

CIENCIA DE DATOS

INFORME GRAFICOS GENERADOS

EDISON ANDRES FORERO RIAÑO

DAMIAN DANILO NARANJO PERILLA

ESCUELA TECNOLÓGICA INSTITUTO TÉCNICO CENTRAL

FACULTAD DE INGENIERÍA

TECNOLOGÍA EN DESARROLLO DE SOFTWARE

BOGOTÁ D.C

2024

## **Informe sobre la Visualización de Datos del Dataset de Iris**

En este informe, se analizan los gráficos generados a partir del dataset de Iris, que incluye las dimensiones de los sépalos y pétalos de tres especies de flores: Iris-setosa, Iris-versicolor e Iris-virginica. Se realizaron visualizaciones utilizando gráficos de dispersión, gráficos de pares y un mapa de calor, lo que permitió observar patrones y relaciones significativas entre las variables.

### **Gráficos de Dispersión y Gráficos de Pares**

Los gráficos de dispersión muestran una clara separación entre las especies de Iris en función de las dimensiones del sépalo y el pétalo. En particular, Iris-setosa se distingue notablemente de las otras dos especies, ya que sus dimensiones de sépalo y pétalo son significativamente más pequeñas. Esta característica hace que Iris-setosa sea fácilmente identificable en comparación con Iris-versicolor e Iris-virginica, que presentan características más similares.

Al observar los gráficos de pares, se pueden identificar relaciones interesantes. Por ejemplo, se observa que hay una fuerte correlación positiva entre la longitud y el ancho del pétalo. A medida que aumenta la longitud del pétalo, también lo hace el ancho del pétalo, lo que sugiere que estas dos dimensiones están interrelacionadas. Esta relación es particularmente evidente en las especies Iris-versicolor e Iris-virginica, donde los puntos en el gráfico se agrupan más estrechamente, indicando que estas especies comparten características similares en sus dimensiones de pétalo.

### **Mapa de Calor**

El mapa de calor revela la matriz de correlación entre las variables numéricas del dataset. Las correlaciones más altas se observan entre la longitud del pétalo y el ancho del pétalo (correlación de aproximadamente 0.96) y entre la longitud del sépalo y el ancho del sépalo (correlación de aproximadamente 0.87). Estas correlaciones indican que las dimensiones del pétalo y del sépalo están estrechamente relacionadas y son fundamentales para la clasificación de las especies.

Además, el mapa de calor muestra que las dimensiones del pétalo son más discriminativas que las del sépalo al clasificar las especies. Esto sugiere que, en un contexto de clasificación, utilizar las medidas del pétalo podría ser más efectivo para identificar las especies de Iris.

## **Conclusión**

En conclusión, las visualizaciones generadas proporcionan información valiosa sobre las relaciones entre las dimensiones de los sépalos y pétalos de las especies de Iris. Los gráficos de dispersión y pares evidencian patrones claros que facilitan la identificación de las especies, mientras que el mapa de calor destaca las correlaciones significativas entre las variables. Estas visualizaciones son herramientas útiles en el proceso de clasificación, permitiendo una mejor comprensión de las características distintivas de cada especie.