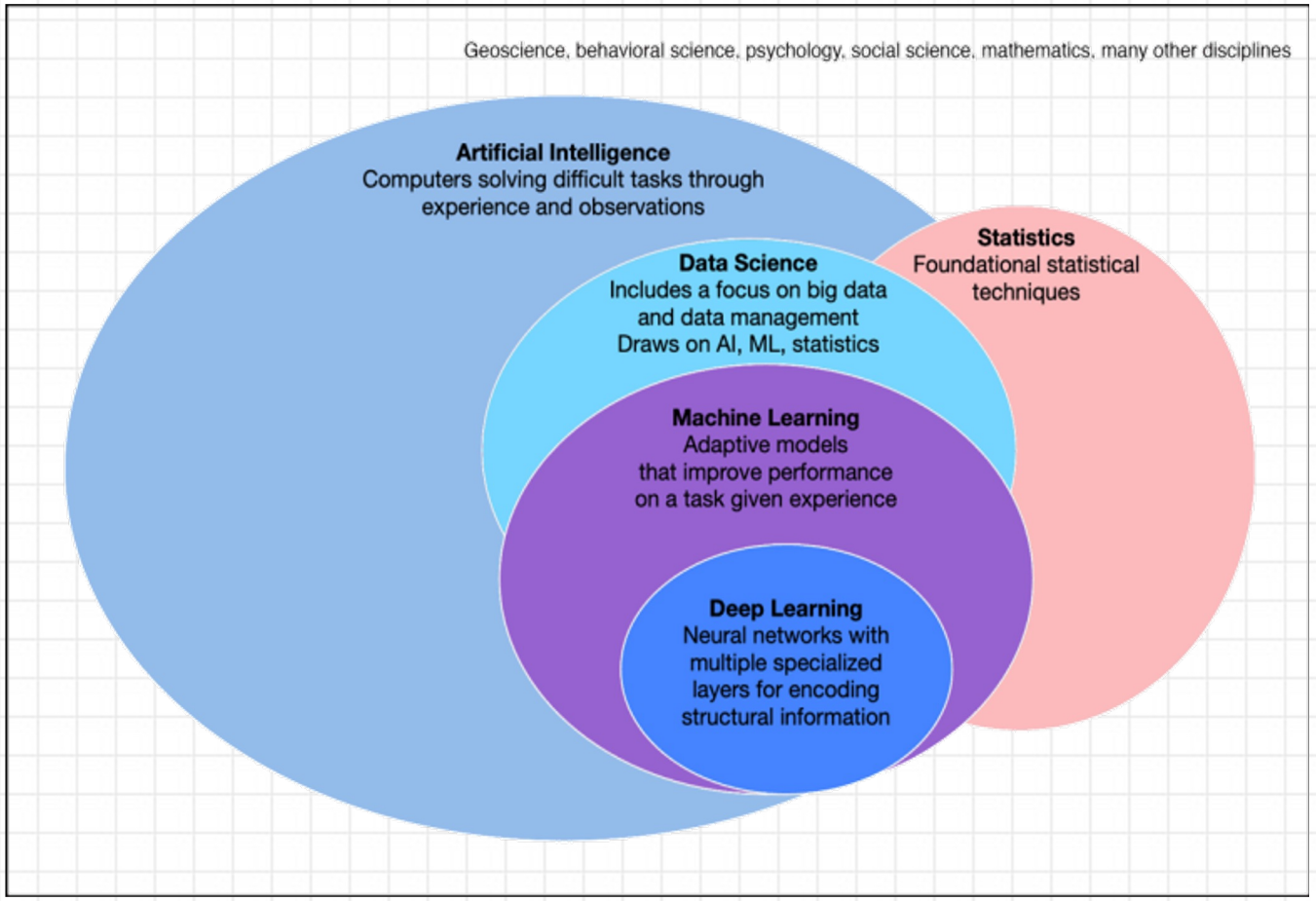
Terminology and Buzzwords

- Data Science:
  - An interdisciplinary academic field that uses statistics, scientific computing, scientific methods, processes, algorithms and systems to extract or extrapolate knowledge and insights from noisy, structured, and unstructured data. (Wikipedia)

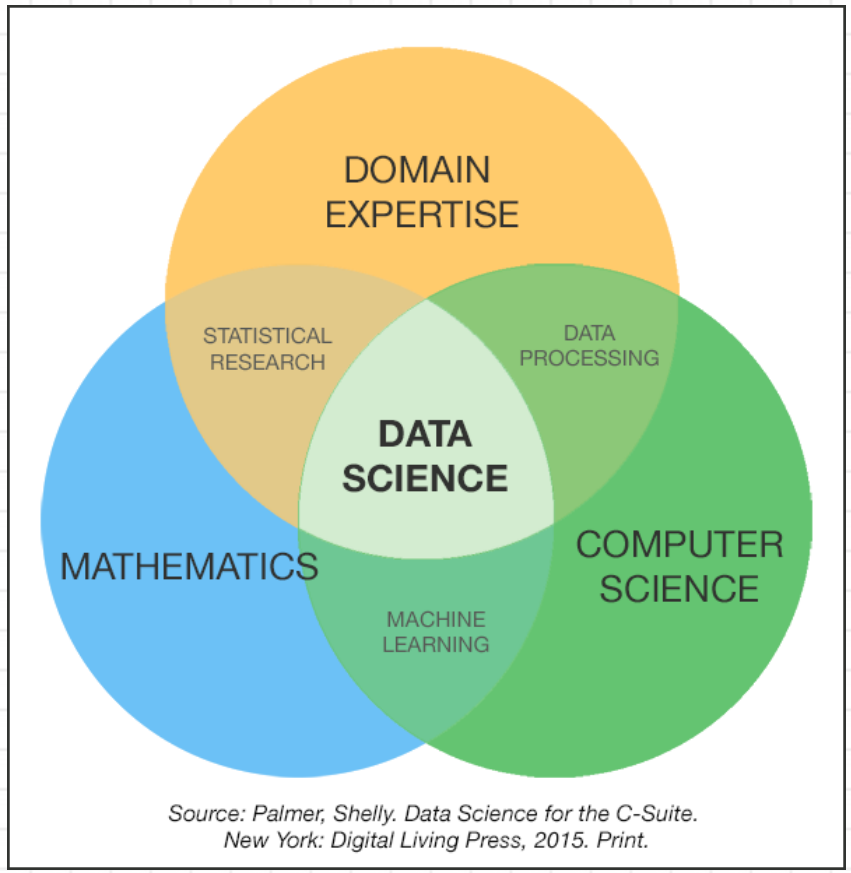
- Data Mining
  - Data mining is the process of extracting and discovering patterns in large data sets involving methods at the intersection of machine learning, statistics, and database systems. (Wikipedia)

Terminology and Buzzwords



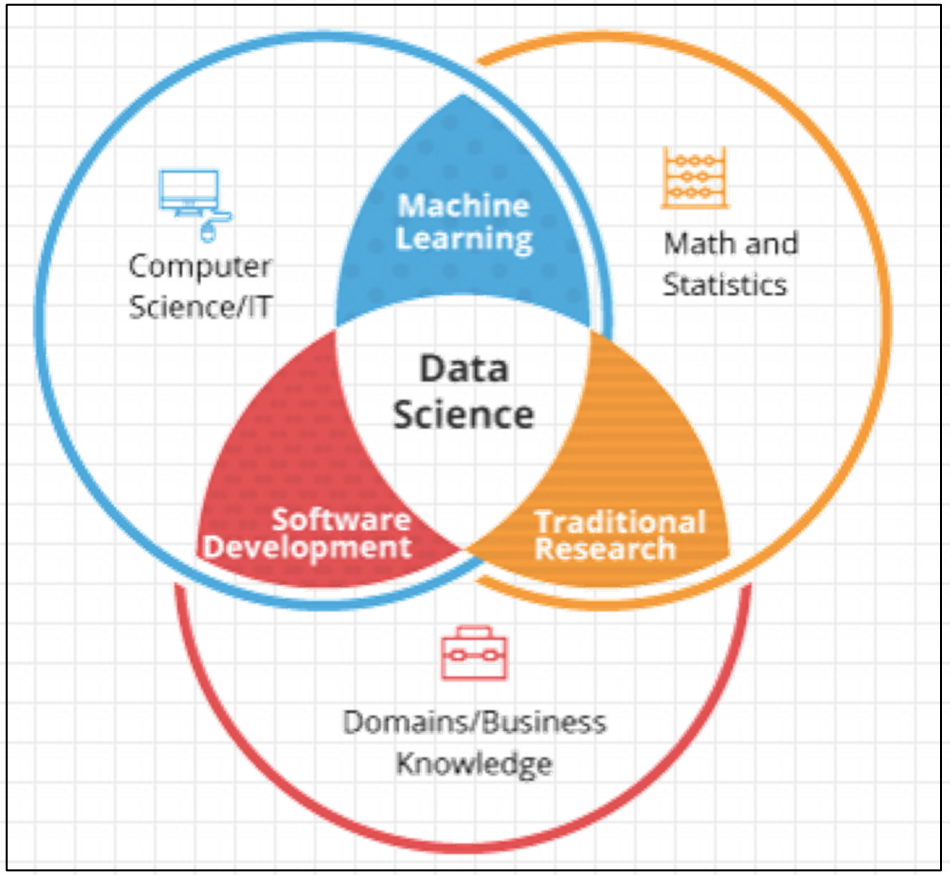


Terminology and Buzzwords

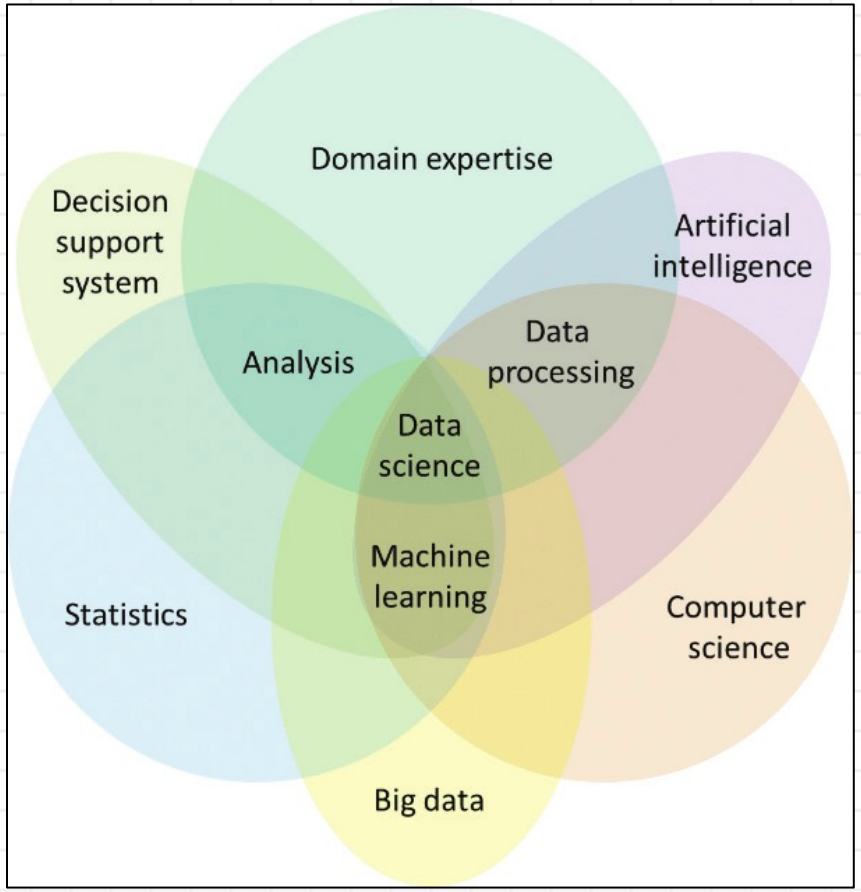


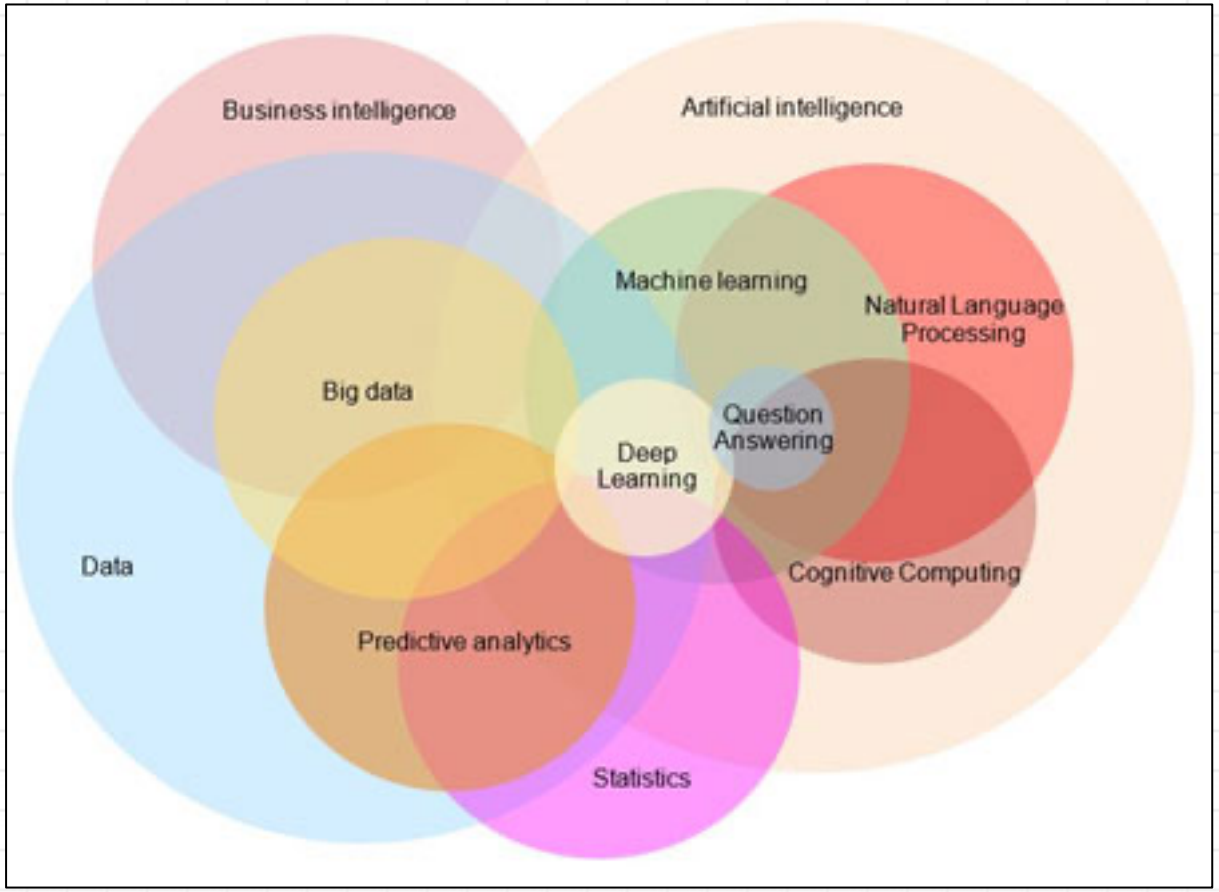
Source: Palmer, Shelly. Data Science for the C-Suite. New York: Digital Living Press, 2015. Print.

More Venn Diagrams?



Terminology and Buzzwords



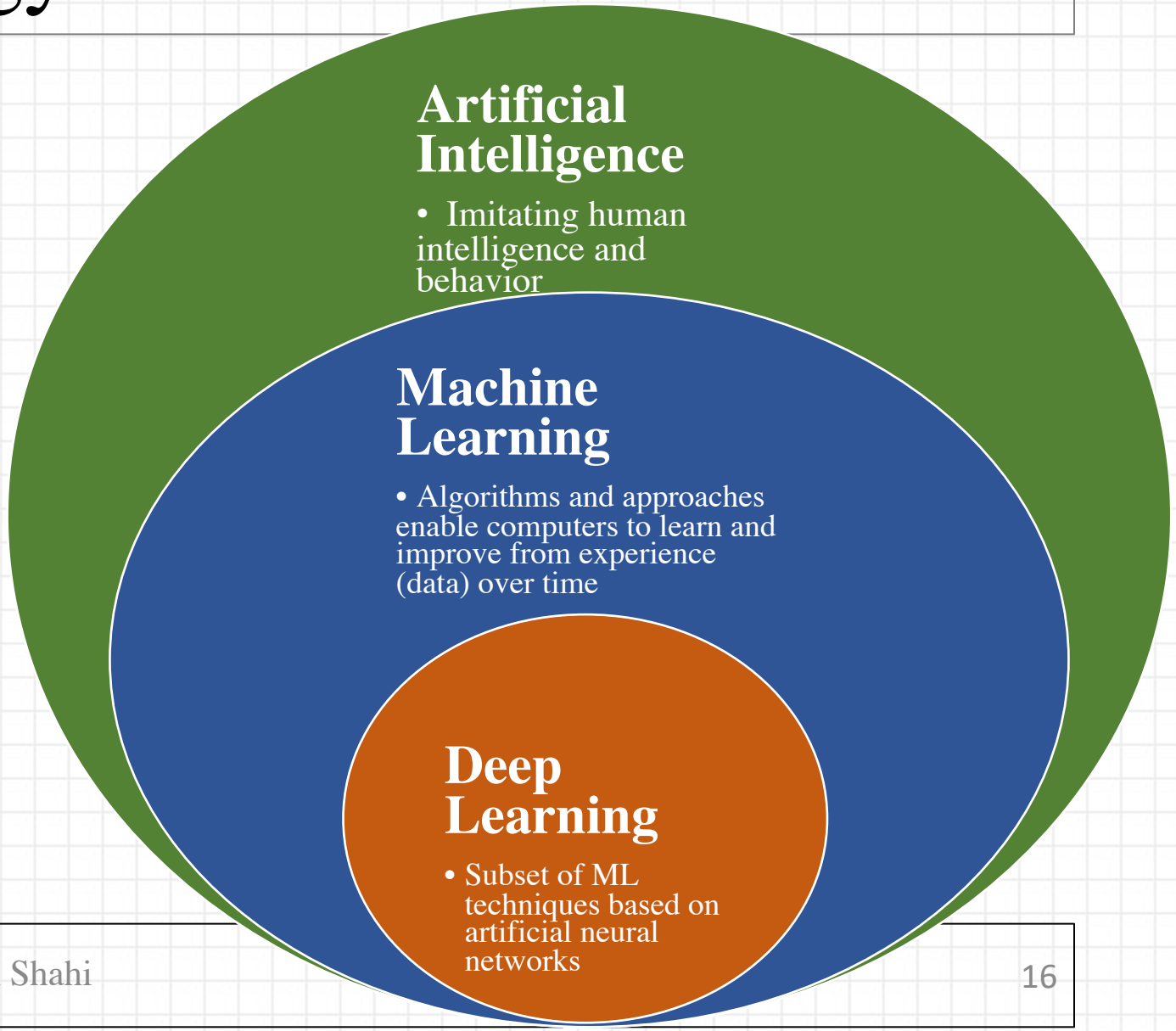


Terminology and Buzzwords

- **Artificial Intelligence (AI):** Simulation of human intelligence in machines to perform tasks that would naturally need human cognitive function, e.g., learning and perception, decision-making.

- **Machine Learning (ML):** A branch of AI that focuses on allowing computer systems to learn and improve over time.

- **Deep Learning (DL):** A subset of ML approaches based on artificial neural networks (ANN)

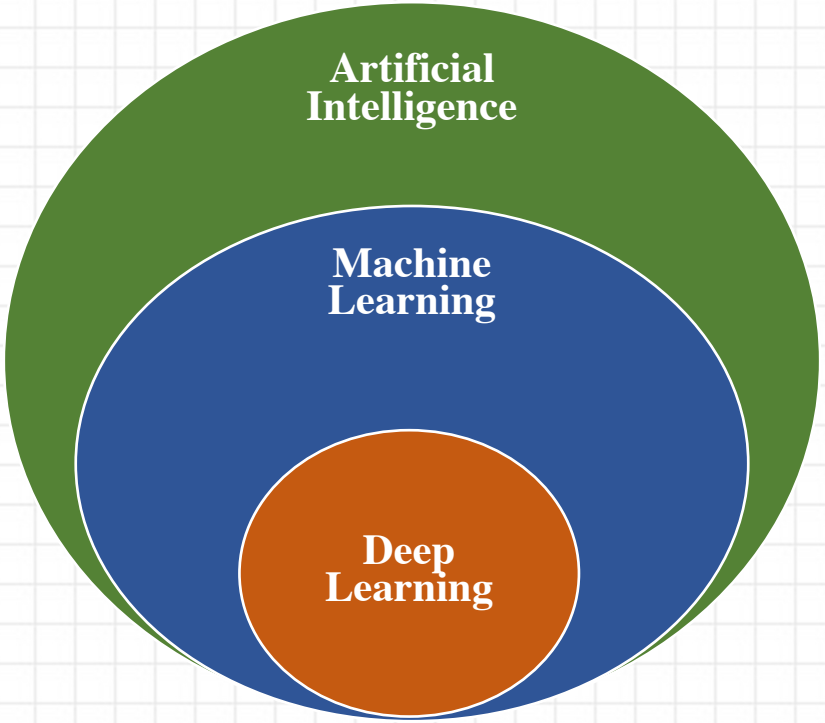


Terminology and Buzzwords

- **Machine Learning (ML):** A branch of AI that focuses on allowing computer systems to learn and improve over time.

- The standard (widely-quoted) definition:

"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***." – Tom Mitchell



Interesting Readings:

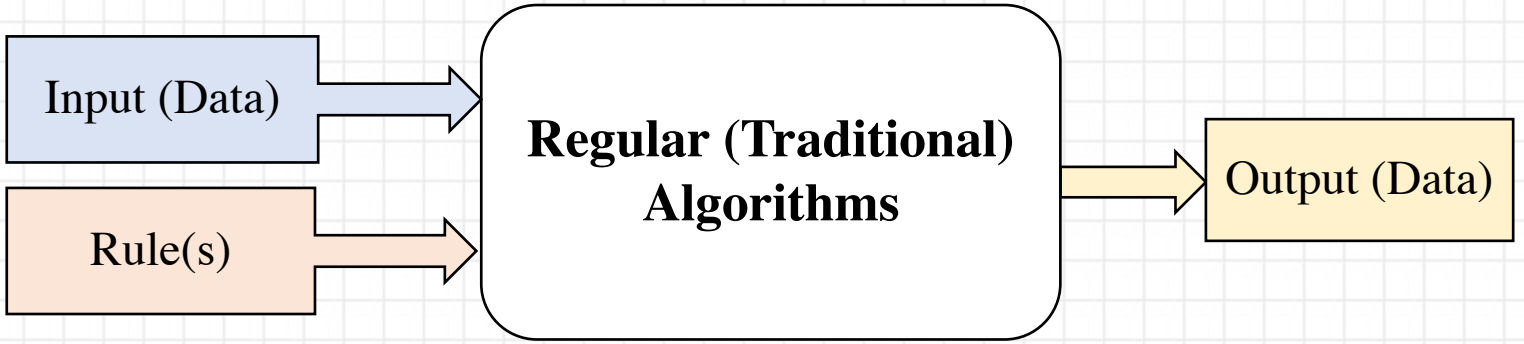
- Timeline of machine learning
- Timeline of artificial intelligence

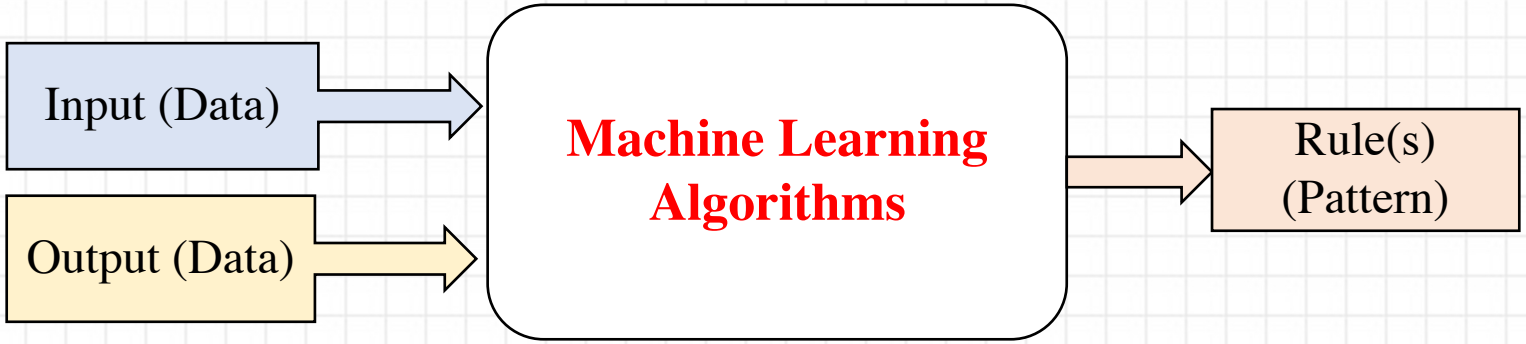
Machine Learning Applications

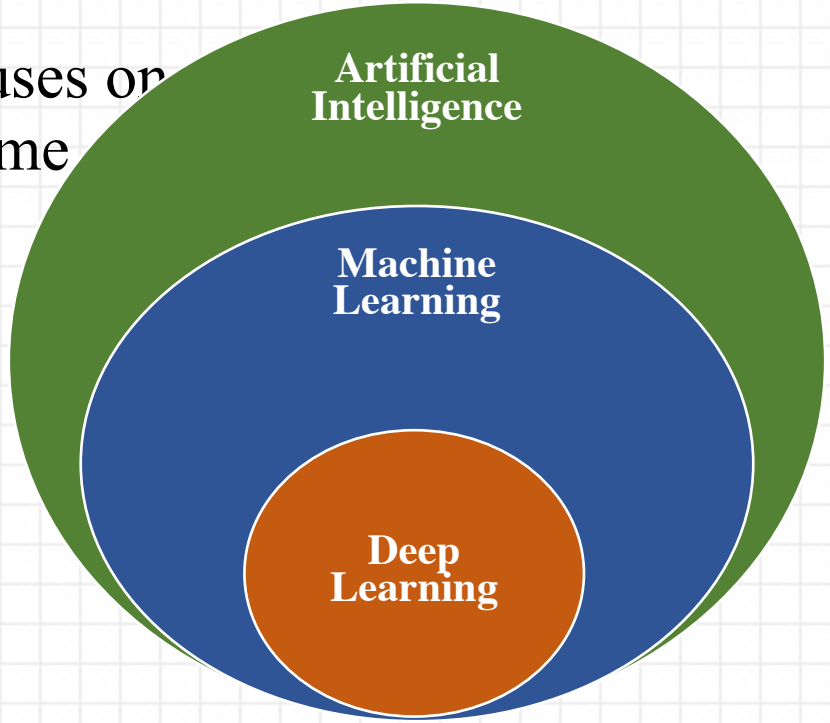
- Recommender systems
  - Encouraging additional purchases: Amazon, eBay, …
  - Increasing user engagements: YouTube, Netflix, Spotify, Social media feed
- Computer Vision, Image and Video Recognition, Speech Recognition
  - A wide range of deep-learning techniques
- Natural Language Processing (NLP)
  - Enabling computers to understand and process natural human language
  - Labeling textual contents, chatbots, voice bots, virtual assistants, ChatGPT!
- Robotics
  - Autonomous cars

Machine Learning Algorithms

- **Machine Learning (ML):** A branch of AI that focuses on allowing computer systems to learn and improve over time





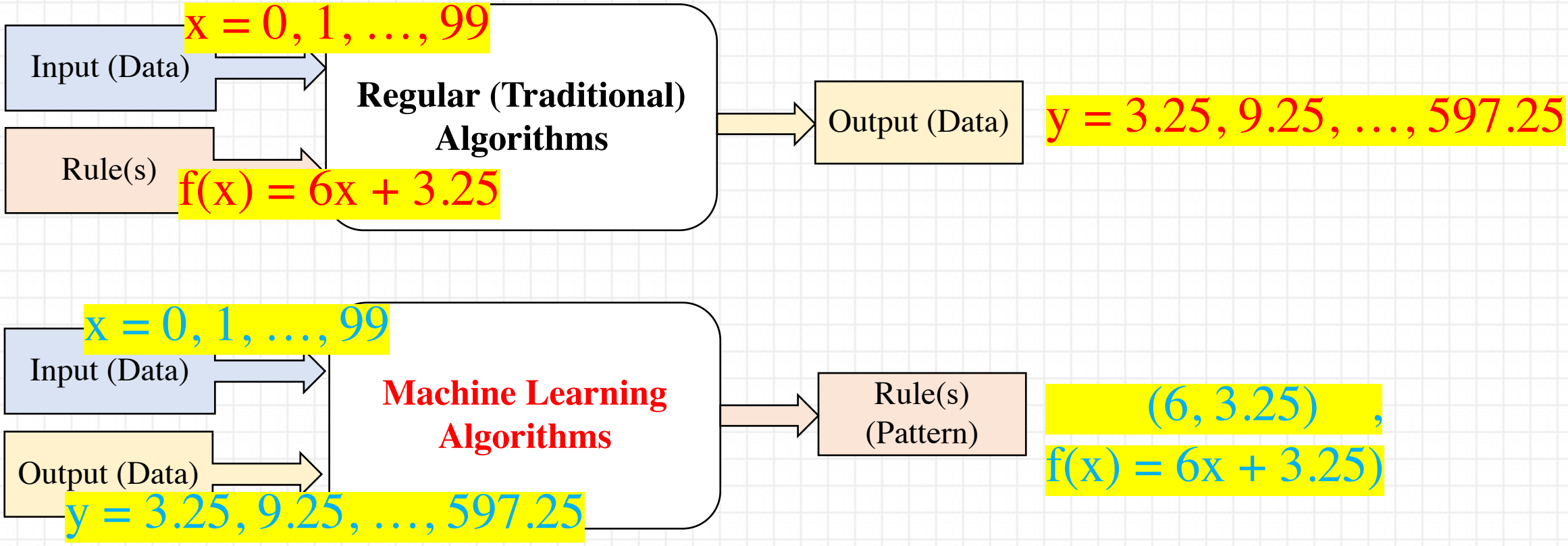


Interesting Readings:

- Timeline of machine learning
- Timeline of artificial intelligence

Machine Learning Algorithms

- **Machine Learning (ML):** A branch of AI that focuses on allowing computer systems to learn and improve over time.



Let's see some demo code!

Types of Machine Learning

- **Supervised machine learning:** Supervised machine learning are types of machine learning that are trained on well-labeled training data. Labeled data means the training data is already tagged with the correct output.
- **Unsupervised machine learning:** Unlike supervised learning, unsupervised learning doesn't have any tagged data. It learned patterns from untagged data. Basically, it creates a group of objects based on the input data/features.
- **Semi-supervised machine learning:** Semi-supervised learning falls between supervised and unsupervised learning. It has a small amount of tagged data and a large amount of untagged data.

https://data-flair.training/

Types of Machine Learning

1. Supervised Learning (SL):
   - Labeled data
     - Continuous value: **regression** task
     - Categorical/discrete: **classification** task
2. Unsupervised Learning (UL):
   - Unlabeled data
     - Clustering
     - Dimensionality reduction, feature extraction
     - Anomaly/novelty detection
3. [Semisupervised Learning]
4. Reinforcement Learning (RL)
   - Learning system is called an *agent*
   - Observe the environment,
   - Select and perform actions, and
   - get *rewards* in return (or *penalties* in the form of negative rewards).
   - It must then learn by itself what is the best strategy, called a *policy*, to get the most reward over time.
   - A policy defines what action the agent should choose when it is in a given situation.

Types of Machine Learning

1. Supervised Learning (SL) → Labeled data ~Task-driven
   - Regression,
   - Classification
2. Unsupervised Learning (UL) → Unlabeled data ~Data-driven
   - Clustering
   - Dimensionality reduction, feature extraction
   - Anomaly/novelty detection
3. [Semisupervised Learning] → SL + UL

4. Reinforcement Learning (RL) → Learning from experience over time

Types of Machine Learning

This lecture!

1. Supervised Learning (SL) → Labeled data ~Task-driven
   - Regression,
   - Classification
2. Unsupervised Learning (UL) → Unlabeled data ~Data-driven
   - Clustering
   - Dimensionality reduction, feature extraction
   - Anomaly/novelty detection
3. [Semisupervised Learning] → SL + UL

4. Reinforcement Learning (RL) → Learning from experience over time

Types of Machine Learning

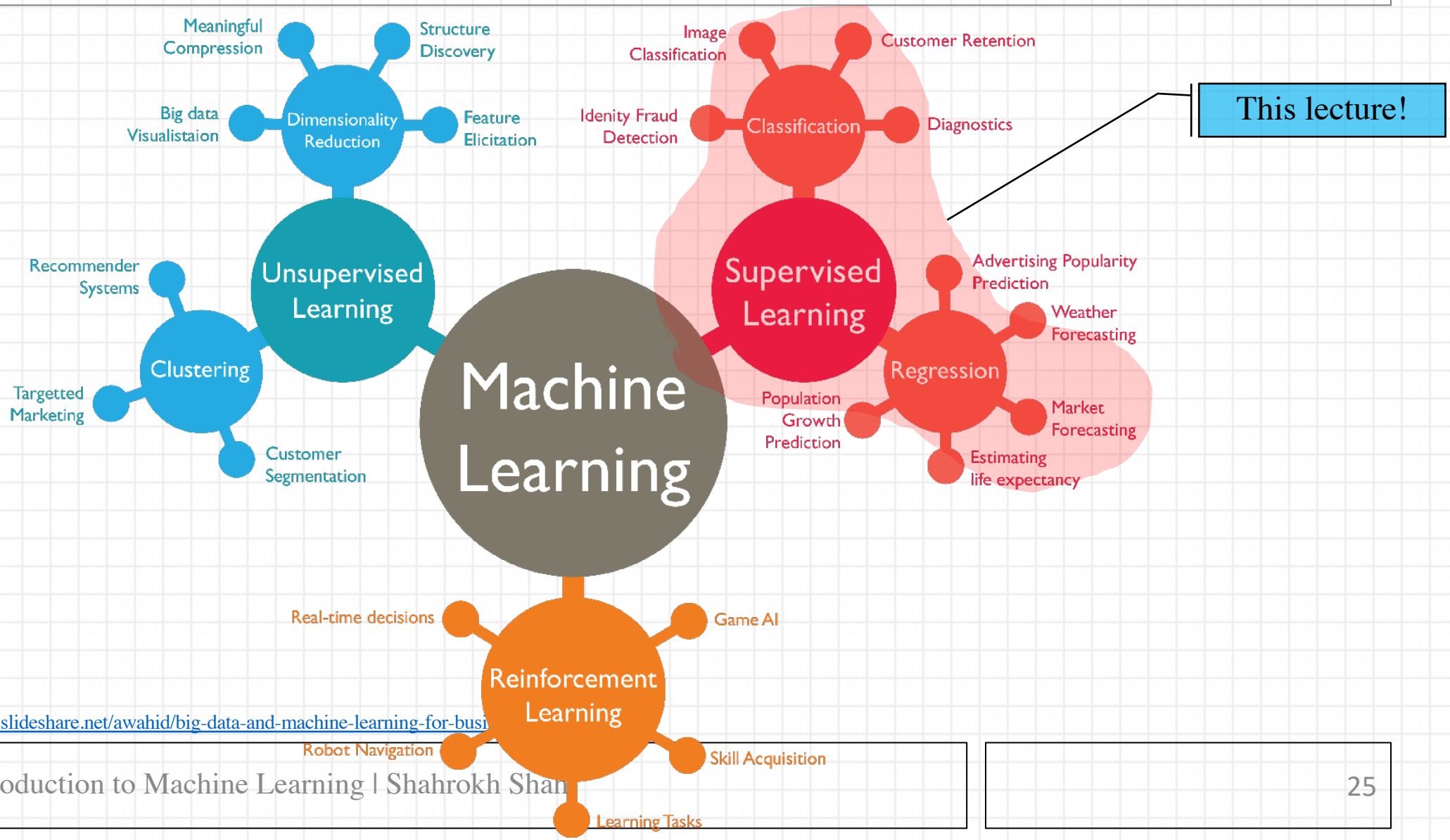


Image: https://www.slideshare.net/awahid/big-data-and-machine-learning-for-busi

Supervised Learning

- Regression
  - Predicting values

Given the California housing dataset including samples with the following attributes, design a predictive system to estimate the house price for any other house in California given theses features.

1. **MedInc** median income in block group
2. **HouseAge** median house age in block group
3. **AveRooms** average number of rooms per household
4. **AveBedrms** average number of bedrooms per household
5. **Population** block group population
6. **AveOccup** average number of household members
7. **Latitude** block group latitude
8. **Longitude** block group longitude

Supervised Learning

- Classification
  - Predicting classes (categories)

Given the iris plants dataset including samples from 3 classes of iris plants (50 instances of each class; 150 in total) with the following attributes, train a classifier model to estimate the type of any other iris plant given these features.

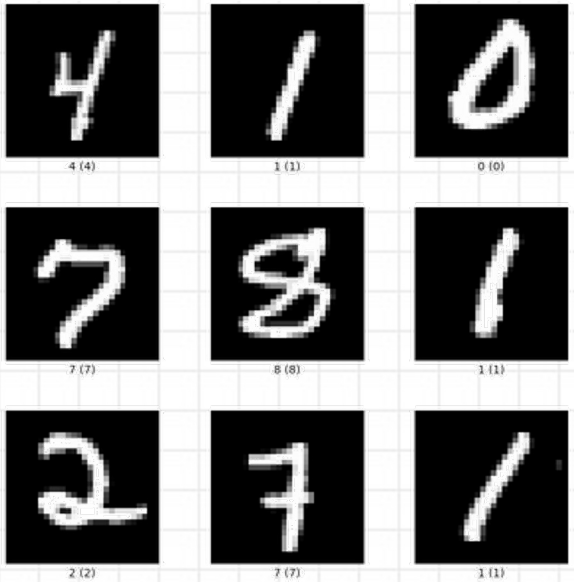
1. sepal length in cm
2. sepal width in cm
3. petal length in cm
4. petal width in cm

- Either binary classification (two classes) or multiclass classification (≥2 classes)
- The input can represent numbers, images, audio, etc.
  - Image classification datasets in TensorFlow package

Supervised Learning

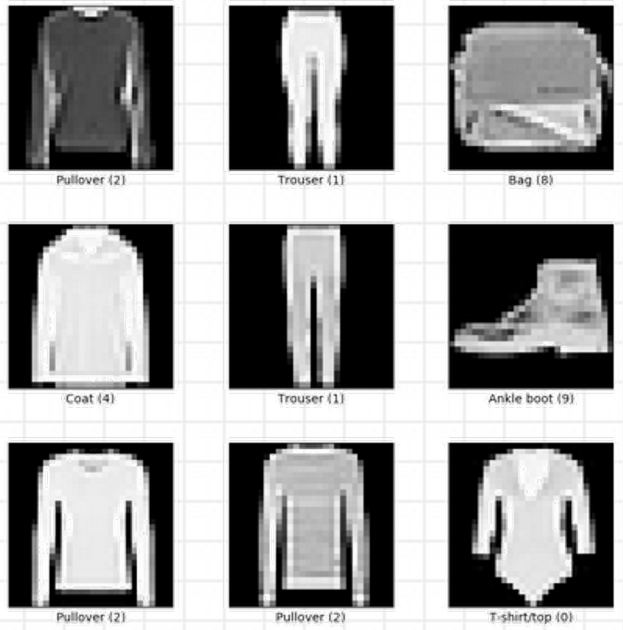
- Classification
  - Predicting classes (categories)

Dataset: mnist
- Handwritten digits
- Grayscale images 28 x 28 pixels
- 10 Labels (0~9)



Dataset: fashion-mnist
- Grayscale images 28 x 28 pixels
- 10 Labels

Label	Description
0	T-shirt/top
1	Trouser
2	Pullover
3	Dress
4	Coat
5	Sandal
6	Shirt
7	Sneaker
8	Bag
9	Ankle boot



Supervised Learning

- Regression
  - Predicting values

Given the California housing dataset including samples with the following attributes, design a predictive system to estimate the house price for any other house in California given theses features.

- Classification
  - Predicting classes (categories)

Given the iris plants dataset including samples from 3 classes of iris plants (50 instances of each class; 150 in total) with the following attributes, train a classifier model to estimate the type of any other iris plant given these features.

1. sepal length in cm
2. sepal width in cm
3. petal length in cm
4. petal width in cm

Steps of an End-to-End ML Project

- Problem Description | Looking at the big picture
- Obtaining the data
- Studying and visualizing the data to gain insights
- Data preprocessing (cleaning, normalizing, …)
- Splitting the data into training and test sets
- Selecting a model (or a set of models) and train it
- Evaluating the trained model
- Fine tuning the trained model
- Launch, monitor, and maintain the ML system

Exploratory data analysis (EDA)

Data Preprocessing

Machine learning modeling

Supervised Learning: Regression

Given the California housing dataset including samples with the following attributes, design a predictive system to estimate the house price for any other house in California given theses features.

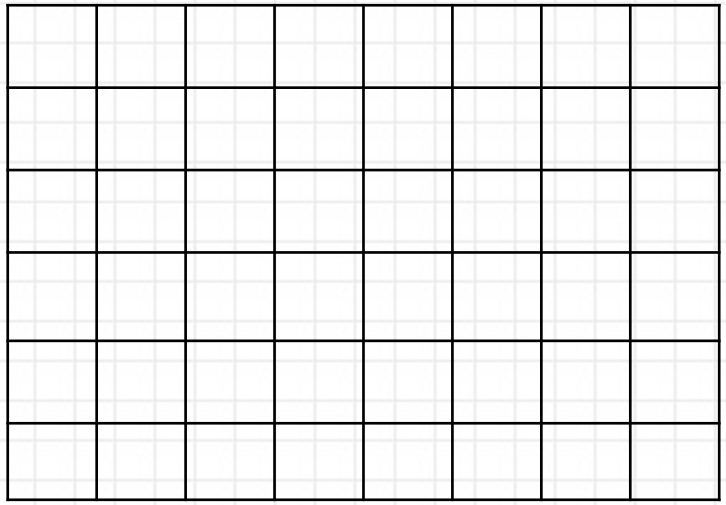
1. **MedInc** median income in block group
2. **HouseAge** median house age in block group
3. **AveRooms** average number of rooms per household
4. **AveBedrms** average number of bedrooms per household
5. **Population** block group population
6. **AveOccup** average number of household members
7. **Latitude** block group latitude
8. **Longitude** block group longitude

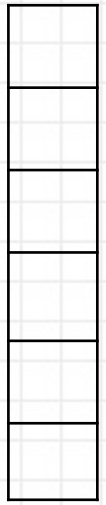
Data Formatting

Inputs:
Feature Matrix, Attributes

Outputs:
Target

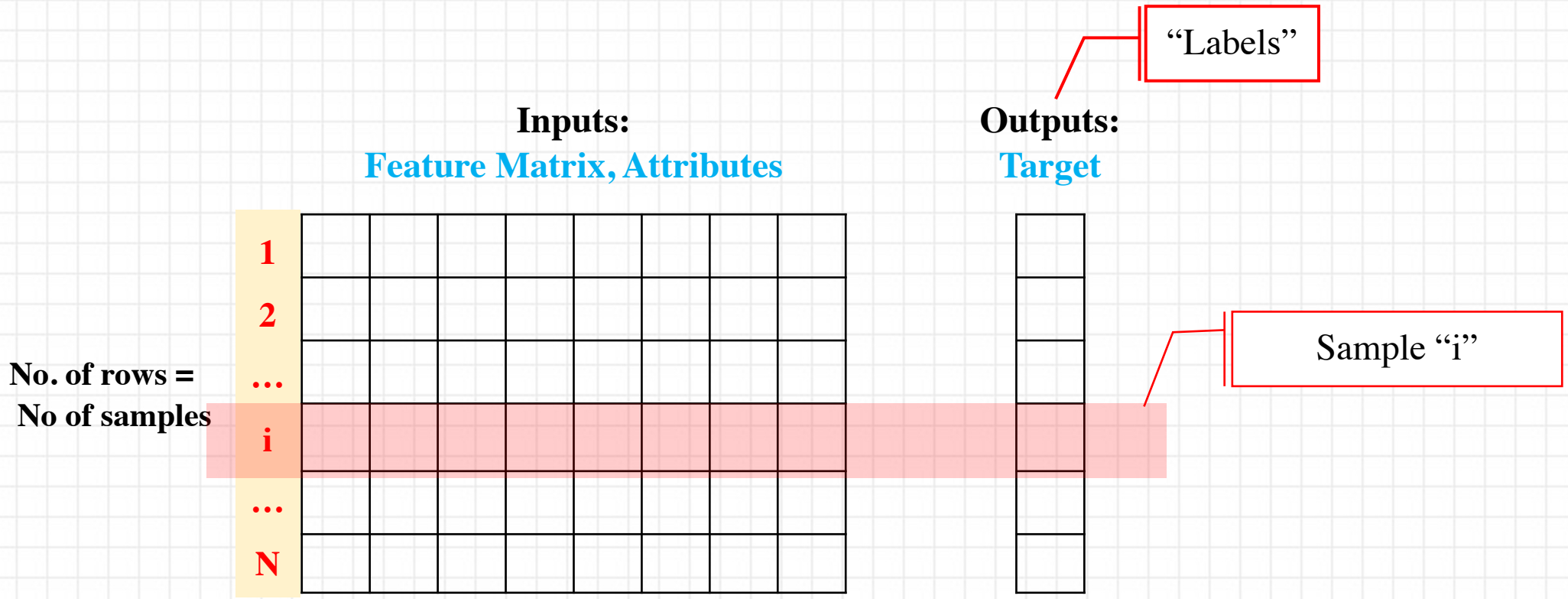
No. of rows =
No of samples





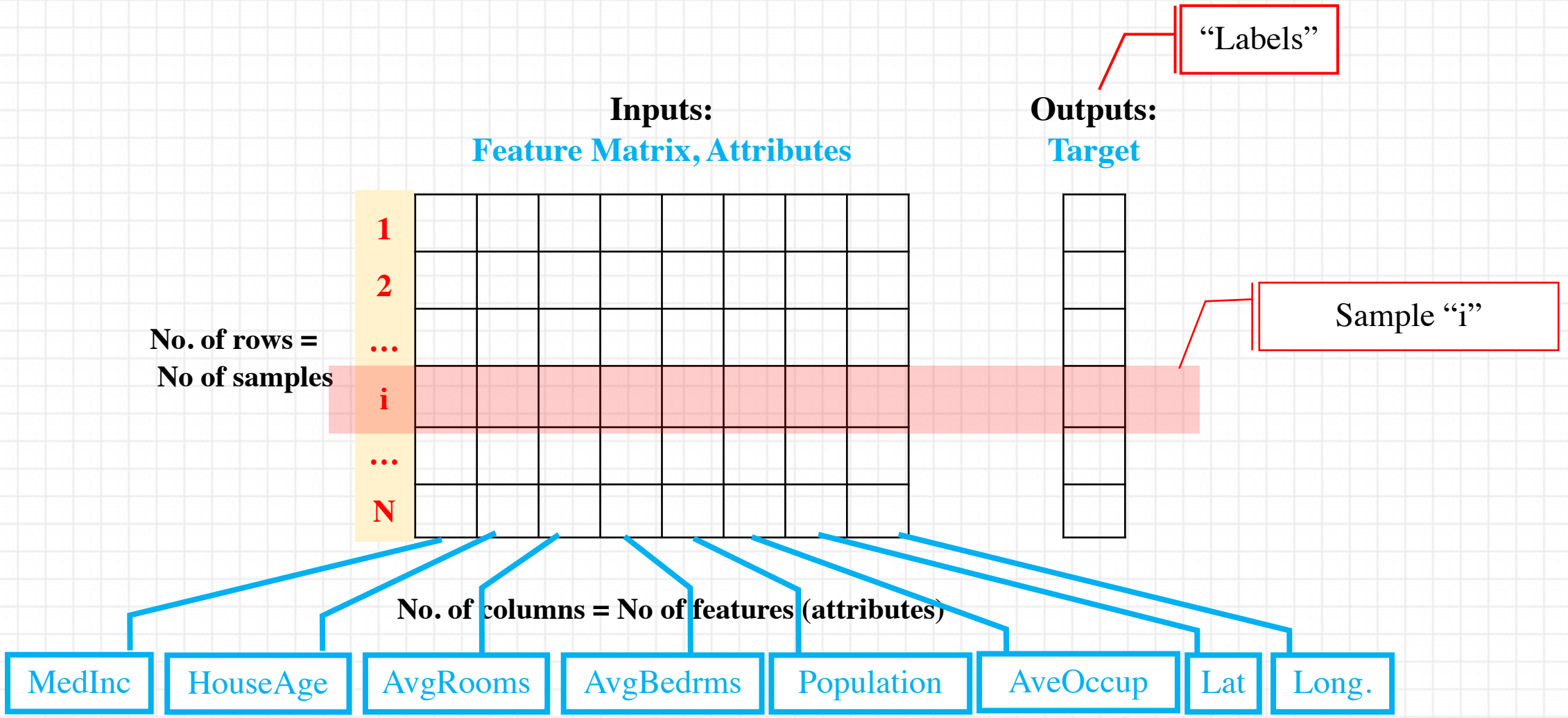
No. of columns = No of features (attributes) = No. of dimensions

Data Formatting



No. of columns = No of features (attributes) = No. of dimensions

Data Formatting



Linear Regression Model

- Fitting a linear model

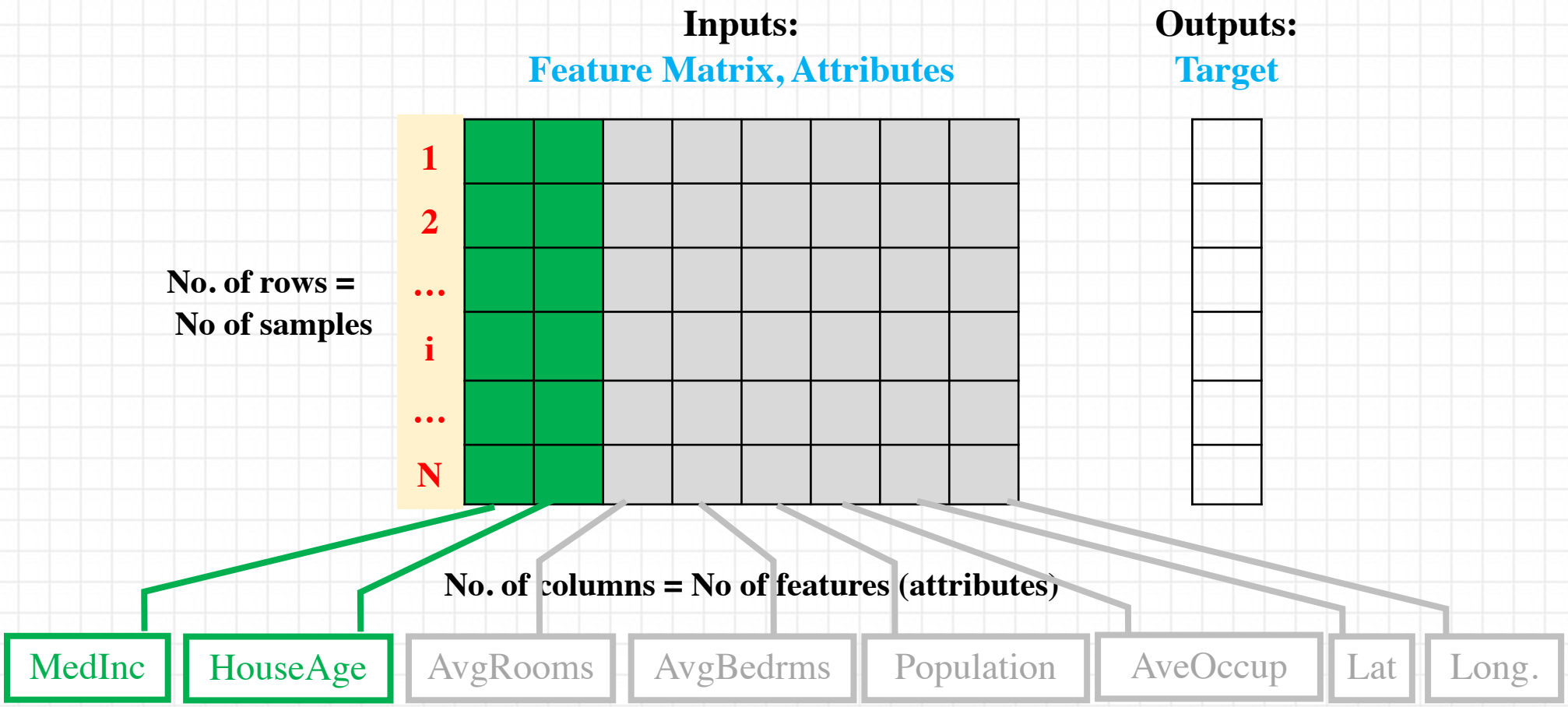
$$y = c_0 + c_1 x_1 + c_2 x_2 + \cdots c_n x_n$$

  - y: Estimation (prediction) of the target value
  - x_i: The value of i-th feature

- For house price prediction:

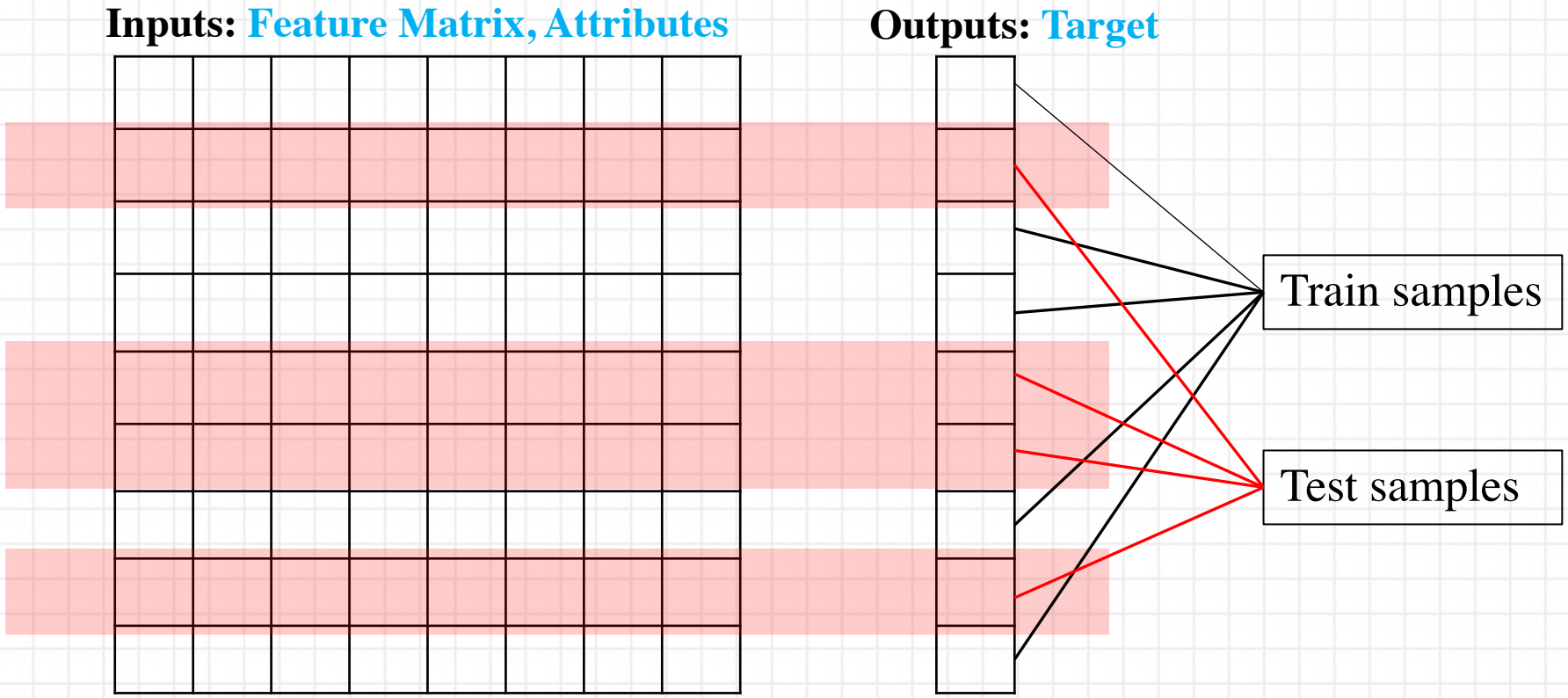
$$\text{house\_price} = c_0 + c_1(\text{MedInc}) + c_2(\text{HouseAge}) + \cdots$$

Feature Selection



Train-Test Split

- Splitting the dataset **randomly** into training and test set
- The model is trained using the train set, and will be evaluated using the unseen test data



More about Data Visualization

- Data Visualization in Python:
  - Matplotlib
  - Pandas
  - Seaborn
  - …

- D3.js
  - D3: Data-Driven Documents
  - A JavaScript library for manipulating documents based on data. **D3** helps you bring data to life using HTML, SVG, and CSS.

- Course CSE6242/CX4242 Data and Visual Analytics

Supervised Learning: Classification

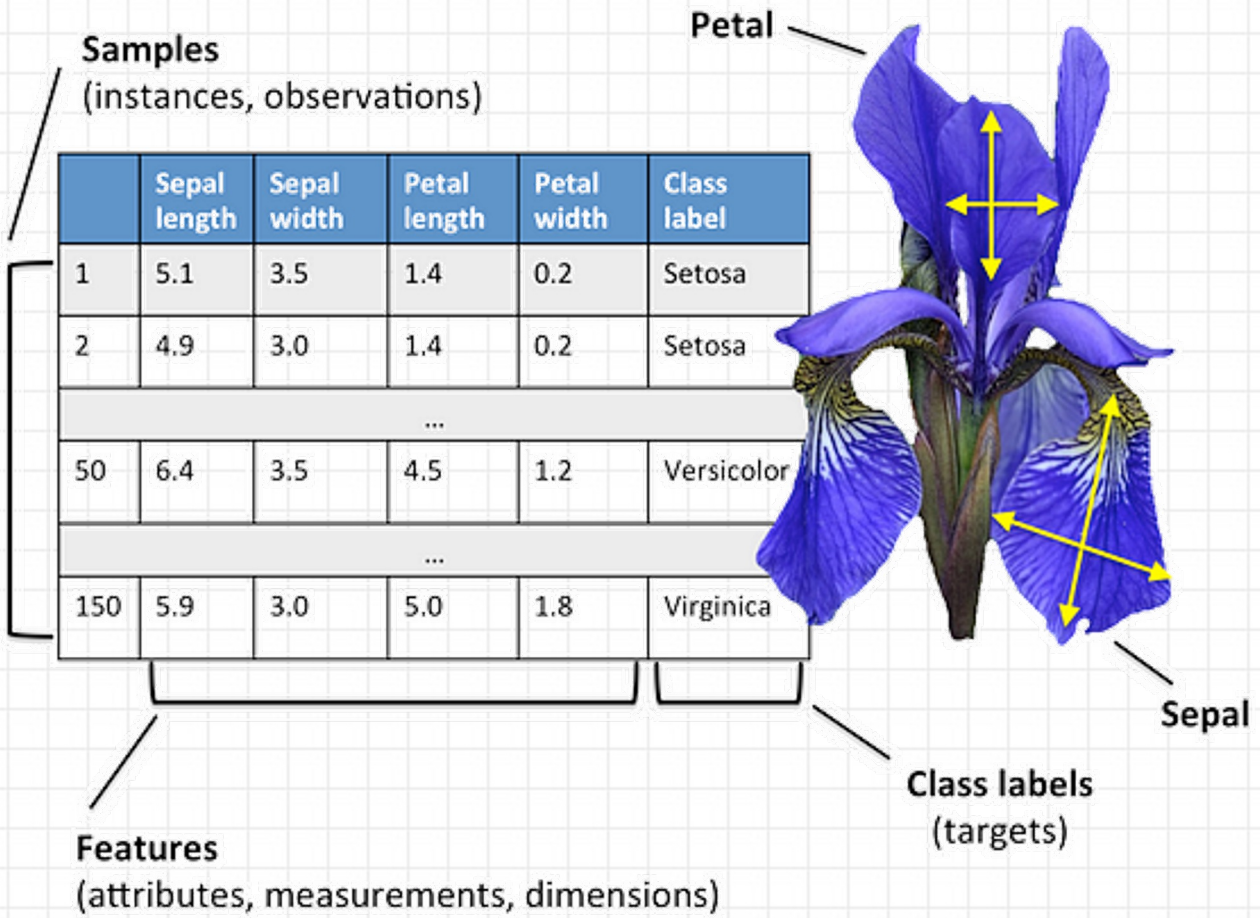
- Iris classification (Iris dataset):
  - One of the best-known datasets to be found in the pattern recognition literature.
  - The data set contains 3 classes of 50 instances each (total 150 samples), where each class refers to a type of iris plant.

- **Features:**
  1. sepal length in cm
  2. sepal width in cm
  3. petal length in cm
  4. petal width in cm
- **Target (class labels):**
  1. Iris Setosa
  2. Iris Versicolour
  3. Iris Virginica

- Task: Train a model which, for any new sample, can **predict the class** of the iris plant given the four input features

Classification Example

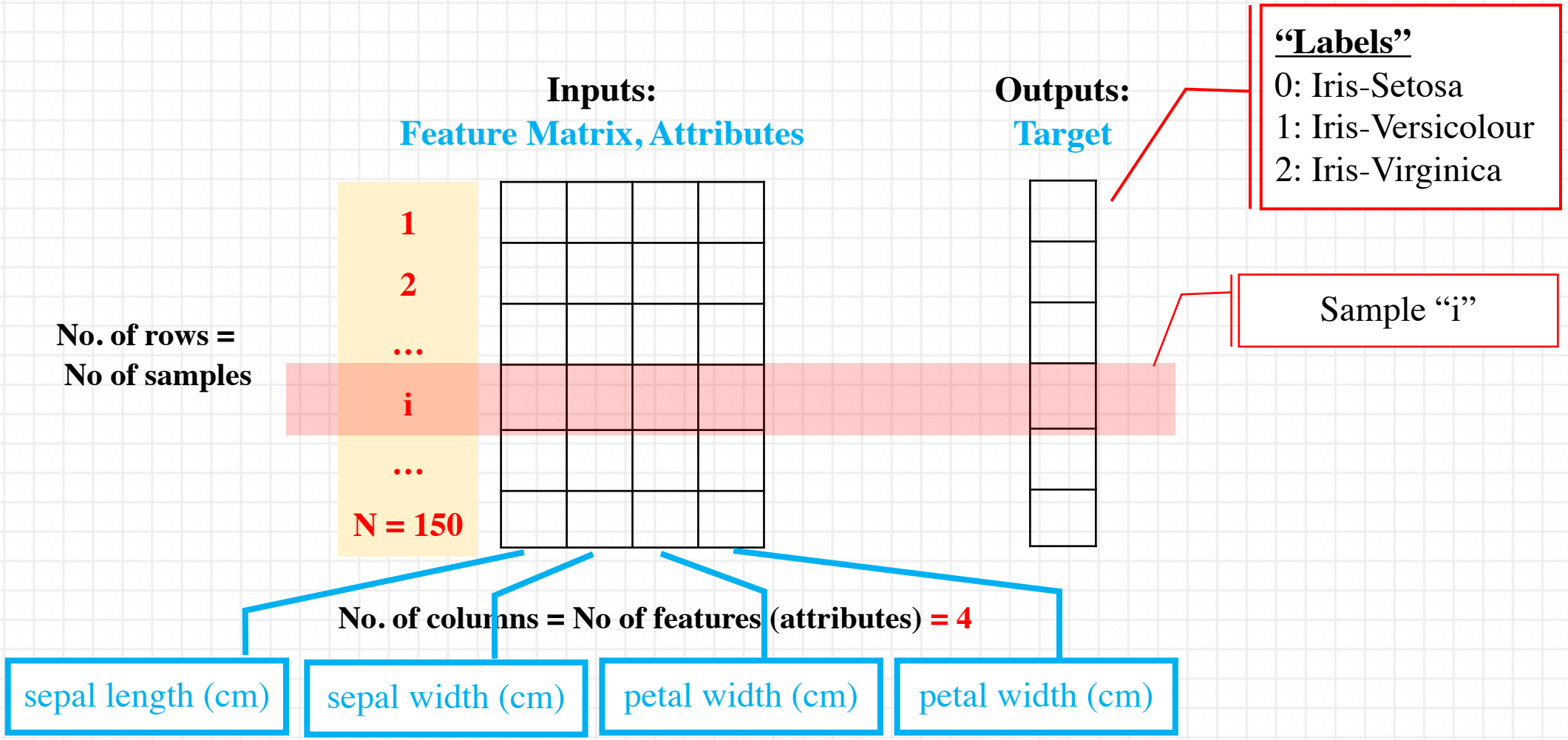
- Iris classification (Iris dataset):
  - **Features:**
    1. sepal length in cm
    2. sepal width in cm
    3. petal length in cm
    4. petal width in cm
  - **Target (class labels):**
    1. Iris Setosa
    2. Iris Versicolour
    3. Iris Virginica



	Sepal length	Sepal width	Petal length	Petal width	Class label
1	5.1	3.5	1.4	0.2	Setosa
2	4.9	3.0	1.4	0.2	Setosa
...					
50	6.4	3.5	4.5	1.2	Versicolor
...					
150	5.9	3.0	5.0	1.8	Virginica

Image: https://medium.com/@jebaseelanravi96/

Data Formatting

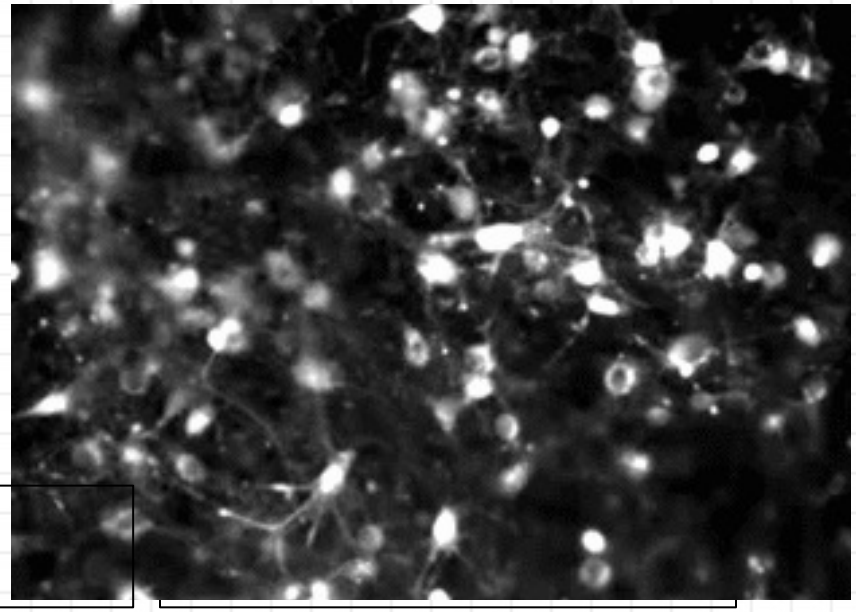


What we did not cover…

- Evaluation metrics
  - Regression: Mean-squared-error (MSE), mean-absolute-error (MAE)
  - Classification: Accuracy, recall, precision, f1-score, confusion matrix, …
- Bias-variance tradeoff
  - Overfitting and underfitting
- Fine tuning
  - The importance of using a validation set for hyperparameter tuning
- Cross-validation and model selection

Neural Networks and Deep Learning

- Artificial Neural Networks (ANNs)
  - Inspired by biological neural networks
  - A collection of connected units known as neurons
  - Neuron
    - Receives a signal
    - Process it by some nonlinear function
    - Sends the output as signals to the connected neurons
    - Connected with edges mimicking synapses in biological neurons
    - May have activation threshold
  - Edges
    - have weights increasing/decreasing the strength of each signal at a connection
    - the weights should [normally] be adjusted in “learning” procedure



Neural Networks and Deep Learning

- Artificial Neural Networks (ANNs)
  - Inspired by biological neural networks
  - A collection of connected units known as neurons
  - Neuron
    - Receives a signal
    - Process it by some nonlinear function
    - Sends the output as signals to the connected neurons
    - Connected with edges mimicking synapses in biological neurons
    - May have activation threshold
  - Edges
    - have weights increasing/decreasing the strength of each signal at a connection
    - the weights should [normally] be adjusted in “learning” procedure

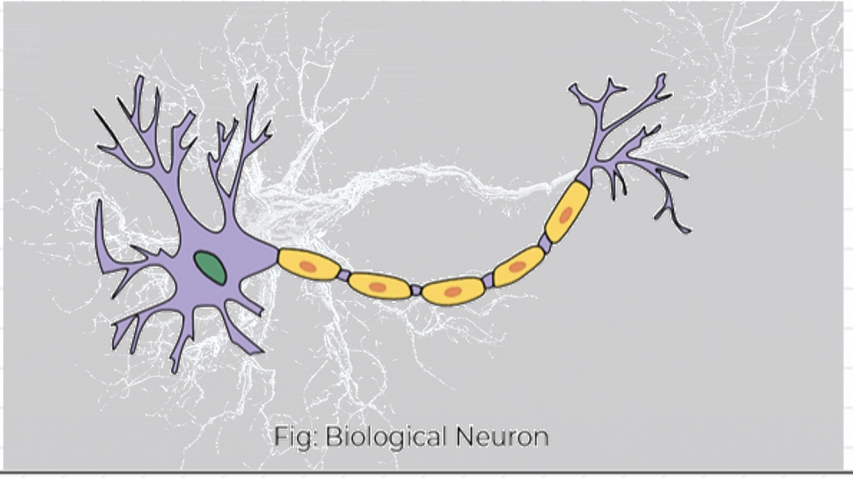


Fig: Biological Neuron

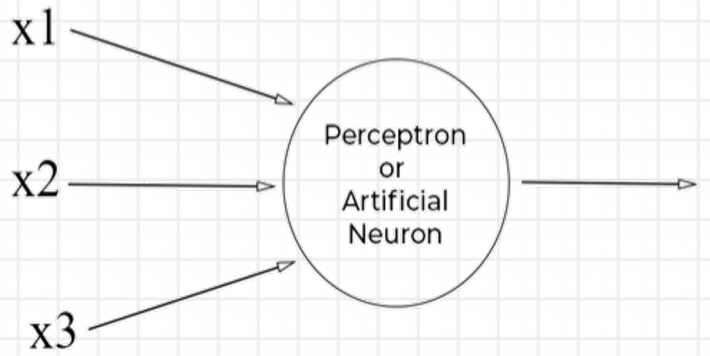
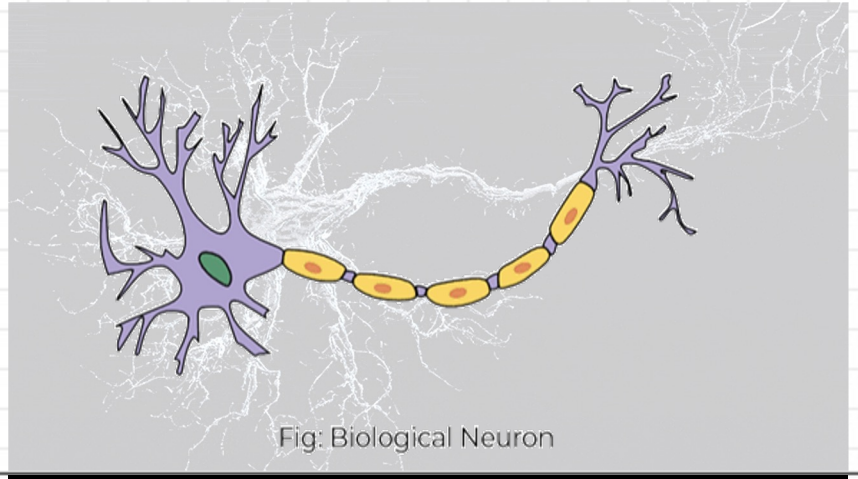


Fig: Artificial Neuron

Machine Learning and Deep Learning

- Artificial Neural Networks (ANNs)
  - Inspired by biological neural networks
  - A collection of connected units known as neurons
  - Neuron
    - Receives a signal
    - Process it by some nonlinear function
    - Sends the output as signals to the connected neurons
    - Connected with edges mimicking synapses in biological neurons
    - May have activation threshold
  - Edges
    - have weights increasing/decreasing the strength of each signal at a connection
    - the weights should [normally] be adjusted in "learning" procedure
- Connections (Architecture of Network)?



Machine Learning and Deep Learning

- Architecture of ANNs
  - Shallow (single-layer)
  - Deep (multi-layer)

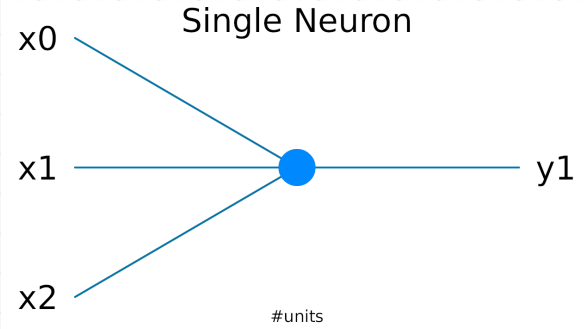
  - Feedforward Neural Networks
  - Recurrent Neural Networks

Machine Learning and Deep Learning

- Architecture of ANNs
  - Shallow (single-layer)
  - Deep (multi-layer)

- Feedforward Neural Networks
- Recurrent Neural Networks

- Simplest architecture
  - One input layer
  - One hidden layer
  - One output layer

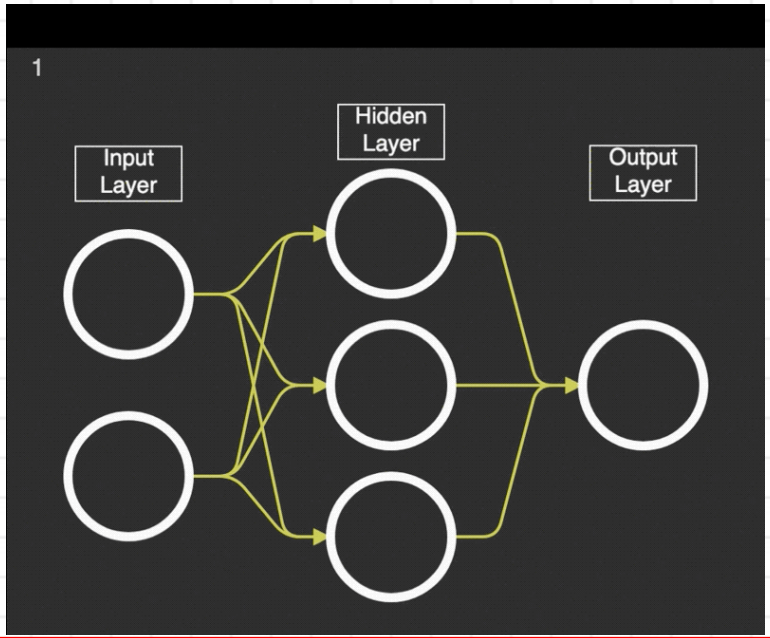


Machine Learning and Deep Learning

- Architecture of ANNs
  - Shallow (single-layer)
  - Deep (multi-layer)

- Feedforward Neural Networks
- Recurrent Neural Networks

- Learning by Backpropagation algorithm
  - Adjusting weights
  - Gradient-based approaches (e.g. gradient descent)
  - Minimizing total cost function (prediction error)

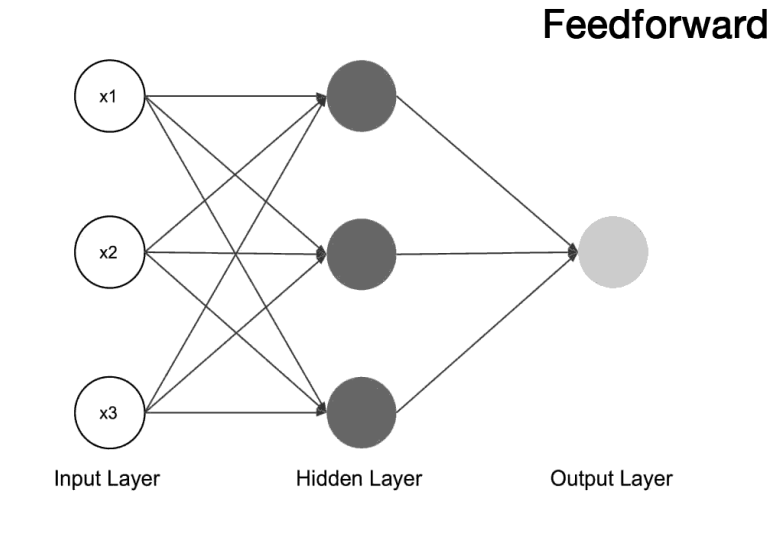


Machine Learning and Deep Learning

- Architecture of ANNs
  - Shallow (single-layer)
  - Deep (multi-layer)

- Feedforward Neural Networks
- Recurrent Neural Networks

- Learning by Backpropagation algorithm
  - Adjusting weights
  - Gradient-based approaches (e.g. gradient descent)
  - Minimizing total cost function (prediction error)

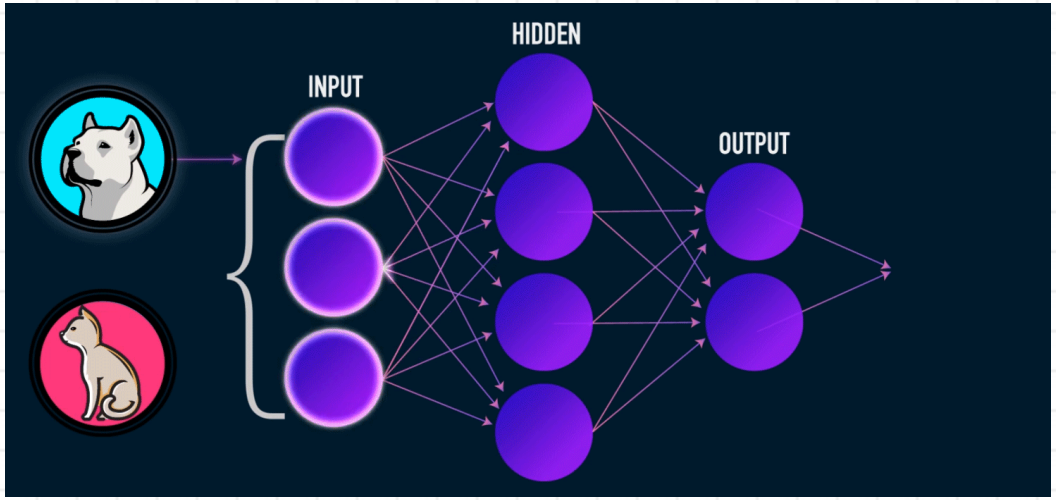


Machine Learning and Deep Learning

- Architecture of ANNs
  - Shallow (single-layer)
  - Deep (multi-layer)

- Feedforward Neural Networks
- Recurrent Neural Networks

- Learning by Backpropagation algorithm
  - Adjusting weights
  - Gradient-based approaches (e.g. gradient descent)
  - Minimizing total cost function (prediction error)
- Predicting unseen data

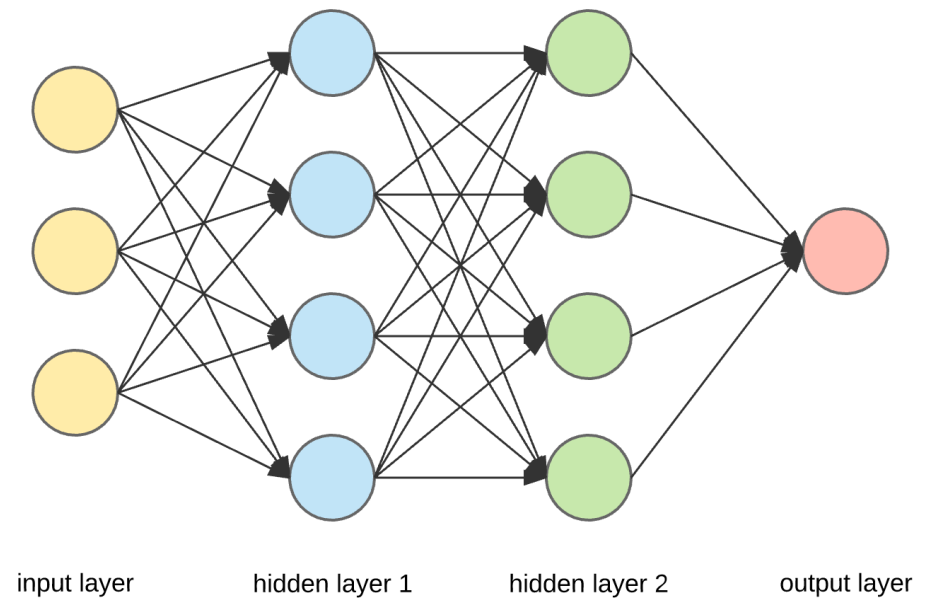


Machine Learning and Deep Learning

- Architecture of ANNs
  - Shallow (single-layer)
  - Deep (multi-layer)

- Feedforward Neural Networks
- Recurrent Neural Networks

- Deep Learning, Deep Structured Learning
- Using multiple layers, more than one hidden layer
- Progressively extract higher-level features

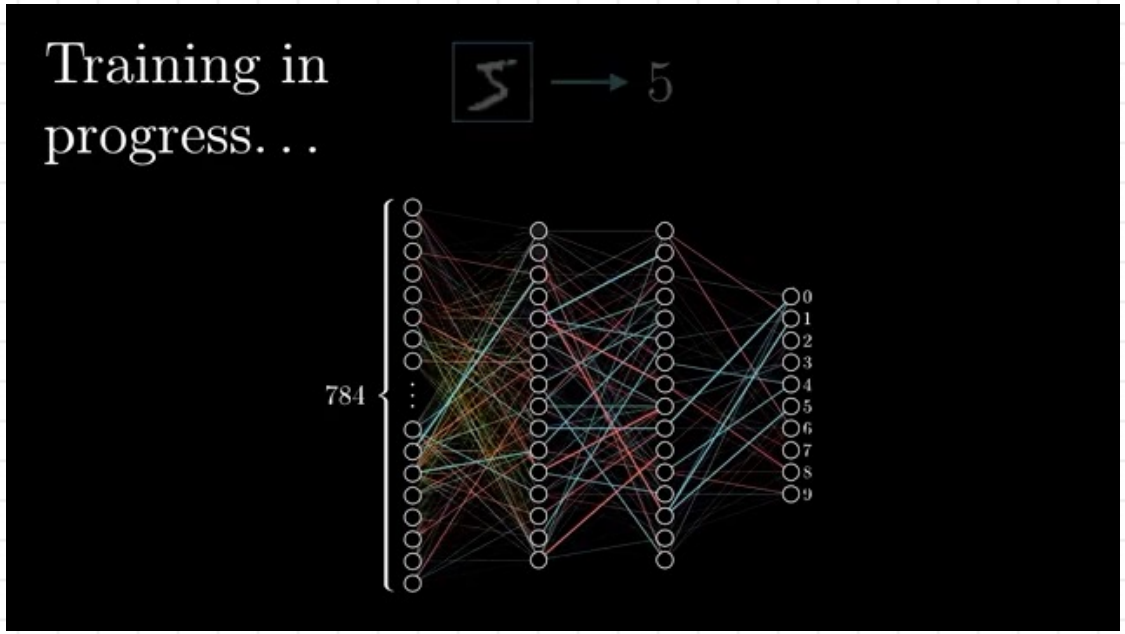


Machine Learning and Deep Learning

- Architecture of ANNs
  - Shallow (single-layer)
  - Deep (multi-layer)

- Feedforward Neural Networks
- Recurrent Neural Networks

- Deep Learning, Deep Structured Learning
- Using multiple layers, more than one hidden layer
- Progressively extract higher-level features
- Training by Backpropagation algorithm:

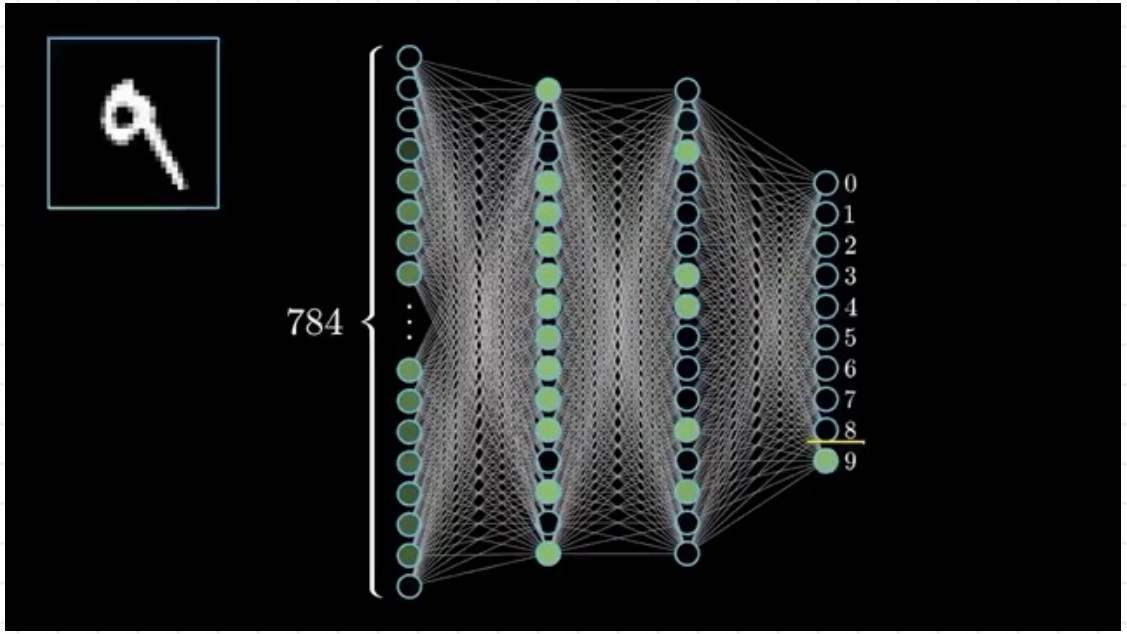


Machine Learning and Deep Learning

- Architecture of ANNs
  - Shallow (single-layer)
  - Deep (multi-layer)

- Feedforward Neural Networks
- Recurrent Neural Networks

- Deep Learning, Deep Structured Learning
- Using multiple layers, more than one hidden layer
- Progressively extract higher-level features
- Prediction:



Machine Learning and Deep Learning

- Architecture of ANNs
  - Shallow (single-layer)
  - Deep (multi-layer)
    - Feedforward Neural Networks
    - Recurrent Neural Networks

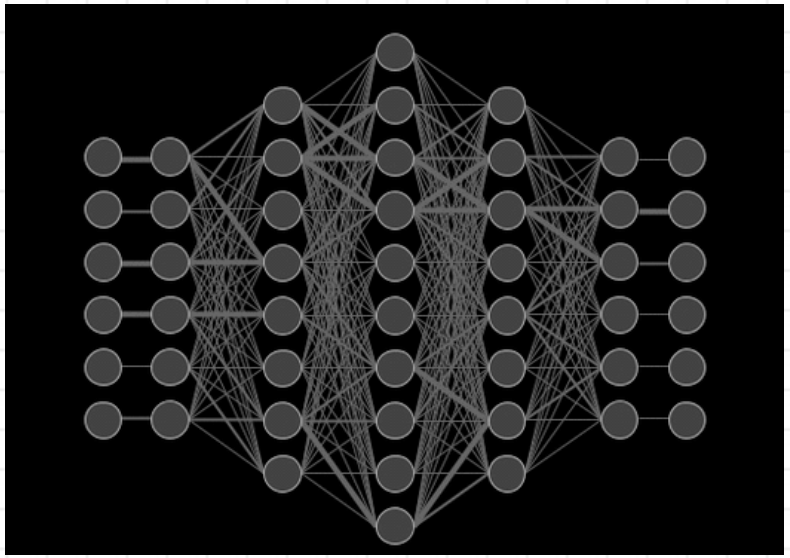
- Applications
  - Image processing
  - Pattern recognition
  - Natural Language Processing
  - Recommendation systems
  - ...

Machine Learning and Deep Learning

- Architecture of ANNs
  - Shallow (single-layer)
  - Deep (multi-layer)

- Feedforward Neural Networks
- Recurrent Neural Networks

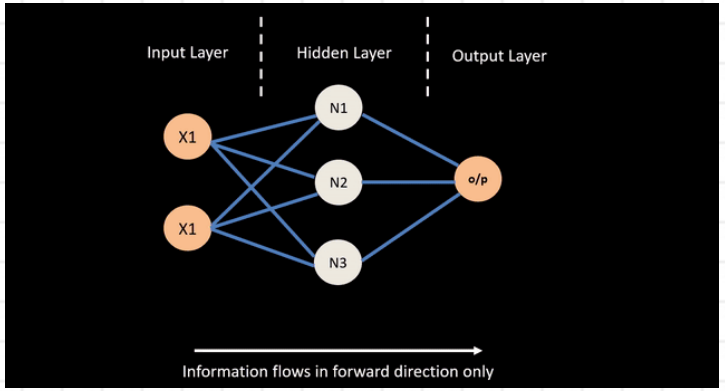
- Information moves in only one direction: **forward**
- From input nodes, through hidden nodes, to the output nodes
- Connections do not form a cycle
- Directed Acyclic Graph (DAG)

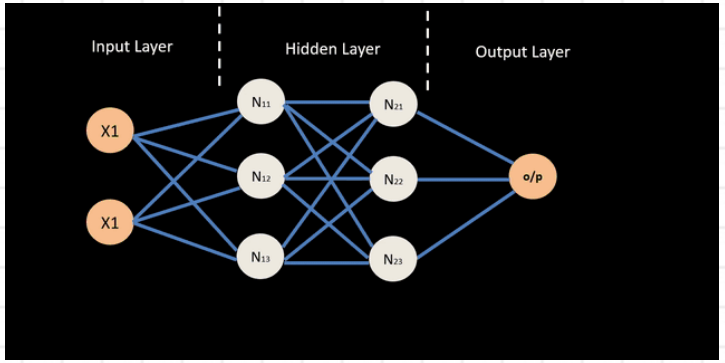


Machine Learning and Deep Learning

- Architecture of ANNs
  - Shallow (single-layer)
  - Deep (multi-layer)
  - Feedforward Neural Networks
  - Recurrent Neural Networks

- Information moves in only one direction: **forward**
- Either single- or multi-layer





Machine Learning and Deep Learning

- Architecture of ANNs
  - Shallow (single-layer)
  - Deep (multi-layer)

- Feedforward Neural Networks
- Recurrent Neural Networks

- Information moves in only one direction: **forward**
- **Non-dynamic** nature, Great for **non-temporal** problems
  - **Universal function approximator**
  - Proven by the universal approximation theorem
  - Every continuous function can be approximated arbitrarily closely by a multi-layer perceptron with just one hidden layer

- What about **temporal** problems?
  - Weather forecast,
  - Predicting dynamical system
  - Predicting financial data
  - Vision and Speech (Recognition, Processing)
  - ..., in short, real-world application

Machine Learning and Deep Learning

- Architecture of ANNs
  - Shallow (single-layer)
  - Deep (multi-layer)

- Feedforward Neural Networks
- Recurrent Neural Networks

- Solving temporal problems using feed-forward structure:
  - Delayed embedding: a chaotic dynamical system can be constructed from sequence of observations of the state of the system
  - Converts the temporal problem to spatial one

- Better solution:
  - Adding recurrent connections (feedback loops) to the forward architecture
  - Transforming the system into a complex dynamical system

Machine Learning and Deep Learning

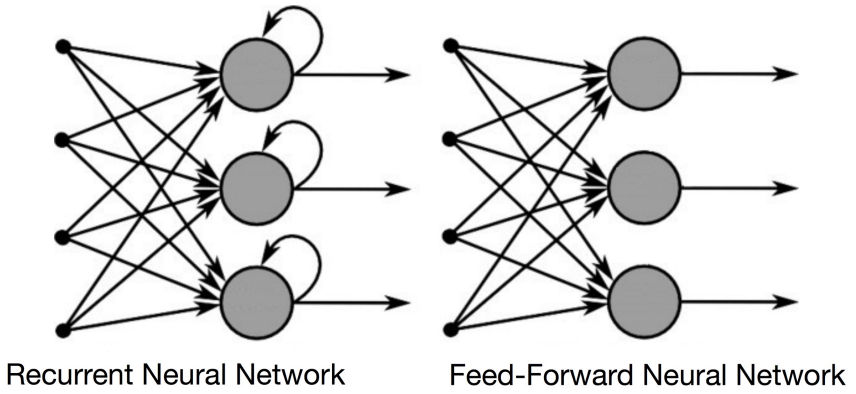
- Architecture of ANNs
  - Shallow (single-layer)
  - Deep (multi-layer)

- Feedforward Neural Networks
- Recurrent Neural Networks

- Solving temporal problems using feed-forward structure:
  - Delayed embedding: a chaotic dynamical system can be constructed from sequence of observations of the state of the system
  - Converts the temporal problem to spatial one

- F

ops)

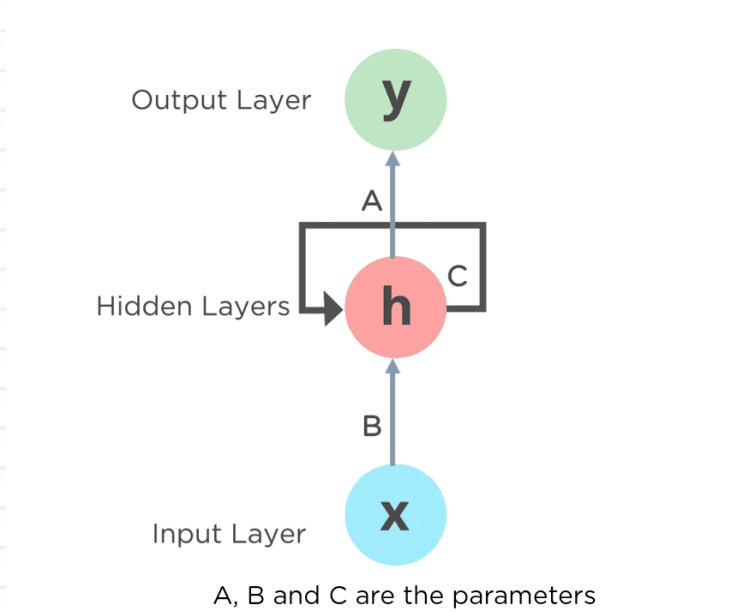


Machine Learning and Deep Learning

- Architecture of ANNs
  - Shallow (single-layer)
  - Deep (multi-layer)

- Feedforward Neural Networks
- Recurrent Neural Networks

- Directed graph along a temporal sequence
- Exhibits temporal dynamic behavior
- Internal state (memory)



Summary

- A very brief introduction to machine learning techniques
- Supervised learning
  - Regression
  - Classification
- Hands-on experiences
  - Google Colab notebook
  - A very brief introduction to Pandas DataFrame and EDA
  - Scikit-learn
  - A very brief introduction to neural networks and deep learning
  - TensorFlow + keras

What's Next?

- Working/practicing with real-world "open" datasets
  - Popular open data repositories
    - UC Irvine ML Repository
    - Kaggle
    - Amazon's AWS datasets
    - A list on ML datasets on Wikipedia

- Recommended resources
  - Georgia Tech: CS-3600, CSE-6740, CS-4641/7641, CS-4644/7643
  - Other online courses: ML Specialization (coursera), Google ML crash course
  - Books: See the last slide (references)
  - Blogs: Towards Data Science, Machine Learning Mastery

References

- [1] Géron, A., 2022. ***Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow.*** " O'Reilly Media, Inc.".
- [2] Albon, C., 2018. ***Machine learning with python cookbook: Practical solutions from preprocessing to deep learning.*** " O'Reilly Media, Inc.".
- [03] Avila, J. and Hauck, T., 2017. ***Scikit-learn cookbook: over 80 recipes for machine learning in Python with scikit-learn.*** Packt Publishing Ltd.
- [04] Animated demonstration are mostly obtained from GIPHY.
- Please let me know if you know the correct references to any animated image or diagram to add the proper citation in this document.