



Instituto Politécnico Nacional
Escuela Superior de Cómputo



Integrantes:

Escalona Bastida Yael Tonatiuh
Salazar Carreón Jeshua Jonatán
Torres Abonce Luis Miguel

Grupo:

6CV1

Materia:

PROC. DE LENG. NAT|MACHINE
LEARNING

Profesor:

Vanessa Alejandra Camacho Vázquez

Trabajo:

Clasificador por el método de Máquinas de Vectores de Soporte
en Python

Maquinas de Soporte Vectorial (SVM)

Las maquinas de Soporte Vectorial (SVM) se posicionan como un algoritmo de aprendizaje automático supervisado fundamental, reconocido por su robustez y sólida base matemática. Su enfoque se centra en la construcción de un hiperplano de separación óptimo que maximiza el margen entre las clases de datos. Este método no solo destaca por su eficacia, sino, que también se extiende a problemas de regresión.

Fundamento Matemático

El corazón del las SVM reside en la búsqueda del hiperplano que mejor separa las clases de datos en espacio multidimensional. Este hiperplano se define como una línea recta (en 2 dimensiones) o un plano en dimensiones superiores, que divida los puntos de datos en 2 categorías. La clave radica en encontrar el hiperplano que maximiza el margen entre las 2 clases, es decir, la distancia entre el hiperplano y los puntos de datos más cercanos a él (conocido como vectores de soporte).

La aplicación de los multiplicadores de Lagrange conduce a una formulación de una función objetivo que se minimiza para encontrar el hiperplano óptimo. Esta función incorpora tanto la maximización del margen como el cumplimiento de las restricciones, logrando una manera de equilibrio entre ambas.

Multiplicadores de Lagrange

Para encontrar el hiperplano de separación óptimo, las SVM emplean multiplicadores de Lagrange, una técnica matemática que permite resolver problemas de optimización con restricciones. En este caso, las restricciones se establecen para garantizar que los vectores de soporte se encuentren

exactamente sobre el margen máximo en el plano.

Al maximizar este margen, las SVM logran una mayor generalización del Modelo reduciendo el riesgo de sobre ajuste (over fitting). En otras palabras, el algoritmo aprende a separar las clases de manera efectiva, incluso con datos nuevos no vistos durante el entrenamiento.

Función Kernel: Superando la Barrera de la separabilidad lineal.

Si bien las SVM sobresalen en la separación lineal de las clases, algunas situaciones presentan datos que no pueden ser separados linealmente. Para abordar este desafío, las SVM introducen el concepto de Funciones Kernel. Estas funciones transforman los datos de entrada a un espacio de mayor dimensiones, donde la separación lineal se vuelve posible.

Existen diferentes tipos de Funciones Kernel, cada una con sus propias características y ventajas. La elección a menudo depende de la naturaleza de los datos y del problema específico a resolver. Entre las funciones Kernel más comunes se encuentra la función lineal, la función polinomial y la función de base radial (RBF).

Ventajas Clave de las Máquinas de Soporte Vectorial.

- Robustez: Las SVM son menos sensibles a valores o ruido en los datos, lo que hace más confiable en comparación con otros métodos.

- **Alta Dimensionalidad:** Las SVM pueden manejar problemas con alto número de características sin sufrir problemas de dimensionalidad.
- **Interpretabilidad:** El hiperplano de separación obtenido por las SVM proporciona una representación intuitiva de la frontera de decisión entre clases.
- **Versatilidad:** Las SVM se pueden extender a problemas de regresión mediante la técnica de regresión epsilon SVR.

Aplicaciones de las Maquinas de Soporte Vectorial.

- **Clasificación de imágenes:** Las SVM se utilizan ampliamente para clasificar imágenes en diferentes categorías, como reconocimiento facial, detección de objetos y análisis de imágenes médicas.
- **Análisis de texto:** Las SVM se emplean para tareas de procesamiento del lenguaje natural, como la clasificación de documentos, la extracción de información y detección de spam.
- **Bioinformática:** Las SVM se aplican en bioinformática para analizar datos genómicos y proteómicos, identificar patrones moleculares y predecir estructuras de proteínas.
- **Finanzas:** Las SVM en el sector financiero para predecir tendencias del mercado, detectar fraudes y evaluar riesgos creditivos.

Consideraciones Adicionales y Áreas de Investigación.

- Selección de la función Kernel: La elección de la función de Kernel adecuada es crucial para el rendimiento de las SVM. Existen diferentes métodos para seleccionar la función Kernel óptima, como la validación cruzada y selección heurística.
- Optimización de parámetros: Las SVM tienen parámetros que pueden ajustarse para mejorar su rendimiento. Estos parámetros incluyen el parámetro de regularización C y el parámetro de la función Kernel.
- Interpretación de modelos: Si bien las SVM son interpretables en cierto grado, la interpretación de los modelos complejos puede ser un desafío. Se han desarrollado técnicas.

Multiplicadores de Lagrange.

Las SVM emplean multiplicadores de Lagrange para resolver problemas con restricciones que surgen al maximizar el margen entre las clases de datos. Las restricciones, en este caso, aseguran que los vectores de soporte se encuentran exactamente sobre el margen máximo.

La técnica de los multiplicadores de Lagrange introduce variables duales asociadas a cada restricción, transformando el problema de optimización restringida en un problema de optimización sin restricciones. Al minimizar la función objetivo que incorpora estas variables duales, se obtienen los valores óptimos de los parámetros del hiperplano y los multiplicadores de Lagrange.

Mapeando datos a Espacio de Mayor Dimensión.

Las SVM se enfrentan a un desafío cuando los datos no son linealmente separables en el espacio original. Para superar este obstáculo, se introducen las funciones Kernel, las cuales transforman los datos de entrada a un espacio mayor de dimensión donde la separación lineal se vuelve posible.

La función Kernel opera como una función de similitud, calculando la similitud entre 2 puntos de datos en el espacio original. Al aplicar la función Kernel, se obtiene un nuevo espacio de características donde los datos se presentan de manera que el hiperplano de separación lineal pueda encontrarlos definitivamente.

Existen diferentes tipos de funciones Kernel, cada uno con sus propias características y ventajas. Algunas son:

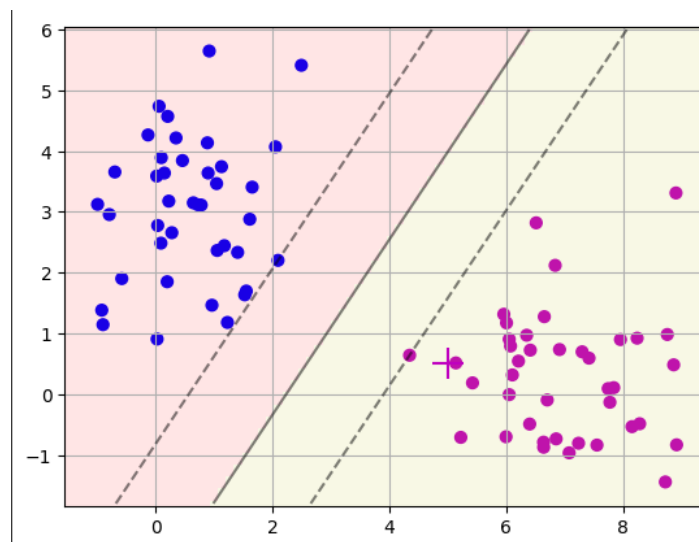
- Lineal: calcula la similitud lineal entre dos puntos de datos.
- Polinomial: Eleva la similitud lineal a una potencia polinomial antes de calcularla.
- Base Radial: Esta función calcula la similitud como una función exponencial de la distancia entre los puntos dados.

Los multiplicadores de Lagrange y las funciones Kernel son componentes esenciales de las SVM que permiten las limitaciones de la separabilidad lineal y optimizar el modelo.

Conjunto de datos	No. de Muestras	No. de características	Características	Uso
Make_blobs	Variable	Variable	-Muestras generadas de manera aleatoria en clusters definidos. -Puede especificarse el número de clusters, muestras por cluster y desviación estándar de cada cluster.	Clasificación simple para demostrar algoritmos de aprendizaje automático.
Make_circles	Variable	2	-Puntos generados en dos círculos concéntricos. -Puede especificarse el número de muestras, ruido y factor de escala	Útil para problemas de clasificación no lineales, donde las clases están distribuidas en forma de círculos concéntricos
Load_digits	1797	64	-Imágenes de dígitos manuscritos, representadas como matrices de píxeles. - Cada imagen tiene 8x8 píxeles, totalizando 64 características.	Clasificación de dígitos manuscritos. Comúnmente utilizado para demostrar algoritmos de clasificación en aprendizaje automático.

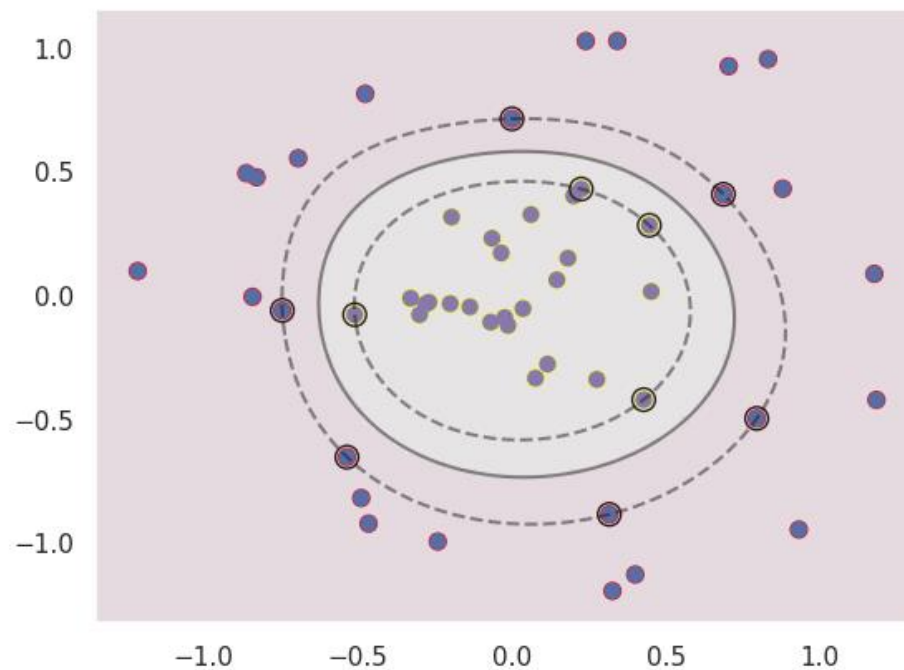
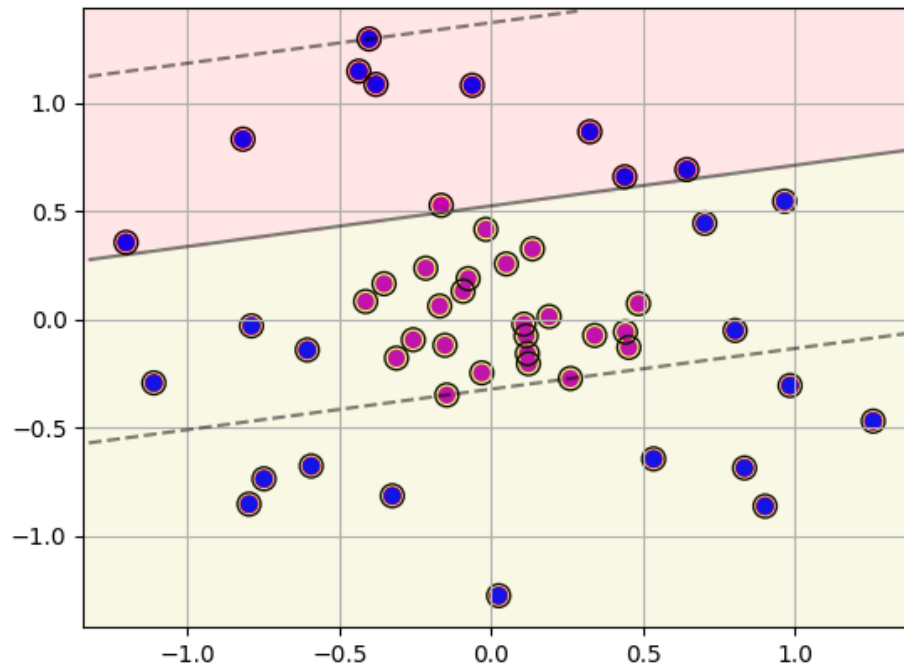
Código

Implementación - Problema de clases linealmente separables En esta parte, deberán realizar los siguientes cambios: usar 80 muestras en total y cambiar los colores de su gráficarlos, así como el de los puntos los puntos los puntos los puntos para las 2 clases diferentes.

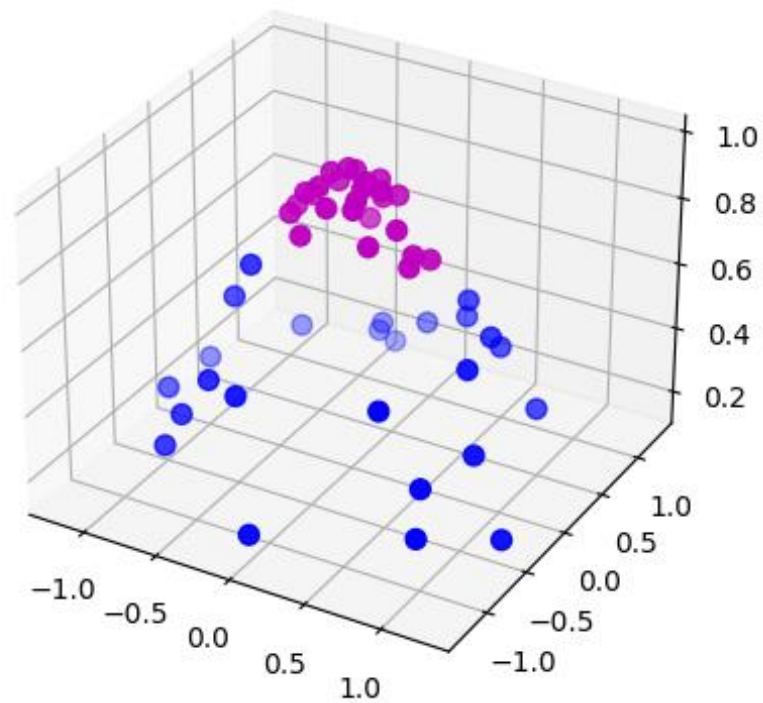


Implementación - Problema de clases linealmente no separables usando el truco Kernel

En esta parte, deberán realizar los siguientes cambios: usar 50 muestras en total y cambiar los colores de su gráfica, así como el de los puntos para las 2 clases diferentes. Incluso, deberán adjuntar la gráfica del truco Kernel.



Truco Kernel



Implementación – Dataset digits

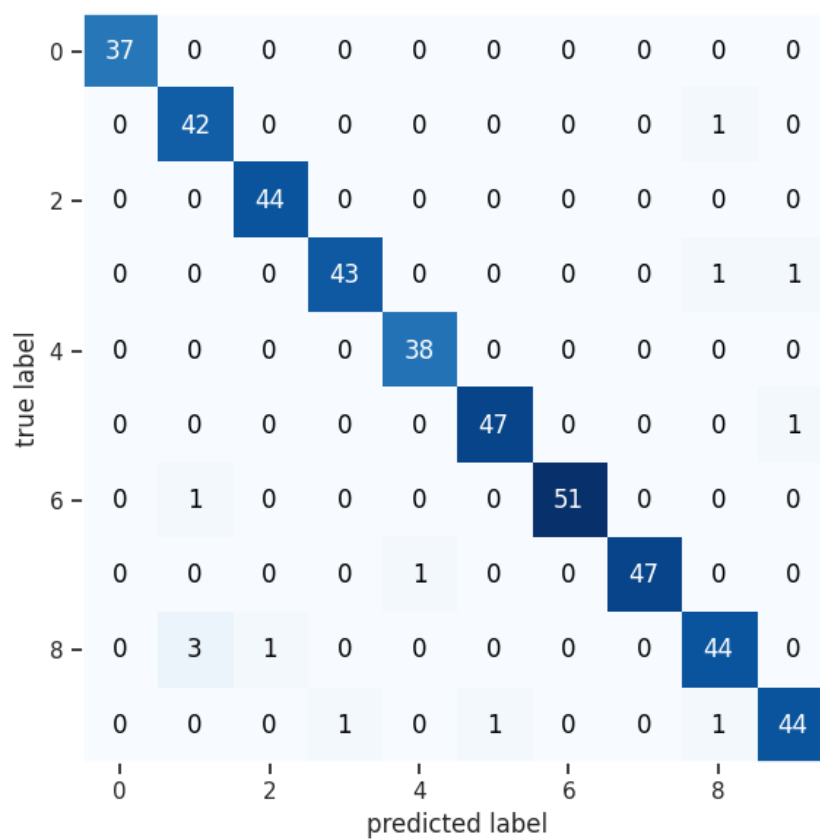
```
%matplotlib inline
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns; sns.set()
import pandas as pd
from sklearn.datasets import load_digits

# cargamos el dataset
digits = load_digits()
digits.data.shape

(1797, 64)
```



Predicción



Maquinas de vectores de soporte

- Ventajas.

- **Eficacia en dimensiones altas:** SVM funciona muy bien en espacios de alta dimensión, como en clasificación de texto y problemas de bioinformática donde el número de características puede ser muy grande en comparación con el número de muestras.
- **Efectividad** en casos donde el número de dimensiones es mayor que el número de muestras. Debido a que utiliza un subconjunto de puntos de entrenamiento en la función de decisión.
- **Versatilidad:** A través del uso de diferentes funciones Kernel, SVM es capaz de adaptarse a diferentes tipos de datos y dominios, lo que permite modelar relaciones no lineales complejas.
- **Generalización:** La regularización en SVM implica que tienen un buen rendimiento de generalización, indicando que tiene un menor riesgo de sobreajuste.
- **Solución de optimización convexa:** El problema de optimización en SVM garantiza la obtención de una única solución global, evitando mínimos globales que podrían no ser óptimos.

- Desventajas.

- Selección del Kernel y ajuste de parámetros: La elección del Kernel y la configuración de sus parámetros puede ser compleja y requiere métodos como la búsqueda en cuadrícula con validación cruzada.
- No proporciona estimaciones de probabilidad directamente: Los SVM no ofrecen estimaciones de probabilidad de manera natural.
- Eficiencia computacional: Aunque los SVM son relativamente eficientes, no escalan bien con el número de muestras.
- Sensibilidad a los datos no escalados: Requiere que las características se normalicen para que el modelo funcione correctamente, ya que la escala puede influir significativamente en la interpretación de los márgenes y los vectores de soporte.
- Resultados difíciles de interpretar: A diferencia de algunos modelos como los árboles de decisión, los modelos SVM son más difíciles de interpretar, lo cual puede ser un problema cuando se necesita entender las razones detrás de las decisiones del modelo.

Conclusiones

Conclusiones

S.T.E.B.

Los SVM son una herramienta poderosa y versátil dentro del campo del aprendizaje supervisado, especialmente adecuada para problemas de clasificación y regresión en espacios de alta dimensionalidad. Una de las principales fortalezas de SVM es su capacidad para manejar eficazmente grandes dimensiones de características, lo que las hace particularmente útiles en dominios como el análisis de texto y la bioinformática. Además, gracias a su enfoque basado en maximizar el margen entre clases, SVM tiende a mostrar una excelente capacidad de generalización, minimizando el sobreajuste a pesar de la complejidad del modelo.

Aunque también el modelo no está exento de desafíos, como lo pueden ser: la elección del kernel adecuado y la configuración óptima de sus parámetros requieren una cuidadosa experimentación y validación cruzada, lo que puede resultar en un proceso computacionalmente intensivo. Además, la falta de escalabilidad con grandes volúmenes de datos y la dificultad para interpretar los modelos resultantes pueden limitar su aplicabilidad en ciertos contextos donde la transparencia y la eficiencia son críticas.