

## **SVM Kernel Functions – ‘Coz your SVM knowledge is incomplete without it**

The kernel functions play a very important role in SVM. Their job is to take data as input and they transform it in any required form. They are significant in SVM as they help in determining various important things.

In this article, we will be looking at various types of kernels. We will also be looking at how the kernel works and why is it necessary to have a kernel function. This will be an important article, as it will give an idea of what kernel function should be used in specific programs.

### **What is Kernel?**

A kernel is a function used in SVM for helping to solve problems. They provide shortcuts to avoid complex calculations. The amazing thing about kernel is that we can go to higher dimensions and perform smooth calculations with the help of it.

We can go up to an **infinite** number of dimensions using kernels.

Sometimes, we cannot have a hyperplane for certain problems. This problem arises when we go up to higher dimensions and try to form a hyperplane.

A kernel helps to form the hyperplane in the higher dimension without raising the complexity.

### **Working of Kernel Functions**

Kernels are a way to solve non-linear problems with the help of linear classifiers. This is known as the **kernel trick** method. The kernel functions are used as parameters in the SVM codes. They help to determine the shape of the hyperplane and decision boundary.

We can set the value of the kernel parameter in the SVM code.

The value can be any type of kernel from linear to polynomial. If the value of the kernel is linear then the decision boundary would be linear and two-dimensional. These kernel functions also help in giving decision boundaries for higher dimensions.

We do not need to do complex calculations. The kernel functions do all the hard work. We just have to give the input and use the appropriate kernel. Also, we can solve the overfitting problem in SVM using the kernel functions.

**Overfitting** happens when there are more feature sets than sample sets in the data. We can solve the problem by either increasing the data or by choosing the right kernel.

There are kernels like **BBF** that work well with smaller data as well. But, RBF is a universal kernel and using it on smaller datasets might increase the chances of overfitting.

Hence, the use of simpler kernels like linear and polynomial is encouraged.

## Rules of Kernel Functions

There are certain rules which the kernel functions must follow.

Also, these rules are the deciding factors of what kernel should be implemented for classification. One such rule is the **moving window classifier** or we can also call it the **window function**.

This function is shown as:

$$f_n(\mathbf{x}) = 1, \text{ if } \sum_{i=1}^n \mathbb{I}(\|\mathbf{x} - \mathbf{x}_i\| \leq h) (y_i=1) > \sum_{i=1}^n \mathbb{I}(\|\mathbf{x} - \mathbf{x}_i\| \leq h) (y_i=0)$$

$$f_n(\mathbf{x}) = 0, \text{ otherwise.}$$

Here, the summations are from  $i=1$  to  $n$ . 'h' is the width of the window. This rule assigns weights to the points at a fixed distance from 'x'.

' $\mathbf{x}_i$ ' are the points nearby 'x'.

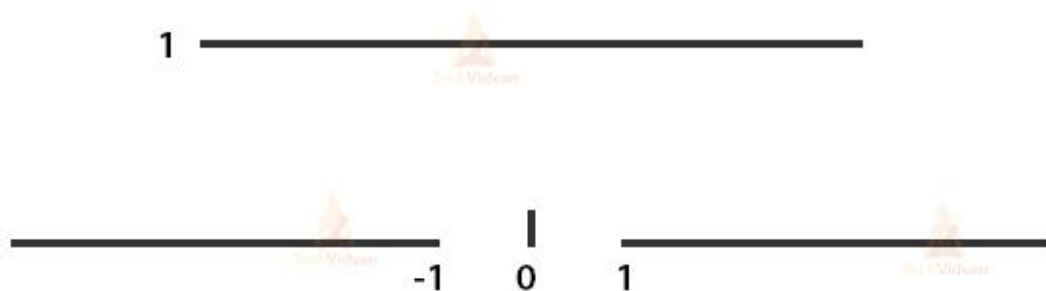
It is essential that the weights should be distributed in the direction of  $\mathbf{x}_i$ . This ensures the smooth working of the weight functions. These weight functions are the kernel functions.

The kernel function is represented as **K:  $\mathbb{R}^d \rightarrow \mathbb{R}$** .

The functions are generally positive and monotonically decreasing. There are some forms of these functions:

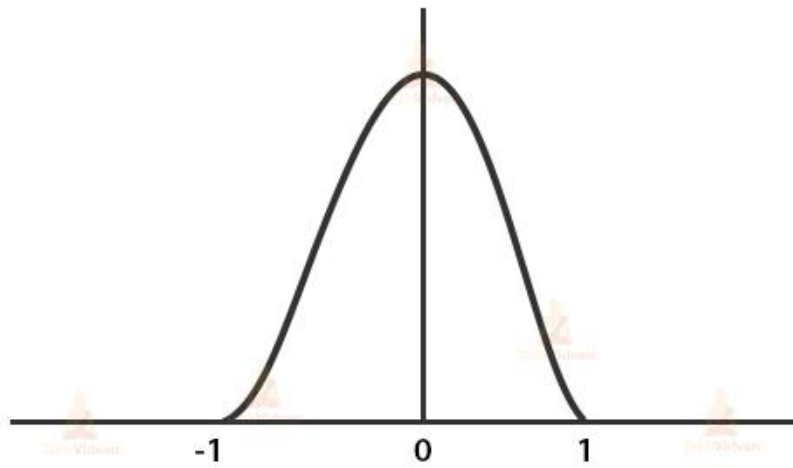
- Uniform kernels:  $K(\mathbf{x}) = \mathbb{I}(\|\mathbf{x}\| \leq 1)$

## Uniform Kernel Function



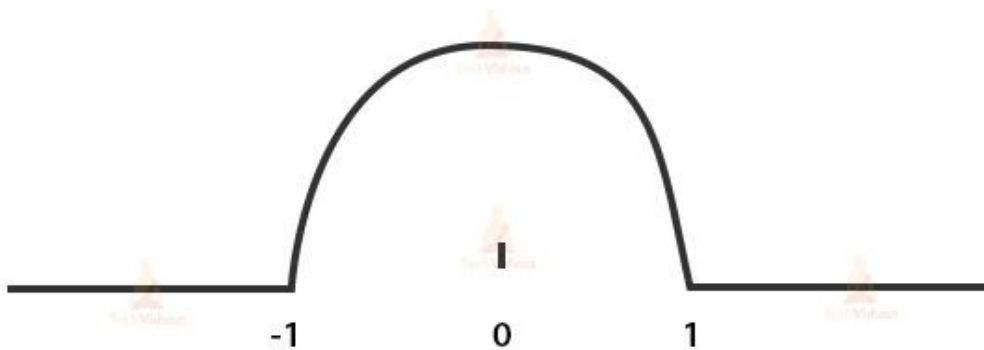
- Gaussian kernels:  $K(\mathbf{x}) = \exp(-||\mathbf{x}||^2)$

## Gaussian Kernel Function

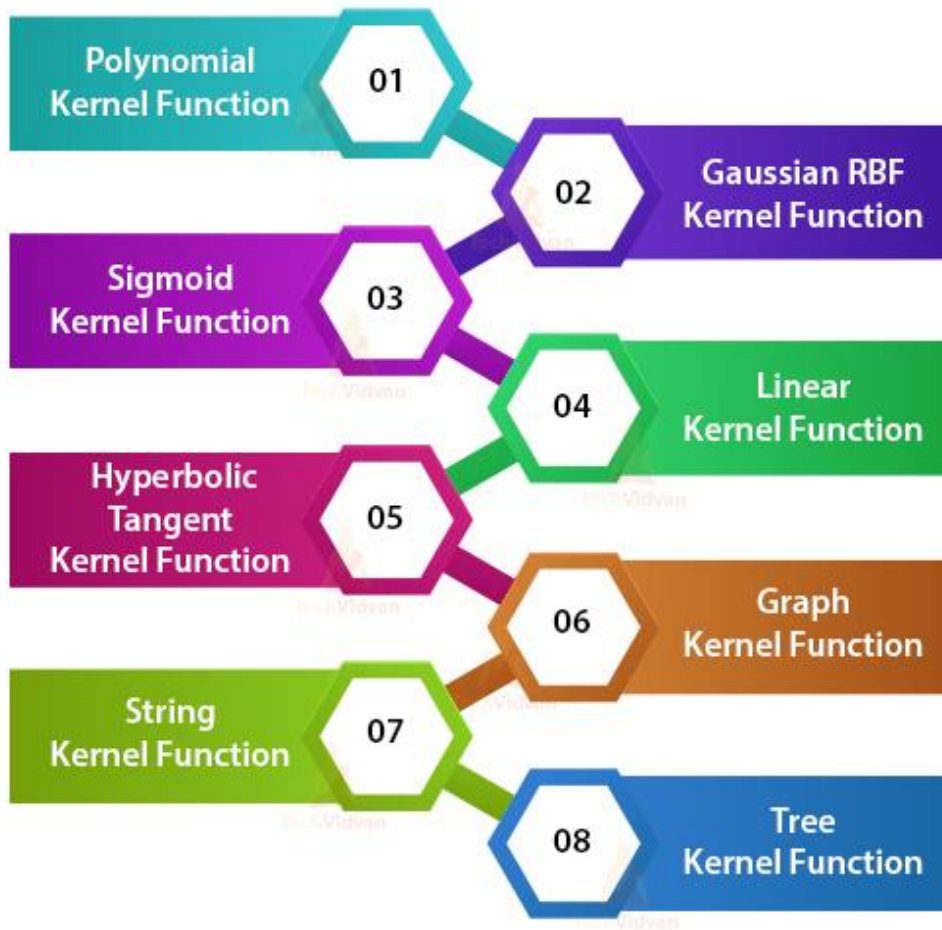


- Epanechnikov kernels:  $\frac{3}{4}(1-||\mathbf{x}||^2)$  ( $||\mathbf{x}|| \leq 1$ )

## Epanechnikov Kernel Function



# Types of Kernel Functions



## Types of Kernel Functions:

### 1. *Polynomial Kernel Function*

The polynomial kernel is a general representation of kernels with a degree of more than one. It's useful for image processing.

There are two types of this:

- **Homogenous Polynomial Kernel Function**

$K(x_i, x_j) = (x_i \cdot x_j)^d$ , where '.' is the dot product of both the numbers and  $d$  is the degree of the polynomial.

- **Inhomogeneous Polynomial Kernel Function**

$K(x_i, x_j) = (x_i \cdot x_j + c)^d$  where  $c$  is a constant.

## 2. *Gaussian RBF Kernel Function*

RBF is the radial basis function. This is used when there is no prior knowledge about the data.

It's represented as:

$$K(x_i, x_j) = \exp(-\gamma ||x_i - x_j||)^2$$

## 3. *Sigmoid Kernel Function*

This is mainly used in neural networks. We can show this as:

$$K(x_i, x_j) = \tanh(\alpha x_i^a y + c)$$

## 4. *Hyperbolic Tangent Kernel Function*

This is also used in neural networks. The equation is:

$$K(x_i, x_j) = \tanh(kx_i \cdot x_j + c)$$

## 5. *Linear Kernel Function*

This kernel is one-dimensional and is the most basic form of kernel in SVM. The equation is:

$$K(x_i, x_j) = x_i \cdot x_j + c$$

## 6. *Graph Kernel Function*

This kernel is used to compute the inner on graphs. They measure the similarity between pairs of graphs. They contribute in areas like **bioinformatics**, **chemoinformatics**, etc.

## 7. *String Kernel Function*

This kernel operates on the basis of strings. It is mainly used in areas like **text classification**. They are very useful in text mining, genome analysis, etc.

## 8. *Tree Kernel Function*

This kernel is more associated with the tree structure. The kernel helps to split the data into tree format and helps the SVM to distinguish between them. This is helpful in language classification and it is used in areas like **NLP**.

## Summary

This was all about the SVM Kernel Functions article. There are various more types of kernels but we have covered the ones that are useful for beginners.

From this article, we understood that kernel is a highly mathematical concept of Machine Learning. We saw how it works and how and where are the various types of kernels used. Then, we also saw what type of problems does the kernel tackles. There were also some kernel rules which are important to understand.