

Introducción a Máquinas de Vectores de Soporte (SVM)

January 6, 2025

¿Qué es SVM?

- ▶ SVM es un modelo de aprendizaje supervisado usado para clasificación y regresión.
- ▶ Busca encontrar un hiperplano óptimo que separe los datos en distintas clases.
- ▶ Funciona tanto en problemas linealmente separables como no separables.

Definición formal de SVM

Dado un conjunto de datos de entrenamiento: $\{(x_i, y_i)\}_{i=1}^n$,
donde:

- ▶ $x_i \in \mathbb{R}^d$: Características.
- ▶ $y_i \in \{-1, 1\}$: Etiquetas.

El objetivo es encontrar un hiperplano:

$$w \cdot x + b = 0$$

- ▶ w : Vector normal al hiperplano.
- ▶ b : Sesgo (bias).

Margen máximo y separación óptima

El margen es la distancia entre el hiperplano y los puntos más cercanos.

- ▶ Maximizar el margen equivale a resolver el problema:

$$\min \frac{1}{2} \|w\|^2 \quad \text{sujeto a} \quad y_i(w \cdot x_i + b) \geq 1, \forall i. \quad (1)$$

- ▶ Esto se resuelve mediante programación cuadrática.

Soft margin: Clasificación no lineal

¿Qué pasa si los datos no son separables?

- ▶ Introducimos variables de holgura ξ_i para permitir violaciones al margen.
- ▶ Resolvemos el problema optimizado:

$$\min \frac{1}{2} \|w\|^2 + C \sum \xi_i \quad \text{sujeto a} \quad y_i(w \cdot x_i + b) \geq 1 - \xi_i. \quad (2)$$

Kernel Trick

SVM puede extenderse a problemas no lineales mediante el uso de kernels.

- ▶ Mapear los datos a un espacio de mayor dimensionalidad.
- ▶ Usar una función kernel $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$ para evitar cálculos explícitos.

Ejemplos de kernels:

- ▶ Lineal: $K(x_i, x_j) = x_i \cdot x_j$.
- ▶ RBF: $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$.
- ▶ Polinomial: $K(x_i, x_j) = (x_i \cdot x_j + c)^d$.

Ejemplo práctico: Clasificación con SVM

Datos separables linealmente.

```
from sklearn.datasets import make_blobs
from sklearn.svm import SVC
import matplotlib.pyplot as plt

# Crear datos
X, y = make_blobs(n_samples=100, centers=2, random_state=0)

# Entrenar SVM
clf = SVC(kernel='linear')
clf.fit(X, y)

# Visualizar
plt.scatter(X[:, 0], X[:, 1], c=y)
plt.show()
```

Curva ROC con SVM

```
from sklearn.metrics import roc_curve, auc

# Generar curva ROC
probas = clf.decision_function(X_test)
fpr, tpr, thresholds = roc_curve(y_test, probas)
roc_auc = auc(fpr, tpr)

# Graficar
plt.plot(fpr, tpr, label=f"AUC = {roc_auc:.2f}")
plt.xlabel("FPR")
plt.ylabel("TPR")
plt.legend()
plt.show()
```

Conclusiones

- ▶ SVM es una herramienta poderosa para clasificación y regresión.
- ▶ Los kernels permiten resolver problemas complejos de forma eficiente.
- ▶ Hiperparámetros como C y γ deben ajustarse cuidadosamente.
- ▶ Se aplica en tareas como clasificación de imágenes, texto y bioinformática.

Referencias

- ▶ Cortes, C., Vapnik, V. (1995). "Support-vector networks." Machine Learning.
- ▶ Documentación de scikit-learn:
<https://scikit-learn.org/>
- ▶ Bishop, C. M. (2006). "Pattern Recognition and Machine Learning."