Kernel Function IN SVM

instructor: Mr. Sophal Cam Chan

Meet the Team

- SEAN VENGNGY
- SANG RITHPORK
- CHEA MAKARA
- HIN BUNRA
- THOU CHANMAKARA
- PHAI RATHA
- LEAT Seanglong

Table of CONTENTS

INTRODUCTION

The choice of kernel function in SVM is another important factor that affects the accuracy of classification. A kernel function is responsible for mapping the input data into a higher-dimensional feature space, where it becomes easier to find a linear hyperplane that separates the classes.

Kernel functions play a crucial role in support vector machines (SVM) by enabling them to handle non-linearly separable data. In SVM, the goal is to find a hyperplane that maximally separates the classes in the feature space. However, in many real-world scenarios, the data may not be linearly separable.

Understanding the mathematical properties of kernel functions

Kernel functions have certain mathematical properties that affect the performance of machine learning algorithms. One important property is positive definiteness, which means that the kernel matrix is always positive semidefinite. This property ensures that the kernel function produces valid similarity measures between data points.

Another key property of kernel functions is symmetry, which means that the similarity between two data points is the same regardless of their order. This property is important for many machine learning algorithms, such as support vector machines, which rely on the symmetry of the kernel function to optimize the decision boundary

One of the key applications of kernel functions in machine learning is in support vector machines (SVMs). SVMs are a type of supervised learning algorithm that can be used for classification or regression tasks. The basic idea behind SVMs is to find a hyperplane that separates the data into different classes, and kernel functions can be used to transform the input data into a higher-dimensional space where this separation is easier to achieve. For example, the radial basis function (RBF) kernel is commonly used in SVMs because it can effectively separate non-linearly separable data.

Here are a few commonly used kernel functions and their mathematical properties:

LINEAR

The linear kernel performs a simple dot product between the input feature vectors.

Formula: $K(x, y) = x * y$

POLYNOMIAL

The polynomial kernel computes the similarity between two vectors as the sum of their polynomial terms.

Formula: $K(x, y) = (x * y + c)^d$

RBF

The RBF kernel calculates the similarity between two vectors based on the Gaussian distribution of their distances.

Formula: $K(x, y) = exp(-gamma * x - y^2)$

SIGMOID

The sigmoid kernel measures the similarity between two vectors using sigmoid functions.

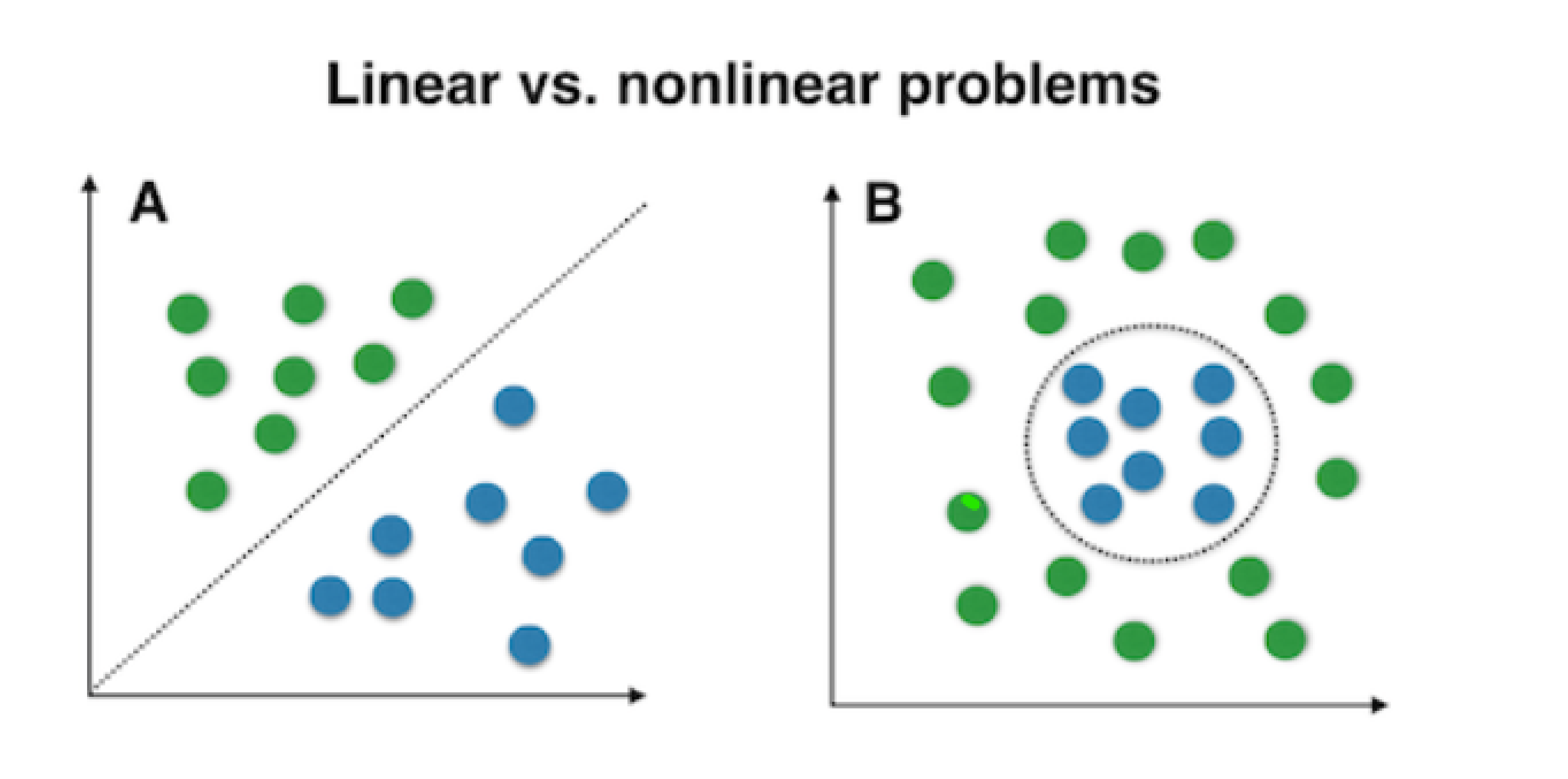
Formula: $K(x, y) = tanh(alpha * (x * y) + c)$

THESE ARE JUST A FEW EXAMPLES OF KERNEL FUNCTIONS USED IN SVMS. EACH KERNEL FUNCTION HAS ITS OWN MATHEMATICAL PROPERTIES AND IMPLICATIONS ON THE SVM'S PERFORMANCE. THE CHOICE OF THE KERNEL DEPENDS ON THE SPECIFIC CHARACTERISTICS OF THE DATASET AND THE PROBLEM

3. Exploring popular kernel functions, such as linear, polynomial, Gaussian (RBF), and sigmoid kernels

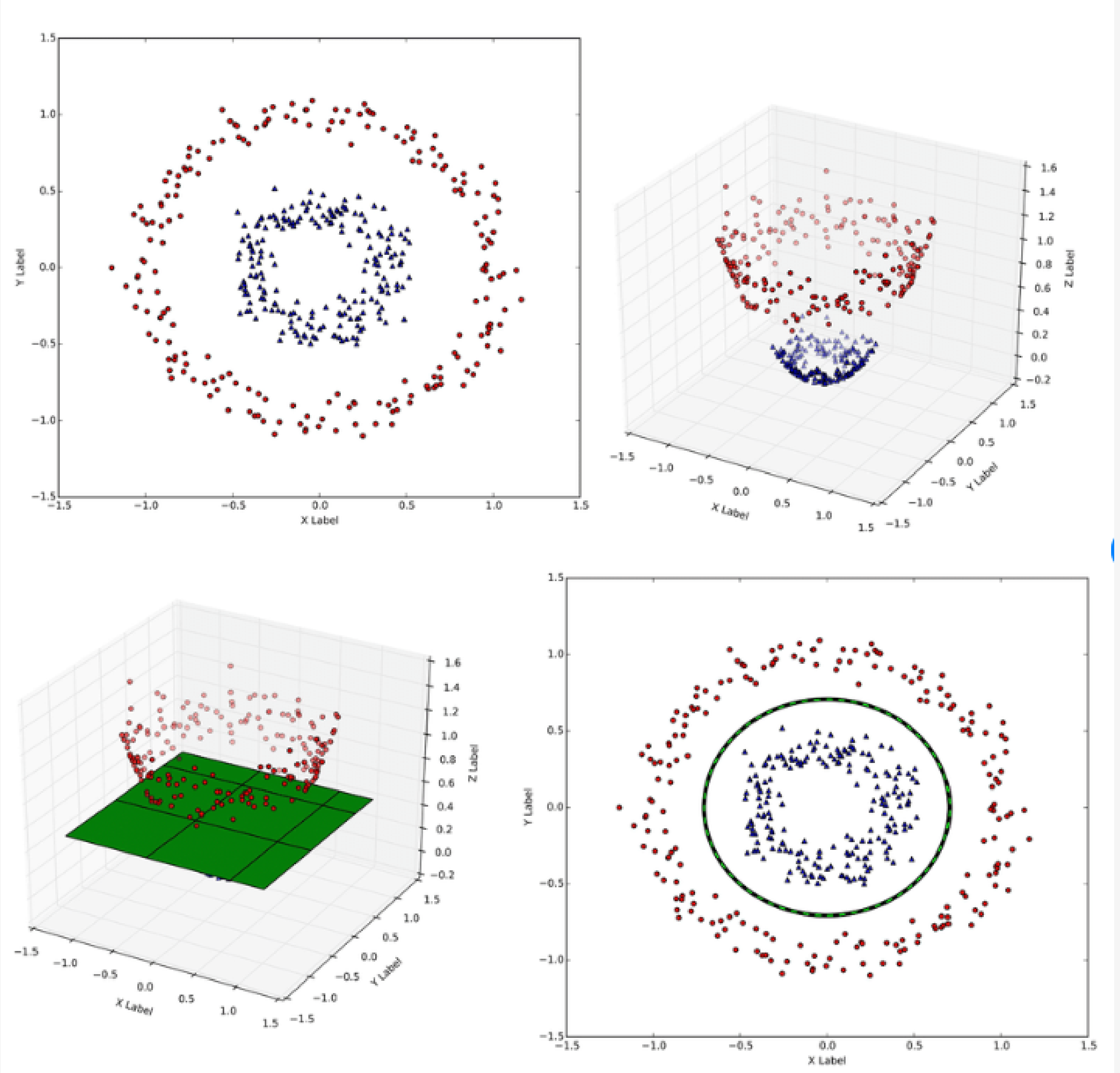
Linear

Linear Kernel: The linear kernel is the simplest kernel function. It represents the dot product of two feature vectors in the input space. Mathematically, the linear kernel can be defined as K(x, y) = x * y, where x and y are feature vectors. It works well when the data is linearly separable.



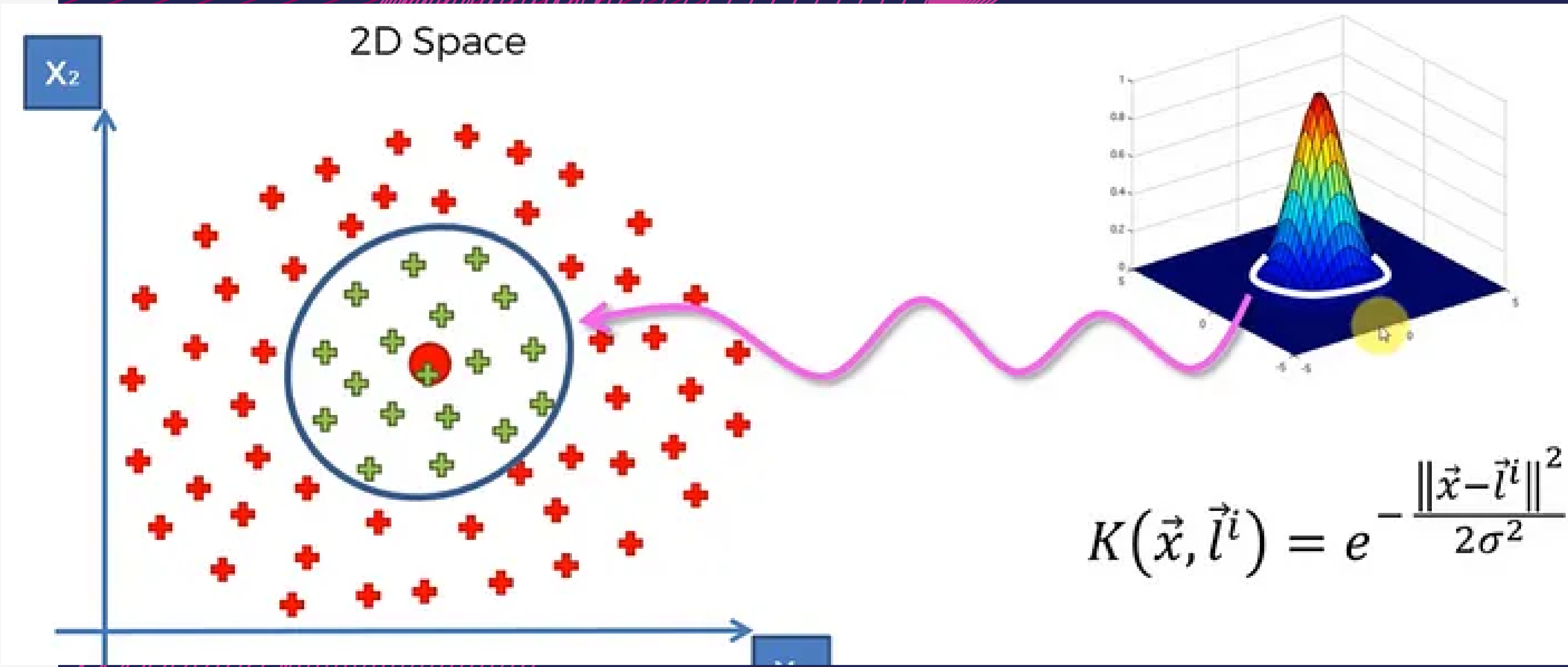
Polynomial

Polynomial Kernel: The polynomial kernel allows for nonlinear decision boundaries by computing the similarity between two samples as the power of the dot product of their feature vectors. Mathematically, the polynomial kernel can be defined as K(x, y) = (alpha * x * y + c)^d, where alpha is a scaling factor, c is a constant term, and d is the degree of the polynomial. Higher degrees can capture complex relationships but may be prone to overfitting.



Gaussian (RBF)

Gaussian (RBF) Kernel: The Gaussian kernel, also known as the radial basis function (RBF) kernel, is commonly used in SVMs. It measures the similarity between samples based on the Euclidean distance between their feature vectors. Mathematically, the Gaussian kernel can be defined as K(x, y) = exp(-gamma * ||x - y||^2), where gamma is a parameter that determines the kernel's influence. A higher value of gamma leads to a narrower peak and more localized decision boundaries.



SIGMOID KERNELS

4. Sigmoid Kernel: The sigmoid kernel is another commonly used kernel function. It computes the similarity between samples based on the hyperbolic tangent function. Mathematically, the sigmoid kernel can be defined as K(x, y) = tanh(alpha * x * y + c), where alpha and c are parameters. The sigmoid kernel can handle nonlinearity and is particularly useful when dealing with neural networks.

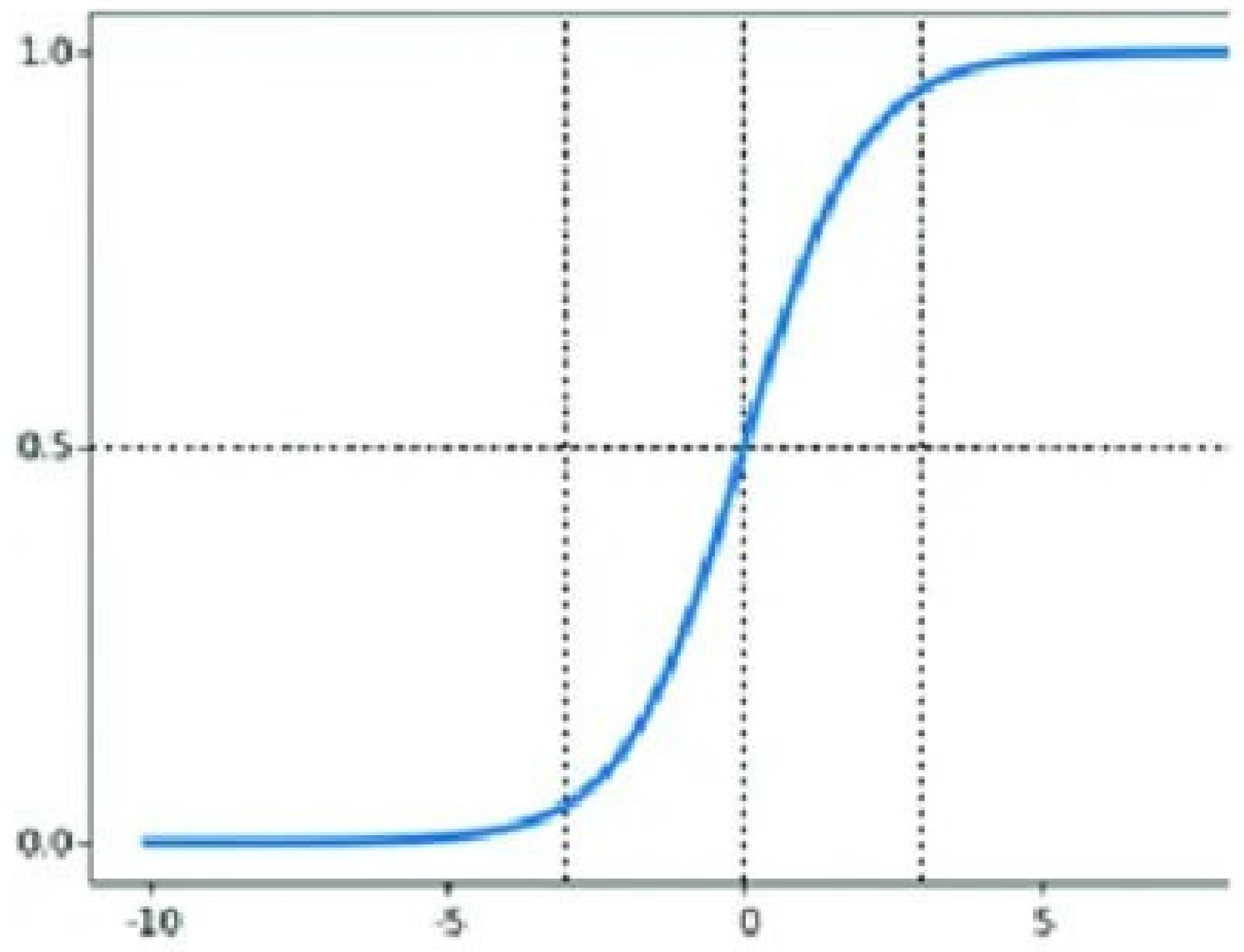


Image Source

Discussing the impact of kernel choice on SVM performance and decision boundaries

The choice of kernel function can have a significant impact on the performance and decision boundaries of an SVM model. Some common kernel functions include:

- Linear kernel: This is the simplest kernel function and it maps the data into a higher-dimensional space where the data can be separated by a linear hyperplane. The linear kernel is only effective if the data is linearly separable in the original space.

- Radial basis function (RBF) kernel: This is a more powerful kernel function that can be used to map data that is not linearly separable in the original space. The RBF kernel works by computing the similarity between each data point and a set of support vectors. The decision boundary is then found by maximizing the margin between the two classes of data.

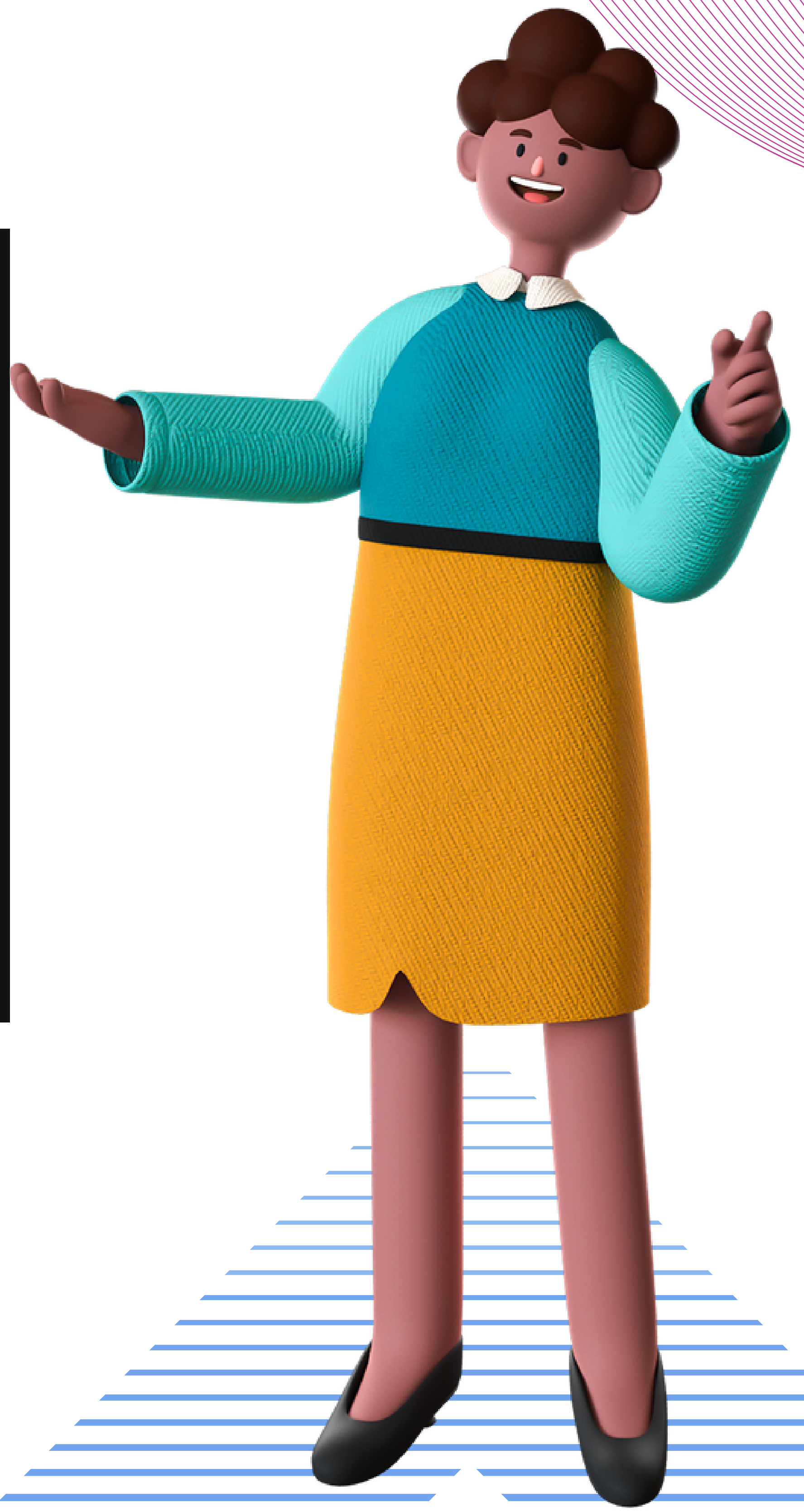
- Polynomial kernel: This kernel function is a generalization of the linear kernel and it can be used to map data that is not linearly separable in the original space. The polynomial kernel works by computing the dot product of each data point and a set of support vectors raised to a certain power. The decision boundary is then found by maximizing the margin between the two classes of data.

- The sigmoid kernel can have a significant impact on the performance and decision boundaries of an SVM model. In general, the sigmoid kernel can achieve better performance than the linear kernel or the polynomial kernel for data that is not linearly separable in the original space. However, the sigmoid kernel can also be more computationally expensive.

The table summarizes of impact and decision boundary.

The following table summarizes the impact of kernel choice on SVM performance and decision boundaries:

Kernel function	Performance	Decision boundary
Linear kernel	Good for linearly separable data	Straight line
RBF kernel	Good for non-linearly separable data	Curve
Polynomial kernel	Good for non-linearly separable data	Straight line, curve, or combination of both



Sigmoid kernel

Feature	Impact
Parameter	gamma
Data	Non-linearly separable data
Computational complexity	More computationally expensive
Decision boundary	Sigmoid curve

Impact summary

In summary, the choice of kernel in SVMs significantly affects the model's performance and the shape of the decision boundaries. Different kernel functions offer various transformations of the data, enabling SVMs to handle different types of classification problems. Understanding the data and selecting an appropriate kernel is essential to achieve accurate and robust results.

WARDIERE INC.

THANK YOU