2025 Semester 1

01211373 Machine Learning and Programming for Industry

Dr. Varodom Toochinda

EMME, Kasetsart University

Lecture 01 : Course Introduction

- Course syllabus/grading policy
- Main text & materials
- Introduction to machine learning
- Python & Jupyter notebook guide

Course Syllabus

Name: Machine Learning and Programming for Industry 3(2-3-6)

Description: (from EMME curriculum) Machining learning in industry. Neural network. Hardware and Software. Perceptron. Supervised and unsupervised learning. Reinforcement learning. Loss functions. Back propagation. Binary and multiclass classification. Hyperparameters tuning. Convolutional neural networks. Transfer learning. Data pipelining. Recurrent neural networks. Sequence models. Natural language processing. Embedded device deployment.

Schedule: Mon 1 – 4 PM

Instructor: Dr.Varodom Toochinda
Mobile: 084-3239613 Email: varodom.t@gmail.com

Texts:

- S. Raschka, Y. Liu and V. Mirjalili. Machine Learning with PyTorch and Scikit-Learn. Packt Publishing. 2022.
- Instructor handouts

Grading:

1) Homework Assignments	20 %	2) Project	40 %
3) Midterm	20 %	4) Final	20 %

Google Class Code : oomwxw2g

https://classroom.google.com/c/NzY5MjM5OTg5OTcx?cjc=oomwxw2g

Github : https://github.com/dewdotninja/mlpi

Course Outline (tentative)

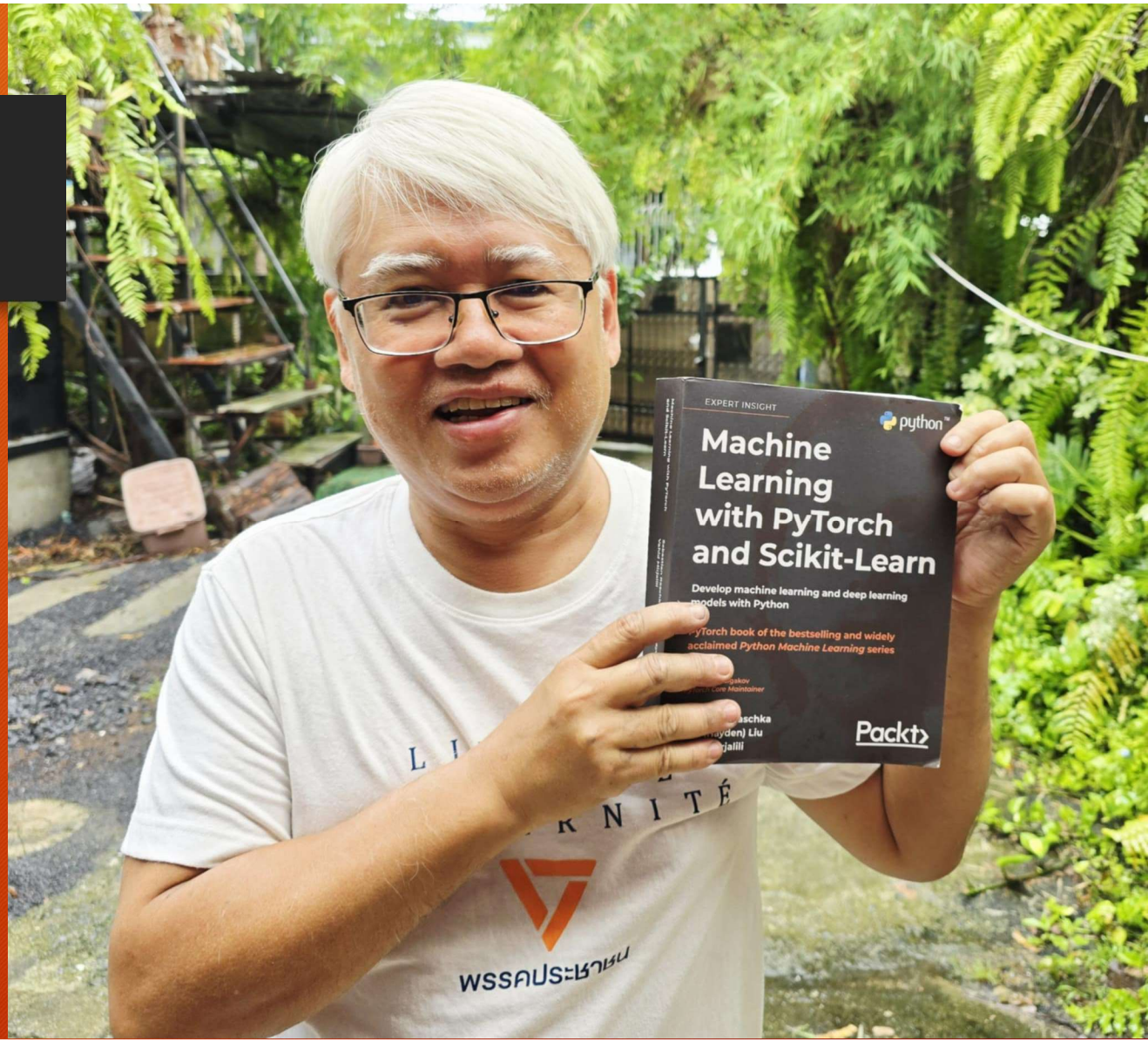
Course Outline (tentative)

Lecture	Topics
1	Course introduction. Math & SW tools (Python & Jupyter notebook)
2	Perceptron. Adaline. Basics of learning.
3	Classifiers. Logistic regression. Regularization.
4	Support vector machines.
5	Decision tree learning. K-nearest neighbors.
6	Data preprocessing.
7	Dimensionality reduction. Principal component analysis.
	Midterm
8	Model evaluation. Hyperparameter tuning.
9	Regression analysis.
10	Clustering Analysis.
11	Multilayer artificial neural networks
12	Convolutional neural networks.
13	Recurrent neural networks.
14	Advanced topics. Course summary.
	Final

Main text :

- S. Raschka, Y. Liu and V. Mirjalili. Machine Learning with PyTorch and Scikit-Learn. Packt Publishing. 2022.

https://sebastianraschka.com/blog/2022/ml-pytorch-book.html



Software

- Python with
  - standard packages : numpy, scipy, matplotlib, panda, etc.
  - scikit-learn
  - pytorch
- Jupyter notebook
  - run on Colab
  - use docker
  - install locally
- Additional software as needed

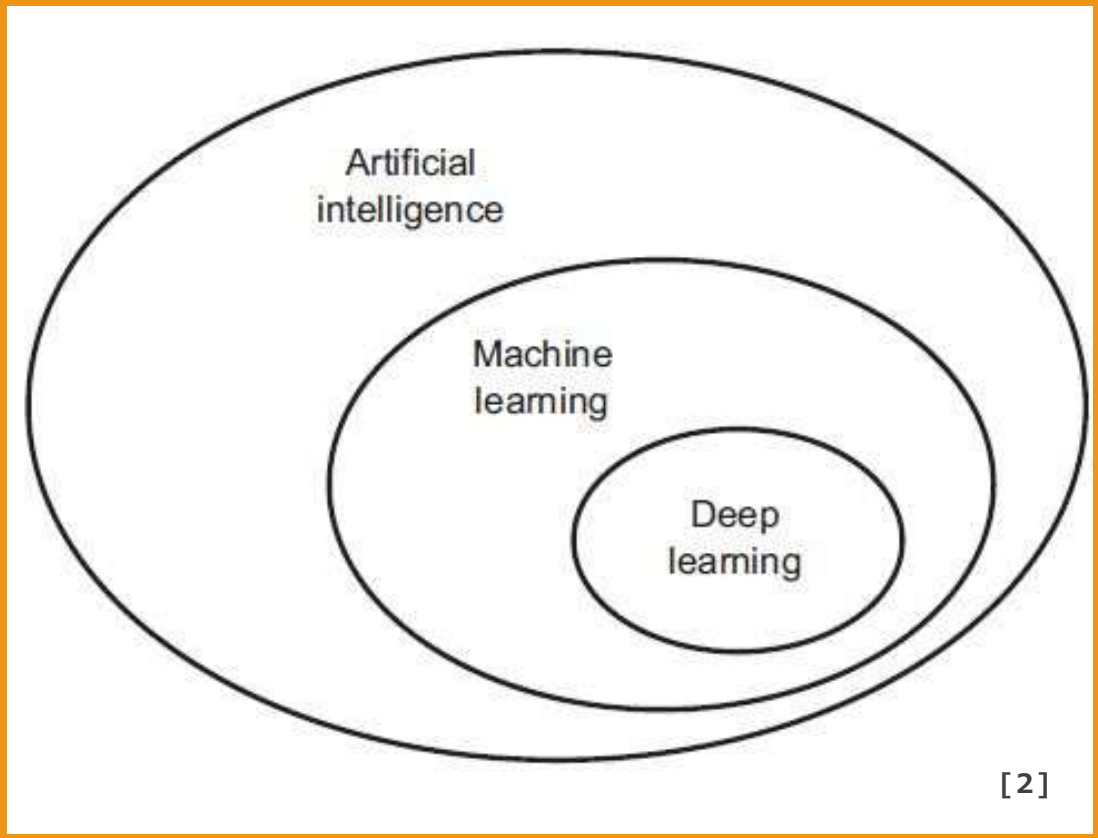
Math

- linear algebra
  - vectors, matrices, tensors
  - eigenvalues/eigenvectors
  - singular value decomposition (SVD)
- calculus
  - derivatives, gradients, jacobians, hessians
- statistics
  - probability, stochastic process

What is machine learning?

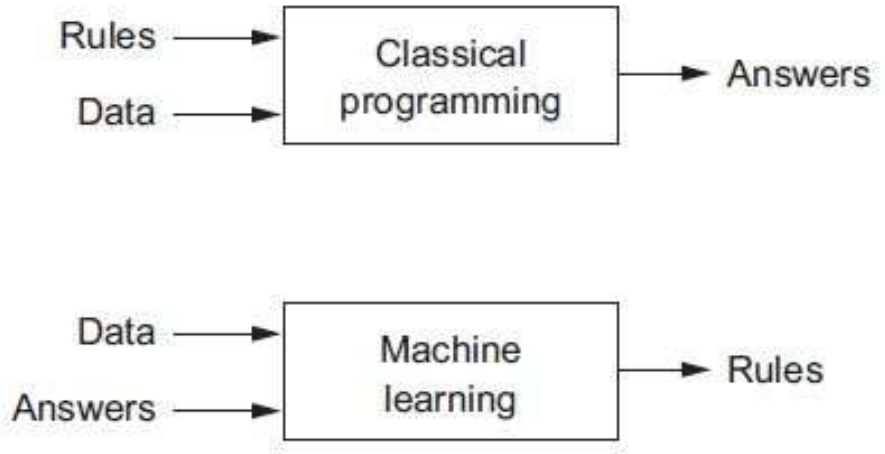
- Algorithms that analyze data and turn into knowledge
  - Detect patterns/structures
  - Classification/regression
  - Predict future events
- Subclass of AI (Artificial Intelligence)
- Trained instead of programmed

AI/ML/DL



[2]

New program paradigm in ML



[2]

3 types of ML

Supervised

- labeled data (“ground truth”)
- direct feedback
- predict outcome/future

Unsupervised

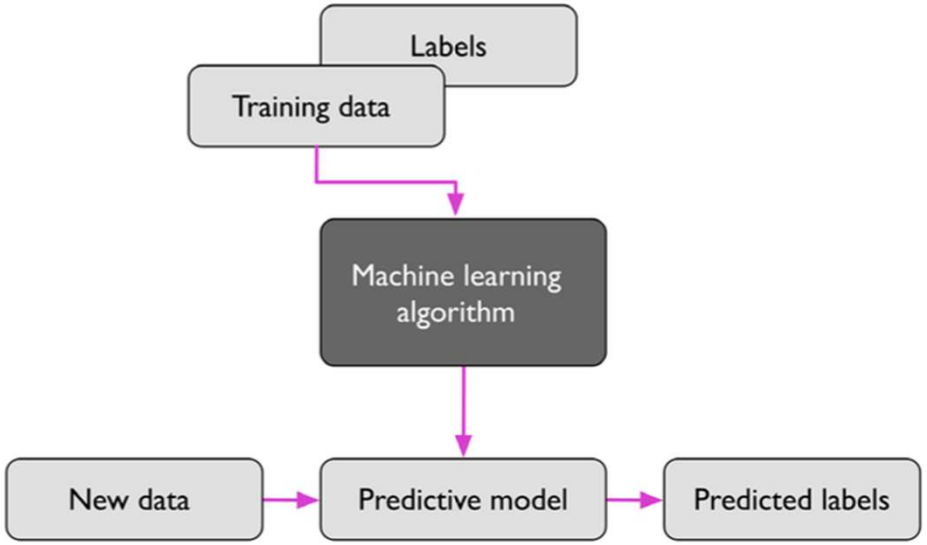
- no labels/targets
- no feedback
- find hidden structure in data

Reinforcement learning (RL)

- decision process
- reward system
- learn series of action

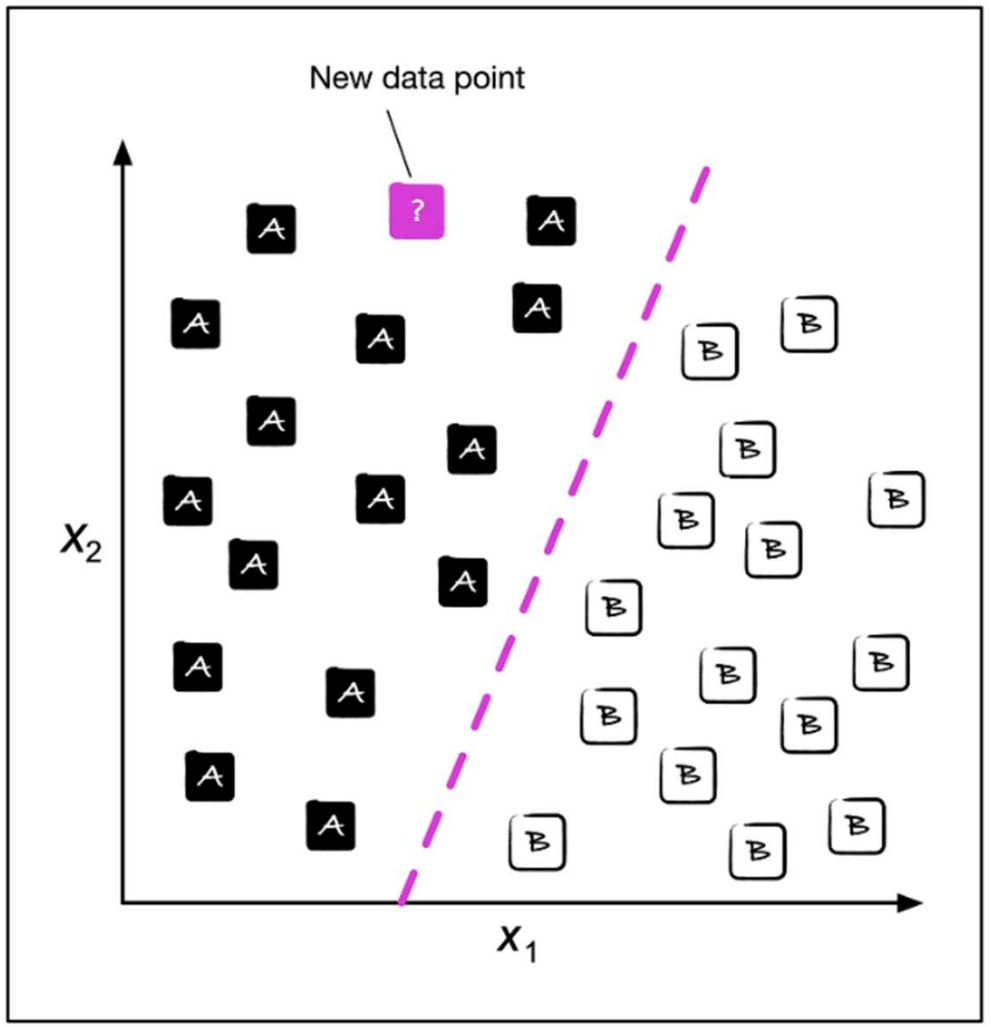
Supervised learning

Supervised learning process



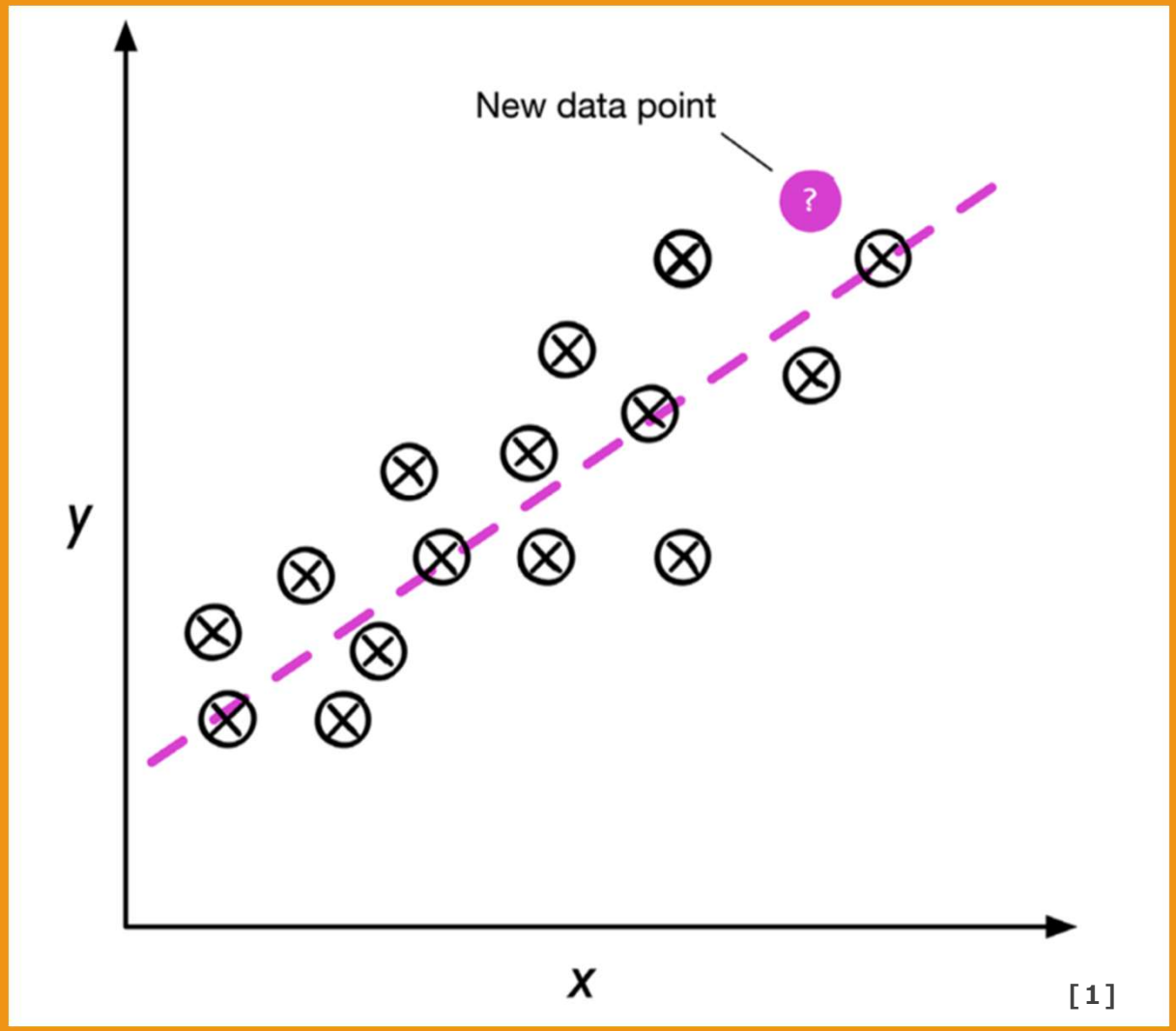
[1]

Classifying a new data point



[1]

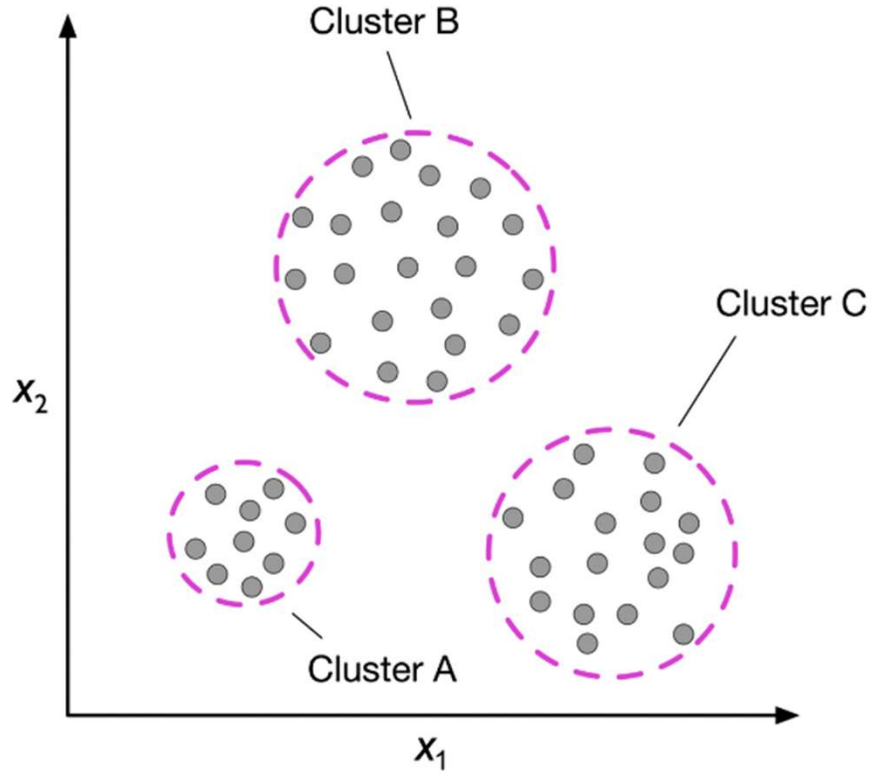
Linear regression



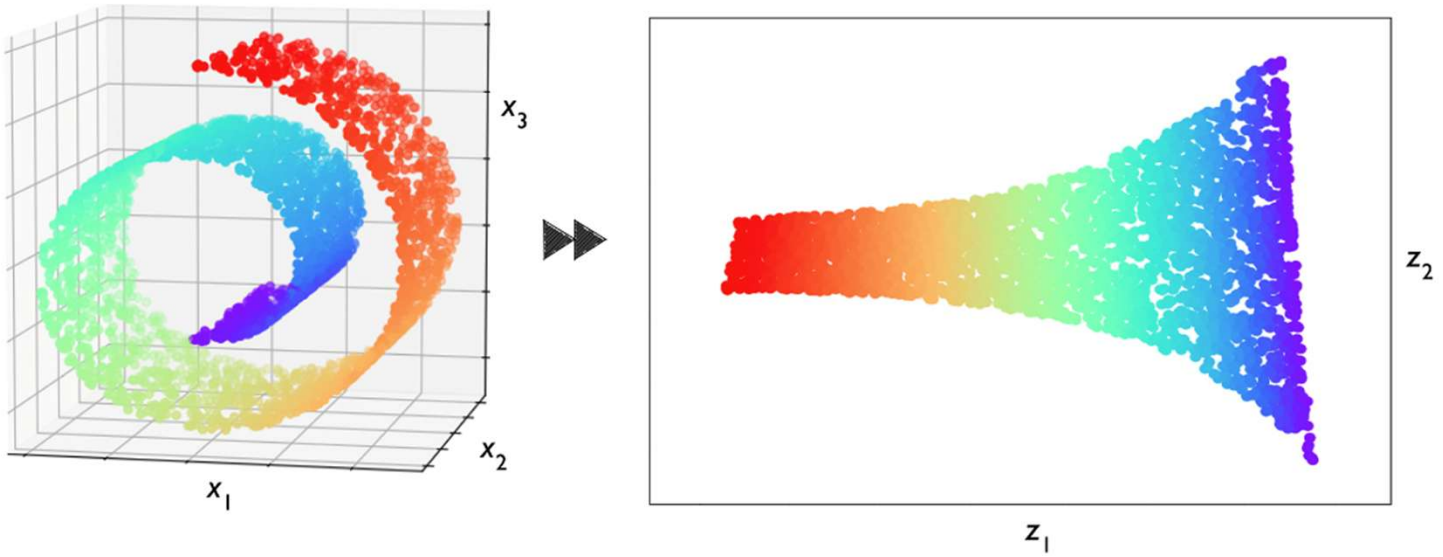
[1]

Unsupervised learning

Unsupervised classification : clustering



[1]

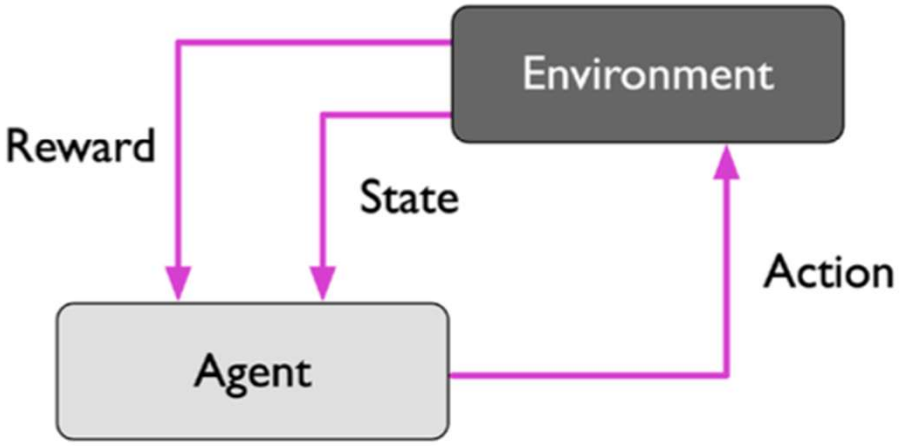


[1]

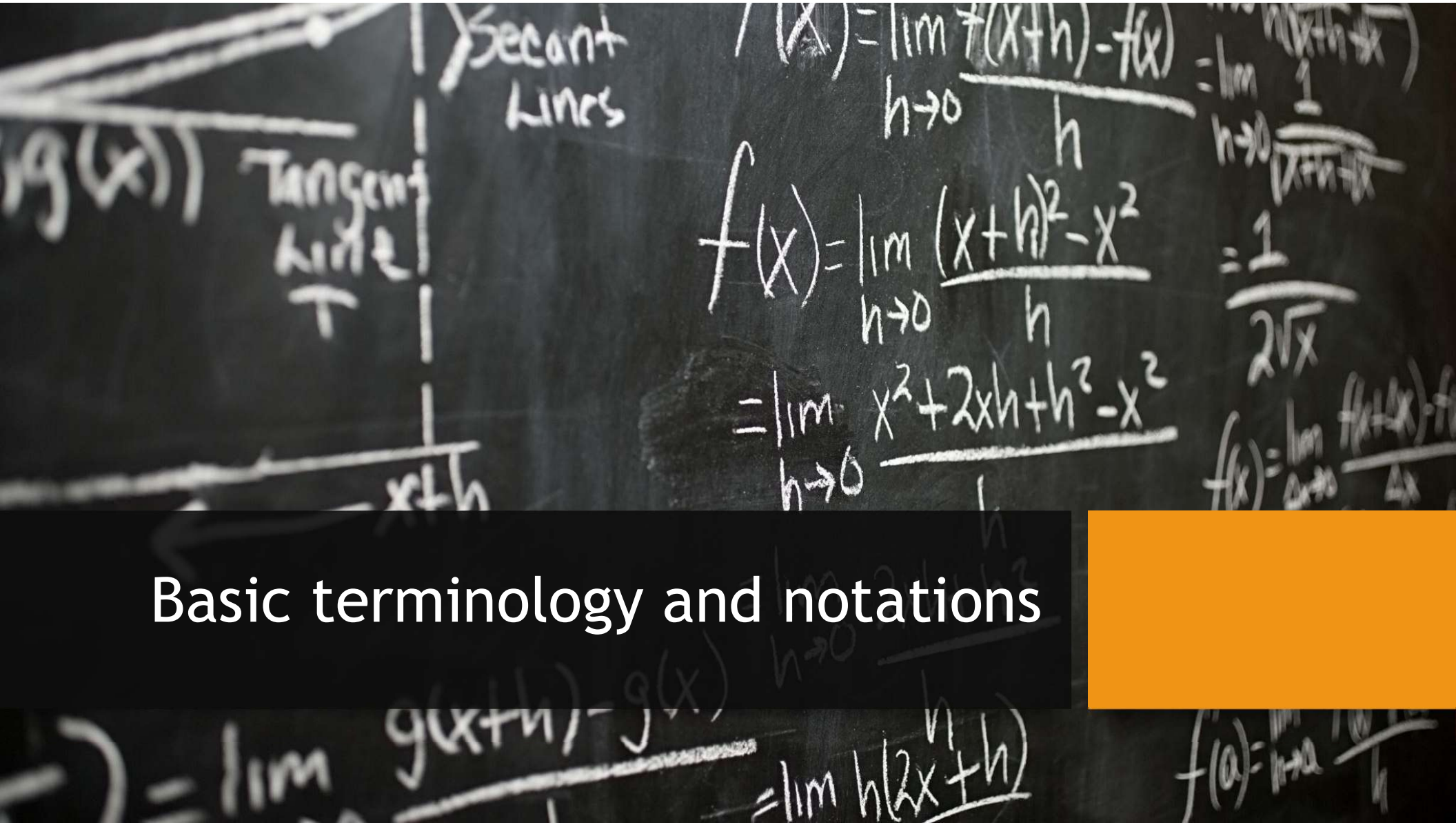
Dimensionality reduction

Reinforcement learning

Reinforcement learning process



[1]



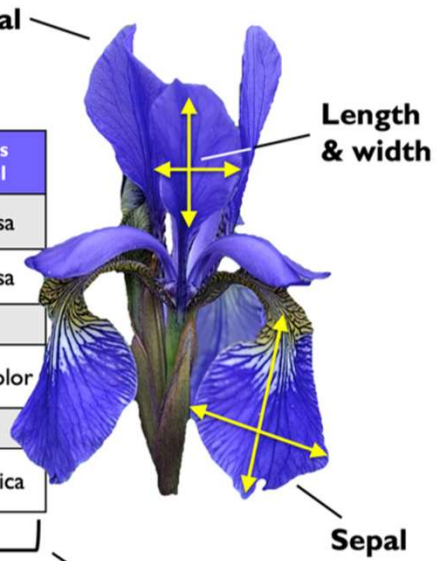
Basic terminology and notations

Samples
(instances, observations)

	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

Features
(inputs, attributes, measurements, dimensions)

Class labels
(targets, outcomes)



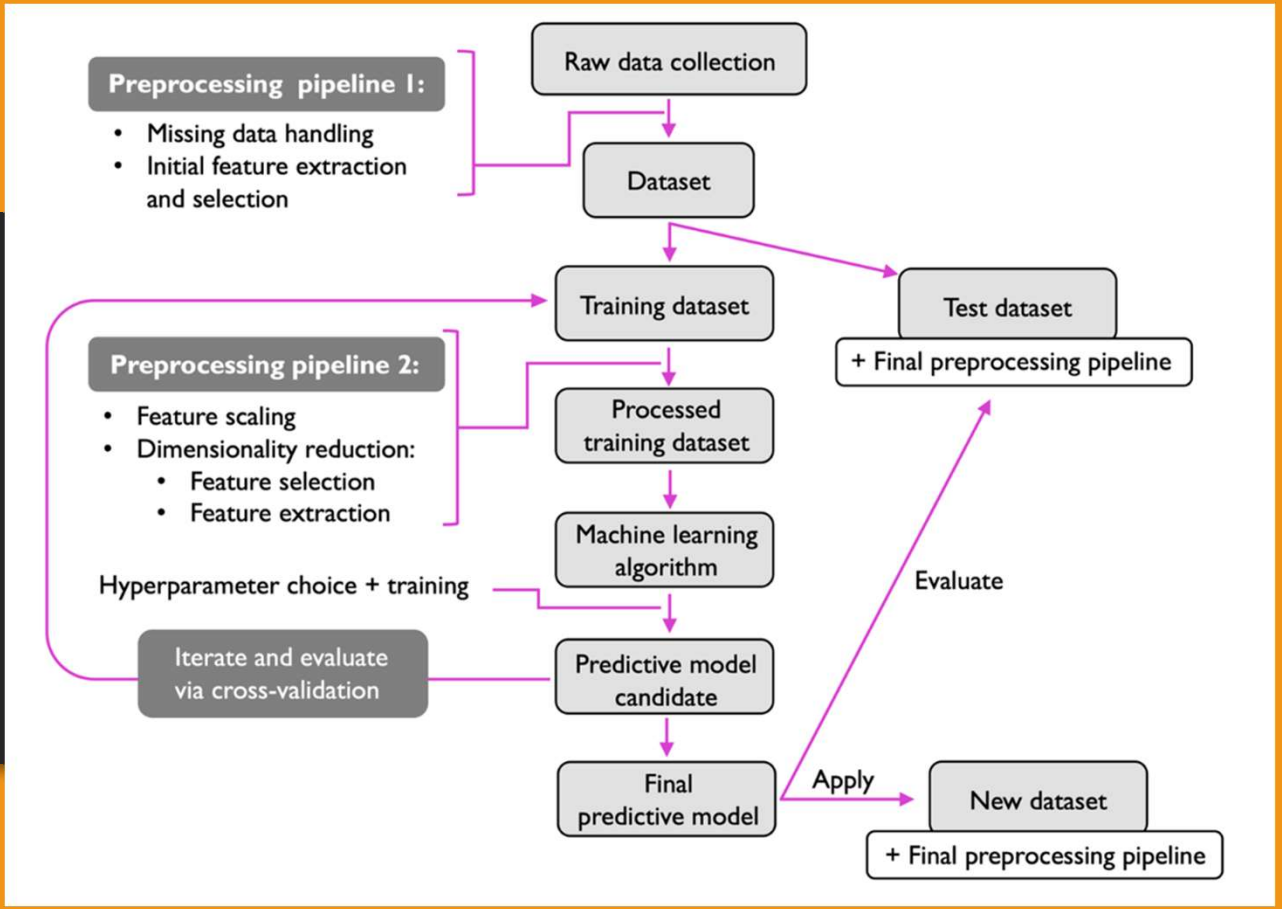
$: \boldsymbol{X} \in \mathbb{R}^{150\times 4}$:

$$\begin{bmatrix} x_1^{(1)} & x_2^{(1)} & x_3^{(1)} & x_4^{(1)} \\ x_1^{(2)} & x_2^{(2)} & x_3^{(2)} & x_4^{(2)} \\ \vdots & \vdots & \vdots & \vdots \\ x_1^{(150)} & x_2^{(150)} & x_3^{(150)} & x_4^{(150)} \end{bmatrix}$$

$$\boldsymbol{y} = \begin{bmatrix} y^{(1)} \\ \cdots \\ y^{(150)} \end{bmatrix}, \text{where } y^{(i)} \in \{\text{Setosa, Versicolor, Virginica}\}$$

Example : the iris dataset

Building machine learning system



Setup Python environment locally

Using conda

- conda install -y jupyter
- conda create -n mlpi
- conda activate mlpi
- conda install -c conda-forge nb_conda
- pip install scikit-learn
- python -m ipykernel install --user --name mlpi --display-name "mlpi”

Then launch jupyter notebook

Using docker

- Install docker desktop
- Create a work folder and put the provided config.yml and Dockerfile in there
- run docker desktop
- cd to that folder and type
  - docker compose up
- Open browser at local:8080
- Specify password to launch jupyter notebook

config.yml

```
services:
notebook:
build: ./
ports:
- 8080:8888
user: root
environment:
- DOCKER_DEFAULT_PLATFORM=linux/amd64
- JUPYTER_TOKEN=yourpassword
- NB_UID=1000
- GRANT_SUDO="yes"
volumes:
- ./:/home/jovyan
```

Dockerfile

```
FROM --platform=linux/amd64 quay.io/jupyter/pytorch-notebook

USER root

RUN pip install --upgrade pip && \
      pip install scikit-learn  && \
       fix-permissions "/home/${NB_USER}"
```

Some useful online courses

- Stanford CS229 (Andrew NG : Autumhn 2018) https://www.youtube.com/playlist?list=PLoROMvodv4rMiGQp3WXShtMGgzqpfVfbU
- Stanford CS229 (Spring 2022) https://www.youtube.com/playlist?list=PLoROMvodv4rNyWOpJg_Yh4NSqI4Z4vOYy
- MIT 6.036 (Fall 2020) https://www.youtube.com/playlist?list=PLxC_ffO4q_rW0bqQB80_vcQB09HOA3ClV

References

1. S. Raschka, Y. Liu and V. Mirjalili. Machine Learning with PyTorch and Scikit-Learn. Packt Publishing. 2022.
2. F. Chollet. Deep Learning with Python 2ed. Manning Publications Co. 2021.
3. A. Zhang, Z. Lipton, M. Li and J.Smola. Dive into Deep Learning. 2022. https://d2l.ai/
4. A. Ng. CS229 Lecture Notes. Stanford University. 2023. https://cs229.stanford.edu/main_notes.pdf
5. T. Broderick. Introduction to Machine Learning course (6.036). Massachusetts Institute of Technology. 2020. https://tamarabroderick.com/ml.html