



Inteligencia Artificial

Clase SVM



Sistema de Evaluación

- Prueba 1: 33,3%
- Prueba 2: 33,3%
- Tareas y Controles: 33,3%



Temario

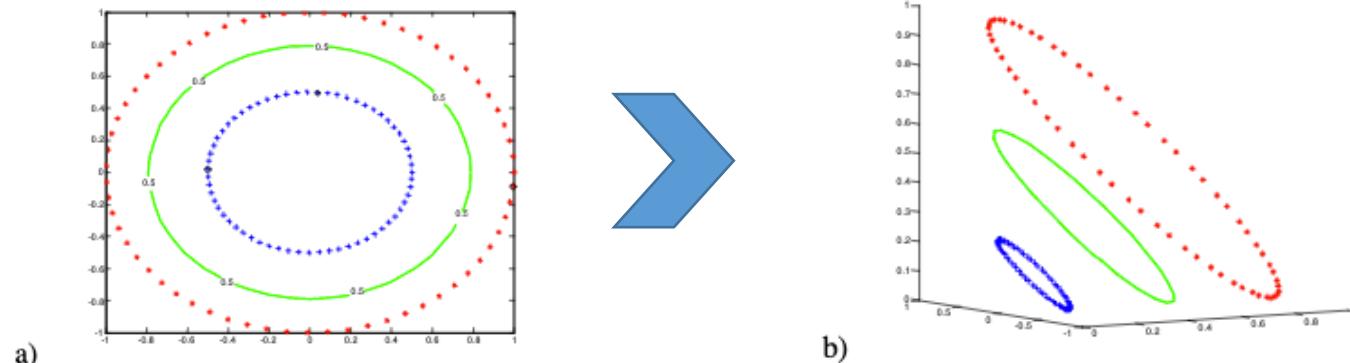
- Definición
- Funcionamiento
- Pasos y Propiedades
- Notación formal
- Funciones Kernel
- Ejemplos de influencia del Parámetro C

Definición

- Las máquinas de vectores soporte (SVM, support vector machine) son estructuras de aprendizaje basadas en la teoría de aprendizaje estadístico y surge inicialmente como un nuevo método para la resolución de problemas de clasificación.
- Este método fue desarrollado por Vapnik en 1964, para dar solución al problema de reconocimiento de patrones.

Funcionamiento (1)

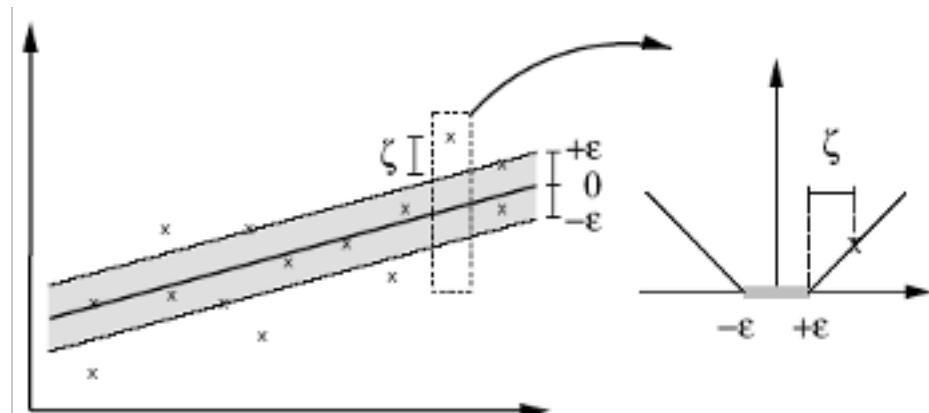
- Se basan principalmente en la transformación del espacio de entrada en otro de mayor dimensión, llevando el problema a una resolución mediante un hiperplano óptimo.



Aplicación de una función kernel modificando el espacio de dimensiones representado por la línea continua en verde, separando el conjunto de datos rojos y azules [Baudat, G. ,Anouar F.]

Funcionamiento (2)

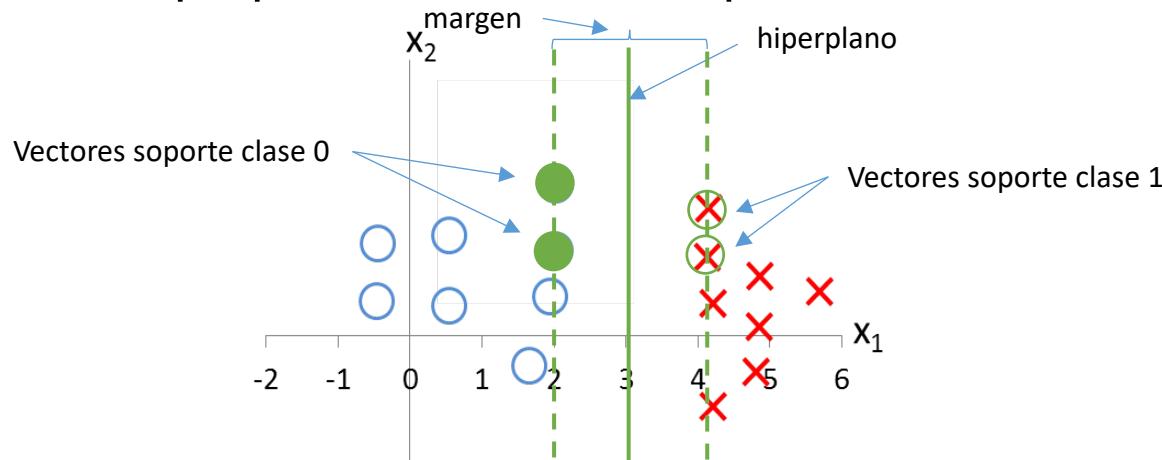
- Encontrar el espacio de margen máximo, el cual se define como la distancia del hiperplano a los valores positivos y negativos más próximos a él.



Ejemplo de un caso ruidoso para una SVM lineal [Smola, 1998].

Pasos y Propiedades (1)

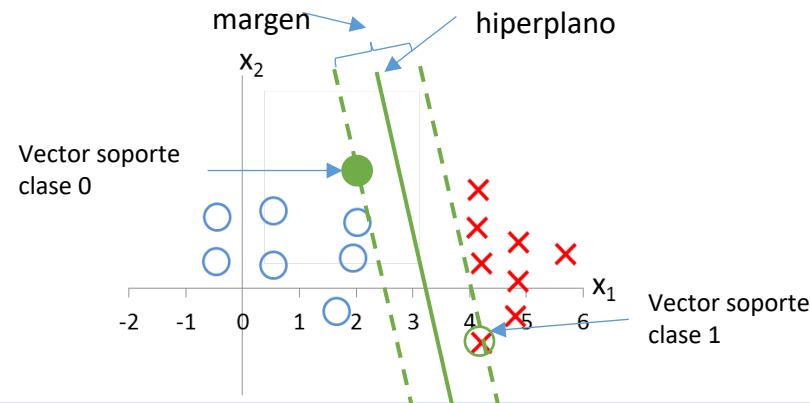
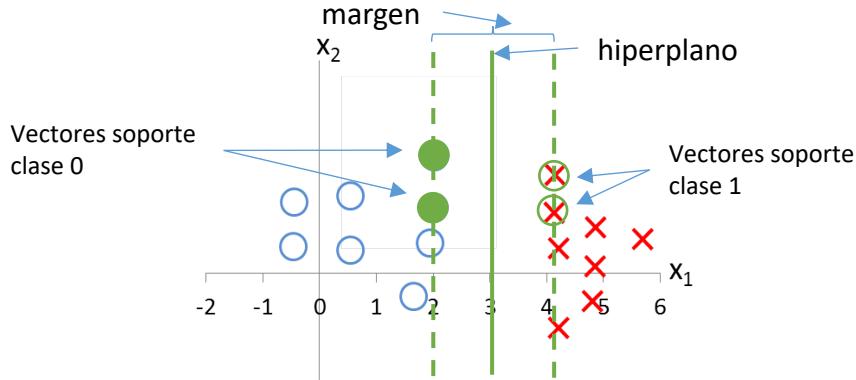
1. Se tiene un conjunto de elementos clasificables en dos grupos de 2^n formas diferentes.
2. Se induce un hiperplano lineal en el espacio de atributos.



Hiperplano separador de conjuntos linealmente separables [Basado en ejemplo de Andrew Ng]

Pasos y Propiedades (2)

3. Se transforma el espacio de atributos de entrada en un espacio de trabajo de dimensión mayor (ver figura [aquí](#)), mediante las denominadas funciones kernel.
4. Se elige el hiperplano de margen máximo entre todos los posibles hiperplanos separadores.





Pasos y Propiedades (3)

- Al elegir un hiperplano de margen máximo, se produce un sesgo inductivo de aprendizaje (learning bias), lo que se complementa a minimizar el error en el conjunto de aprendizaje, contribuyendo a minimizar el riesgo de sobreaprendizaje (overfitting).
- 5. El aprendizaje representa un problema de optimización con restricciones que puede ser resuelto usando una programación cuadrática, esto asegura una solución única debido a su convexidad.

Notación Formal (1)

- El hiperplano separador es representado por la ecuación

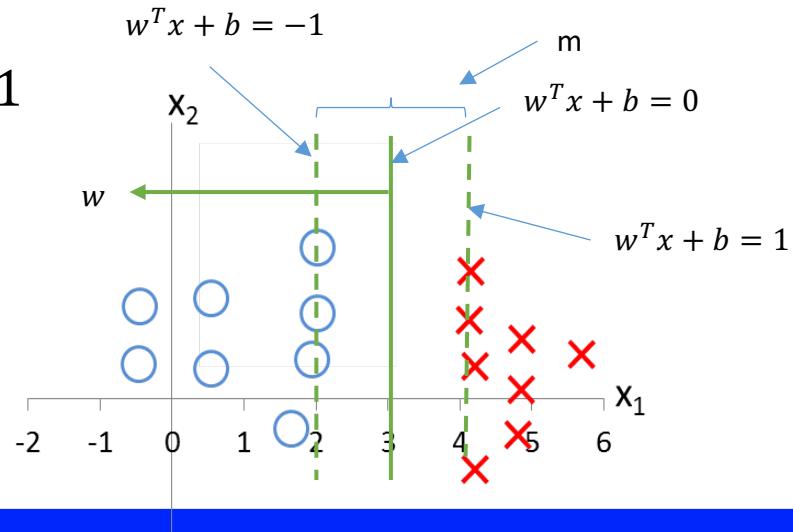
$$w^T x + b = 0$$

- Condiciones para el margen:

$$w^T x + b = 1$$

$$w^T x + b = -1$$

- m es la distancia entre los hiperplanos



Notación Formal (2)

- Considerando la norma del vector w , el problema de optimización es factible minimizando:

$$\frac{1}{2} \|w\|^2$$

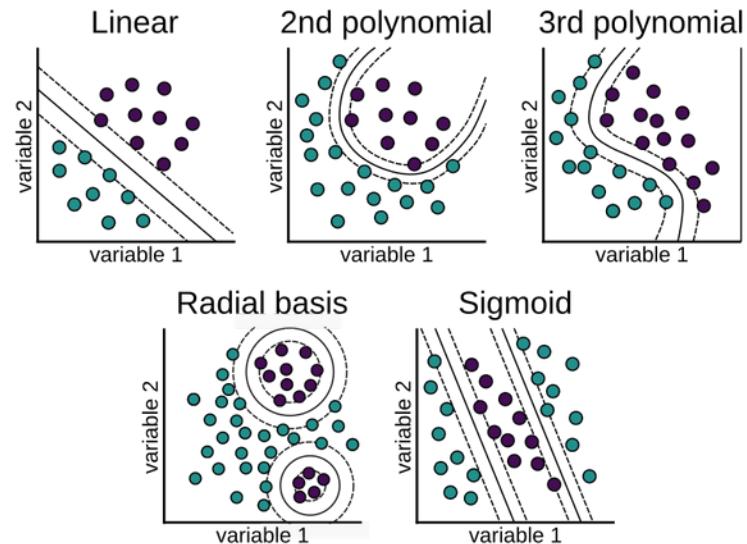
- En caso de que aún no sea posible ajustar una función óptima se incluyen variables de holgura (ξ_i), considerando la siguiente ecuación:

$$\frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i)$$

- Donde $C > 0$ determina el balance entre la influencia del error y la inclinación de la función, considerando la holgura del margen.

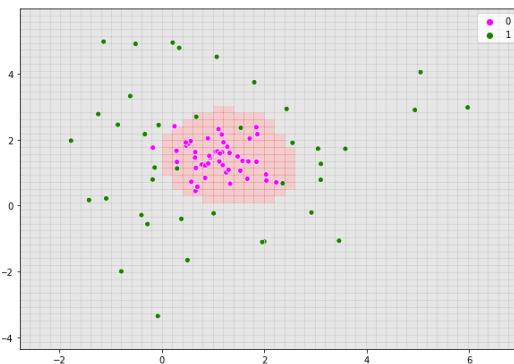
Funciones Kernel

Nombre	Ecuación
Lineal	$K(\vec{x}_i, \vec{x}_j) = \vec{x}_i \cdot \vec{x}_j$
Polinomial (grado m)	$K(\vec{x}_i, \vec{x}_j) = (\vec{x}_i \cdot \vec{x}_j + b)^m$
Radial Basis Function (RBF)	$K(\vec{x}_i, \vec{x}_j) = e^{\frac{-\ \vec{x}_i - \vec{x}_j\ ^2}{2\sigma^2}}$
Sigmoidal	$K(\vec{x}_i, \vec{x}_j) = \tanh(\kappa \vec{x}_i \cdot \vec{x}_j + \theta)$

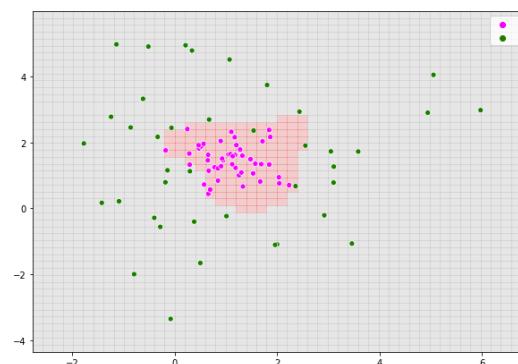


Gráficas de Funciones kernel
 [hefinioanrhys, Support Vector Machines with the mlr package]

Ejemplos de la influencia del parámetro C

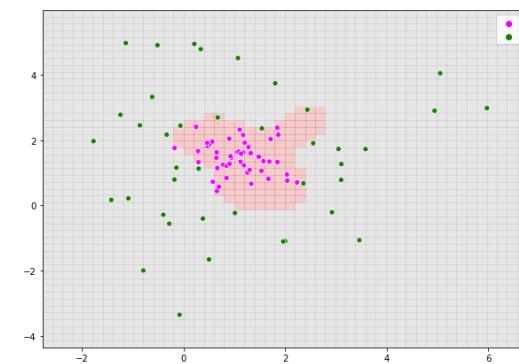


$C=1$

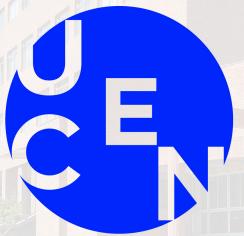


$C=100$

*Ejemplos de uso e influencia del parámetro C
[Ejemplos basado en el código de SVM InteractiveChaos]*



$C=1000000$



Universidad
Central

www.ucentral.cl



Comisión Nacional
de Acreditación
CNA-Chile

4

AÑOS ACREDITADA
GESTIÓN INSTITUCIONAL | DESDE DICIEMBRE 2017
DOCENCIA DE PREGRADO | HASTA DICIEMBRE 2021
VINCULACIÓN CON EL MEDIO