Introduction to Machine Learning

Course Logistics

林彥宇 教授
Yen-Yu Lin, Professor

國立陽明交通大學 資訊工程學系
Computer Science, National Yang Ming Chiao Tung University

About Yen-Yu Lin

- Work Experience
  - Professor, *CS, NCTU,* August 2019 ~ present
  - Associate research fellow, *CITI, Academia Sinica*, 2015 ~ 2019
  - Assistant research fellow, *CITI, Academia Sinica*, 2011 ~ 2015

- Research interests
  - Computer Vision (**CV**):

    Let computers see, recognize, and interpret the world like humans
  - Machine Learning (**ML**):

    Provide a statistical way to learn how human visual system works
  - Goal: *Design **ML** methods to facilitate **CV** applications*

Today's agenda

- Course logistics
- Course overview

Today's agenda

- Course logistics
- Course overview

How to choose and take this course?

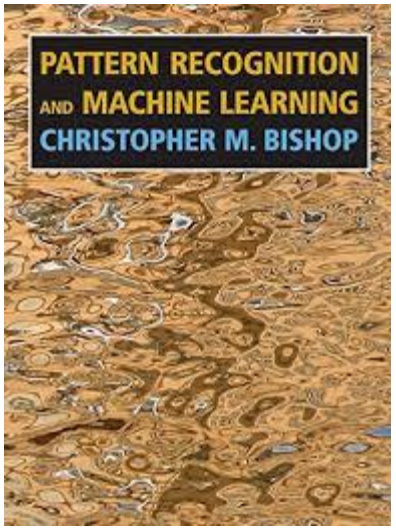
- Please use the online course management system
  - Max number: 90 -> 105 students
- I do not plan to add additional students
  - The size of the classroom
  - The loading of our TAs
  - Considering taking the same course offered by another professor
  - If you have some reason why you must take this course, send me an email with the reason
- Be a guest student?
  - Yes. Send TAs an email with your student ID. We will add you to the student list on E3

Instructor and teaching assistants

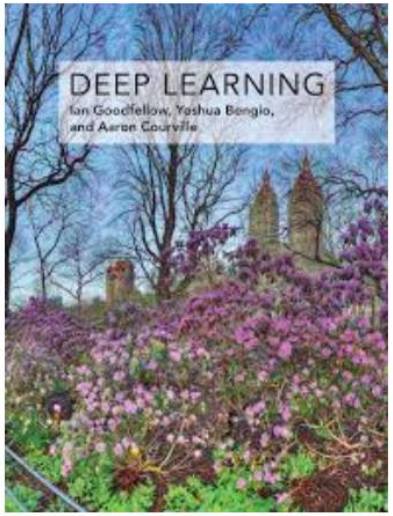
- Instructor: Yen-Yu Lin 林彥宇
  - Email: lin@cs.nctu.edu.tw
  - Office: EC706 (please email me first)
- Teaching assistants:
  - Jui-Che Chiang 江睿哲 Email: benchiang.cs07@nctu.edu.tw
  - Wei-Hsiang Yu 游為翔 Email: weihsiang.yu@gmail.com
  - Ji-Jia Wu 吳季嘉 Email: jijiawu.cs@gmail.com
  - Si-Yu Huang 黃思瑜 Email: stella900604@gmail.com
- Office hour (email first)
  - 4:20 pm ~ 5:20 pm on Tuesdays at EC701 and EC234-C

Textbook

- Pattern Recognition and Machine Learning
  - Christopher Bishop
  - Springer-Verlag, Berlin, 2006
  - Free online at
    https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf



- Deep learning (optional)
  - I. Goodfellow, Y. Bengio, and A. Courville
  - MIT Press, 2016
  - Free online at
    https://www.deeplearningbook.org/



Grading policy

- Four homework assignments: 72% (= 18% x 4)
- For each assignment
  - You are required to implement machine learning algorithms and complete some short answer questions
  - Late policy: 20% off per late day
- Final project: 28%
  - Join a competition on Kaggle

Syllabus

	1	2023-09-12(二)	Introduction
	2	2023-09-19(二)	Linear Model for Regression
HW1	3	2023-09-26(二)	Linear Model for Classification I
	4	2023-10-03(二)	No lecture: The instructor will attend a conference.
	5	2023-10-10(二)	No lecture: Holiday
	6	2023-10-17(二)	Linear Model for Classification II
HW2	7	2023-10-24(二)	Neural Networks
	8	2023-10-31(二)	Ensemble Model I
	9	2023-11-07(二)	Ensemble Model II
HW3	10	2023-11-14(二)	Kernel Method I
	11	2023-11-21(二)	Kernel Method II
HW4	12	2023-11-28(二)	Deep Neural Networks (DNN)
Final Project	13	2023-12-05(二)	Convolutional Neural Networks (CNN) I
	14	2023-12-12(二)	Convolutional Neural Networks (CNN) I and Attention and Transformers I
	15	2023-12-19(二)	Guest lectures: 1. Prof. Chen Yu with UT Austin; 2. David J. Crandall with Indiana University
	16	2023-12-26(二)	Attention and Transformers II

Pre-requisite

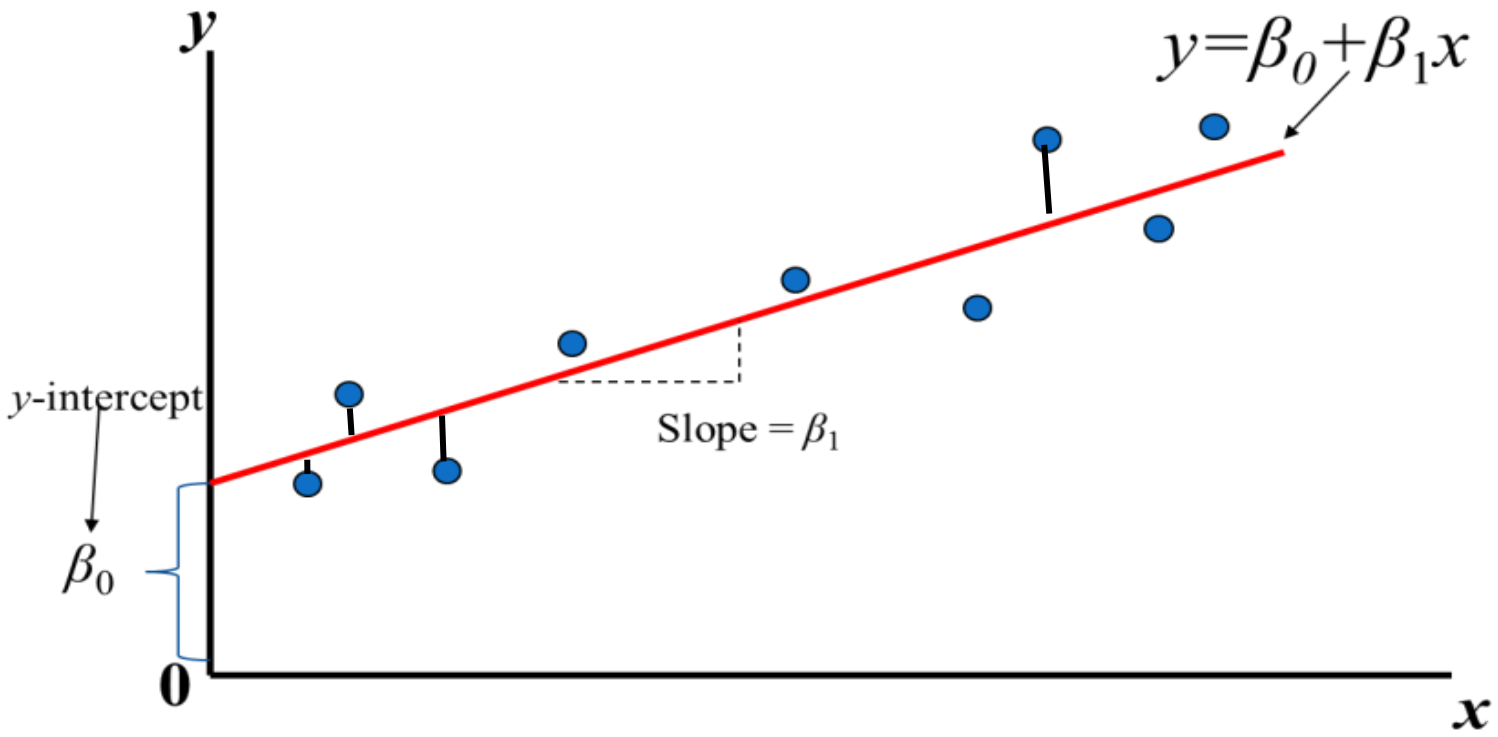
- Linear algebra, probability, calculus, and programming

- Python
  - We strongly encourage students who are not familiar with Python to complete the following tutorial first
  - http://cs231n.github.io/python-numpy-tutorial/

- One deep learning framework, Pytorch or Keras
  - Pytorch: https://pytorch.org/tutorials/
  - Keras: https://elitedatascience.com/keras-tutorial-deep-learning-in-python

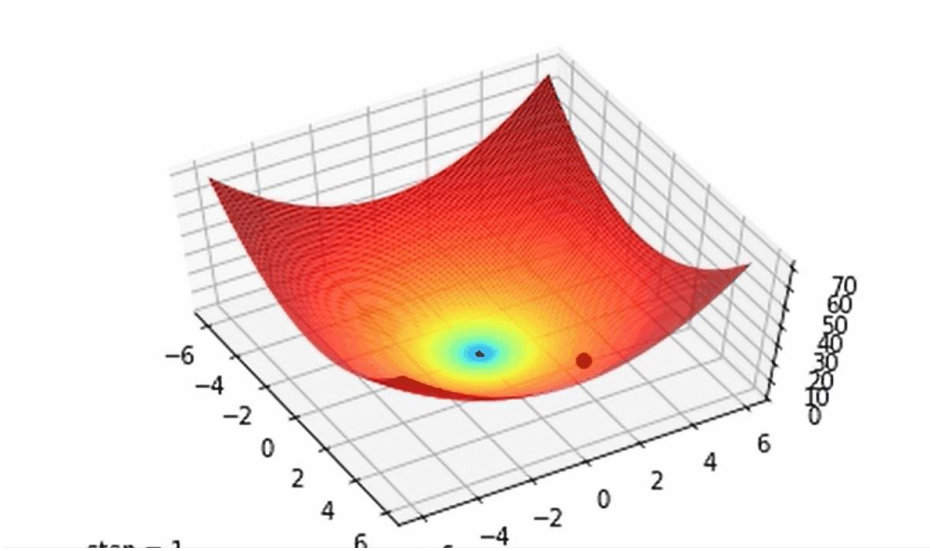
Homework 1: Linear regression (last year)

- Find the value of β0 and β1



Gradient descent

- x-axis and y-axis represent the values of two variables
- z-axis represents the loss of the corresponding variables
- Targets: Find the variable values that minimize the loss



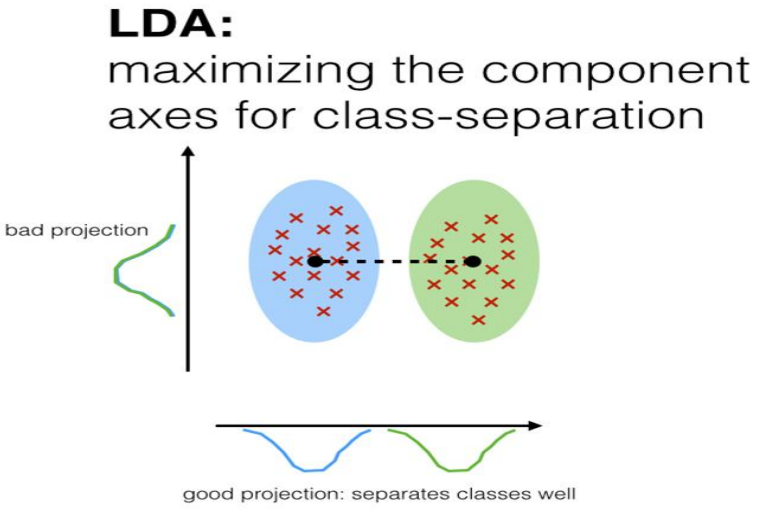
Gradient descent pseudo code

Algorithm

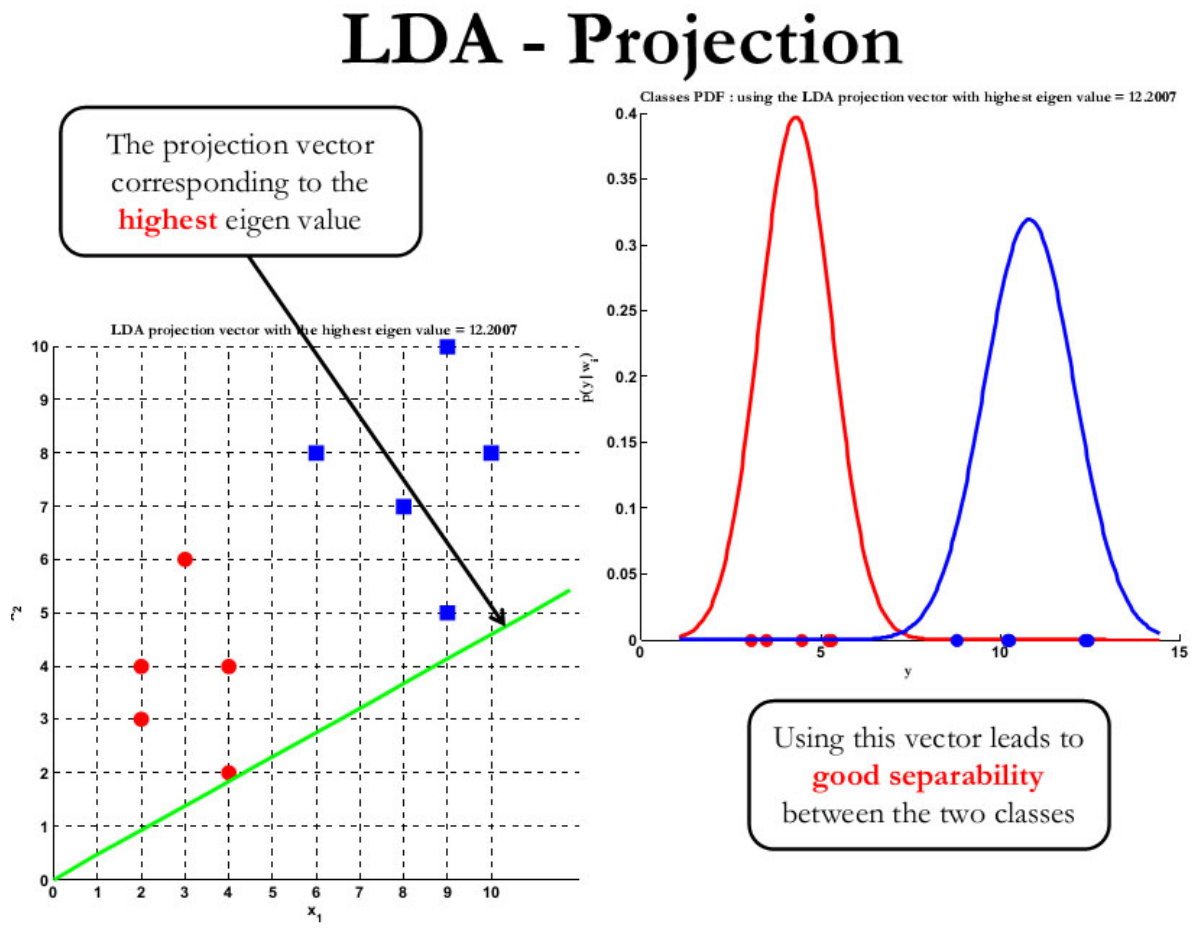
1. Initialize weights randomly $\sim N(0, \sigma^2)$
2. Loop until convergence:
   i. Pick batch of B data points
   ii. Compute gradient. $\frac{\partial J(\Theta)}{\partial \Theta} = \frac{1}{B}\sum_{k=1}^{B} \frac{\partial J_k(\Theta)}{\partial \Theta}$
   iii. Update weights $\theta <- \theta - \eta \frac{\partial J(\Theta)}{\partial \Theta}$
3. Return weights

Homework 2: Fisher's linear discriminant (last year)

- FLD (or LDA) is a "supervised" method and computes the directions representing the axes that maximize the separation between multiple classes.
- FLD seeks the projection **w** that gives a large distance between the projected data means while giving a small variance within each class

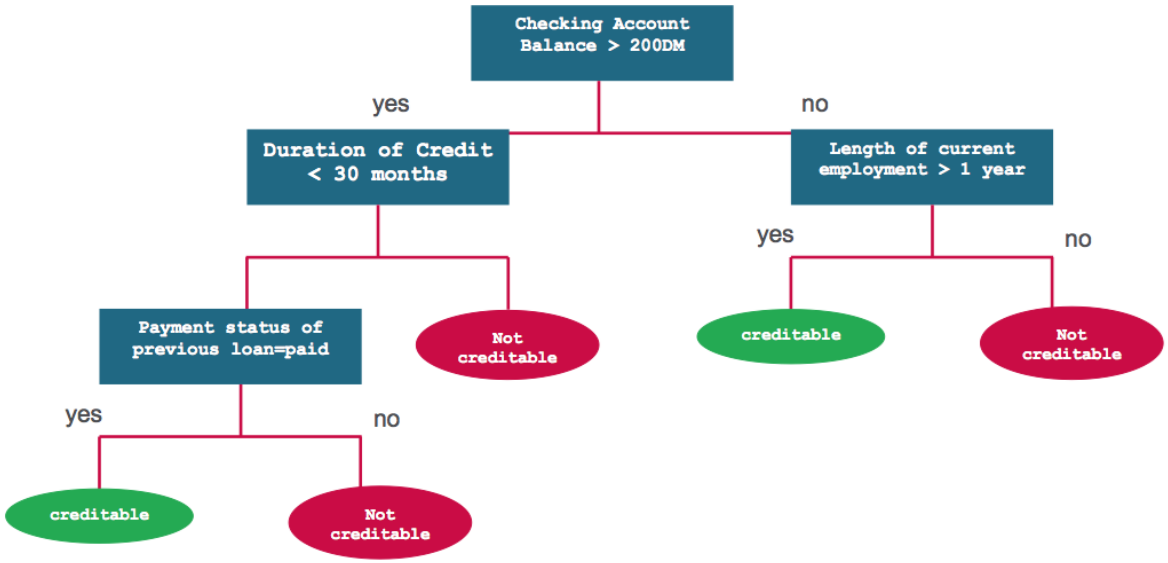


Eigenvalue problem



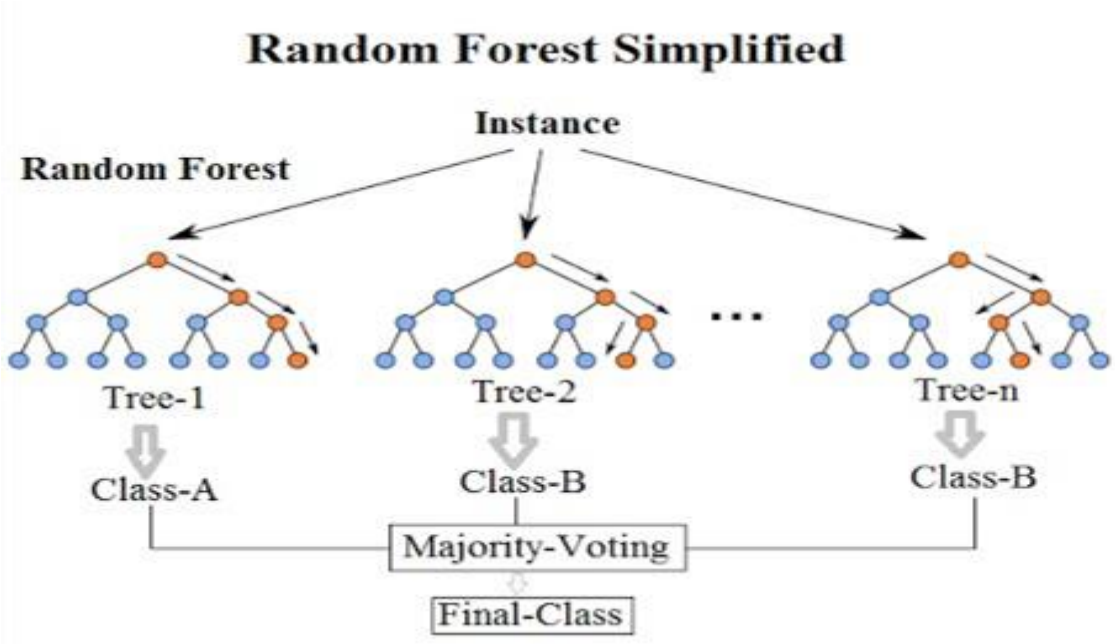
Homework 3: Decision tree algorithm (last year)

- How to find the feature for making decisions? What's the value of feature?
- Find the features to separate data that the class at the resulting nodes are as **pure** as possible



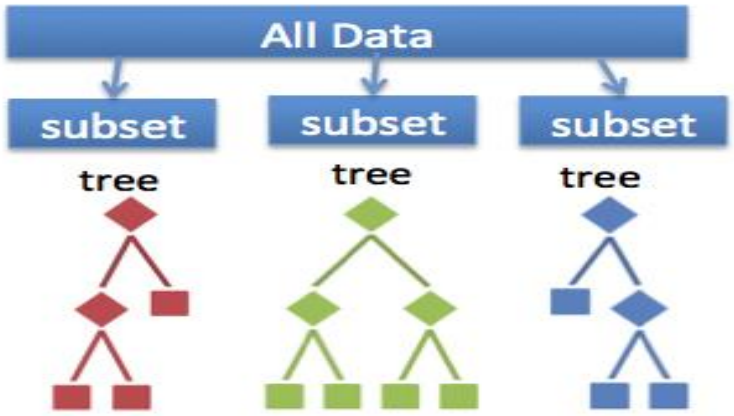
Ensemble method of decision trees: Bagging

- **Bagging (Bootstrap aggregating)**: Fit many large trees to bootstrap-resampled versions of the training data, and classify by majority vote



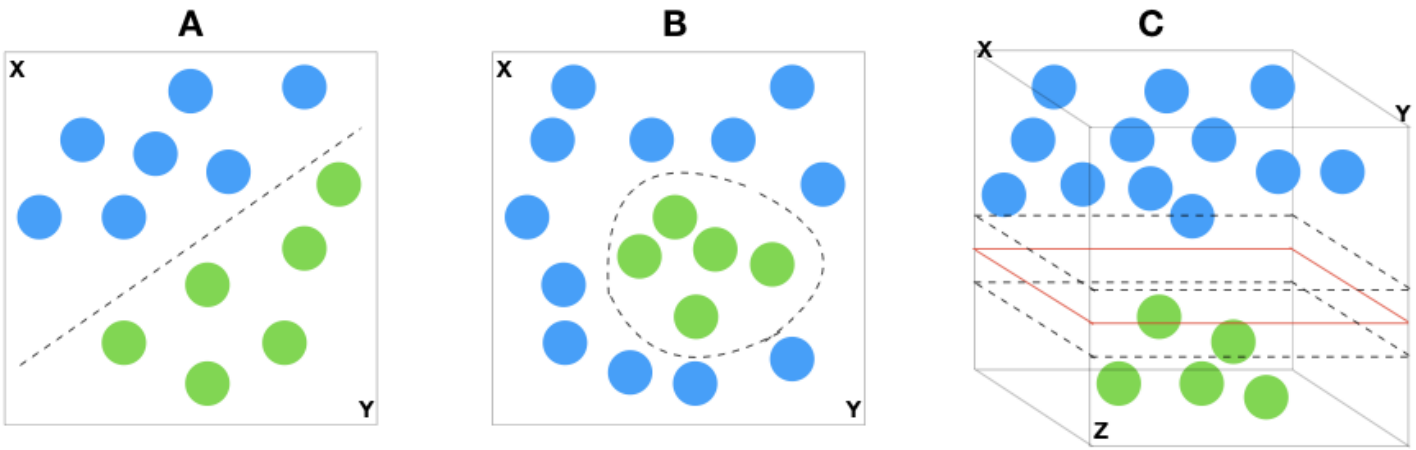
Another ensemble method: Random Forest

- Bootstraped dataset
- Each tree in the forest may grow with different data and features
- Which features or data to be used is **randomly** sampled to grow the tree



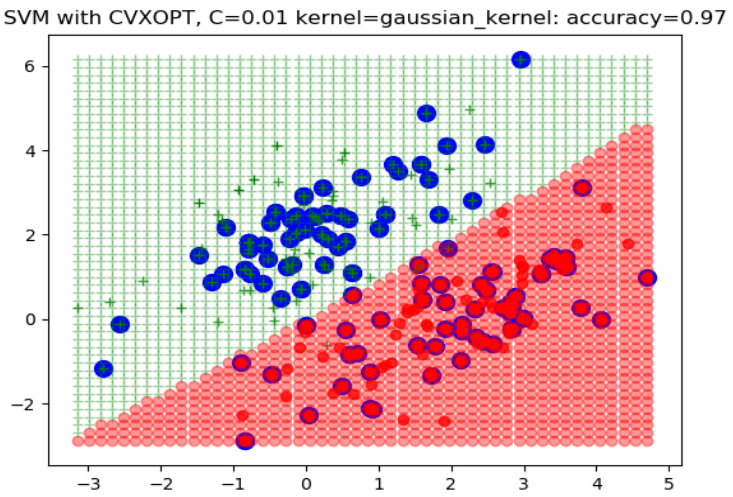
Homework 4: Support vector machines (last year)

- Support Vector Classifier tries to find the best hyperplane to separate the different classes by maximizing the distance between sample points and the hyperplane



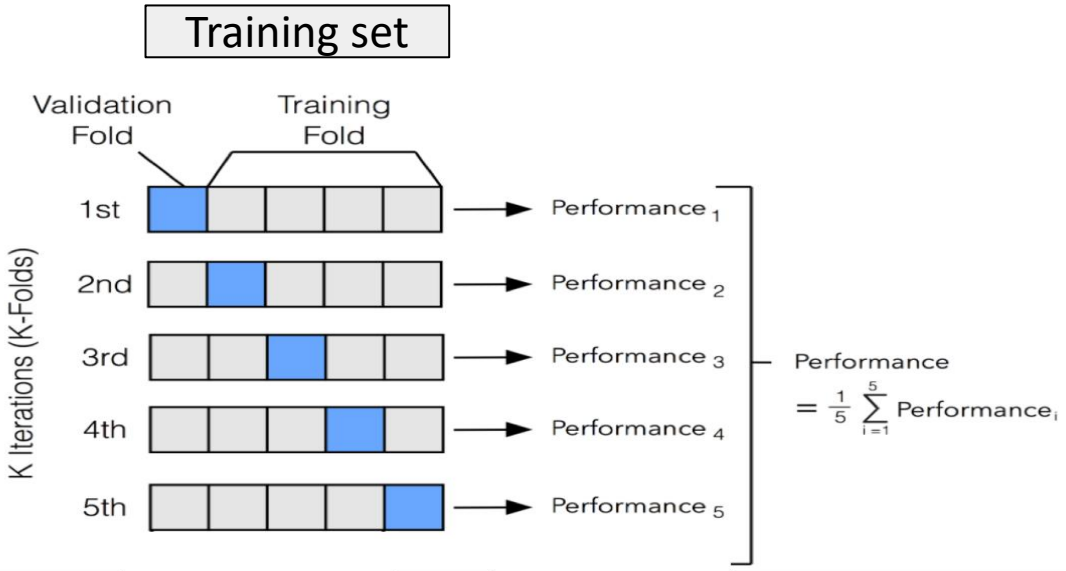
Hyperparameter searching

- Suppose we want to find the best values of two hyperparameters for an RBF kernel SVM namely C and gamma.
- Many hyperparameter combinations to be considered!



K-fold Cross-validation

- We split the dataset into K parts: one part is used for validation, and the remaining K-1 parts are merged into a training subset. This process repeats K times, with each part used exactly once as the validation data



Thank You for Your Attention!

Yen-Yu Lin (林彥宇)

Email: lin@cs.nctu.edu.tw

URL: https://www.cs.nycu.edu.tw/members/detail/lin