



Module: Machine Learning

Live Session-5

Agenda:

Hard Margin SVM

Soft Margin SVM

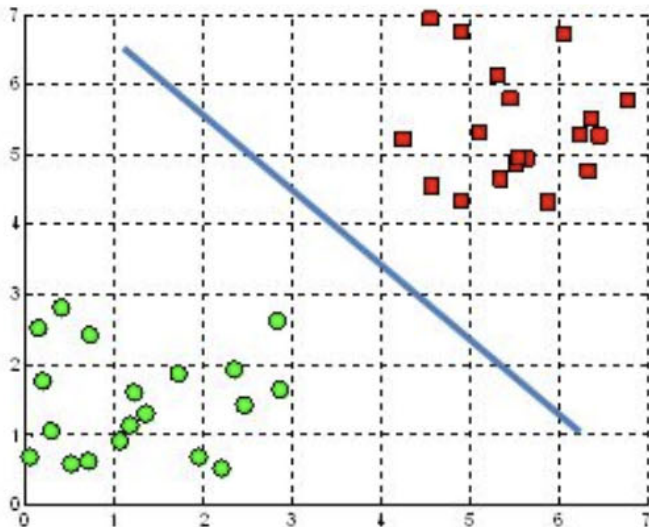
Nonlinear SVM



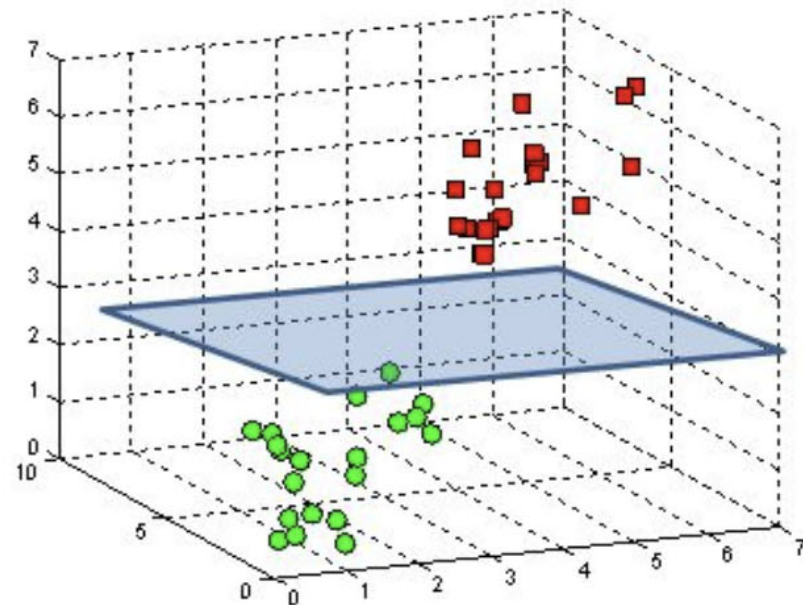
Hyperplane

In geometry, a **hyperplane** is a subspace whose dimension is one less than that of its ambient space. If a space is 3-dimensional then its **hyperplanes** are the 2-dimensional planes, while if the space is 2-dimensional, its **hyperplanes** are the 1-dimensional lines.

A hyperplane in \mathbb{R}^2 is a line



A hyperplane in \mathbb{R}^3 is a plane



Hyperplane: Mathematical representation

2D: $w_1 x_1 + w_2 x_2 + b = 0$

$w^T x + b = 0 \checkmark$

3 No.

3D: $w_1 x_1 + w_2 x_2 + w_3 x_3 + b = 0$

$w^T x + b = 0 \checkmark$

4 No.

nD: $w_1 x_1 + w_2 x_2 + \dots + w_n x_n + b = 0$

$w^T x + b = 0 \checkmark$

n+1

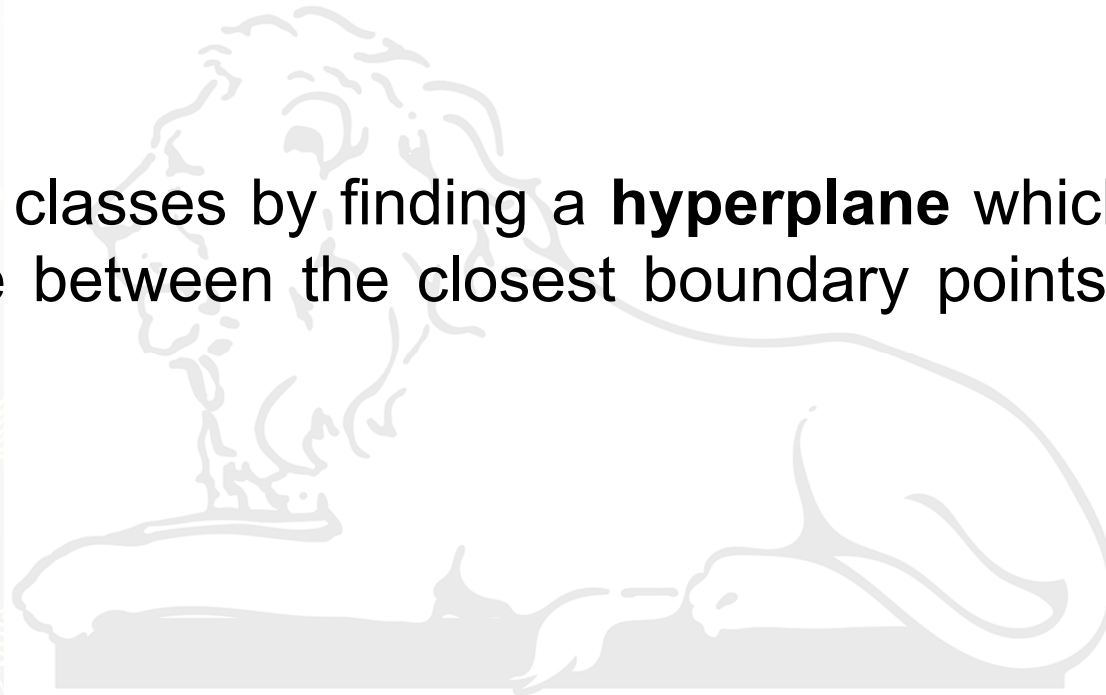
$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$w \quad x$

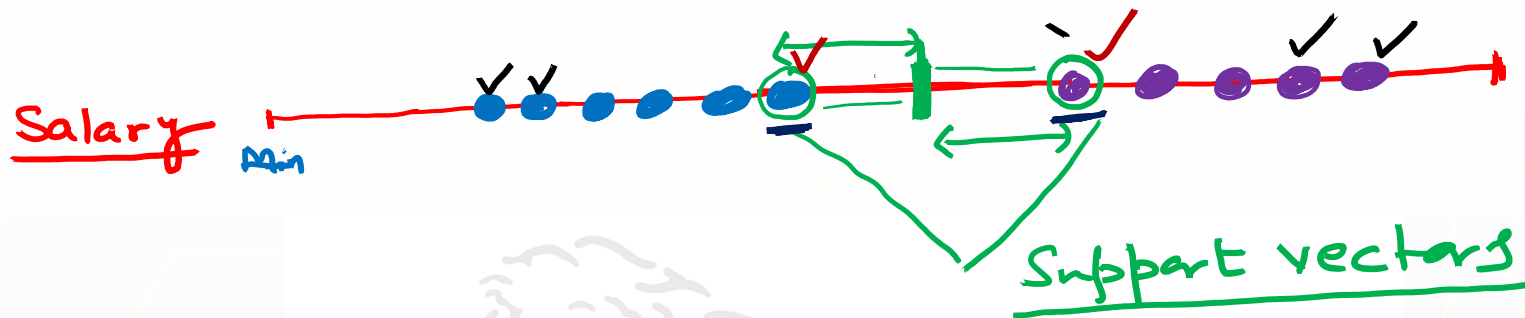
SVM: Concept

A support vector machine (SVM) is a supervised machine learning classification algorithm. SVMs were introduced initially in 1960s and were later refined in 1990s.

To separate classes by finding a **hyperplane** which maximizes the distance between the closest boundary points (vectors) of each class.



SVM: Concept

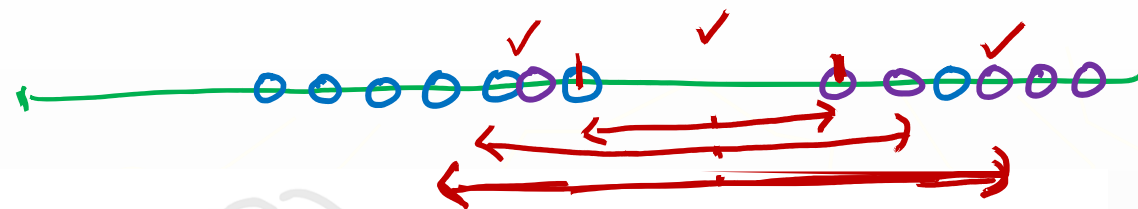


Objective: To find a hyperplane which maximize the margin from the closest boundary points of each class.

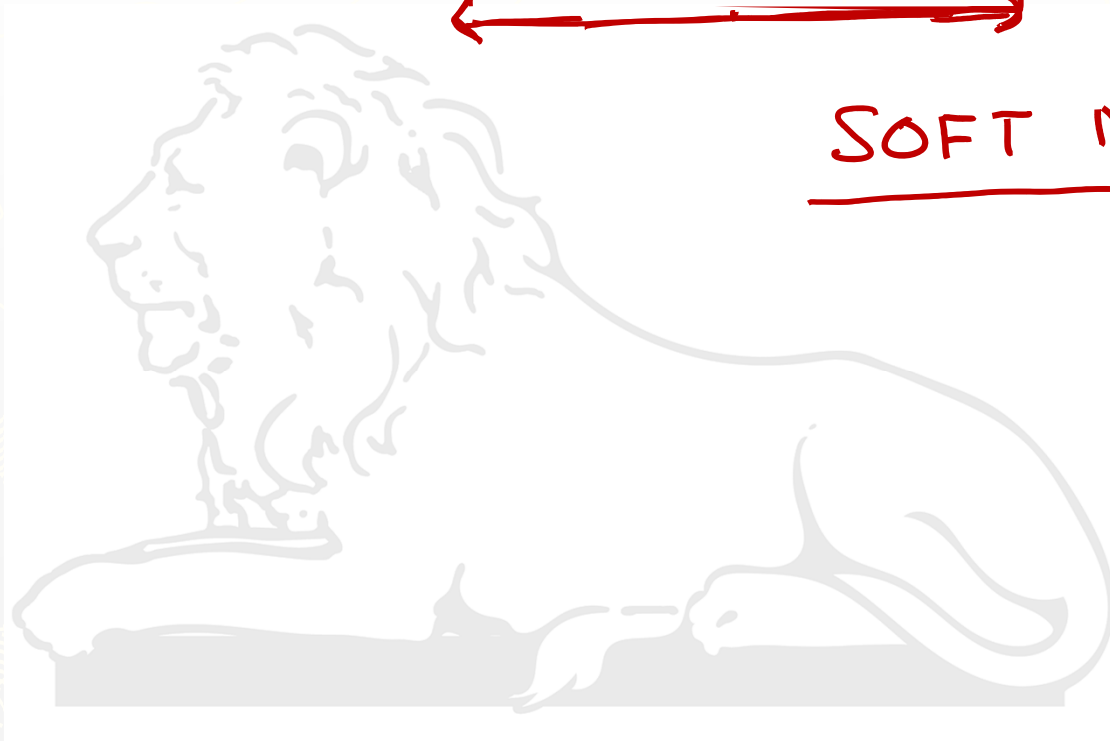
$\{x_i, y_i\}_{i=1}^n \longrightarrow$ Hyperplane

HARD MARGIN CLASSIFIER

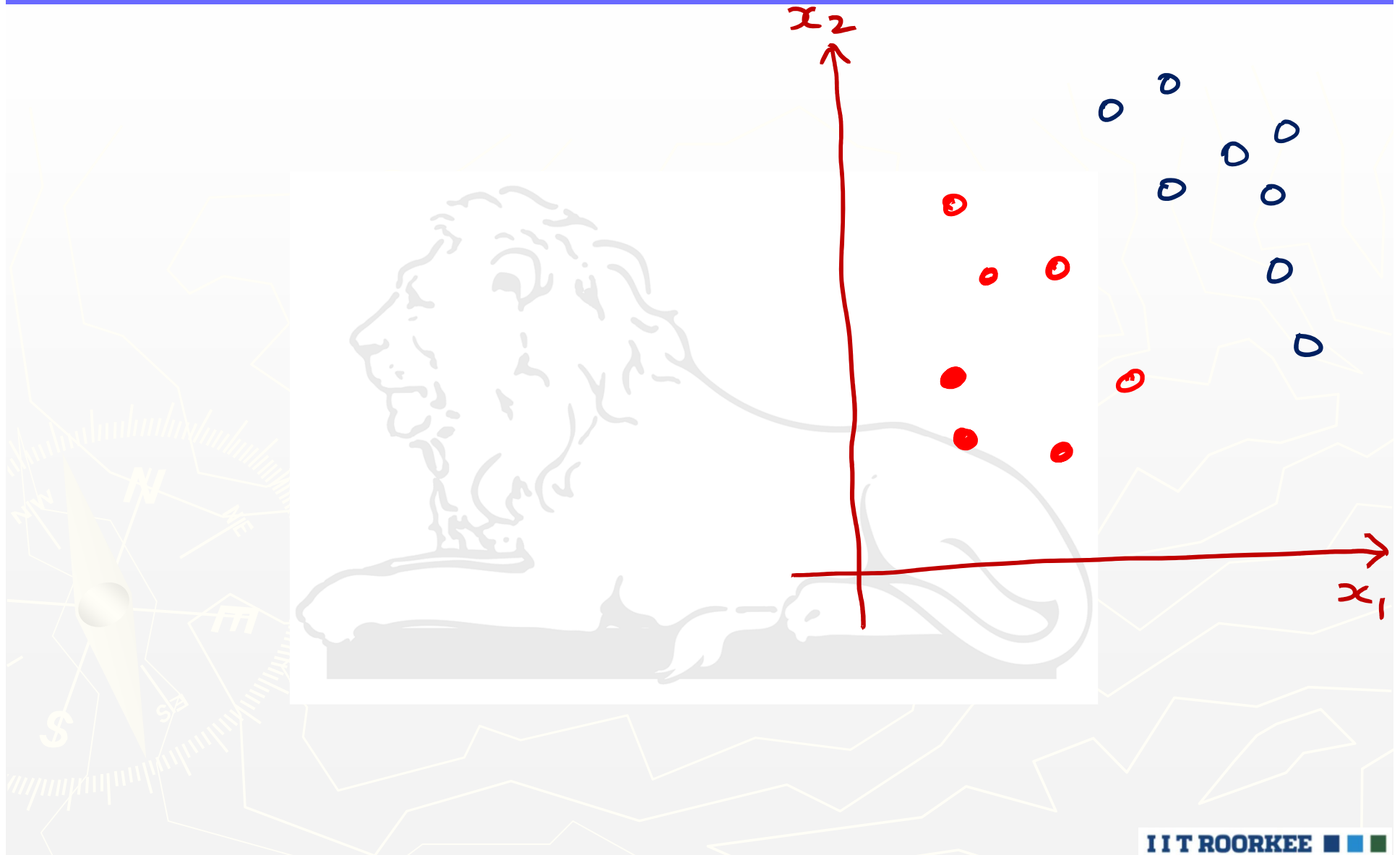
SVM: Concept



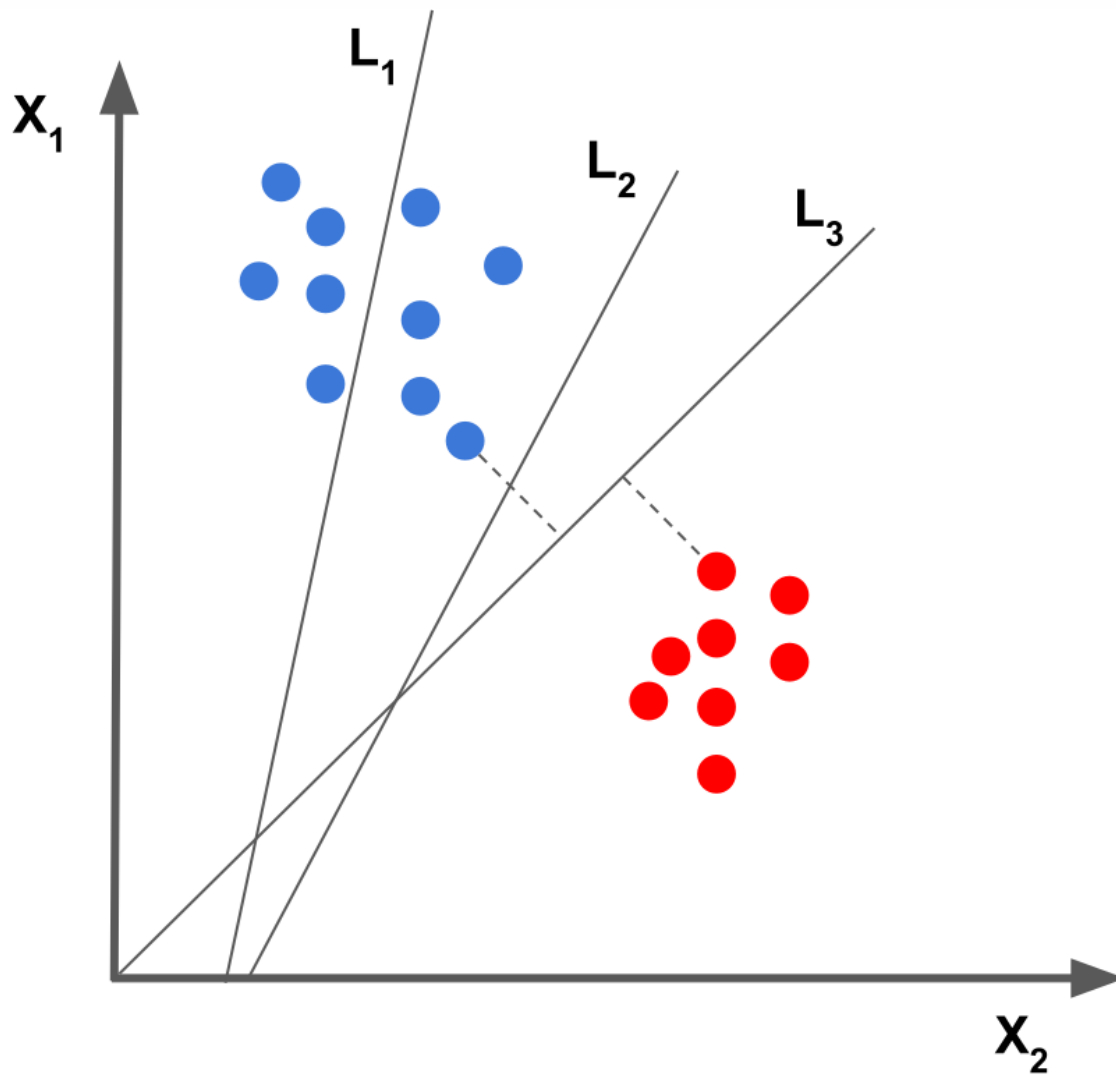
SOFT MARGIN



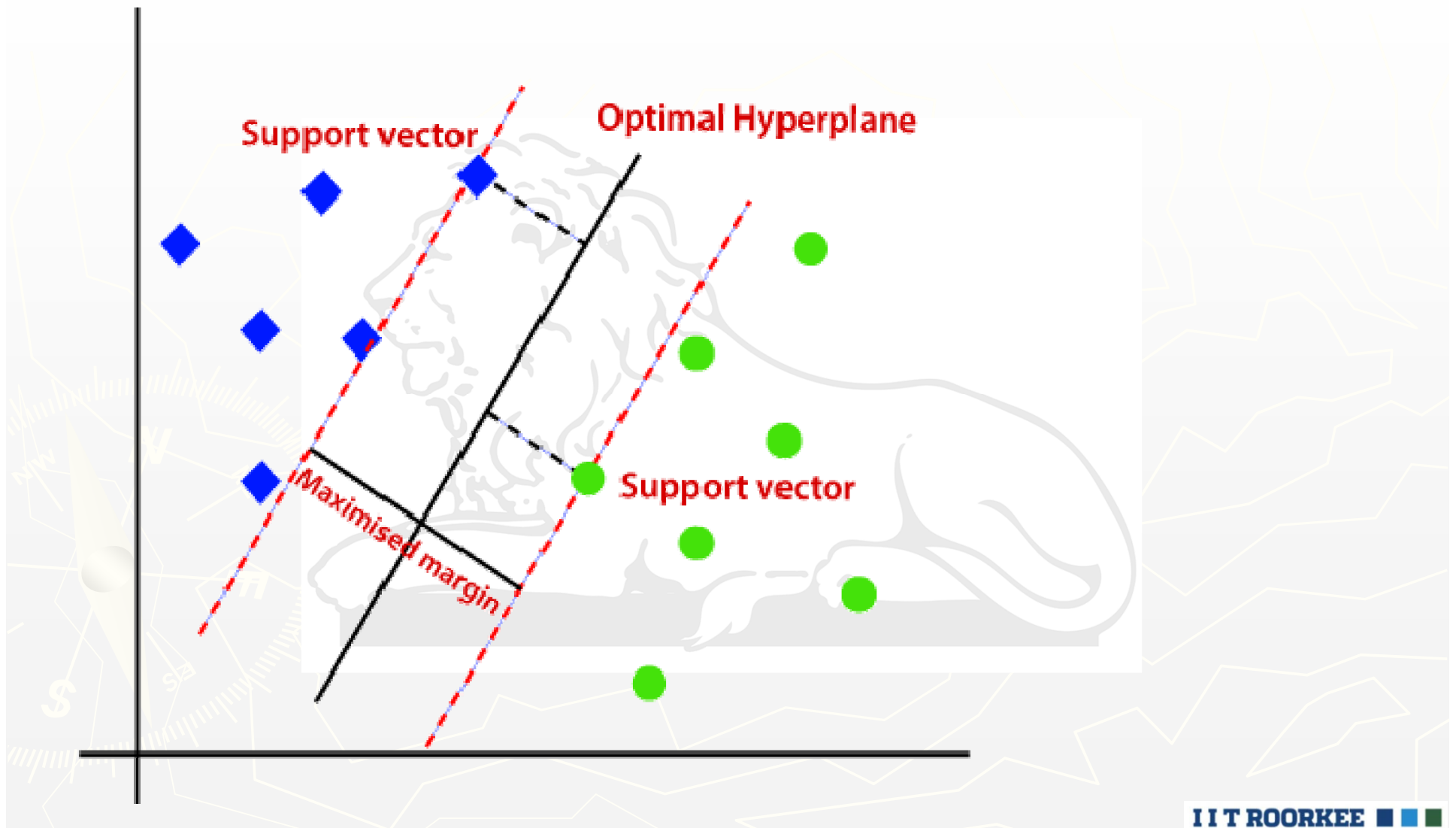
SVM: Concept



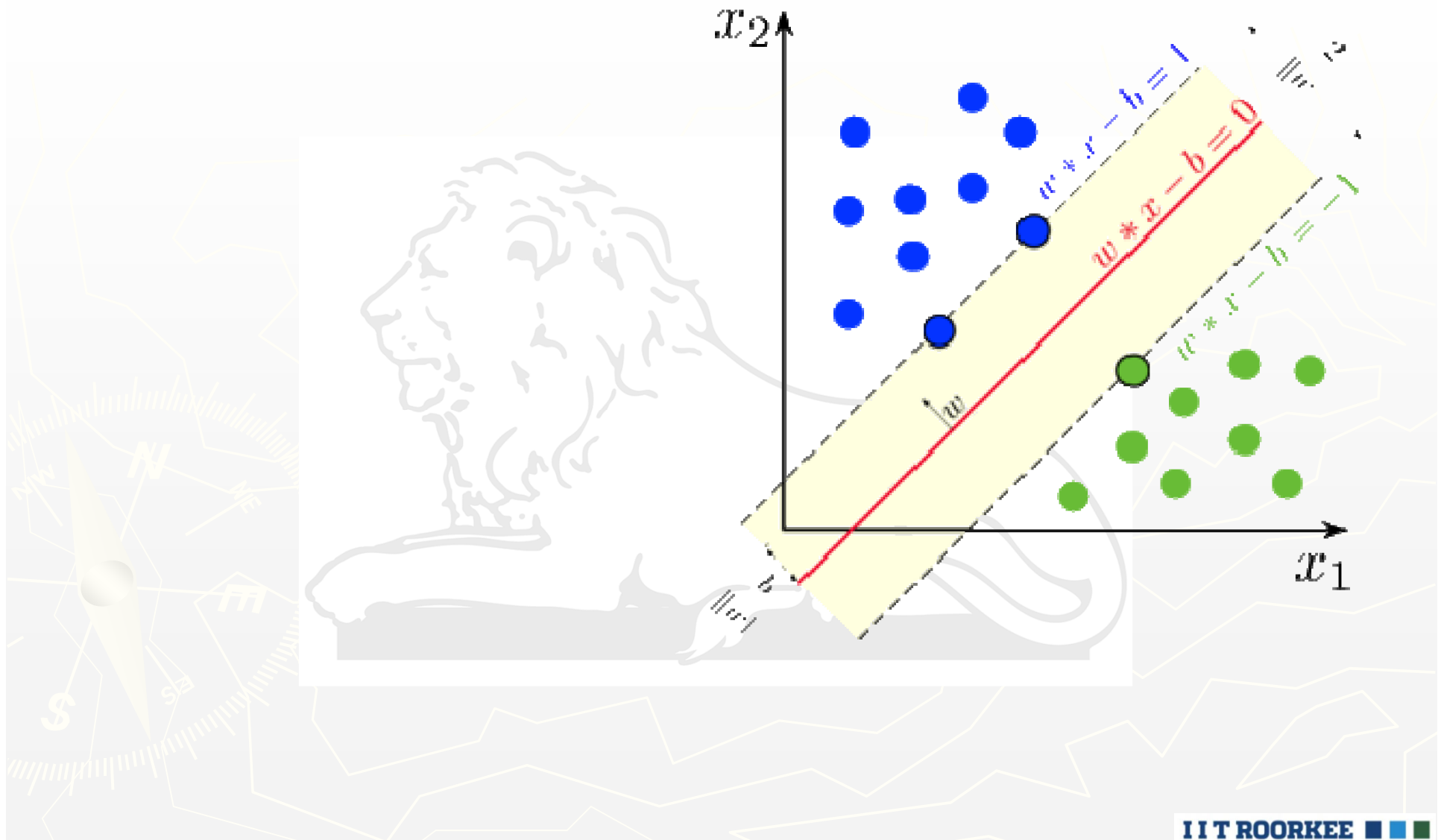
SVM: Concept 2D case



SVM: Concept 2D case



SVM: Concept



SVM: Formulation



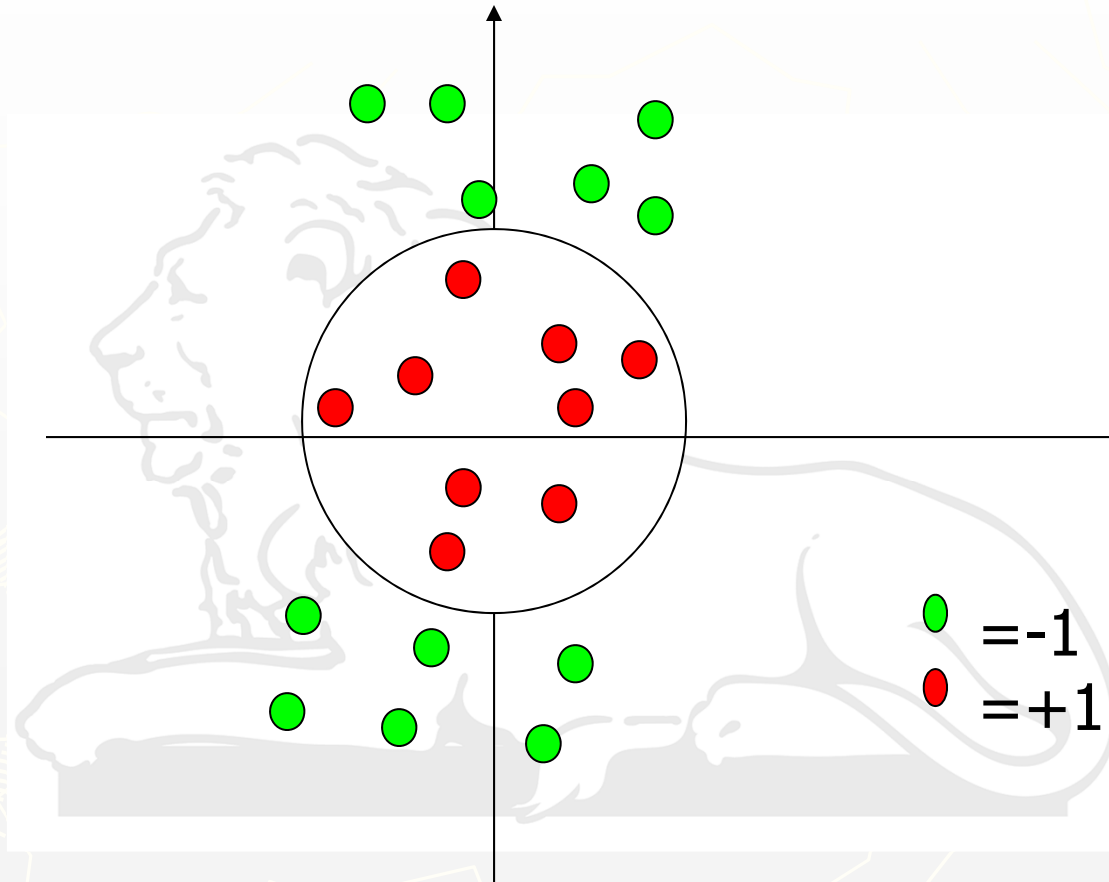
SVM: Formulation



SVM: Formulation



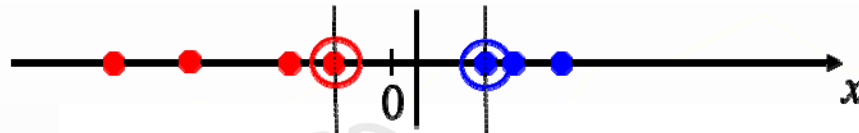
Problem with Linear SVM



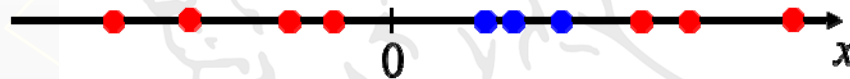
What if the decision function is not a linear?

Non-linear SVMs

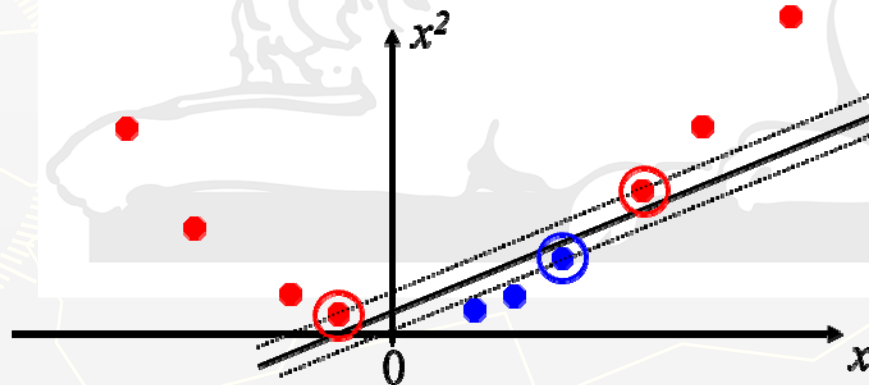
- Datasets that are linearly separable with some noise work out great:



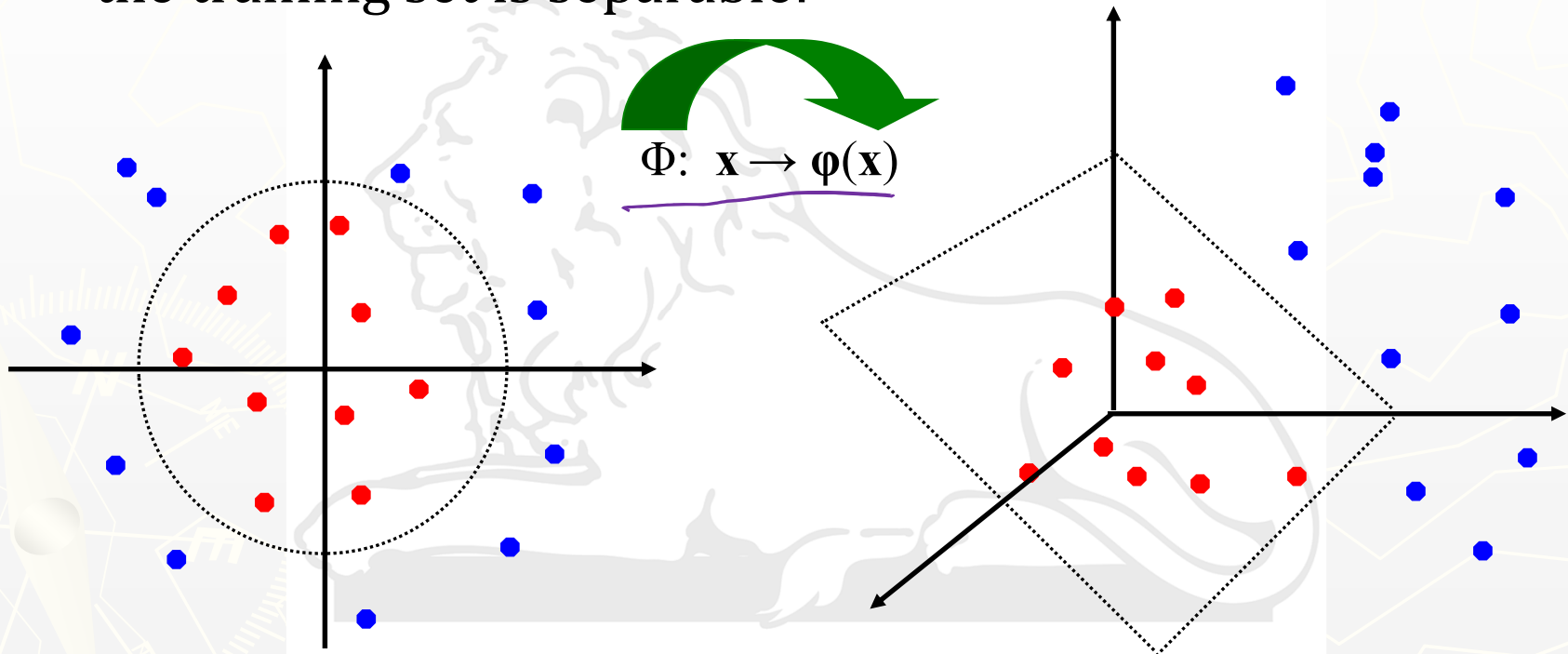
- But what are we going to do if the dataset is just too hard?



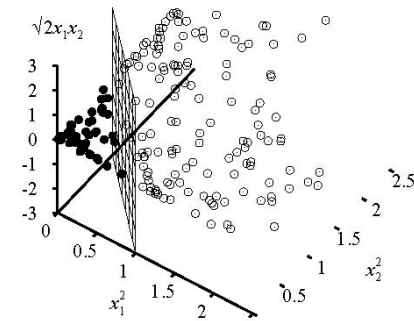
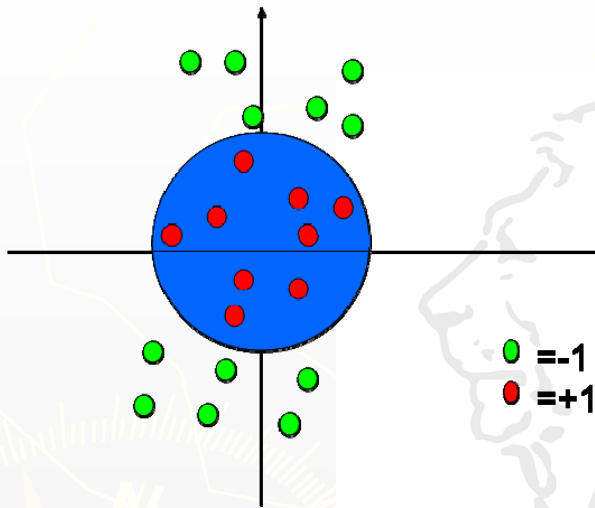
- How about... mapping data to a higher-dimensional space:



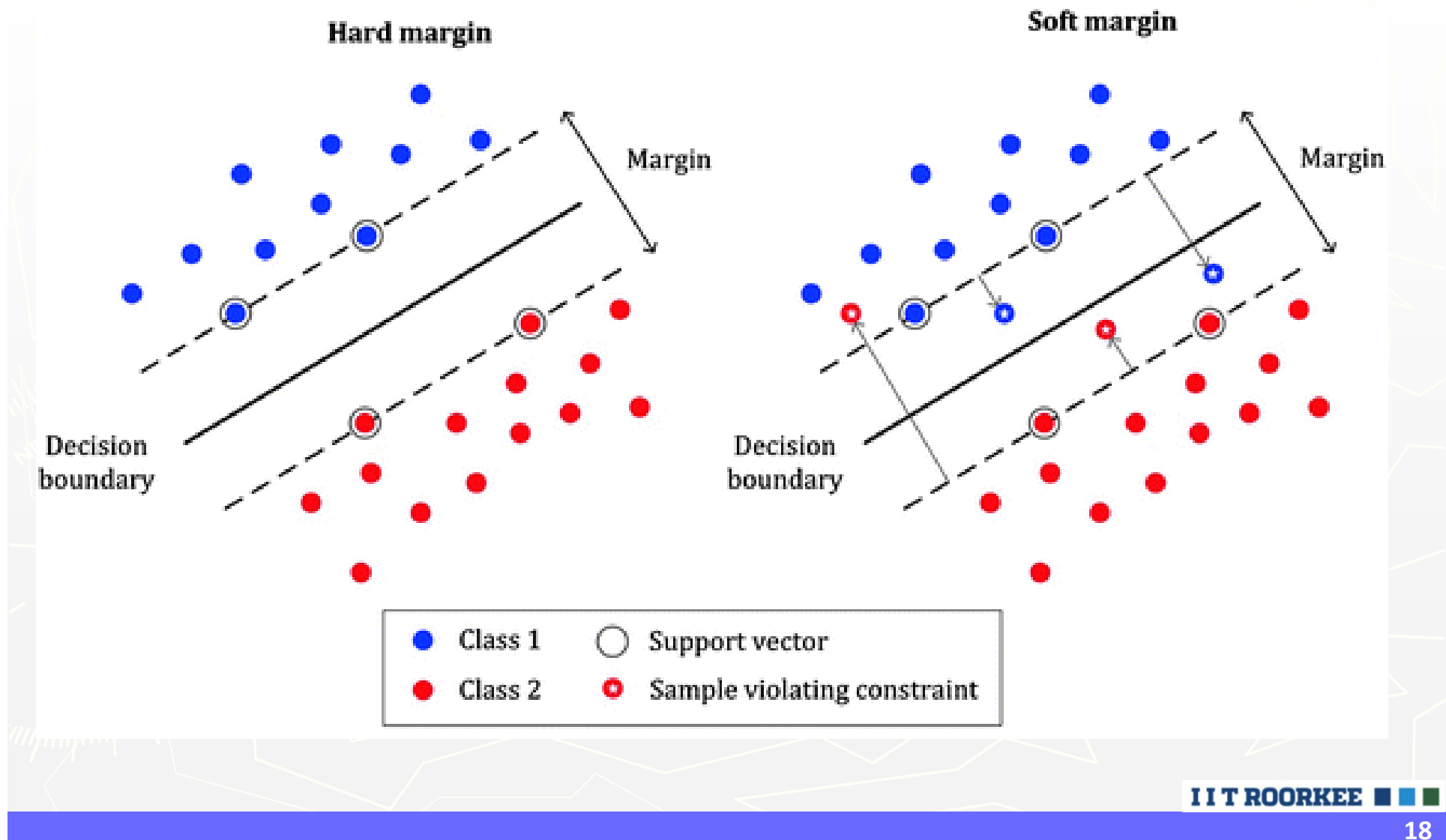
- General idea: the original feature space can always be mapped to some higher-dimensional feature space where the training set is separable:



Data points are linearly separable in the space $(x_1^2, x_2^2, \sqrt{2}x_1x_2)$



SVM: Concept 2D case



SVM and Logistic Regression

- ▶ The risk of overfitting is less in SVM, while Logistic regression is vulnerable to overfitting.
- ▶ SVM works well with unstructured and semi-structured data like text and images while logistic regression works with already identified independent variables.
- ▶ SVM is based on geometrical properties of the data while logistic regression is based on statistical approaches.

The background of the slide features a light gray gradient with faint, stylized topographic contour lines in a pale yellow color. In the bottom-left corner, there is a compass rose graphic. The compass rose has a yellow needle pointing towards the top-left, with a silver-colored circular base. The cardinal and ordinal directions are labeled: 'N' for North, 'NE' for Northeast, 'SE' for Southeast, and 'SW' for Southwest. A dollar sign (\$) is positioned near the bottom-left of the compass rose. The text 'THANK YOU' is centered in the middle of the slide in a bold, black, sans-serif font.

THANK YOU