ALDENAIRE & PARTNERS

SVM KERNEL TRICK AND NON LINEAR SVM

MADE BY: GROUP 3

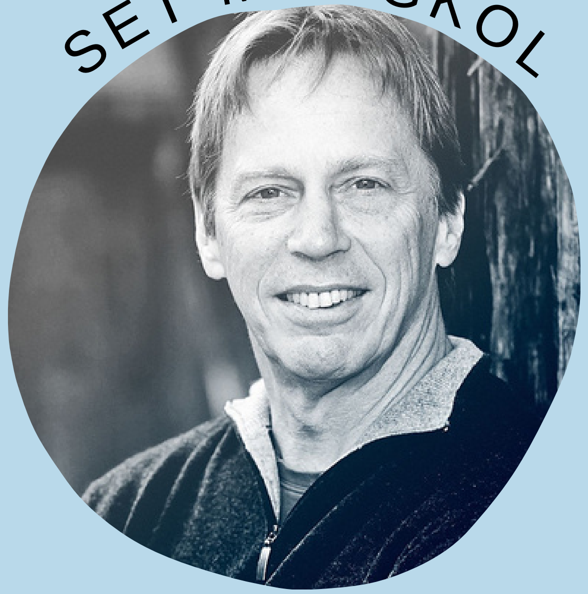
MEET OUR TEAM GROUP 3

HONG KIMMENG

HON RATANA

MEN CHANCHHORPORN

SET MONGKOL



VINLAY ANUSAR

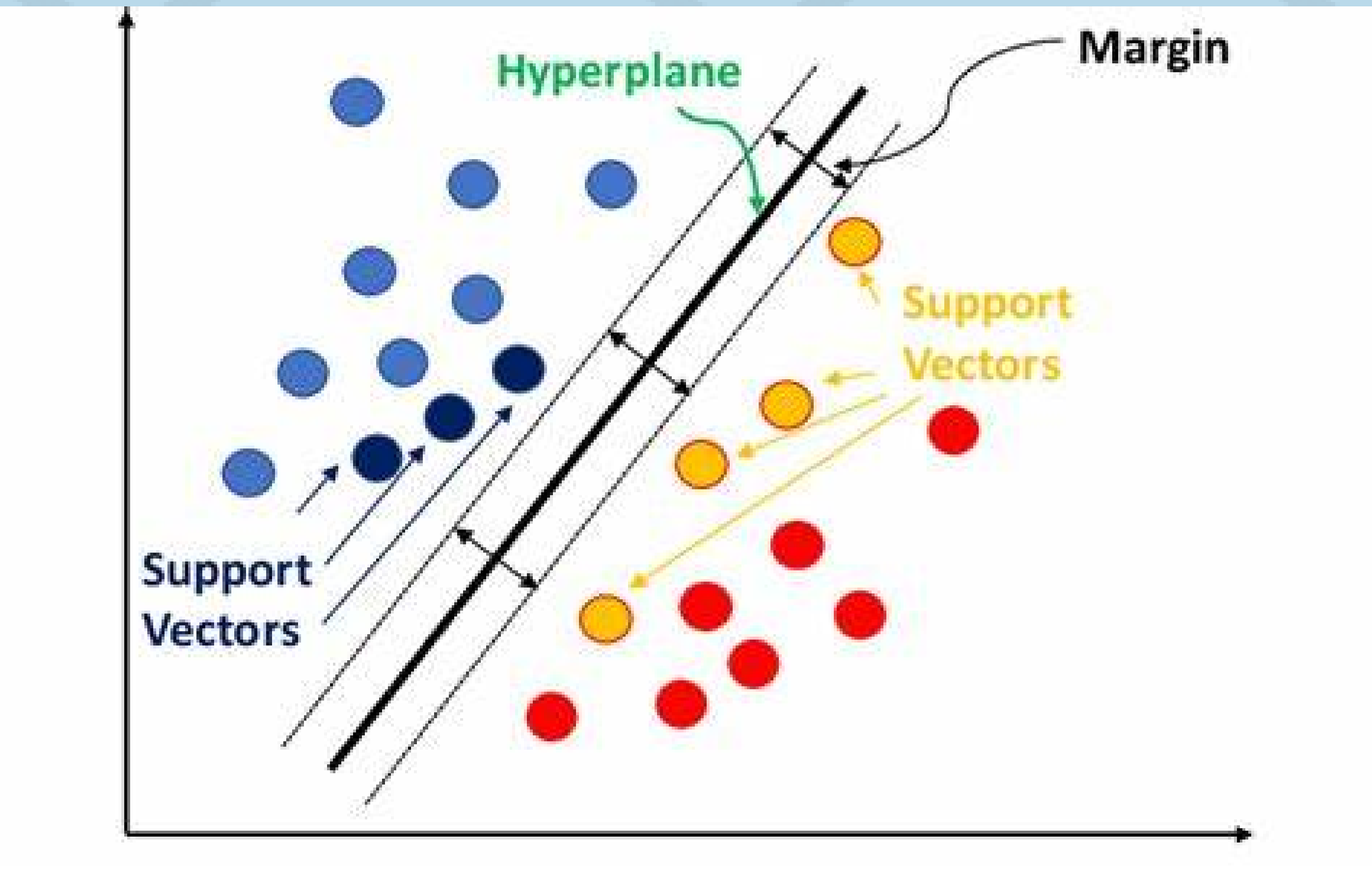
TAING KIMMENG

YOU PHAKKORN

TABLE OF CONTENT

1. Introduction of SVM

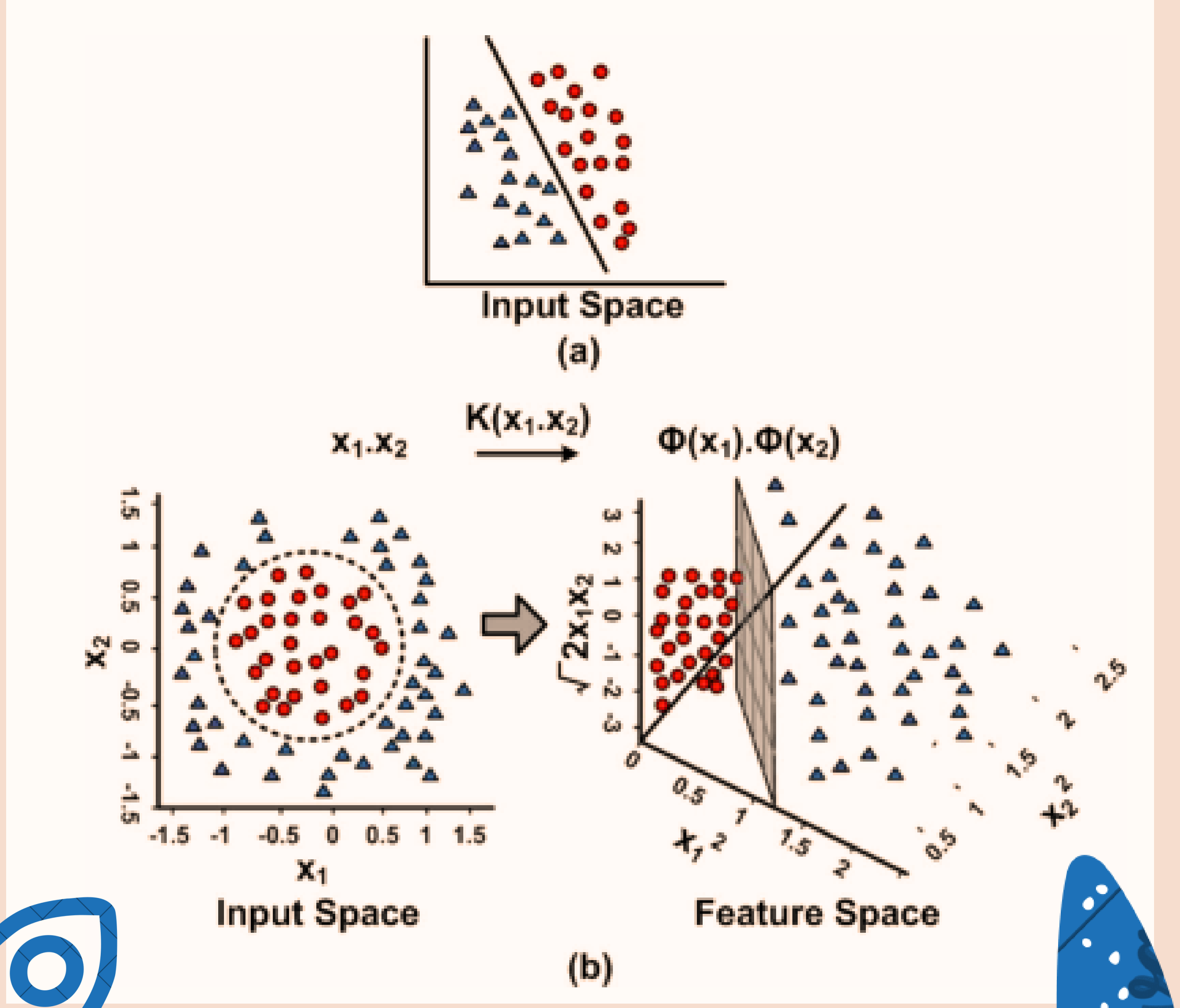
Support Vector Machines (SVM) is a powerful machine learning algorithm used for classification and regression analysis. It belongs to the family of supervised learning methods and is particularly well-suited for solving complex problems with non-linear decision boundaries. SVMs have gained popularity due to their ability to handle high-dimensional data, robustness to outliers, and theoretical foundations.



2. MOTIVATION BEHIND USING KERNEL FUNCTIONS FOR NONLINEAR SVM

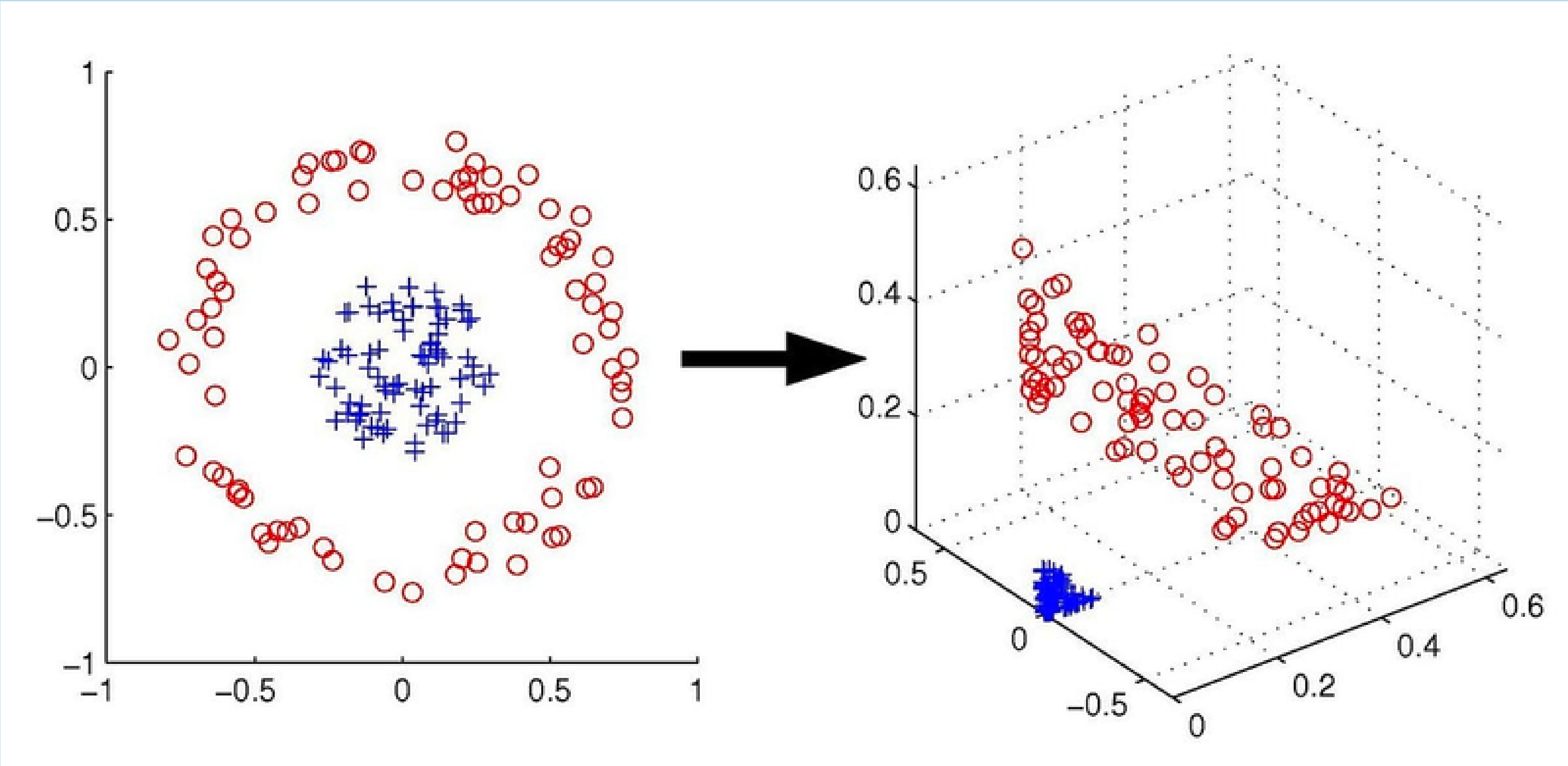
THE MOTIVATION BEHIND USING KERNEL FUNCTIONS FOR NONLINEAR SUPPORT VECTOR MACHINES (SVM) IS TO ENABLE SVMS TO EFFICIENTLY HANDLE NONLINEARLY SEPARABLE DATA.

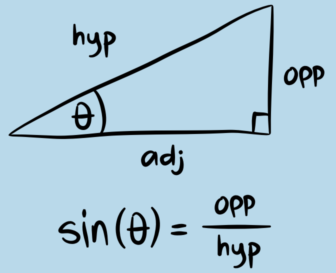
In real SVM applications, classes cannot be separated linearly, so we are working with nonlinear SVM to work around this problem, **That** is to say, by applying a nonlinear transformation to the data to change dimension and easily find a hyperplane classification in this new space, and also to give the classifier more freedom to correctly classify the points even if they are initially points on the wrong side of the initial hyperplane (non-separable categories).

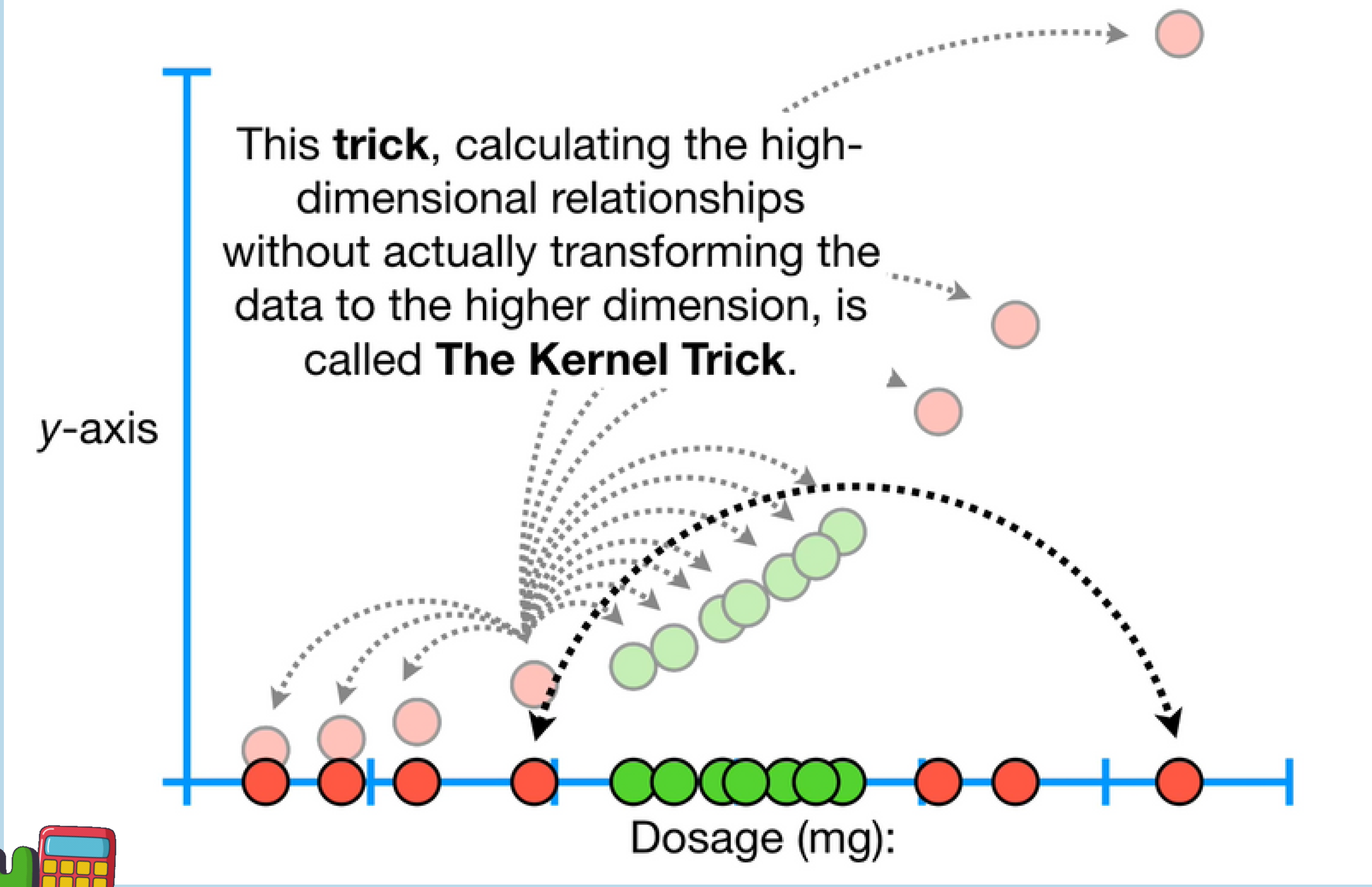


3. EXPLORING THE CONCEPT OF THE KERNEL TRICK AND ITS ROLE IN SVM.

- THE KERNEL TRICK IS A FUNDAMENTAL CONCEPT IN SUPPORT VECTOR MACHINES (SVM), A POPULAR MACHINE LEARNING ALGORITHM USED FOR CLASSIFICATION AND REGRESSION TASKS.
- IT ALLOWS SVM TO EFFICIENTLY HANDLE NON-LINEARLY SEPARABLE DATA BY IMPLICITLY MAPPING THE INPUT FEATURES TO A HIGHER-DIMENSIONAL SPACE.



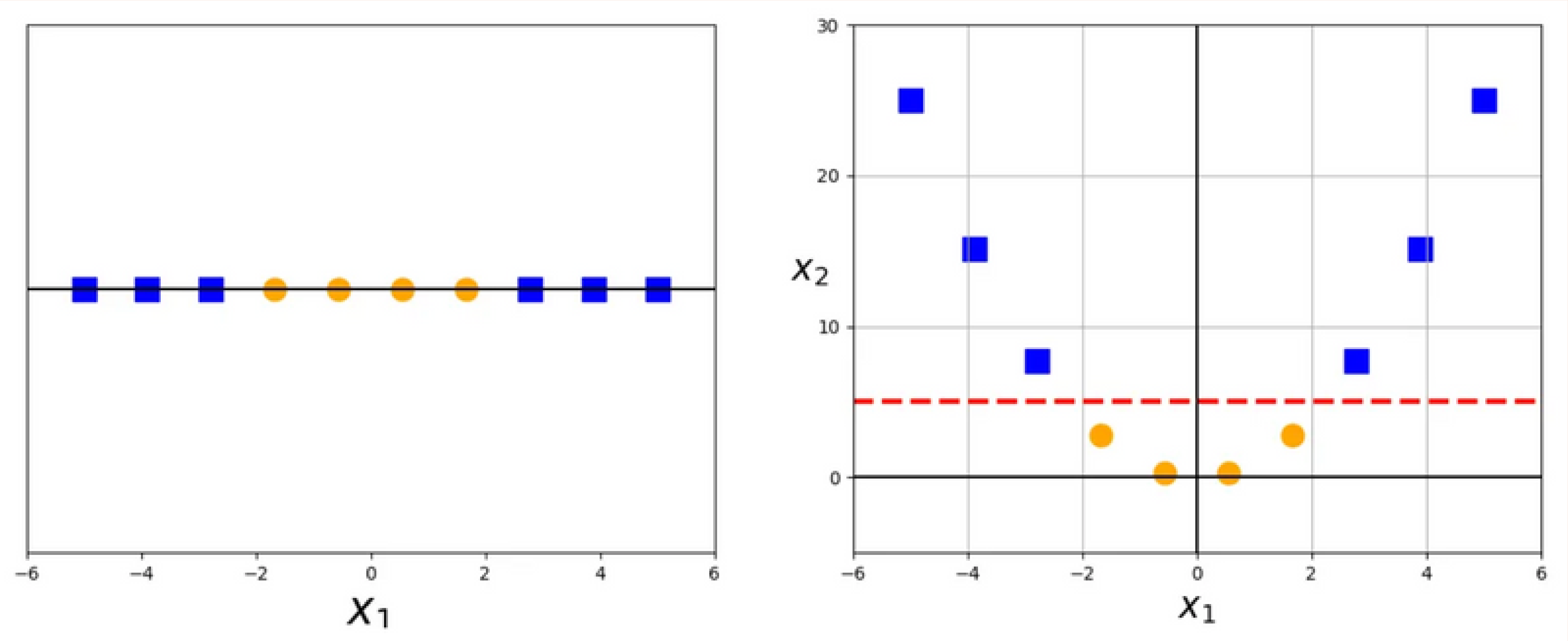




This **trick**, calculating the high-dimensional relationships without actually transforming the data to the higher dimension, is called **The Kernel Trick**.
y-axis
Dosage (mg):

Q. WHAT IS ROLE OF THE KERNEL TRICK IN SVM?

- TO INCREASE THE TRAINING TIME OF SVM
- TO MAP THE INPUT FEATURES TO A HIGHER-DIMENSIONAL SPACE
- TO REDUCE THE COMPLEXITY OF SVM
- TO ELIMINATE THE NEED FOR A KERNEL FUNCTION

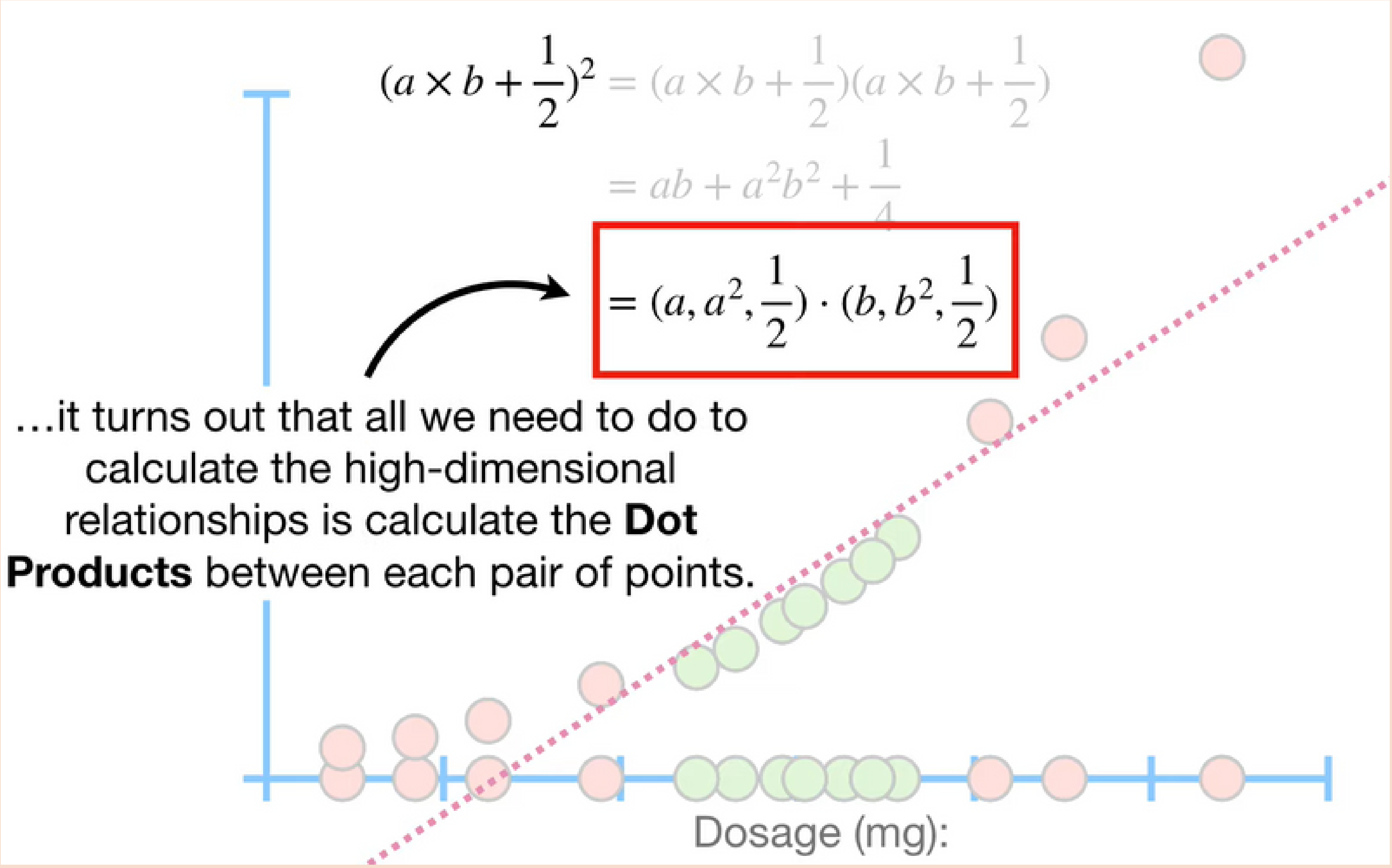


This data becomes linearly separable after a quadratic transformation to 2-dimensions.

HOW KERNEL FUNCTION MAP

THE KERNEL FUNCTION CALCULATES THE SIMILARITY OR INNER PRODUCT BETWEEN TWO DATA POINTS IN THE HIGHER-DIMENSIONAL SPACE. IT ALLOWS SVMS TO OPERATE IN THIS HIGHER-DIMENSIONAL SPACE WITHOUT EXPLICITLY COMPUTING THE COORDINATES. THE KEY IDEA IS THAT IF TWO DATA POINTS ARE SIMILAR IN THE HIGHER-DIMENSIONAL FEATURE SPACE, THEIR INNER PRODUCT WILL BE RELATIVELY HIGH. ON THE OTHER HAND, IF THEY ARE DISSIMILAR, THEIR INNER PRODUCT WILL BE LOW.

$$K(x, y) = \langle \phi(x), \phi(y) \rangle$$

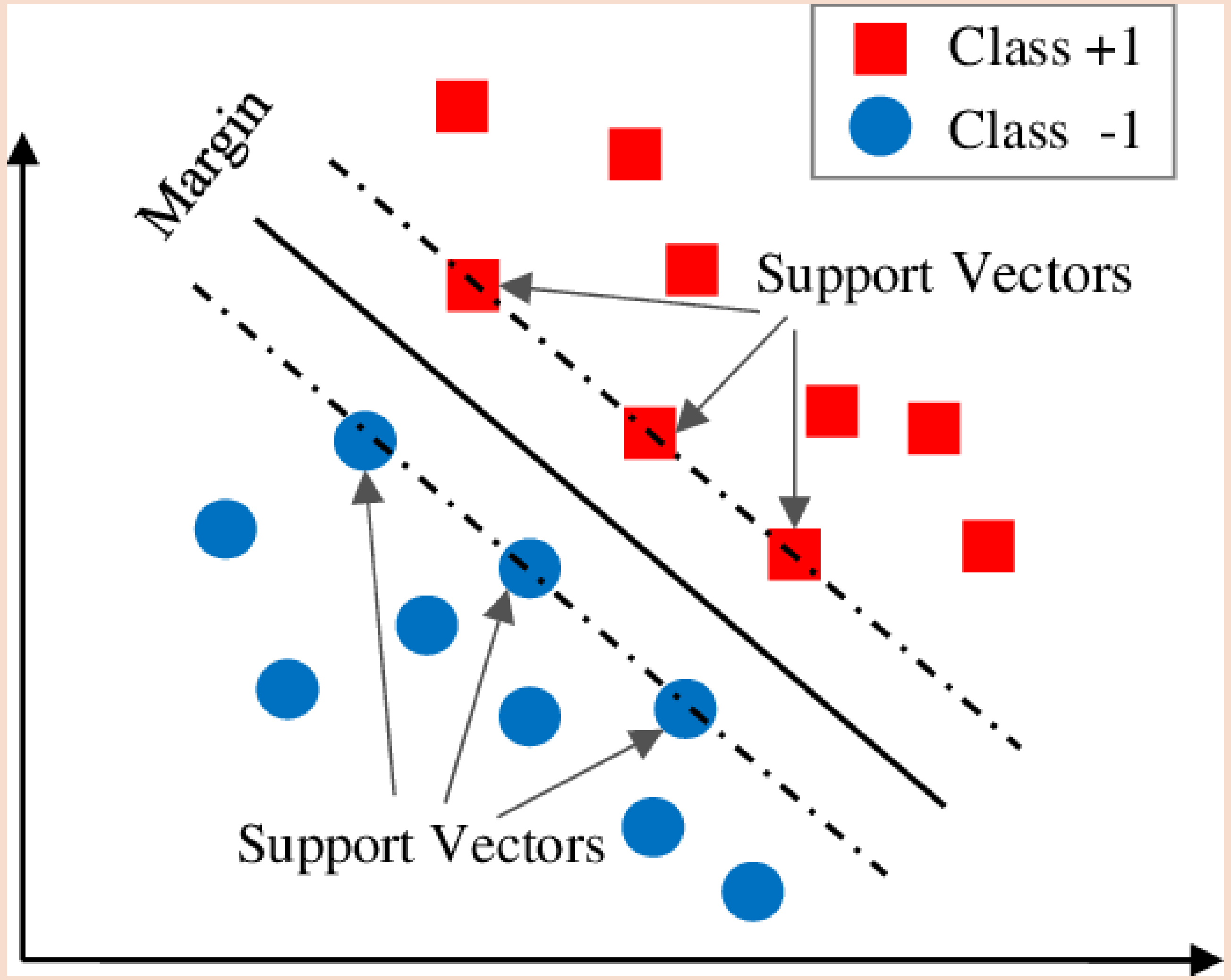


The mapping of the popular kernel functions

Linear Kernel

LINEAR KERNEL: THE LINEAR KERNEL SIMPLY COMPUTES THE INNER PRODUCT OF THE ORIGINAL FEATURE VECTORS, WHICH CORRESPONDS TO NO TRANSFORMATION OR MAPPING INTO A HIGHER-DIMENSIONAL SPACE.

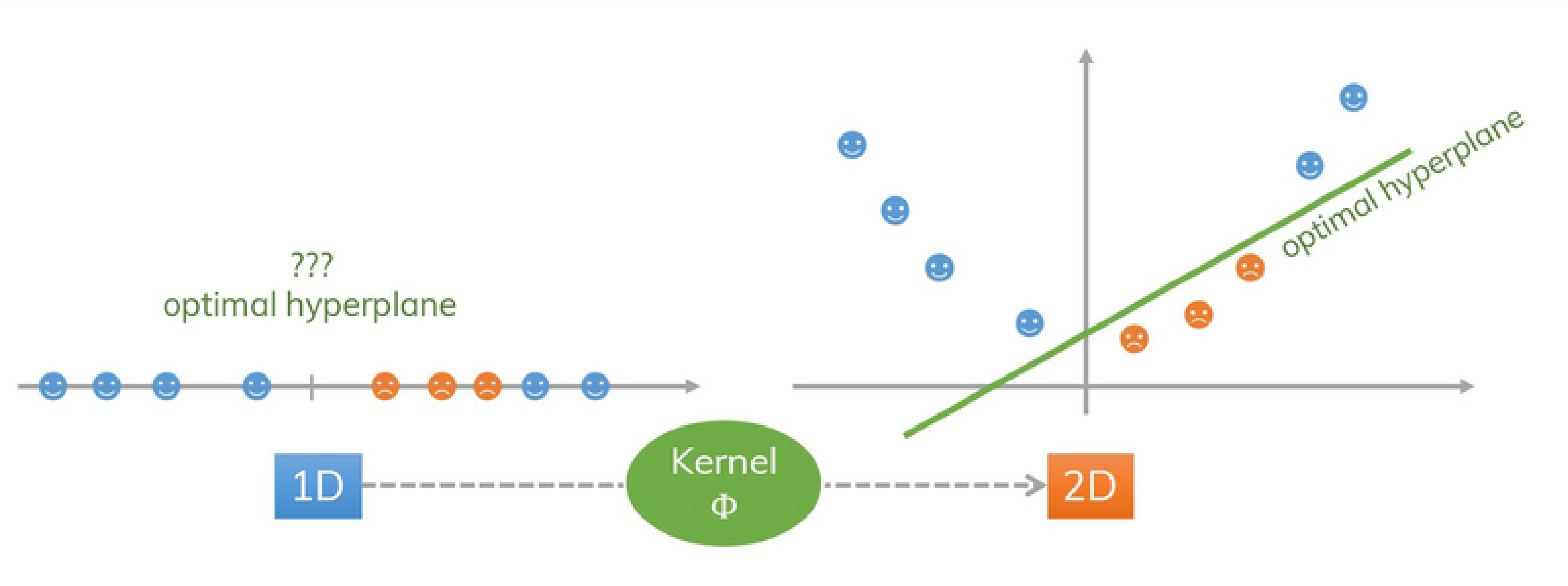
$$\kappa(x, y) = x^T y$$



Polynomial Kernel

POLYNOMIAL KERNEL: THE POLYNOMIAL KERNEL FUNCTION MAPS THE DATA INTO A HIGHER-DIMENSIONAL FEATURE SPACE USING POLYNOMIAL FUNCTIONS. IT CAPTURES INTERACTIONS BETWEEN FEATURES BY CONSIDERING ALL POSSIBLE POLYNOMIAL COMBINATIONS OF THE ORIGINAL FEATURES

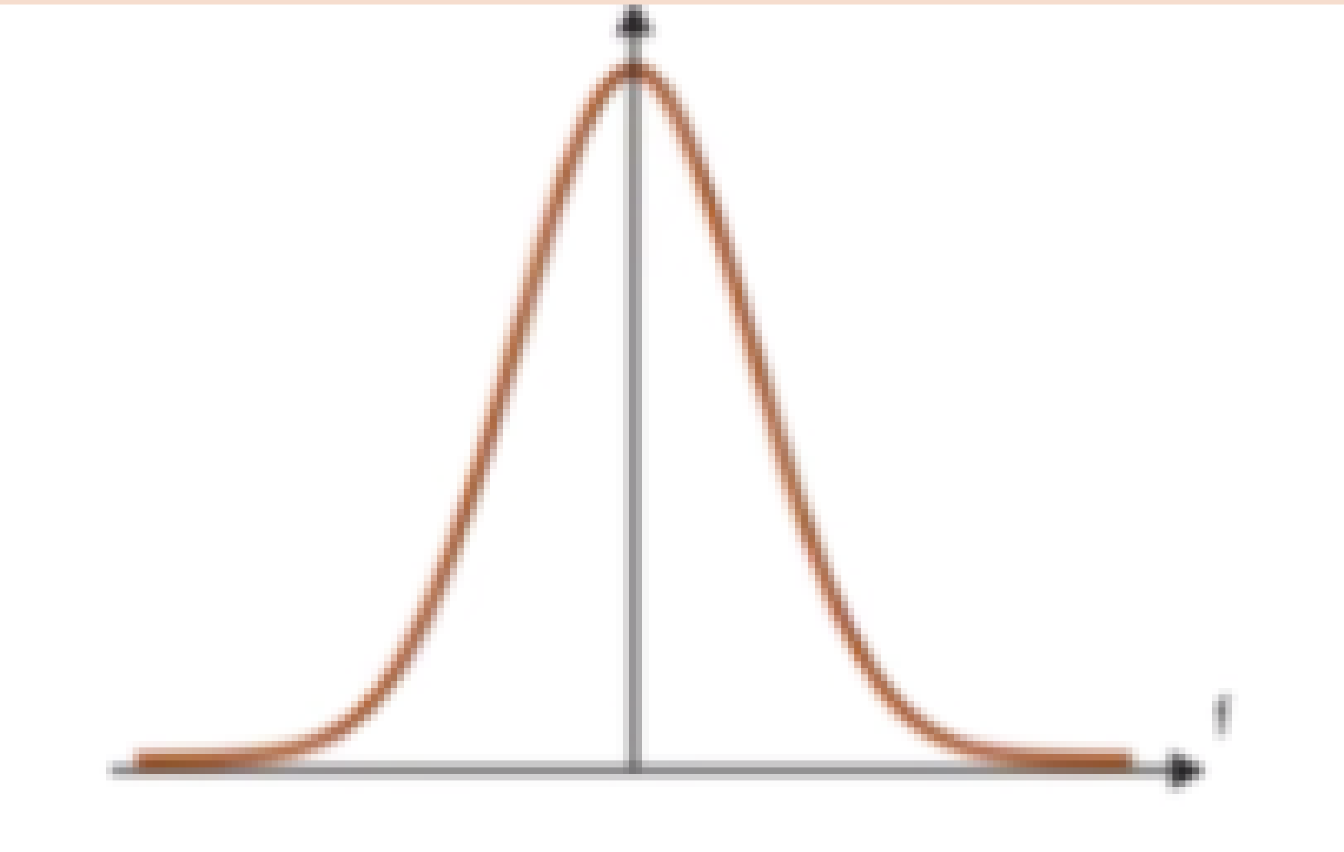
$$\kappa(x, y) = (x^T y)^p$$



RBF Kernel

RADIAL BASIS FUNCTION (RBF) KERNEL: THE RBF KERNEL USES A GAUSSIAN FUNCTION TO MAP THE DATA INTO AN INFINITE-DIMENSIONAL FEATURE SPACE. IT ASSIGNS A SIMILARITY SCORE BASED ON THE DISTANCE BETWEEN DATA POINTS IN THE ORIGINAL FEATURE SPACE. THE RBF KERNEL IS A POPULAR CHOICE BECAUSE IT ALLOWS SVMS TO MODEL COMPLEX NONLINEAR DECISION BOUNDARIES.

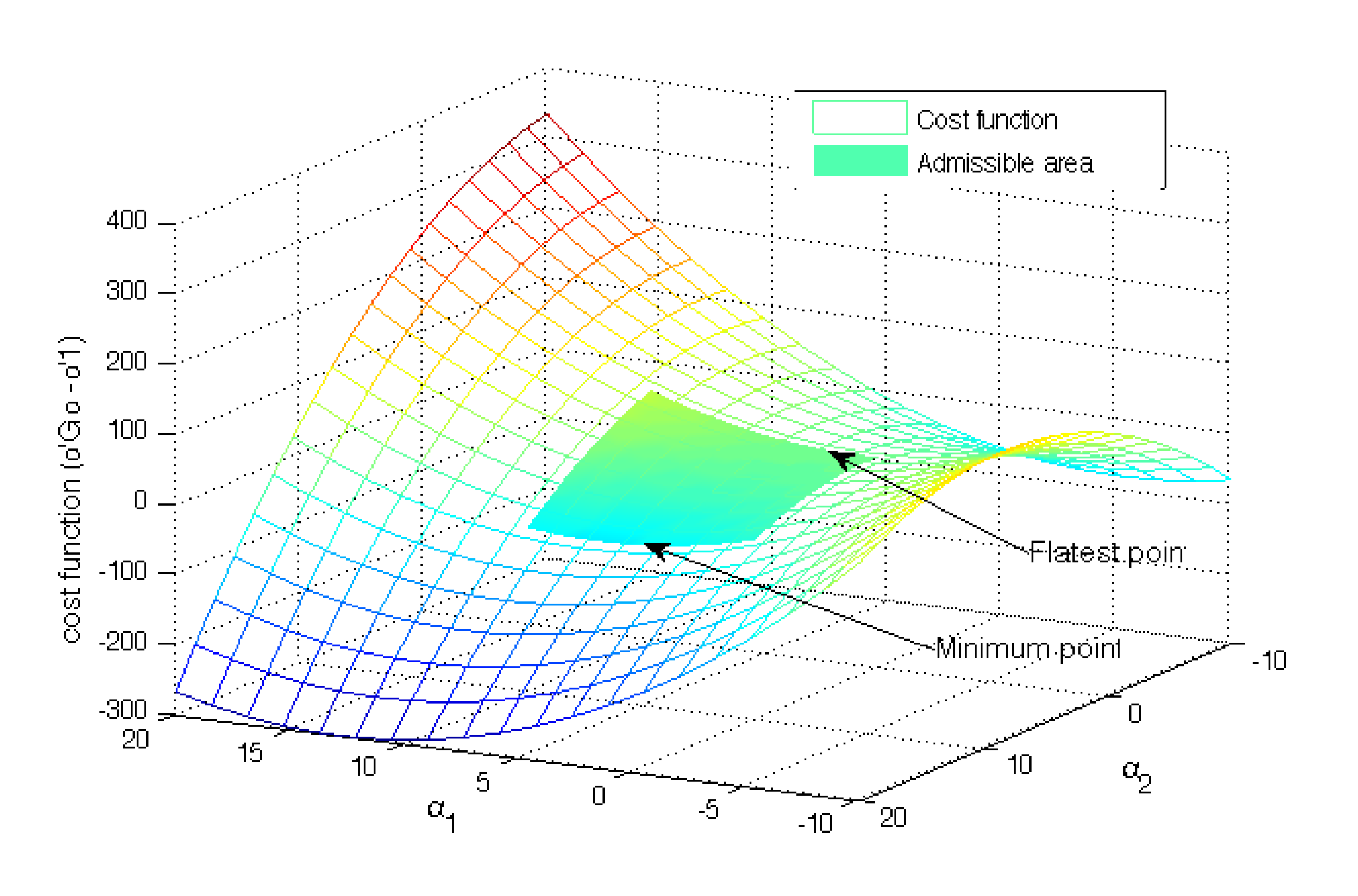
$$\kappa(x, y) = \exp\left(\frac{\|x - y\|^2}{2\sigma}\right)$$



Sigmoid Kernel

SIGMOID KERNEL: THE SIGMOID KERNEL MAPS THE DATA INTO A HIGHER-DIMENSIONAL FEATURE SPACE USING A SIGMOID FUNCTION. IT IS PRIMARILY USED FOR BINARY CLASSIFICATION PROBLEMS AND IS LESS COMMONLY USED COMPARED TO THE LINEAR, POLYNOMIAL, AND RBF KERNELS.

$$K(X, Y) = \tanh(\gamma \cdot X^T Y + r)$$



5. Advantages of Kernel trick in SVM

The kernel trick in Support Vector Machines (SVMs) offers several advantages that make it a valuable tool for handling nonlinear data:

- Nonlinear Decision Boundaries: The primary advantage of the kernel trick is its ability to create nonlinear decision boundaries in the original input space.
- Avoiding Explicit Feature Mapping: The kernel trick avoids the explicit computation and storage of the transformed feature space.
- Efficiency and Scalability: By implicitly operating in the transformed feature space, the kernel trick reduces the computational complexity of SVMs.
- Kernel Flexibility: The kernel trick offers a wide range of kernel functions to choose from, such as the Gaussian (RBF) kernel, polynomial kernel, sigmoid kernel, and more.
- Generalization Performance: The kernel trick allows SVMs to find the best decision boundary while minimizing the risk of overfitting.

limitation of Kernel trick in SVM

.

- KERNEL SELECTION: THE CHOICE OF THE KERNEL FUNCTION IS CRITICAL, AS DIFFERENT KERNELS CAPTURE DIFFERENT PATTERNS AND RELATIONSHIPS IN THE DATA.
- SENSITIVITY TO HYPERPARAMETERS: THE PERFORMANCE OF THE MODEL CAN BE SENSITIVE TO THE CHOICE OF THESE HYPERPARAMETERS, AND FINDING THE RIGHT COMBINATION OFTEN INVOLVES A TRIAL-AND-ERROR PROCESS.
- OVERFITTING: THIS FLEXIBILITY ALSO MAKES SVMS PRONE TO OVERFITTING, ESPECIALLY WHEN THE DIMENSIONALITY OF THE FEATURE SPACE IS VERY HIGH OR THE DATASET IS SMALL.
- MEMORY REQUIREMENTS: IN SOME CASES, THE KERNEL TRICK CAN RESULT IN A SIGNIFICANT INCREASE IN MEMORY REQUIREMENTS. THE IMPLICIT MAPPING OF DATA INTO A HIGHER-DIMENSIONAL SPACE MAY REQUIRE ADDITIONAL STORAGE

THANK YOU