Data Science Academy

www.datascienceacademy.com.br

Machine Learning

Kernel RBF x Kernel Linear x Kernel Polinomial

As funções de Kernel são usadas para mapear o conjunto de dados original (linear / não linear) em um espaço dimensional mais alto, com vista a torná-lo linear.

A fórmula abaixo representa a definição padrão para classificar um ponto de dado (o que é usado por diversos algoritmos):

$$g(\vec{X}) = \vec{W}^T \vec{X} + b$$

Mas essa definição padrão pode não funcionar com dados linearmente inseparáveis. Nós podemos então adicionar uma função para permitir a separação dos dados, conforme mostrado na fórmula abaixo:

$$g(\vec{X}) = \vec{W} \cdot \vec{f}(\vec{X}) + b$$

A função f(x) fornece superfícies de decisão não linear implícitas para os dados originais. **A f(x) é uma função chamada Kernel.**

O Kernel Linear, Polinomial e RBF (ou Gaussiano) são simplesmente diferentes opções para estabelecer o limite de decisão do hiperplano entre as classes.

Geralmente, os Kernels Lineares e Polinomiais consomem menos tempo e fornecem menos precisão do que os Kernels RBF.

O Kernel Linear é o que você esperaria, um modelo linear. O RBF usa curvas normais em torno dos pontos de dados e as soma para que o limite de decisão possa ser definido por um tipo de condição de topologia, como curvas em que a soma está acima de um valor de 0,5.

Em termos matemáticos, nós temos:

Kernel Linear:

$$K(X,Y) = X^T Y$$

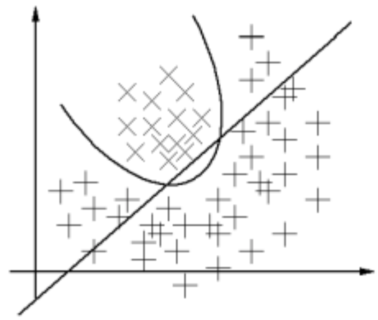
Kernel Polinomial

$$K(X,Y) = (\gamma \cdot X^T Y + r)^d, \gamma > 0$$

Kernel RBF:

$$K(X,Y) = \exp(\|X - Y\|^2 / 2\sigma^2)$$

Observe que em todos os casos o que temos é uma transformação aplicada aos dados.



Formalmente, uma "função de kernel" é qualquer função que satisfaça a condição de Mercer. Uma função, k (x, y), satisfaz a condição de Mercer se, para todas as funções quadradas e integráveis f (x),

$$\iint k(x,y)f(x)f(y)dxdy \geq 0$$

Esta condição é satisfeita por produtos internos (dot products):

$$\vec{W}^T\vec{X} = \sum_{d=1}^{D} w_d x_d$$

As funções de Kernel fornecem um espaço implícito de recurso. Isso nos permitirá usar Kernels para espaços dimensionais infinitos, além de dados não numéricos e simbólicos!

Aplicaremos as funções de Kernel em Python e R nas próximas aulas. A Matemática dessas funções é estudada no curso de Matemática Para Machine Learning aqui na DSA.