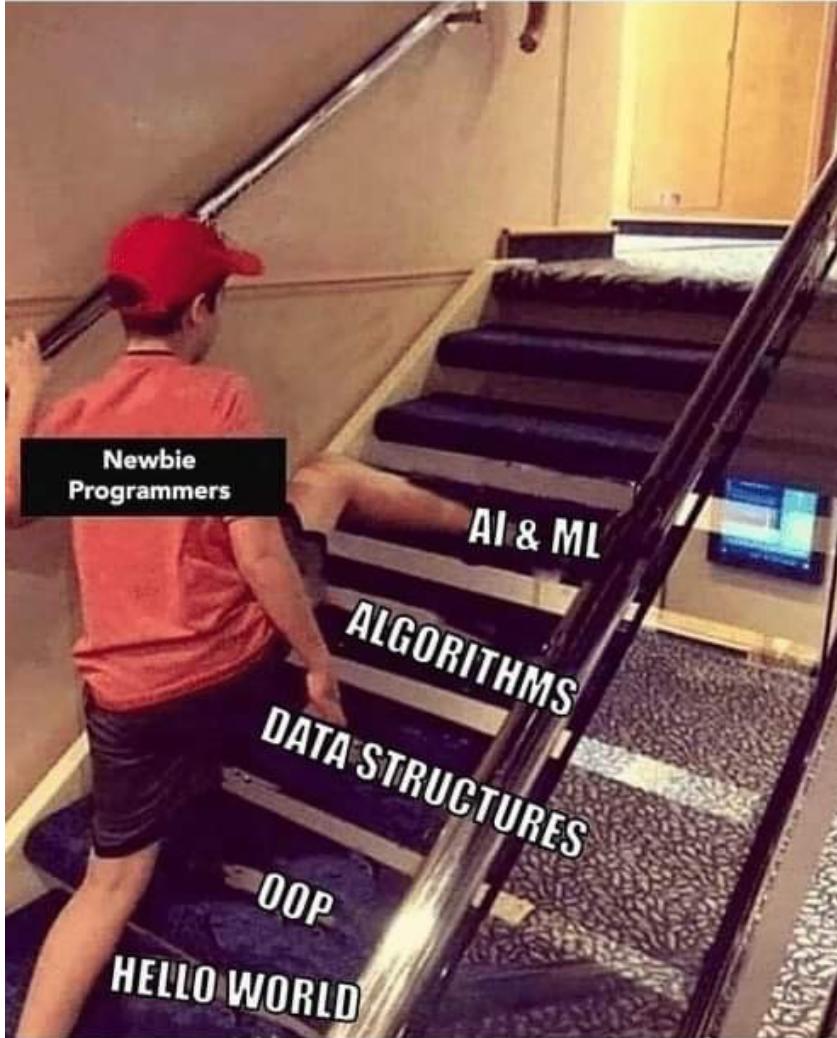


Machine Learning Basics

Prof. Kuan-Ting Lai

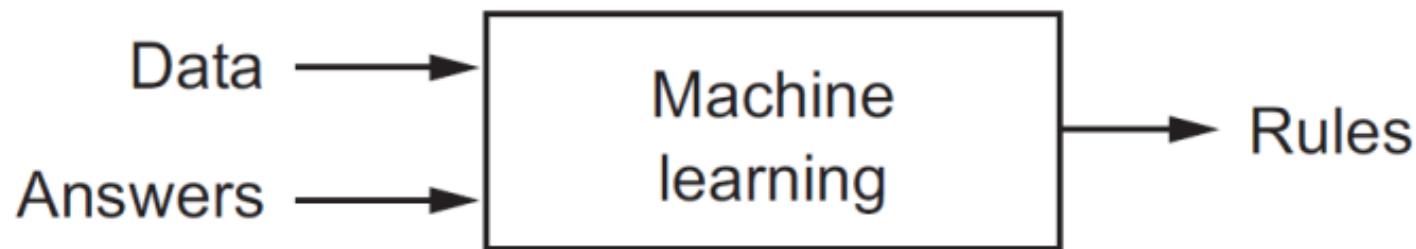
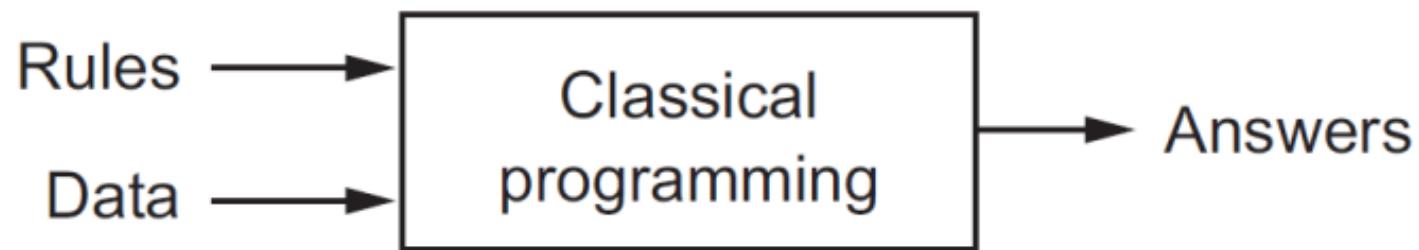
TA Alan Tian

2020/07/15



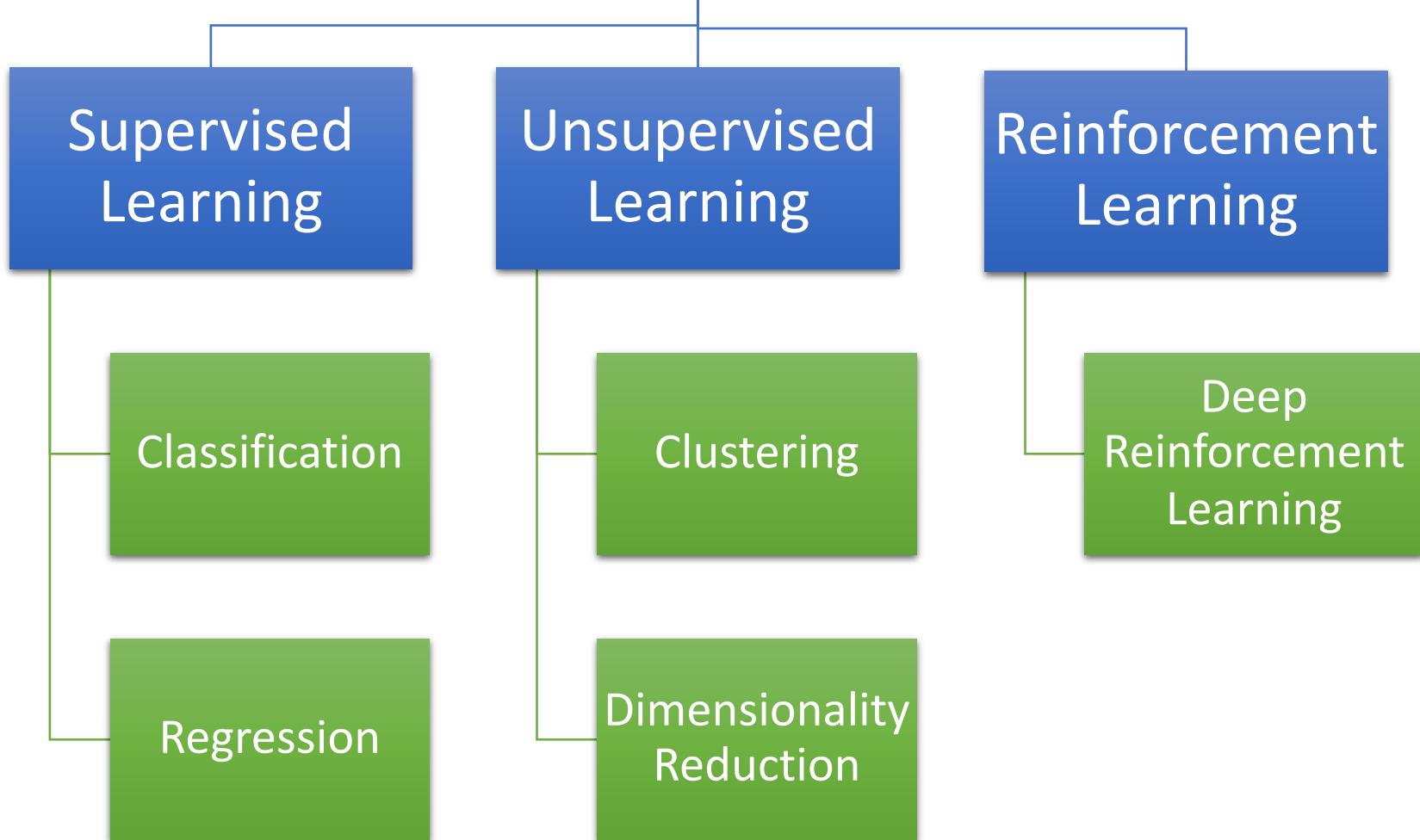
ML/AL expert without basic
knowledge?

Machine Learning

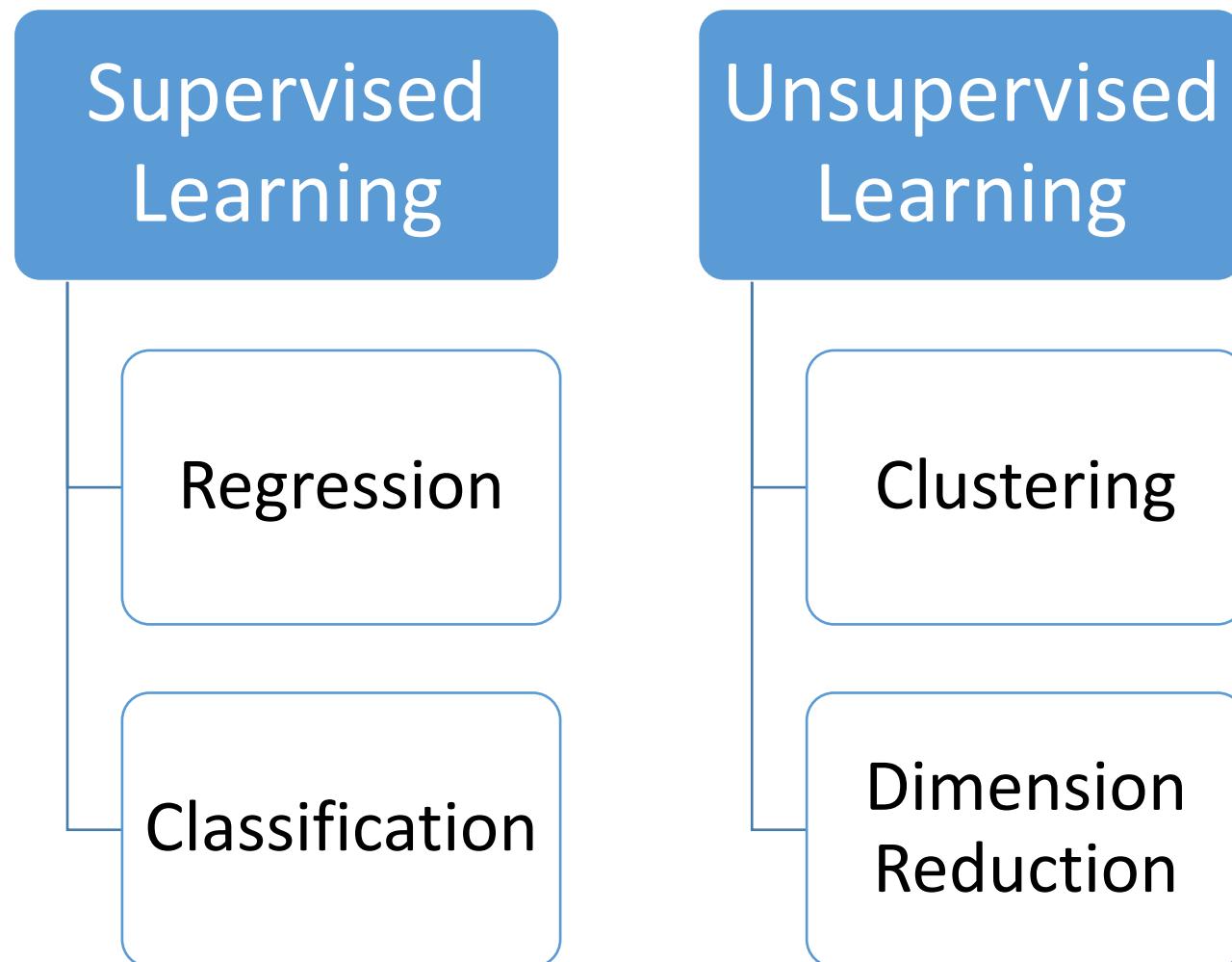


Francois Chollet, "Deep Learning with Python," Manning, 2017

Machine Learning

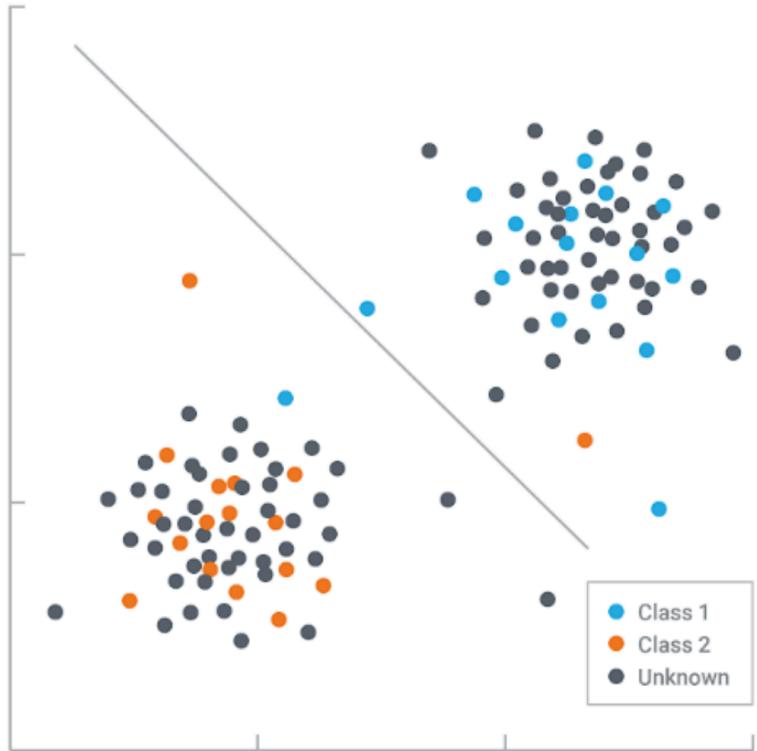


Supervised and Unsupervised Learning

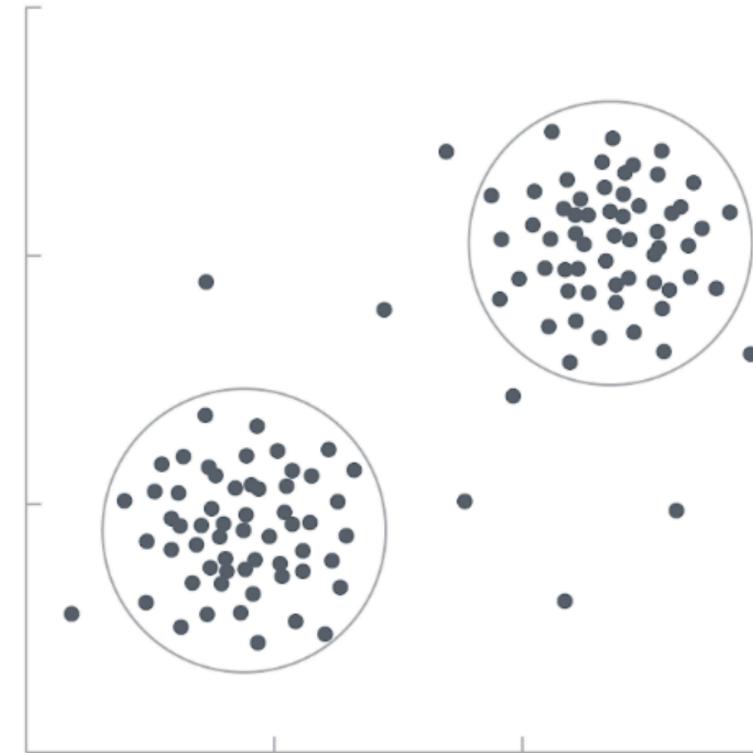


Supervised and Unsupervised Learning

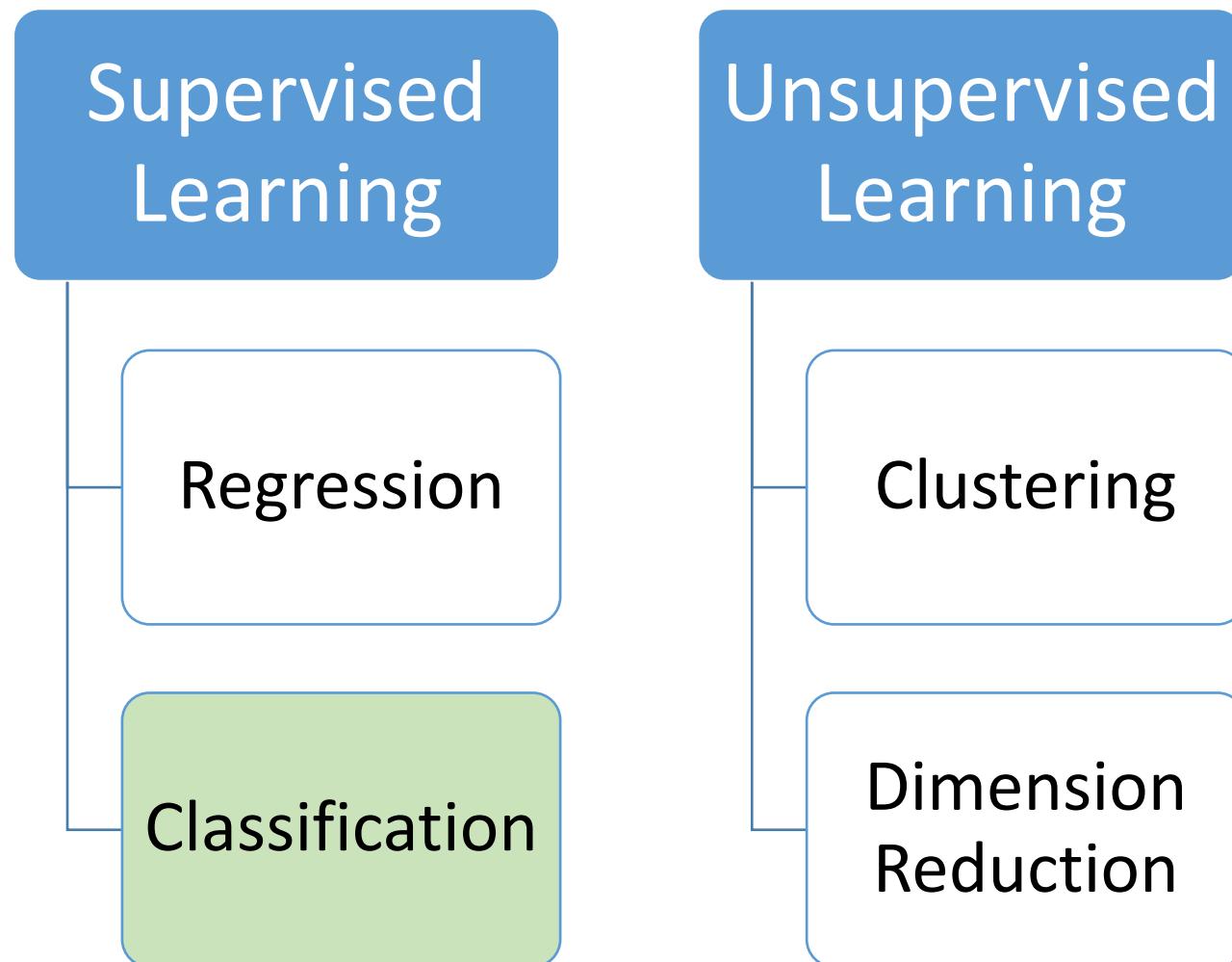
SUPERVISED



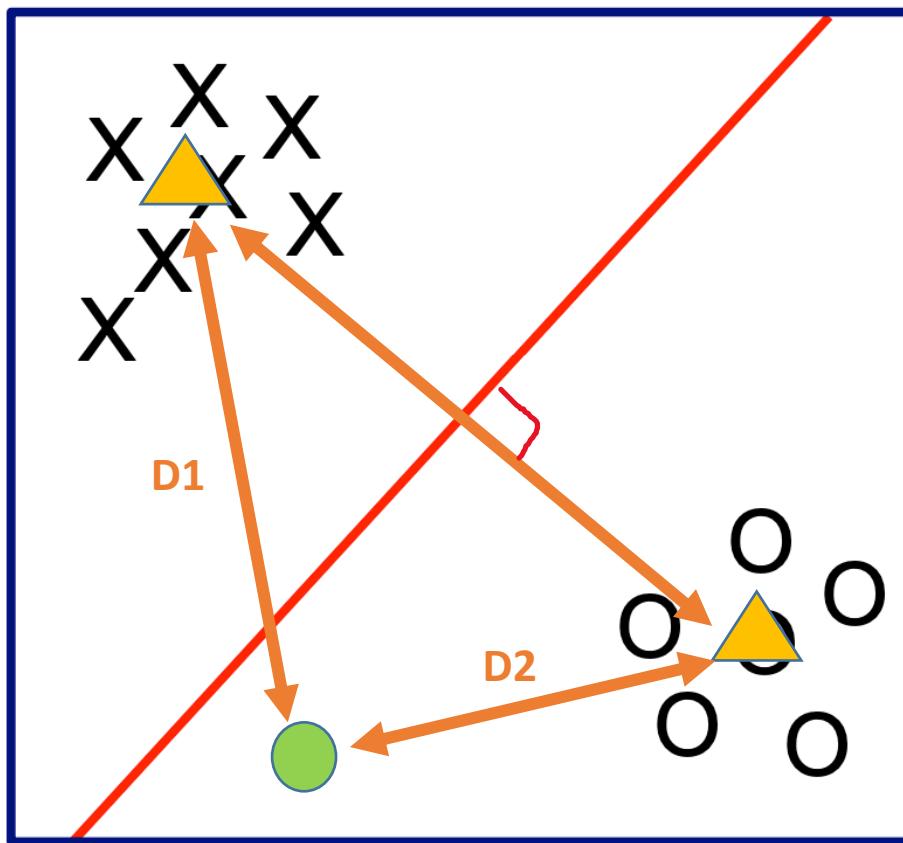
UNSUPERVISED



Supervised and Unsupervised Learning



Supervised – Minimum Distance Classifier



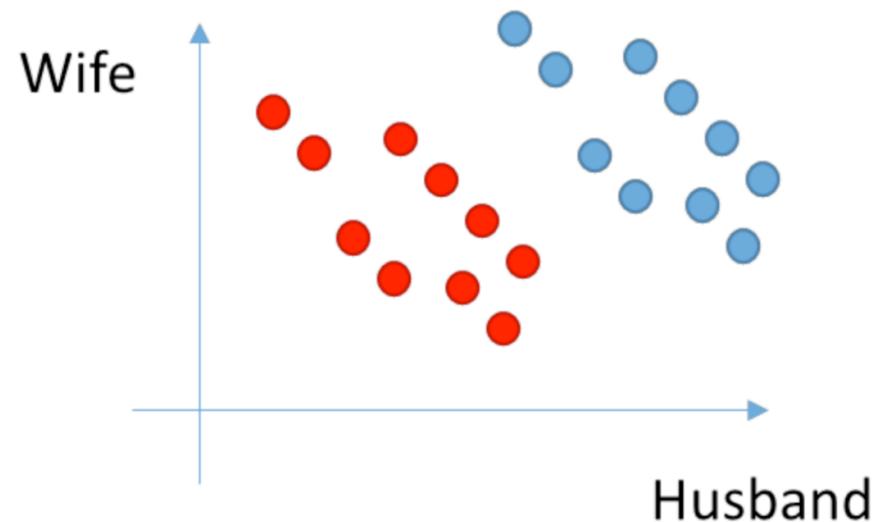
if $D_2 < D_1$:
We classify the
green point to O

Supervised – Minimum Distance Classifier

- Calculate the distance between data to 2 centroids respectively.
- Classify the unknown data by the distance.

[Practice 1 on Colab](#)

- Couple income and stressed level
- [detail](#)

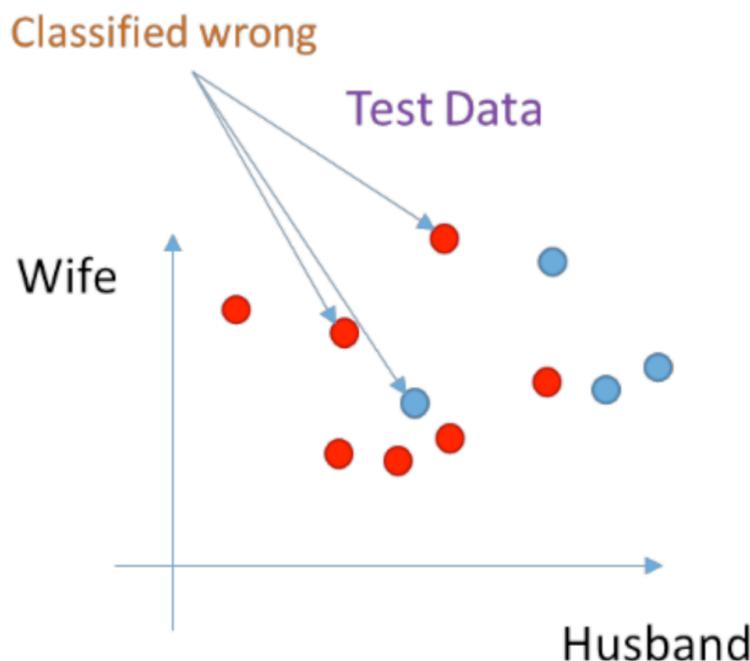


Practice 1 on Colab

- Couple income and stressed level
- Practice Detail:<https://drive.google.com/file/d/1n4NIwAz4YvWagh3pJbZe0guFVsB7Rlo/view?usp=sharing>
- Colab code:
https://colab.research.google.com/drive/1rCoDKLgEvswl1NeaSx_TwusfxbrlhJyp
- Data1 :<https://drive.google.com/file/d/1NvV6BQ9NZSAeKDA9Mb721xboG2ZH-fjL/view?usp=sharing>
- Data2:<https://drive.google.com/file/d/17e80Gdhjxhf9dSL-gocO1CHimegYSrb9/view?usp=sharing>

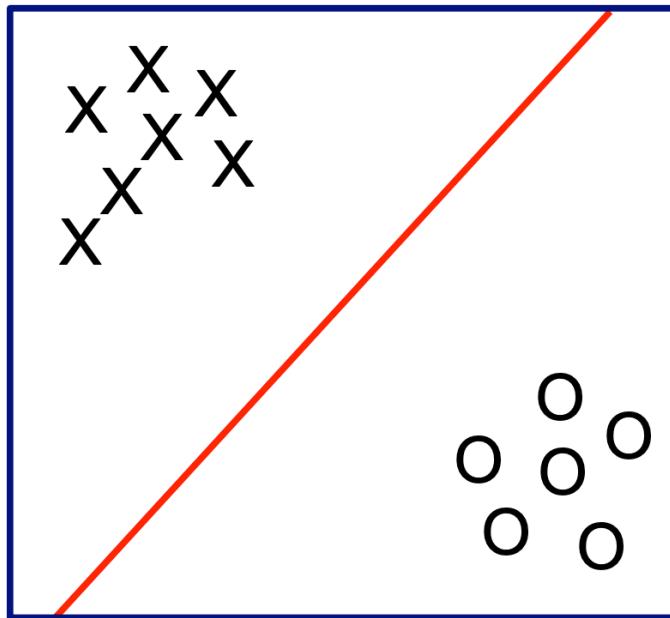
Supervised – Minimum Distance Classifier

Find means of two classes in the training data. Use the two means and classify the test data. Plot your classification result and specify the number of points that were classified wrong.

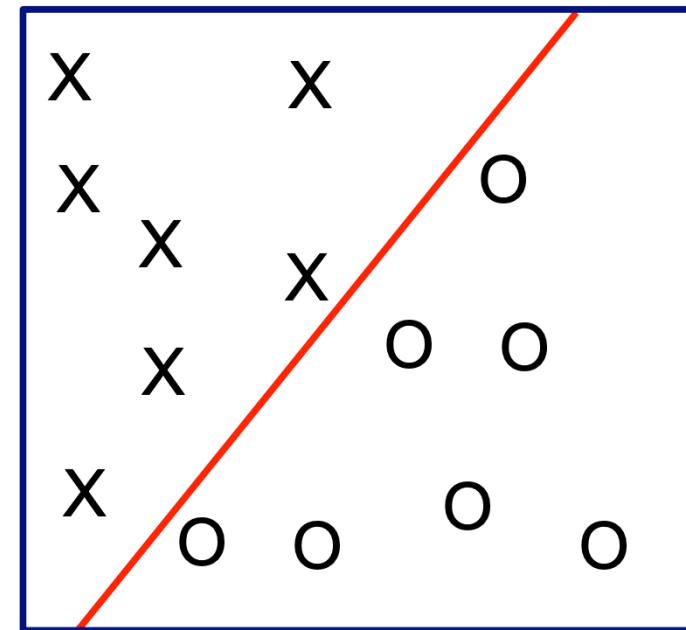


Supervised – Minimum Distance Classifier

Minimum distance classifiers work well if interclass distances are large compared to class spreads.



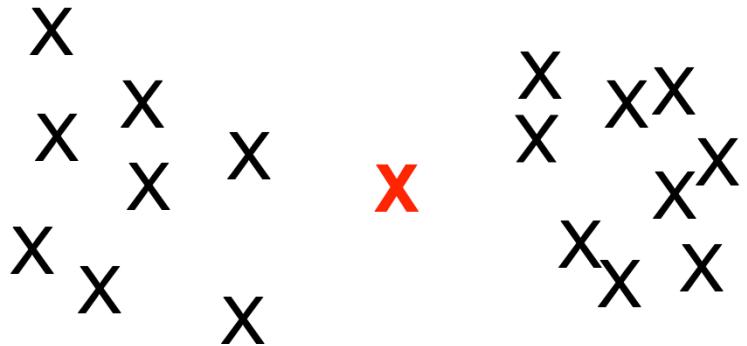
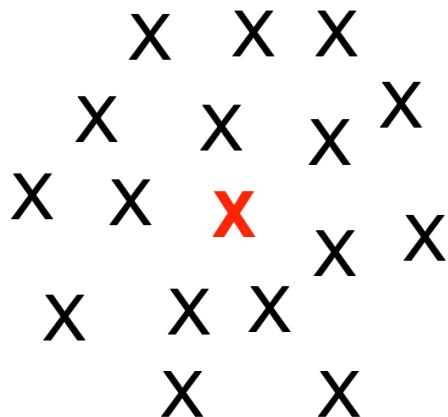
Well-separated classes



Cutting it close – new data may fall on the wrong side.

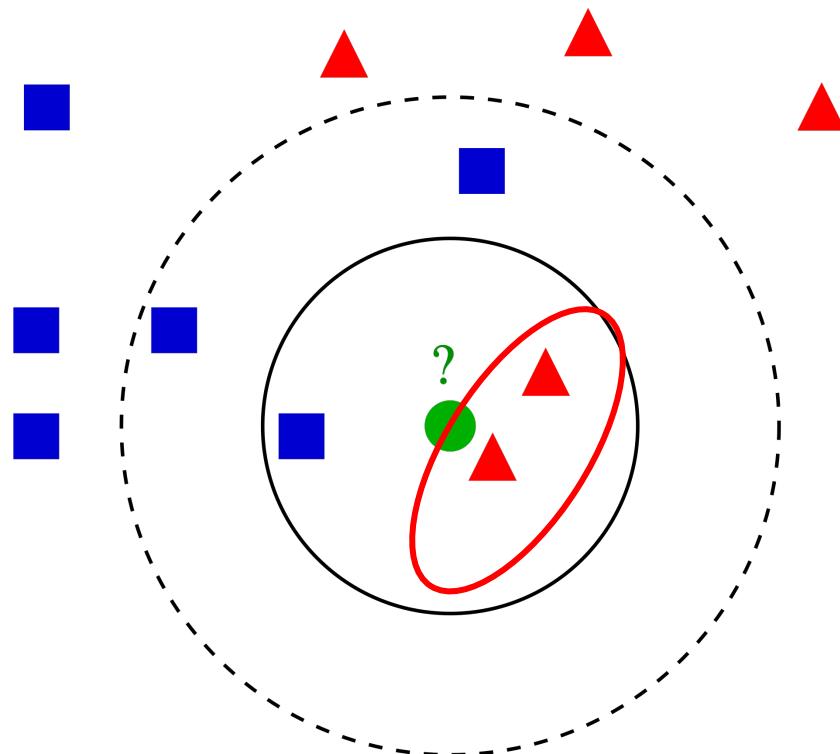
Supervised – Minimum Distance Classifier

The MDC approach assumes that the **class mean**, is a good representative of the class.



The class mean is a good exemplar of the class.

K Nearest-Neighbors (KNN)



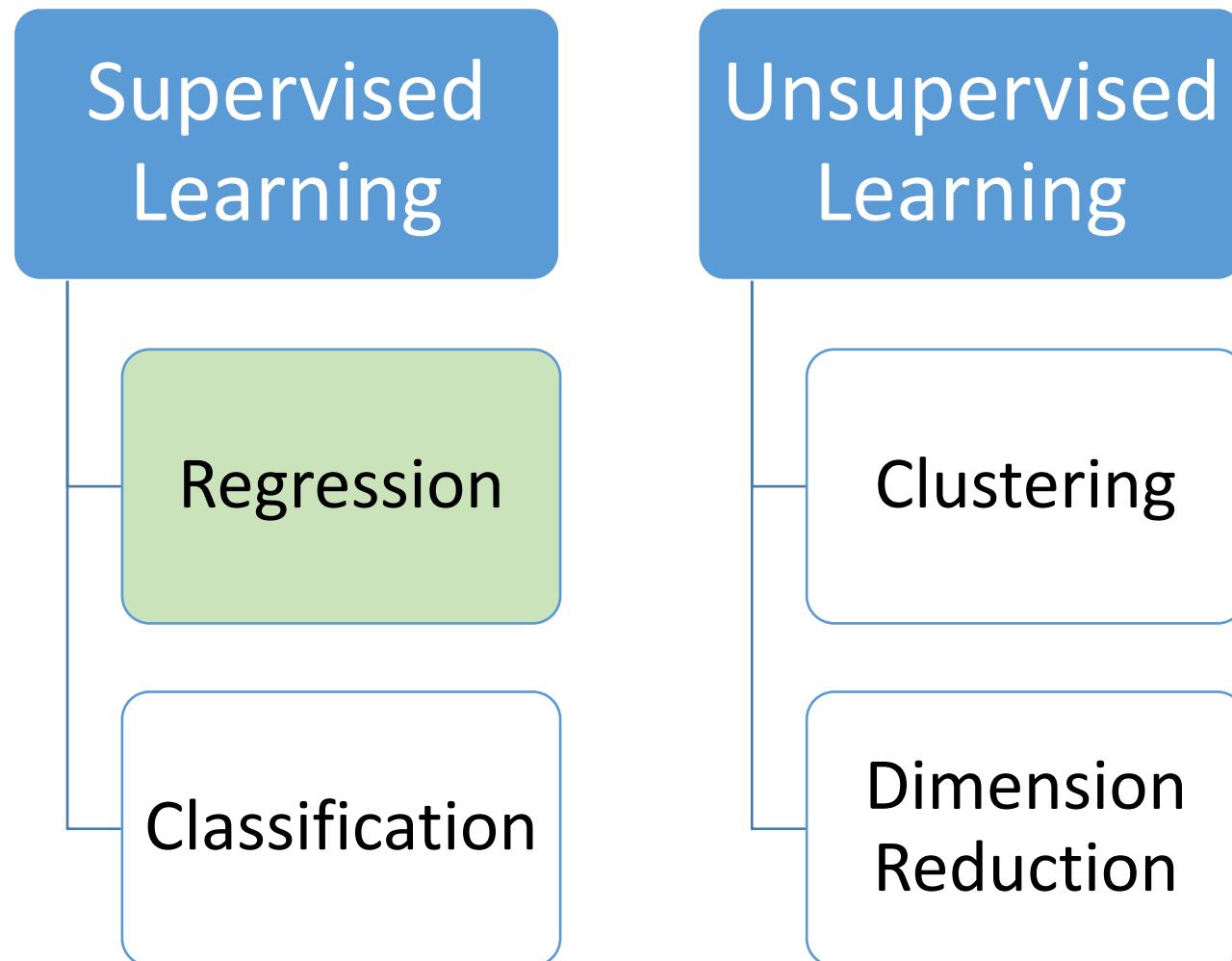
- Pick a number of K
- Find K points that is nearest to the green point
- Vote the class

If pick K = 3

Totally 2 ▲ 1 ■

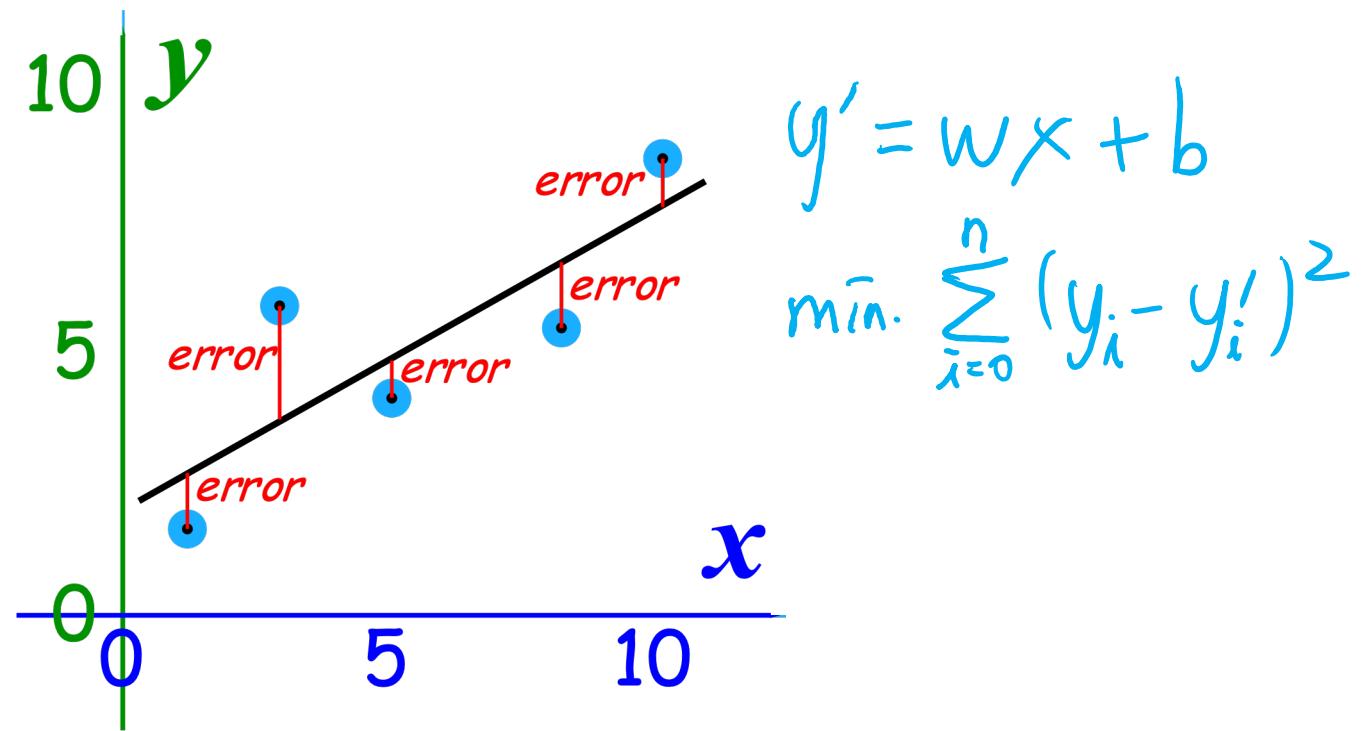
$$\rightarrow \text{●} = \text{▲}$$

Supervised and Unsupervised Learning



Linear Regression (Least squares)

- Find a "line of best fit" that minimizes the total of the square of the errors



References

- Francois Chollet, “Deep Learning with Python”, Chapter 4
- Pedro Domingos, “A Few Useful Things to Know about Machine Learning,” Commun. ACM, 2012
- <https://ml-cheatsheet.readthedocs.io/en/latest/index.html>
- <https://machinelearningmastery.com/roc-curves-and-precision-recall-curves-for-classification-in-python/>
- <https://towardsdatascience.com/data-types-in-statistics-347e152e8bee>
- https://en.wikipedia.org/wiki/Naive_Bayes_classifier