**Kernel**

A kernel is a function that computes a similarity measure between two data points in a higher-dimensional space without explicitly mapping the data space. The kernel trick allows us to implicitly calculate this similarity without the need for computationally expensive transformations.

Kernels are widely used in algorithms like Support Vector Machines (SVM), Gaussian Processes and other models that rely on , measuring similarity between data points.

Key Points about Kernels:

1. Mapping to Higher Dimensions.

Many machine learning algorithms (like SVM) work better when the data is transformed into a higher-dimensional space where the data points become linearly.

However, directly mapping the data to this higher-dimensional space can be computationally expensive. Instead, the kernel trick allows us to compute the dot product of data points in this higher-dimensional space without actually performing the mapping.

1. Kernels as Functions:

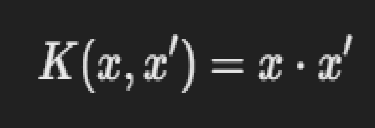
A kernel is a function that computes the dot product between two data points in this higher-dimensional space. The beauty of the kernel trick is that we don’t need to computer the mapping explicitly.

In essence, the kernel computes the similarity between two data points in a higher-dimensional space, allowing algorithms to work as though the were using a higher-dimensional representation.

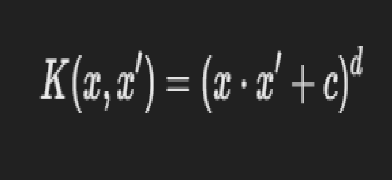
Kernel function in machine learning include

1. Linear kernel
2. Polynomial kernel
3. Radial basis function (RBF)
4. Sigmoid

Linear Kernel :  
 this is the simplest kernel and is essentially just the dot product between two vectors. It’s used when the data is linearly separable.

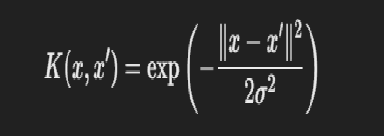


 **Polynomial Kernel**:  
This kernel computes a similarity measure using a polynomial function of the dot product. It is suitable for problems where the decision boundary between classes is non-linear.



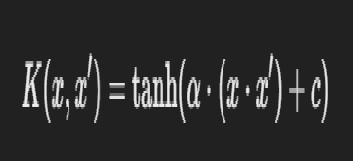
Where ccc is a constant and ddd is the degree of the polynomial.

 **Gaussian (RBF) Kernel**:  
The Radial Basis Function (RBF) or Gaussian kernel is one of the most popular kernels. It computes the similarity between two data points based on their Euclidean distance. It is often used in problems where the data is not linearly separable.



Where ∥x−x′∥\|x - x'\|∥x−x′∥ is the Euclidean distance between xxx and x′x'x′, and σ\sigmaσ is a parameter that controls the spread of the kernel.

 **Sigmoid Kernel**:  
The sigmoid kernel is based on the hyperbolic tangent function. It’s often used in neural networks but can also be used for SVMs.



Where α\alphaα and ccc are kernel parameters.

**Linear vs Non-linear Kernels**:

* A **linear kernel** assumes that the data is linearly separable.
* **Non-linear kernels**, like the **Gaussian kernel**, can be used when the data is not linearly separable and needs to be mapped into a higher-dimensional space for better classification.