

# Inteligencia Artificial

Juan Pablo Restrepo Uribe

Ing. Biomedico - MSc. Automatización y Control Industrial

[jprestrepo@correo.iue.edu.co](mailto:jprestrepo@correo.iue.edu.co)

2024

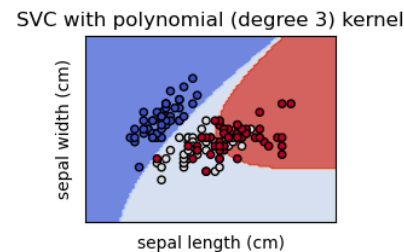
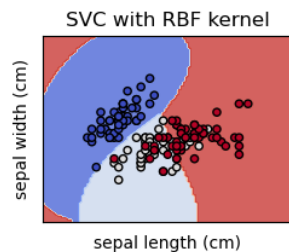
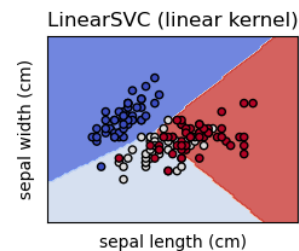
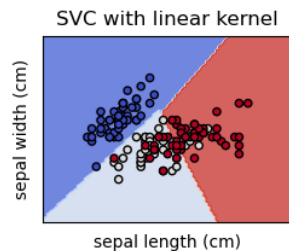
Institución Universitaria de Envigado

## Exposición final

- **Semana del 20/05/2024**
- Los estudiantes contarán con 15 minutos para exponer sus resultados y antecedentes del trabajo realizado
- Los antecedentes hacen referencia a artículos de índole científico donde se aborden temáticas similares (determinar comportamiento financiero en el tiempo, clasificación, etc)
- Se deben presentar diapositivas.
- Este trabajo incluye dos notas: una relacionada con el trabajo práctico y una segunda que corresponde a la búsqueda bibliográfica
- Para ello realice diapositivas.
- No quiero ver código en las diapositivas, únicamente resultados

## SVM

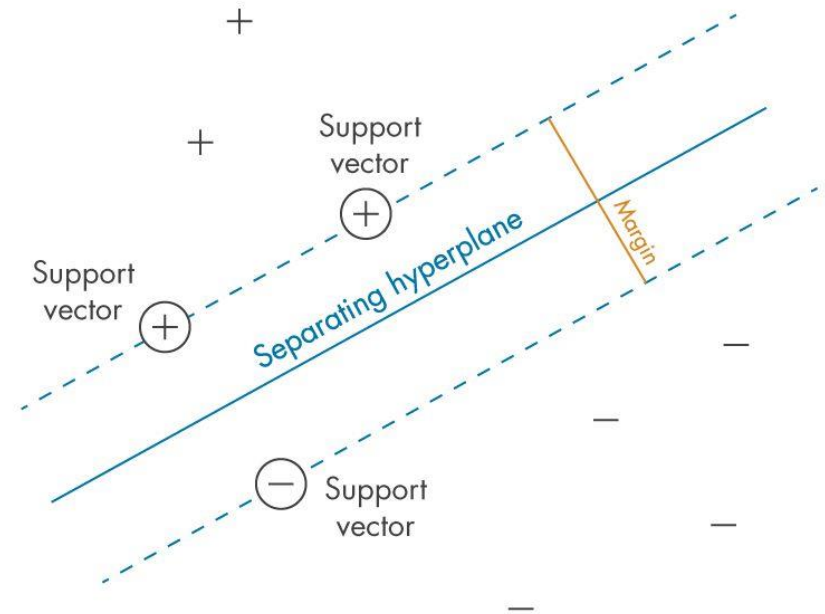
Es un algoritmo de aprendizaje supervisado que se utiliza en muchos problemas de clasificación y regresión, incluidas aplicaciones médicas de procesamiento de señales, procesamiento del lenguaje natural y reconocimiento de imágenes y voz.



## SVM

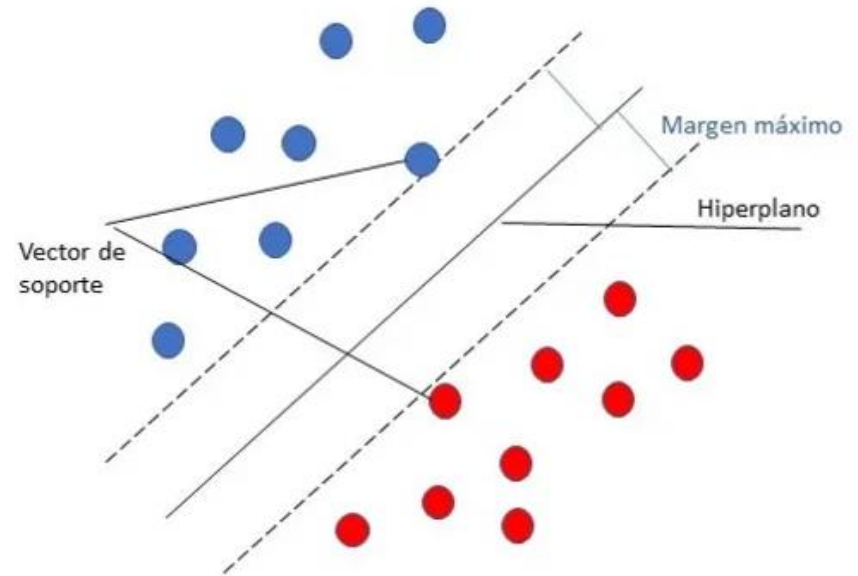
El objetivo del algoritmo SVM es encontrar un hiperplano que separe de la mejor forma posible dos clases diferentes de puntos de datos.

El algoritmo solo puede encontrar este hiperplano en problemas que permiten separación lineal; en la mayoría de los problemas prácticos, el algoritmo maximiza el margen flexible permitiendo un pequeño número de clasificaciones erróneas.



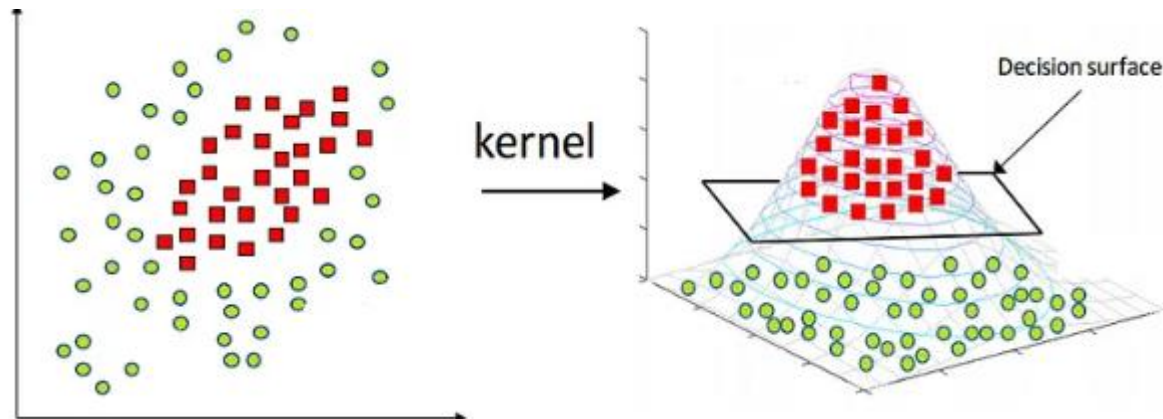
## SVM

Los vectores de soporte hacen referencia a un subconjunto de las observaciones de entrenamiento que identifican la ubicación del hiperplano de separación. El algoritmo SVM estándar está formulado para problemas de clasificación binaria; los problemas multiclase normalmente se reducen a una serie de problemas binarios.



## SVM - Kernel

- Mapeo a un espacio de características de mayor dimensión:
  - Los kernels permiten mapear los datos originales a un espacio de características de mayor dimensión donde pueden ser linealmente separables, incluso si no lo son en el espacio original.
  - Este proceso de mapeo se realiza de manera implícita, lo que significa que no es necesario calcular explícitamente las coordenadas de los puntos en el espacio de características de mayor dimensión.



# SVM

Tipo de SVM	Kernel de Mercer	Descripción
Función de base radial (RBF) o gaussiana	$K(x_1, x_2) = \exp\left(-\frac{\ x_1 - x_2\ ^2}{2\sigma^2}\right)$	Aprendizaje de una clase. $\sigma$ representa la anchura del kernel.
Lineal	$K(x_1, x_2) = x_1^T x_2$	Aprendizaje de dos clases.
Polinómica	$K(x_1, x_2) = (x_1^T x_2 + 1)^\rho$	$\rho$ representa el orden del polinomio.
Sigmoide	$K(x_1, x_2) = \tanh(\beta_0 x_1^T x_2 + \beta_1)$	Representa un kernel de Mercer solo para determinados valores $\beta_0$ y $\beta_1$ .

## SVM

- Support vector machines son muy populares y logran un buen rendimiento en muchas tareas de clasificación y regresión.
- Aunque los algoritmos SVM están formulados para la clasificación binaria, los algoritmos SVM multiclase se construyen combinando varios clasificadores binarios.
- Los kernels hacen que los SVM sean más flexibles y capaces de gestionar problemas no lineales.
- Para construir la superficie de decisión, solo se requieren los vectores de soporte seleccionados a partir de los datos de entrenamiento. Una vez terminado el entrenamiento, el resto de los datos de entrenamiento es irrelevante, produciendo una representación compacta del modelo que es adecuada para generar código de forma automatizada.

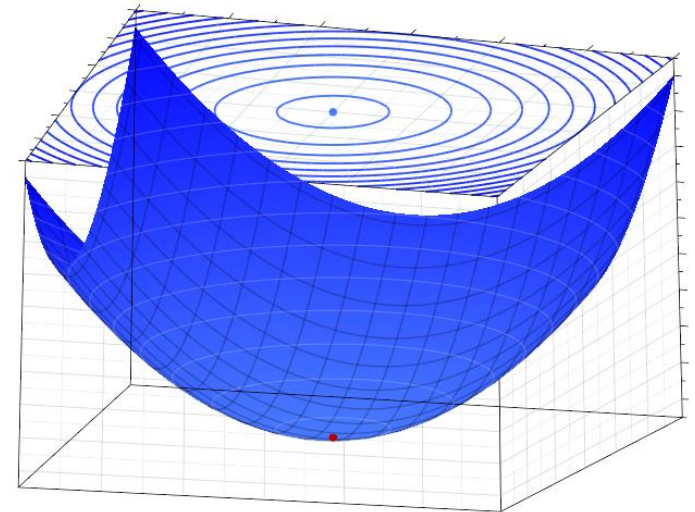


## SVM

El método de clasificación de vectores de soporte se puede extender para resolver problemas de regresión. Este método se llama Regresión de vectores de soporte. El modelo producido por la clasificación de vectores de soporte (como se describió anteriormente) depende solo de un subconjunto de los datos de entrenamiento, porque la función de costo para construir el modelo no se preocupa por los puntos de entrenamiento que se encuentran más allá del margen. Análogamente, el modelo producido por Support Vector Regression depende solo de un subconjunto de los datos de entrenamiento, porque la función de costo ignora las muestras cuya predicción está cerca de su objetivo.

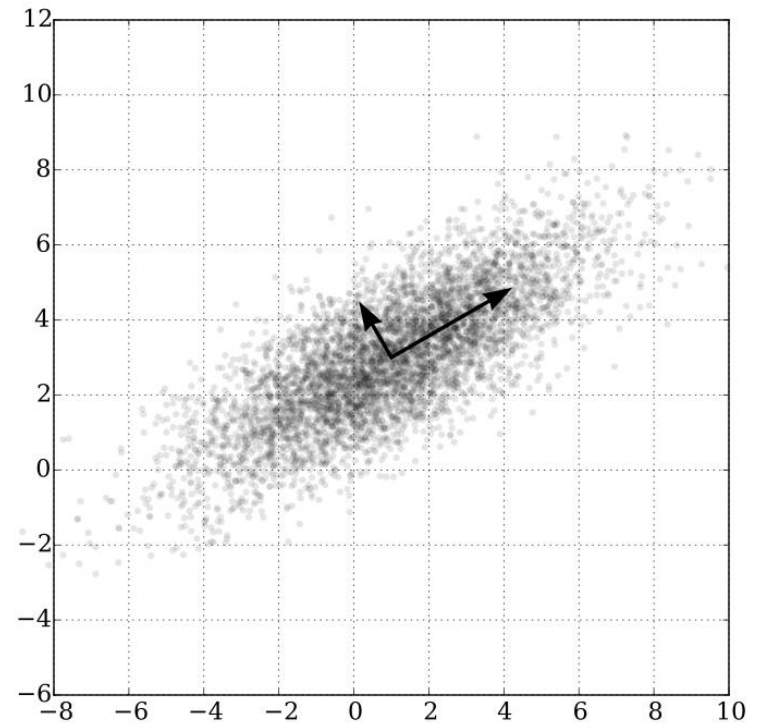
## SVM

- Eficacia en espacios de alta dimensión
- Buena generalización
- Manejo eficaz de espacios no lineales
- Robustez frente a la presencia de ruido
- Flexibilidad en la elección de la función de kernel
- Optimización convexa



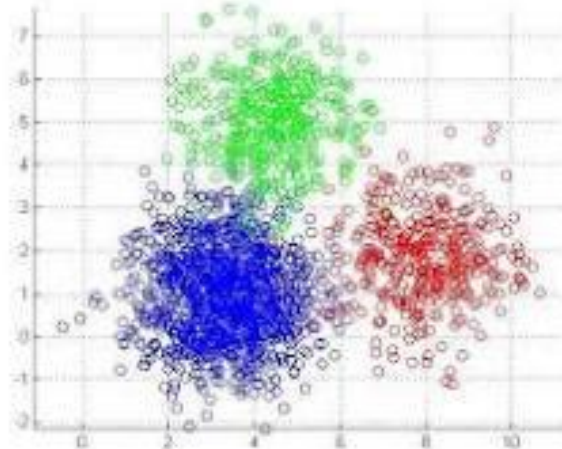
## PCA

es una técnica utilizada para describir un conjunto de datos en términos de nuevas variables no correlacionadas. Los componentes se ordenan por la cantidad de varianza original que describen, por lo que la técnica es útil para reducir la dimensionalidad de un conjunto de datos.



# Clustering

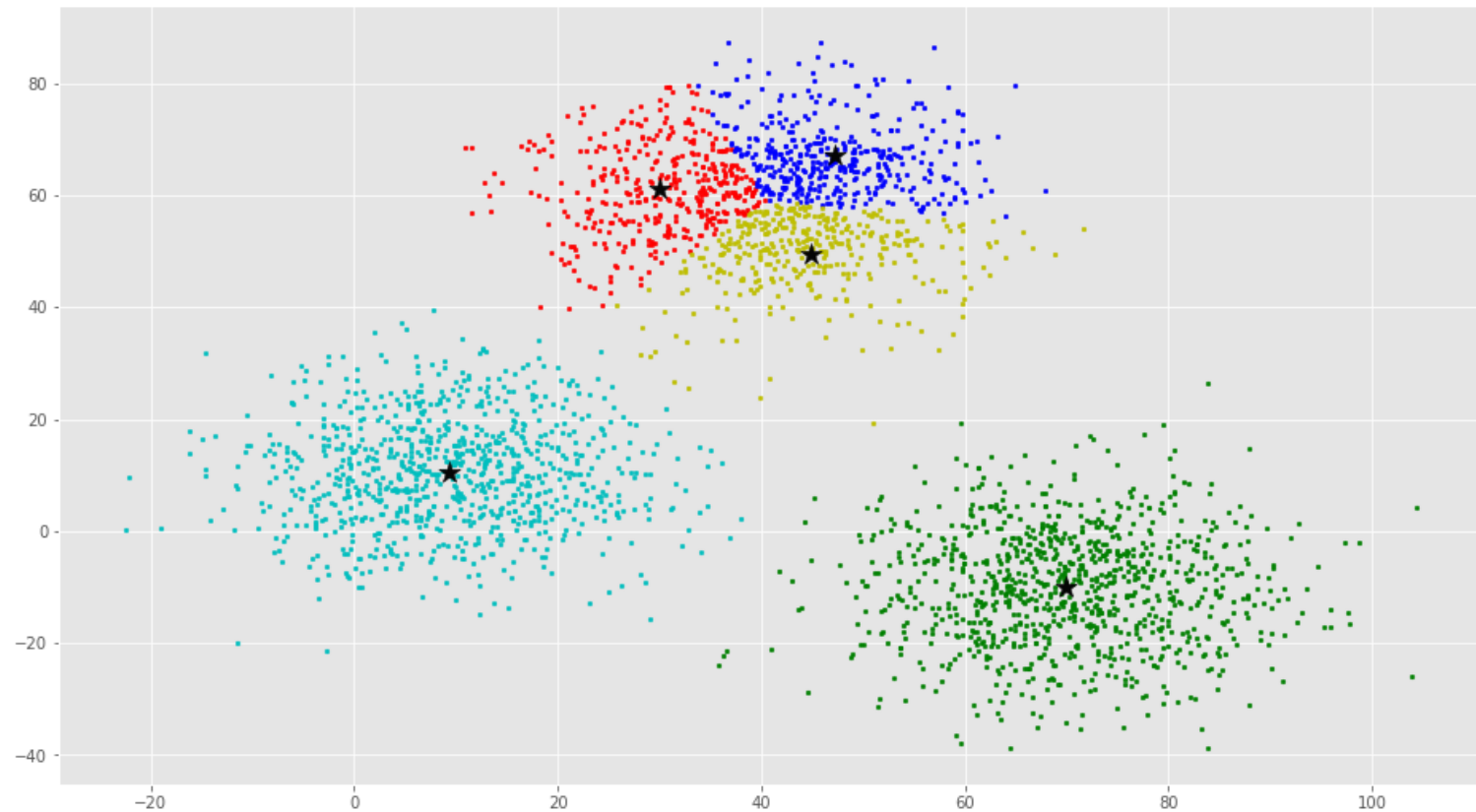
Este proceso desarrolla una acción fundamental que le permite a los algoritmos de aprendizaje automatizado entrenar y conocer de forma adecuada los datos con los que desarrollan sus actividades.



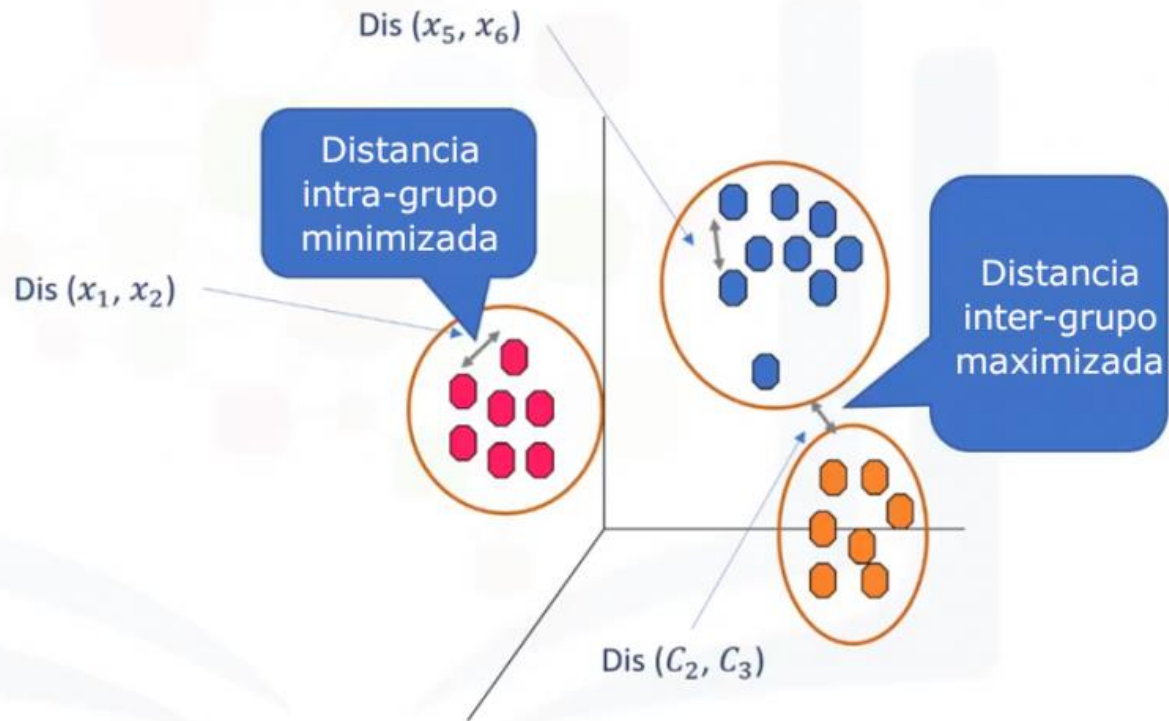
## Algoritmo de k-medias

Este proceso desarrolla una acción fundamental que le permite a los algoritmos de aprendizaje automatizado entrenar y conocer de forma adecuada los datos con los que desarrollan sus actividades. Es tal vez el método clásico para aplicar y entender el proceso de agrupamiento. Se establece un número de grupos previamente determinado. En este caso el algoritmo buscará los mejores centroides para realizar el agrupamiento, de manera que los miembros de cada grupo estén lo más cerca posible de sus centroides. El algoritmo funciona de forma iterativa, actualizando el centro de los clústeres de manera de ir reduciendo las distancias entre los miembros de cada cluster y su centro.

# Algoritmo de k-medias

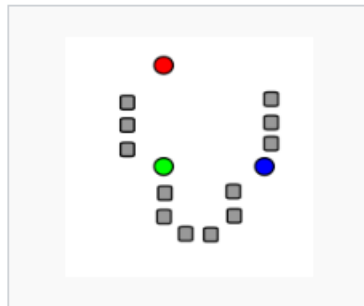


# Algoritmo de k-medias

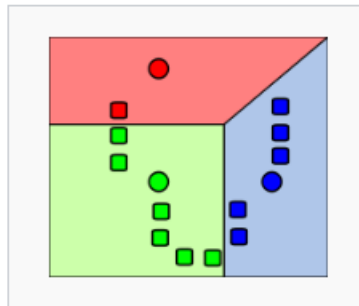


# Algoritmo de k-medias

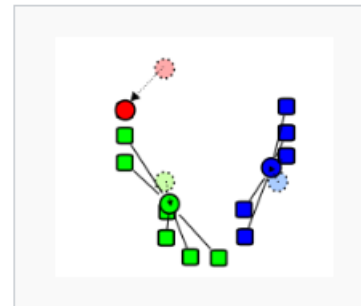
## Demostración del algoritmos estándar



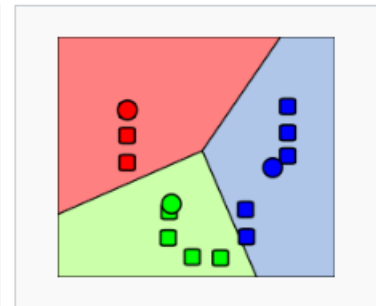
1)  $k$  centroides iniciales (en este caso  $k=3$ ) son generados aleatoriamente dentro de un conjunto de datos (mostrados en color).



2)  $k$  grupos son generados asociándole el punto con la media más cercana. La partición aquí representa el [diagrama de Voronoi](#) generado por los centroides.



3) EL [centroide](#) de cada uno de los  $k$  grupos se recalcula.



4) Pasos 2 y 3 se repiten hasta que se logre la convergencia.



## Clustering jerárquico

Uno de los métodos más utilizados, debido a la visualización práctica en forma de dendrograma que se obtiene. El clustering jerárquico puede realizarse tanto en forma divisiva o aglomerativa. Este método nos permite analizar alternativas para distintos números de grupos. Para entender un poco acerca de su funcionamiento si bien su procedimiento es bastante simple, fijémonos por ejemplo en el caso aglomerativo:

- Se parte de tantos grupos como individuos haya.
- De acuerdo a la medida de similitud previamente seleccionada, unimos los dos grupos con mayor similitud para formar uno solo.
- Continuamos de la misma forma hasta formar un solo grupo.