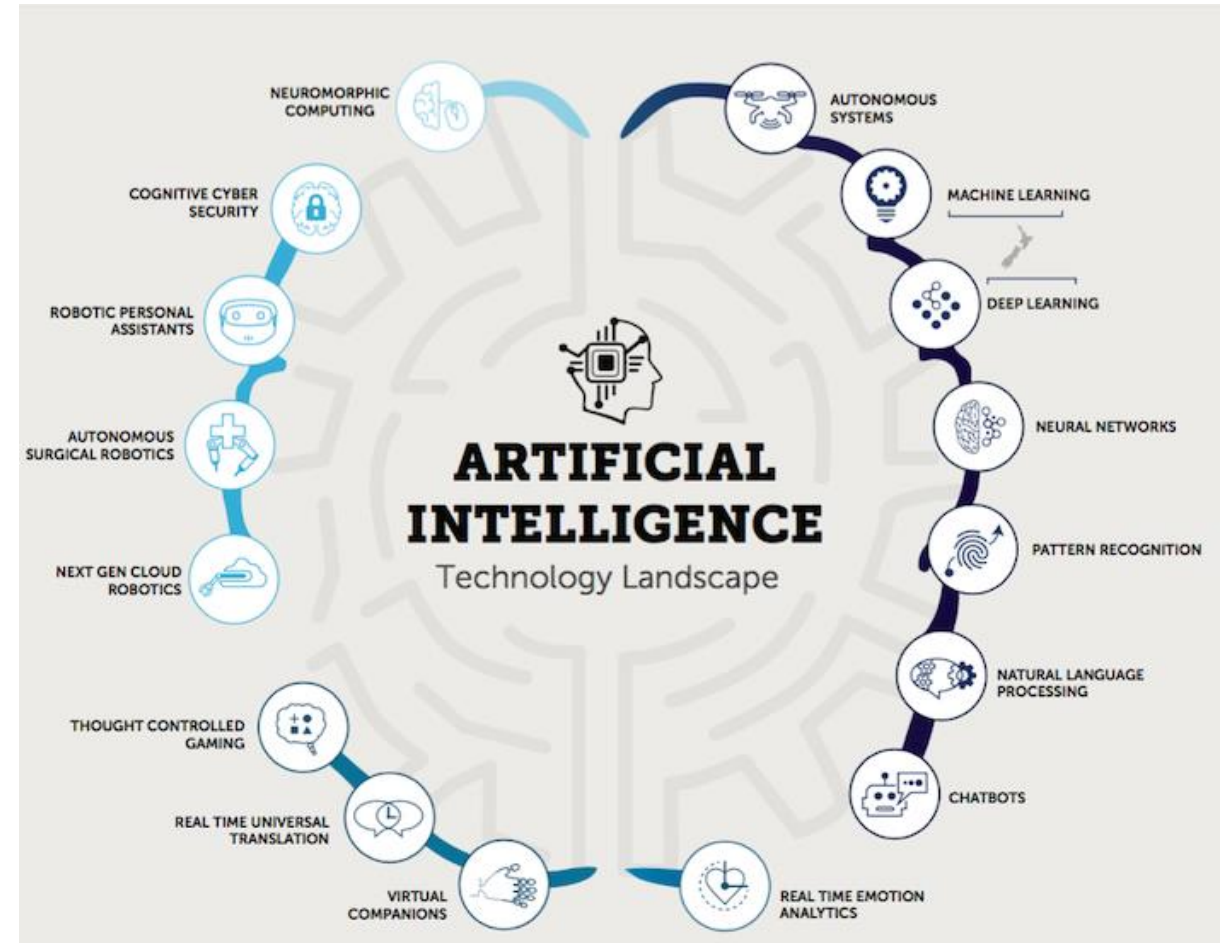## Basic Concepts in Machine Learning

Machine Learning Introduction Presentation
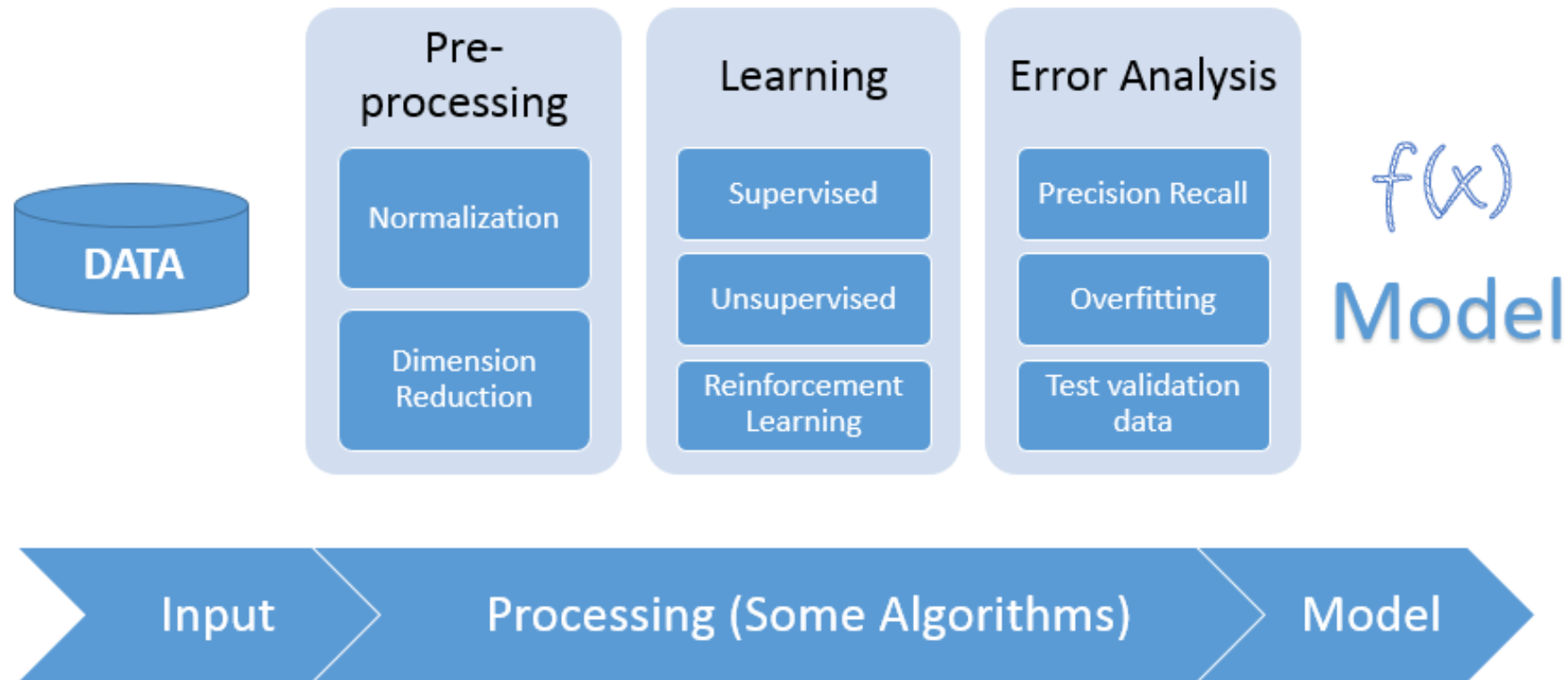
E. Ortiz Escoriza, Prof. Adham Atyabi

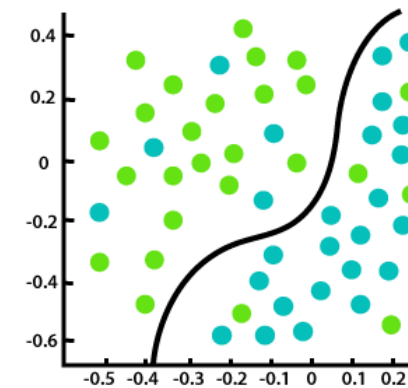**Overview meeting**

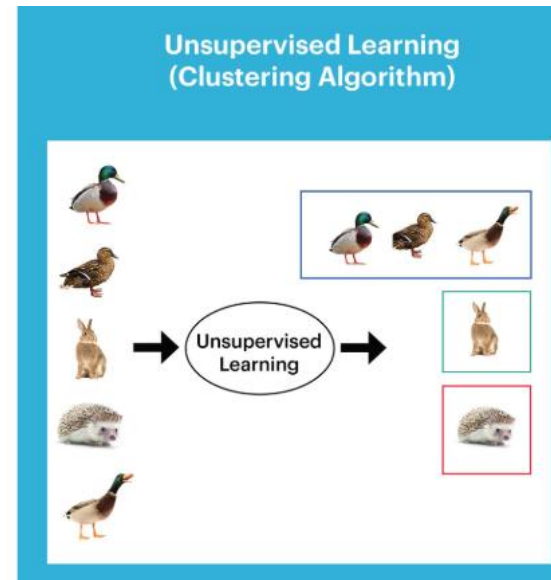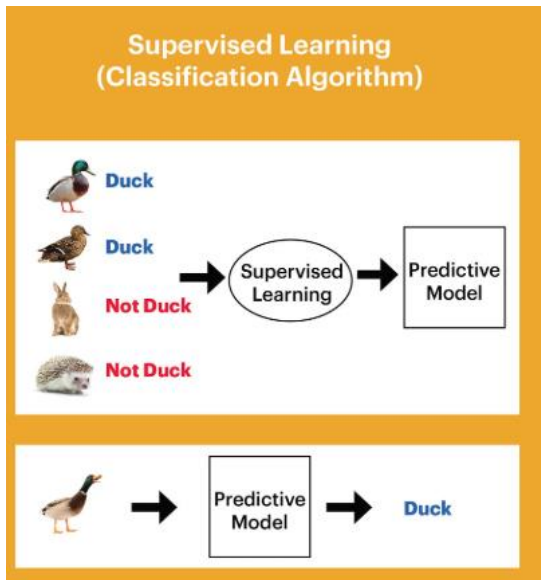*June 25th, 2020*

Introduction

UCCS ENGINEERING

- Machine Learning [1]

Machine Learning Algorithms
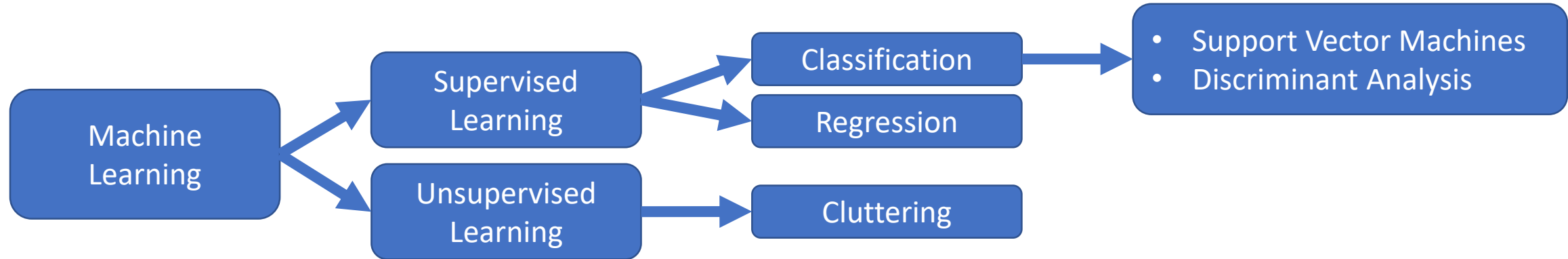
- Supervised Learning & Unsupervised Learning [2]
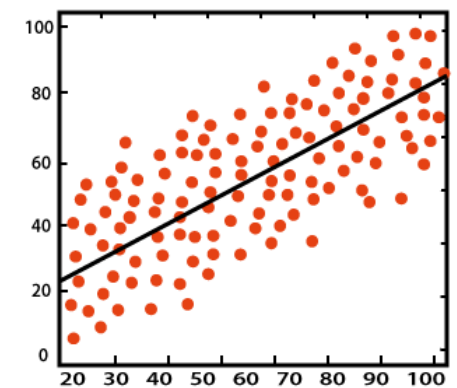
- Supervised Learning Vs Unsupervised Learning [3] [4]

	Supervised	Unsupervised
Method	Input & Output variables will be given	Only input data will be given
Goal	Determinate the function so well that when new input data set given, can predict the output	Model the hidden patterns or underlying structure in the given input data in order to learn about the data
Input Data	Algorithms are trained using labeled data	Algorithms are used against data which is not labeled
Uses	Often used for export system in image recognition, speech recognition, forecasting, financial analysis and training neural networks and decision trees	Often used to pre-process the data, during exploratory analysis or to pre-train supervised learning algorithms

	Supervised Learning	Unsupervised Learning
Discrete	classification or categorization	clustering
Continuous	regression	dimensionality reduction

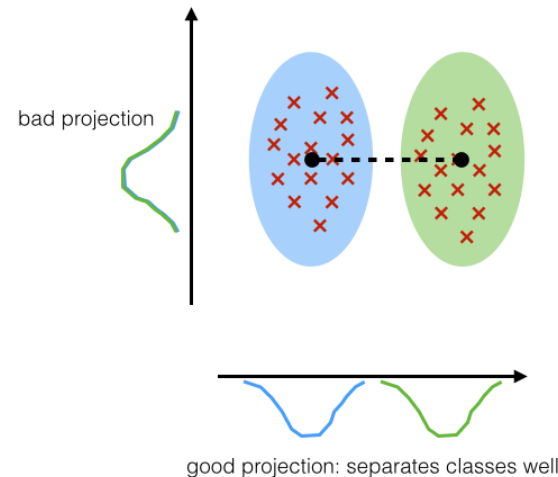**Supervised Learning – Classification Analysis**

UCCS ENGINEERING

- Linear Discriminant Analysis (LDA) [5]

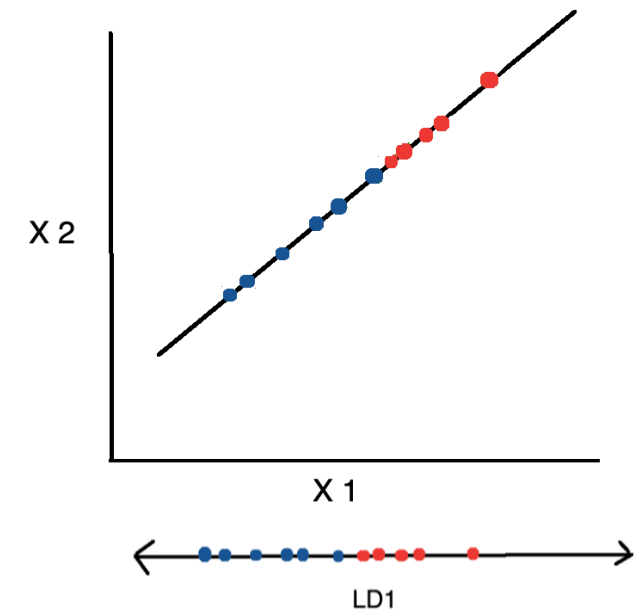

PCA:
component axes that
maximize the variance
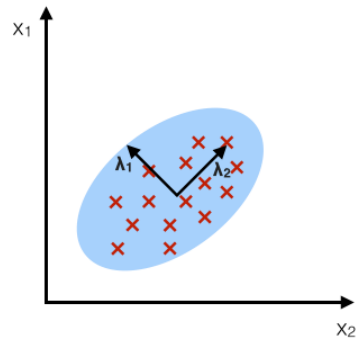
LDA:
maximizing the component
axes for class-separation

bad projection

good projection: separates classes well
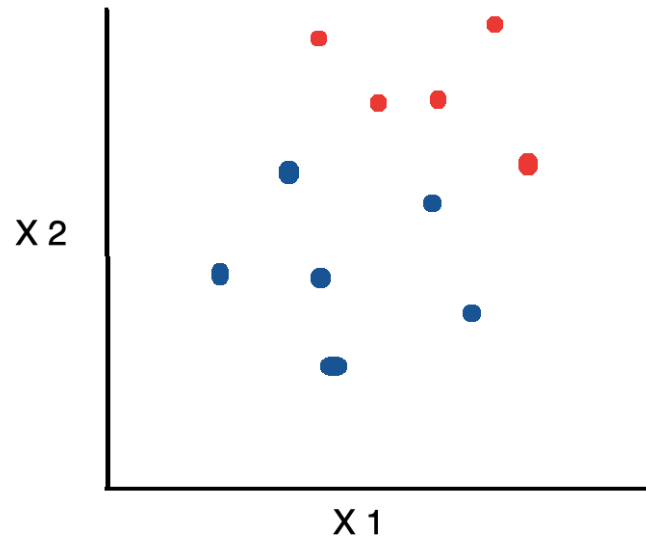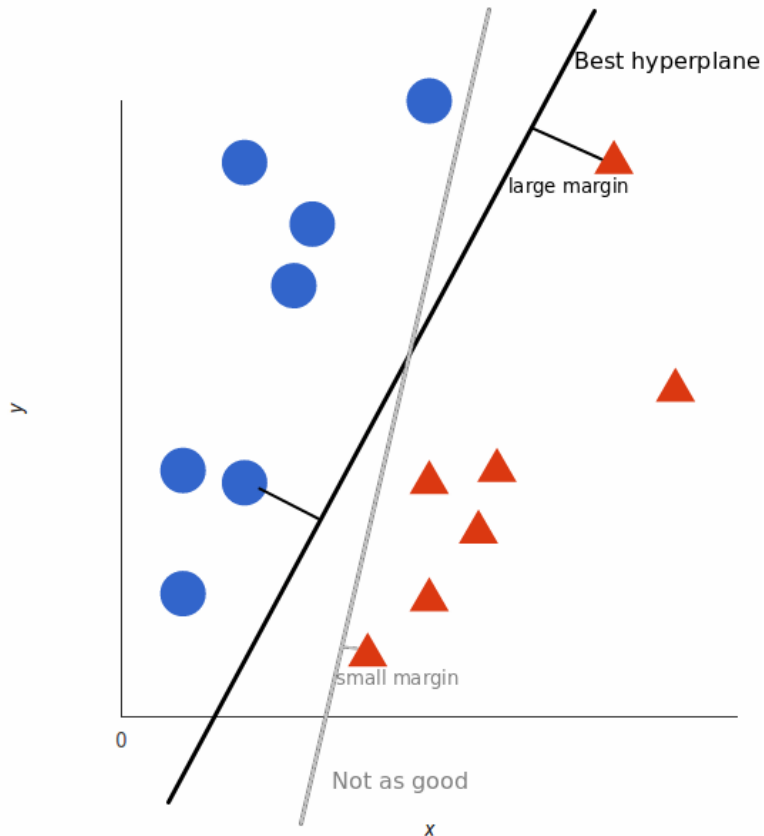
Summarizing the LDA approach in 5 steps:

1. Compute the d-dimensional mean vectors for the different classes from the dataset.

2. Compute the scatter matrices (in-between-class and within-class scatter matrix).

3. Compute the eigenvectors ($e_1, e_2, \dots, e_d$) and corresponding eigenvalues ($\lambda_1, \lambda_2, \dots, \lambda_d$) for the scatter matrices.

4. Sort the eigenvectors by decreasing eigenvalues and choose k eigenvectors with the largest eigenvalues to form a d x k dimensional matrix W (where every column represents an eigenvector).

5. Use this d x k eigenvector matrix to transform the samples onto the new subspace. This can be summarized by the matrix multiplication: $Y = X$ x W (where X is a n x d-dimensional matrix representing the n samples, and y are the transformed n x k-dimensional samples in the new subspace).

Supervised Learning – Classification Analysis

UCCS ENGINEERING

- Linear Discriminant Analysis (LDA) [6]

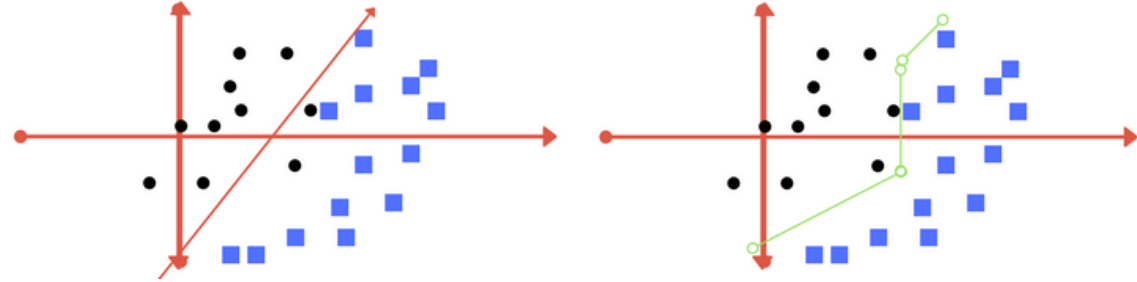**Supervised Learning – Classification and Regression analysis**
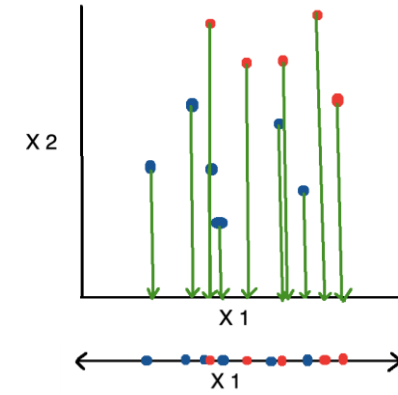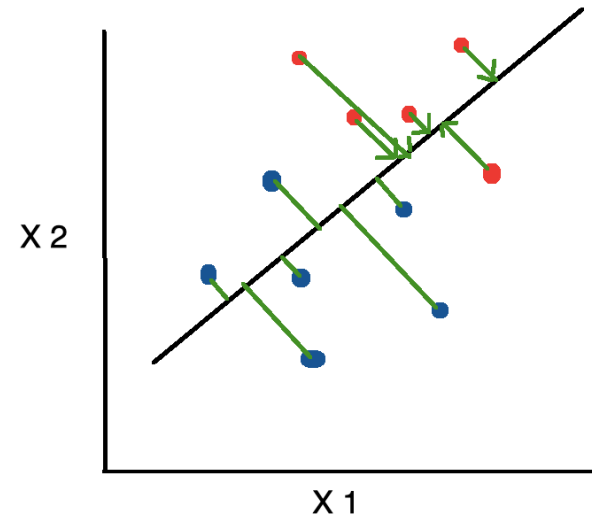
• Support Vector Machines (SVM) [7]
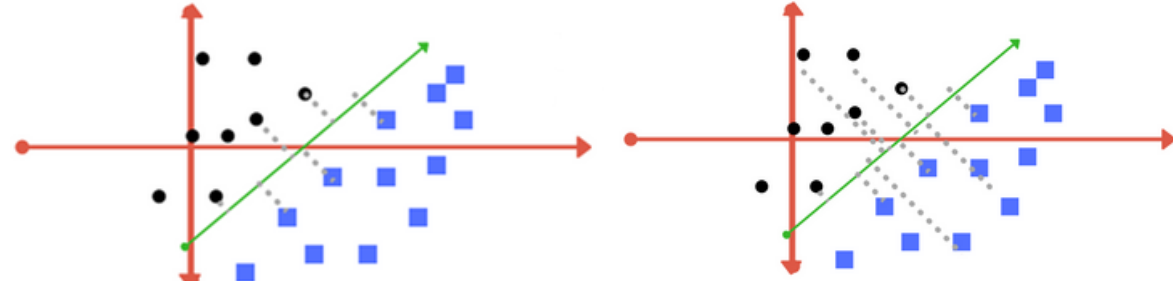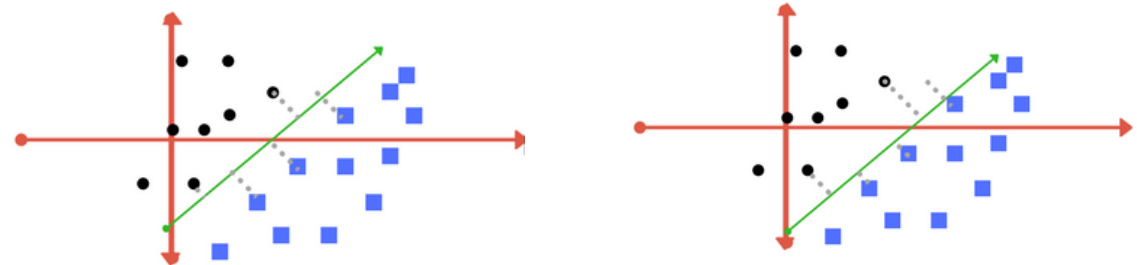

Tuning Parameters:
- Low and high regularization value

- High and low Gamma

- Good and bad Margin

Thank you for your attention

UCCS ENGINEERING

[1]   Jaiswal, Sujeet Kumar. "Machine Learning : An Introduction." *Medium*, Developer's Blog, 24 Sept. 2016,
        https://blog.sujeetjaiswal.com/machine-learning-an-introduction-de88d85ebc5d

[2]   Zhou, Linda. "Simplify Machine Learning Pipeline Analysis with Object Storage." *Western Digital Corporate Blog*, 11 Jan. 2019,
        https://blog.westerndigital.com/machine-learning-pipeline-object-storage/

[3]   "Supervised vs Unsupervised Learning: Key Differences." *Guru99*, https://www.guru99.com/supervised-vs-unsupervised-learning.html

[4]   "Supervised Learning vs Unsupervised Learning: Top 7 Amazing Difference." *EDUCBA*, 7 May 2020,
        https://www.educba.com/supervised-learning-vs-unsupervised-learning/

[5]   "Linear Discriminant Analysis." *Dr. Sebastian Raschka*, 3 Aug. 2014, https://sebastianraschka.com/Articles/2014_python_lda.html

[6]   Maklin, Cory. "Linear Discriminant Analysis In Python." *Medium*, Towards Data Science, 4 Aug. 2019,
        https://towardsdatascience.com/linear-discriminant-analysis-in-python-
        76b8b17817c2#:~:text=Linear%20Discriminant%20Analysis%20(LDA)%20is,as%20much%20information%20as%20possible.

[7]   Patel, Savan. "Chapter 2 : SVM (Support Vector Machine) - Theory." *Medium*, Machine Learning 101, 4 May 2017,
        https://medium.com/machine-learning-101/chapter-2-svm-support-vector-machine-theory-f0812effc72