**Teoría de Support Vector Machines (SVM)**

SVM busca encontrar un hiperplano que separa las clases de la mejor manera posible

* 2d una línea
* 3d un plano
* En dimensiones mayores es un hiperplano

Lo importante es que no solo separa, sino que lo hace maximizando el margen, es decir, la distancia entre el hiperplano y los puntos más cercanos de cada clase; esos puntos cercanos que definen el margen se llaman vectores de soporte

**Conceptos:**

1. **Margen**: distancia entre el hiperplano y los vectores de soporte. Cuánto más grande, mejor generalización
2. **Vectores de soporte**: puntos que quedan en el borde del margen o más cercanos al hiperplano; son los que sostiene al hiperplano
3. **Clasificación lineal**: si los datos son linealmente separables, SVM encuentra la línea/hiperplano ideal
4. **Clasificación no lineal**: si los datos no se pueden separar con una línea, SVM usan la técnica del kernel

**Kernel trick**

El truco del kernel permite transformar los datos a un espacio de mayor dimensión para que se vuelvan separables; ejemplo:

* En 2d no se puede separar con una línea una forma circular
* SVM con kernel RBF (Radial Basic Functión) transforma los datos a una dimensión superior donde su es separable con un hiperplano

Tipos de kernel

* Lineal: buenos si los datos son linealmente separables
* Polinómico: crea fronteras curvas
* RBF (gaussiano): el más usado, maneja relaciones complejas

**Parámetros**:

1. C (regularización)
   1. Controla cuanto tolera errores
   2. Valor grande -> intenta clasificar todo bien (poco margen, riesgo de sobreajuste)
   3. Valor pequeño -> permite más errores, pero genera mayor margen (mejor generalización)
2. Y (gamma, en kernel RBF/polinómicos):
   1. Determina la influencia de un punto de entrenamiento
   2. Alto -> cada punto tiene influencia cercana (frontera muy compleja -> sobreajuste)
   3. Bajo -> influencia amplia (frontera más suave)

SVM para clasificación y regresión

* **Clasificación (SVC**): encuentra el hiperplano con máximo margen para separar clases
* **Regresión (SVR):** busca un hiperplano que prediga valores continuos con un margen de tolerancia (e)

**SVM en 2D**

Es una línea recta:

w1​x1​+w2​x2​+b=0; donde las x son las variables de entrada las w son los pesos y b el sesgo.

Las ecuaciones se igualan a 0 es por la forma en que se definen los hiperplanos. El 0 actúa como línea divisoria, valores + a un lado y – al otro lado. El hiperplano se define como el conjunto de puntos x que cumplen:

w⋅x+b=0

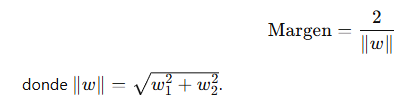
donde

* w es el vector de pesos
* x es el vector de características
* b es el sesgo bias

si en la ecuación un punto > 0 entonces está en un lado del hiperplano -> clase +1

se es < 0, entonces está en el lado opuesto -> clase -1

El margen:



**SVM en 3D**

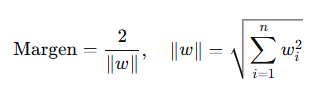
Es un plano

w1​x1​+w2​x2​+w3​x3​+b=0

**SVM n-dimensiones**

Hiperplanos

w1​x1​+w2​x2​+⋯+wn​xn​+b=0



El SVM busca maximizar el margen. Esto equivale a minimizar:

