

# Support Vector Machine

# Classification with SVM

- Imagine that you obtain the data **containing characteristics of thousand of human cell samples** extracted from patients who were believed to be at risk of developing cancer.
- Analysis of the original data showed that many of the characteristics differed significantly between **benign** and **malignant samples**

ID	Clump	UnifSize	UnifShape	MargAdh	SingEpiSize	BareNuc	BlandChrom	NormNucl	Mit	Class
1000025	5	1	1	1	2	1	3	1	1	benign
1002945	5	4	4	5	7	10	3	2	1	benign
1015425	3	1	1	1	2	2	3	1	1	malignant
1016277	6	8	8	1	3	4	3	7	1	benign
1017023	4	1	1	3	2	1	3	1	1	benign
1017122	8	10	10	8	7	10		7	1	malignant
1018099	1	1	1	1	2	10	3	1	1	benign
1018561	2	1	2	H	2	1	3	1	1	benign
1033078	2	1	1	1	2	1	1	1	5	benign
1033078	4	2	1	1	2	1	2	1	1	benien

ID	Clump	UnifSize	UnifShape	MargAdh	SingEpiSize	BareNuc	BlandChrom	NormNucl	Mit	Class
1000015	6	1	1	1	7	1	3	1	1	

modeling

prediction

Accuracy 89 %

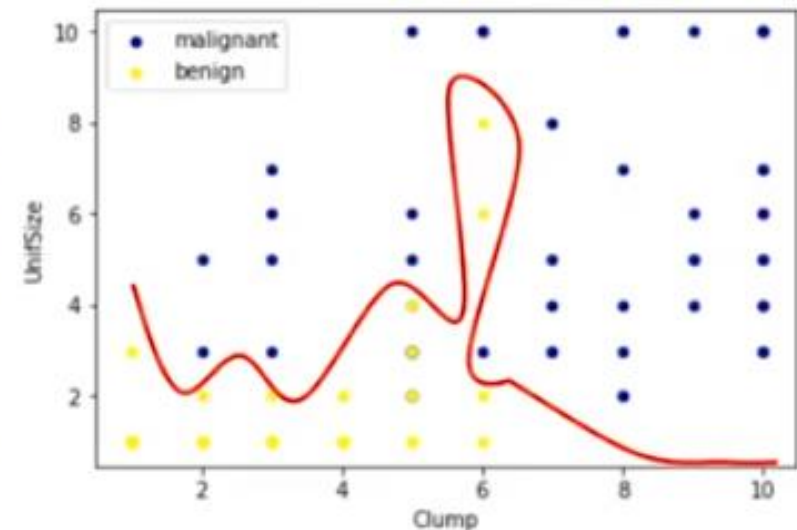


# What is SVM?

SVM is a supervised learning algorithm that classifies cases by finding a separator.

1. Mapping data to a high-dimensional feature space
2. Find a separator.

Clump	UnifSize	UnifShape	MargAdh	SingEpiSize	BareNuc	BlandChrom	NormNucl	Mit	Class
5	1	1	1	2	1	3	1	1	benign
5	4	4	5	7	10	3	2	1	benign
3	1	1	1	2	2	3	1	1	malignant
6	8	8	1	3	4	3	7	1	benign
4	1	1	3	2	1	3	1	1	benign
8	10	10	8	7	10		7	1	malignant
1	1	1	1	2	10	3	1	1	benign
2	1	2	H	2	1	3	1	1	benign
2	1	1	1	2	1	1	1	5	benign
4	2	1	1	2	1	2	1	1	benign

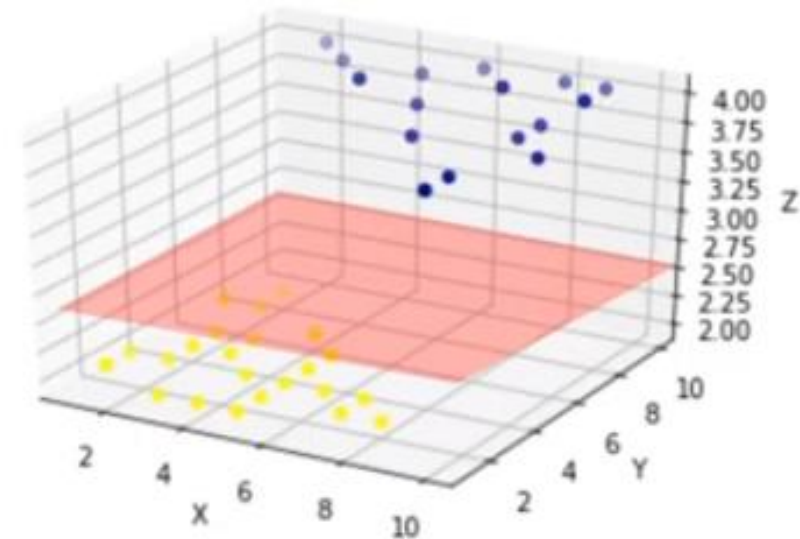


# What is SVM?

SVM is a supervised learning algorithm that classifies cases by finding a separator.

1. Mapping data to a high-dimensional feature space
2. Find a separator.

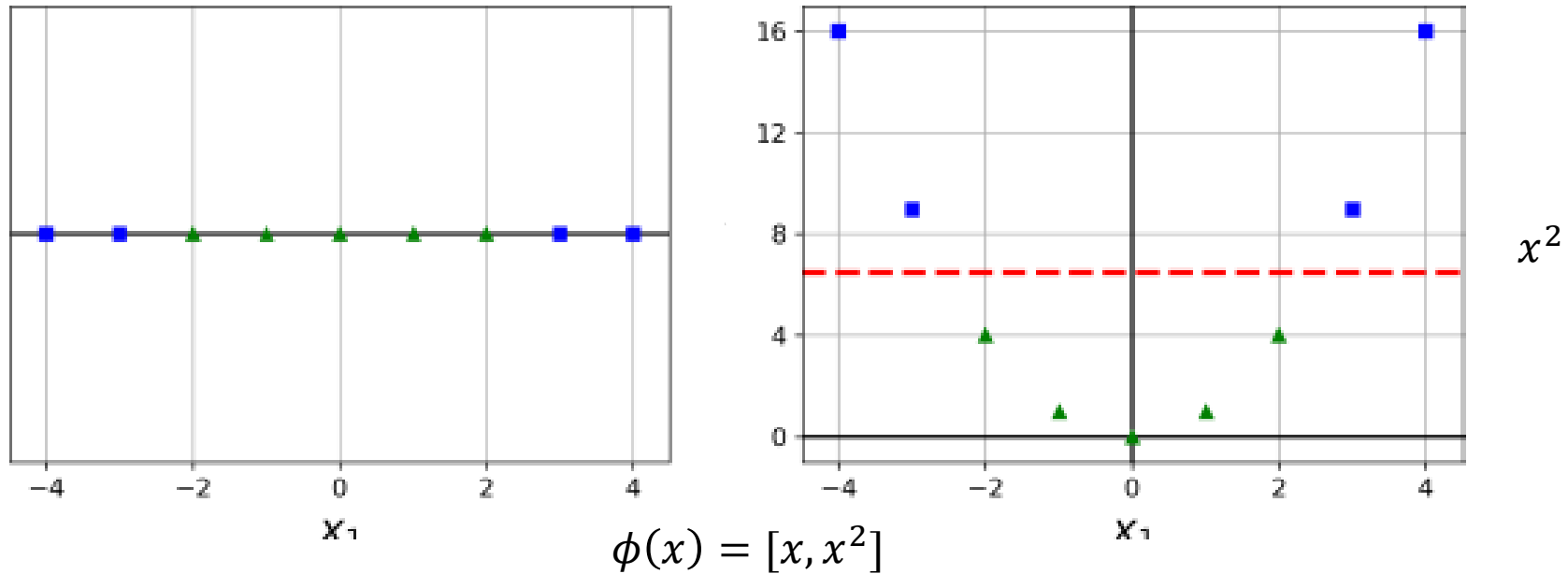
Clump	UnifSize	UnifShape	MargAdh	SingEpiSize	BareNuc	BlandChrom	NormNucl	Mit	Class
5	1	1	1	2	1	3	1	1	benign
5	4	4	5	7	10	3	2	1	benign
3	1	1	1	2	2	3	1	1	malignant
6	8	8	1	3	4	3	7	1	benign
4	1	1	3	2	1	3	1	1	benign
8	10	10	8	7	10		7	1	malignant
1	1	1	1	2	10	3	1	1	benign
2	1	2	H	2	1	3	1	1	benign
2	1	1	1	2	1	1	1	5	benign
4	2	1	1	2	1	2	1	1	benign



# Two challenging questions

- How to we transfer data in such the way that the separator could be drawn such as a hyperplane?
- How can we find the best or optimized hyperplane separator after transformation?

# Data Transformation

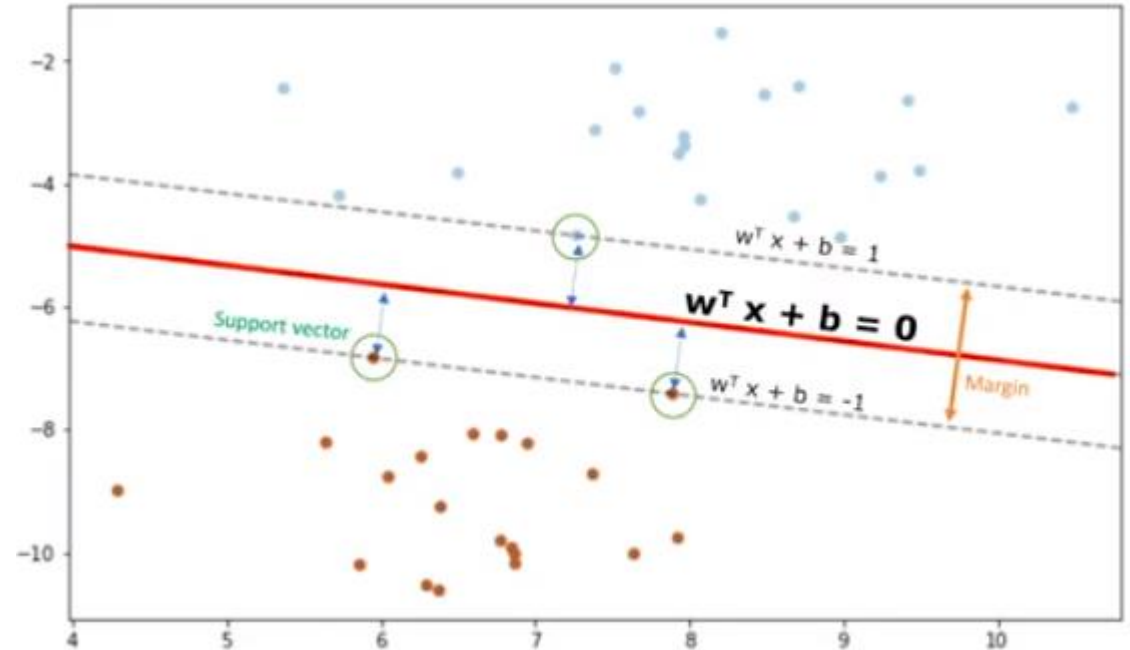


Mapping data into a higher-dimensional space is called **Kernelling**

- Linear
- Polynomial
- RBF
- Sigmoid

# Using SVM to find the hyperplane

Find  $\mathbf{w}$  and  $b$  such that  $\Phi(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{w}$  is minimized, and for all  $\{\mathbf{x}_i, \mathbf{y}\}$ :  $\mathbf{y}_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1$



# Pros and cons of SVM

- Advantages:
  - Accuracy in high-dimensional spaces
  - Memory efficient
- Disadvantages:
  - Prone of over fitting if the number of feature is greater than the number of samples.
  - No probability estimation
  - Small datasets



# SVM Applications

- Image recognition
- Text category assignment
- Detecting spam
- Sentiment analysis
- Gene expression classification
- Regression, outlier detection and clustering and more...