

PRÁCTICA 4

AGRUPAMIENTO

APRENDIZAJE AUTOMÁTICO



Profesor: José Anastacio Hernández Saldaña

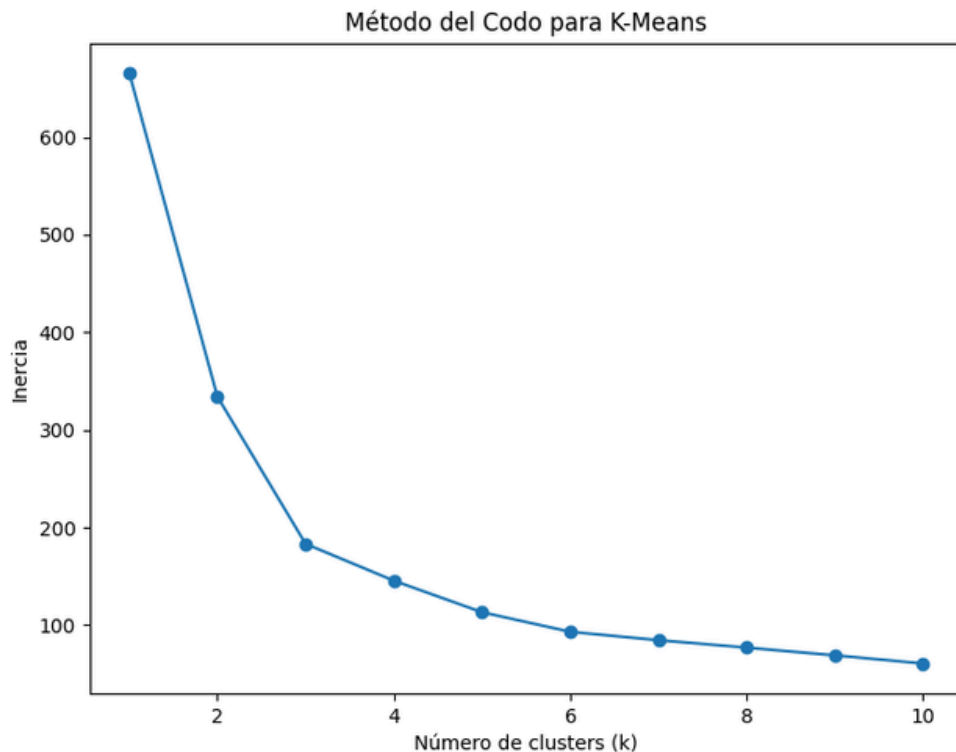
Alumna: Judith Carolina Lucio Garza

Matrícula: 1877415

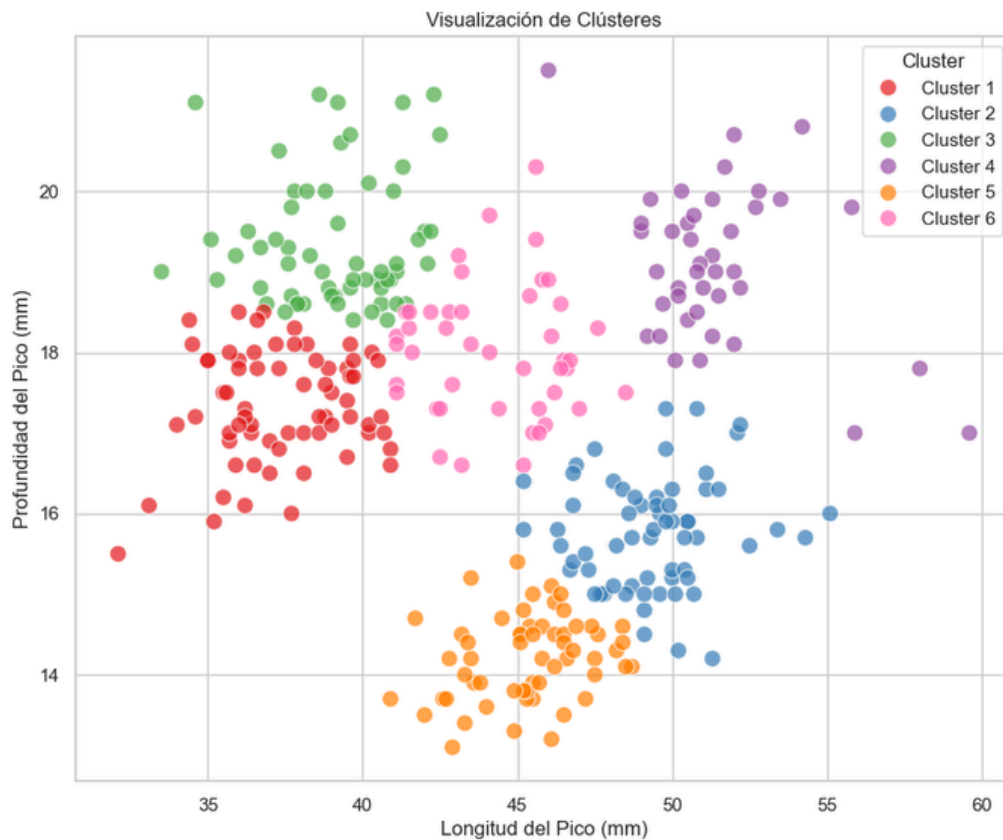
PENGUINS

Continuando con la base de datos "Penguins" trabajada en la práctica anterior, se busca realizar ahora mediante aprendizaje no supervisado un agrupamiento de los especímenes de pingüinos a través de K-Means. Para ello, primeramente se grafica el método de Codo para determinar la cantidad adecuada de clusters.

Como se muestra en la gráfica, a partir del cluster 6 se aprecia una línea recta, esto es coherente con lo registrado en la base de datos, ya que no solo se toma en cuenta la especie del pingüino, sino también su sexo.



En la gráfica de la página siguiente se aprecian los 6 clusters en los que el modelo ha agrupado a los pingüinos a partir de las longitudes y profundidad de sus picos



Scores del Modelo

Homogeneidad: 67%

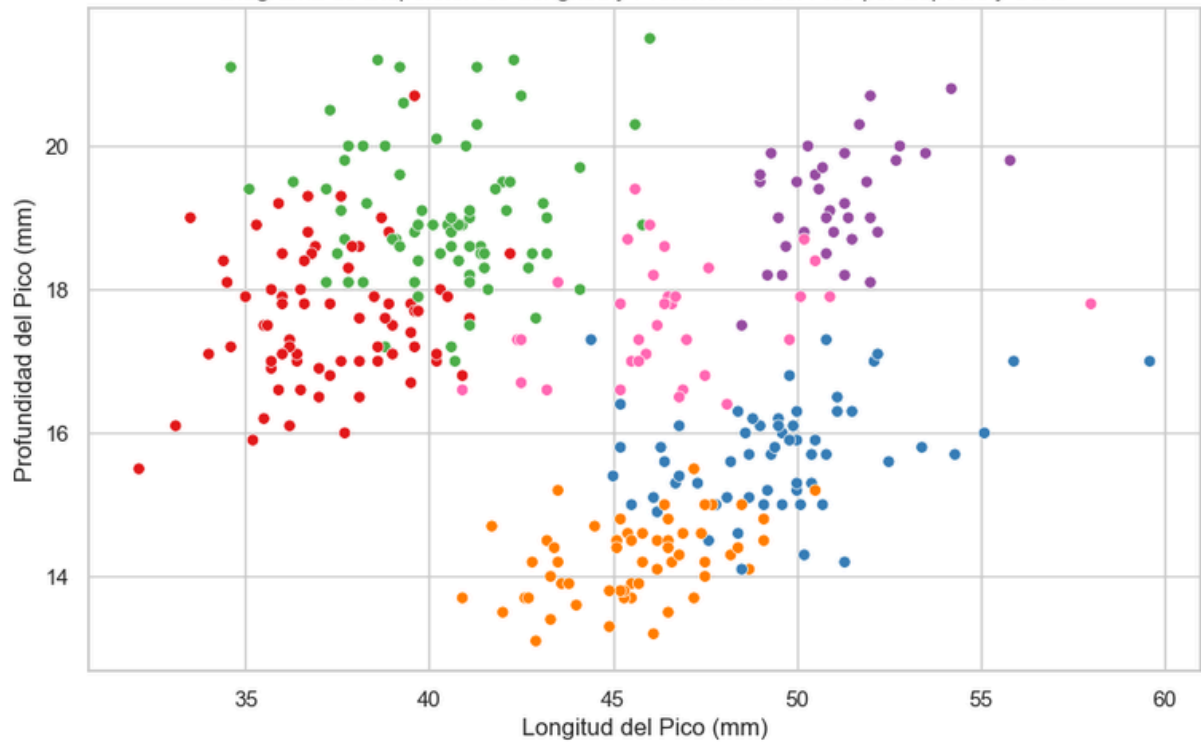
Compleitud: 66%

V-Measure: 66%

Posteriormente, se realizó una comparación con un modelo de clasificación KNN tomando en cuenta las seis distintas categorías de pingüinos: Hembras y Machos de Adelie, Chinstrap y Gentoo. Se tuvo una precisión del 81%.

Resulta relevante destacar que sin las etiquetas correspondientes, KMeans permitió agrupar los ejemplares registrados. Al comparar la gráfica de clusters con la que se muestra en la siguiente página, se puede apreciar una similitud notoria respecto a la manera en que los grupos se distinguen.

Diagrama de Dispersión de Longitud y Profundidad del Pico por Especie y Sexo



Especie_Sexo

- Adelie_MALE
- Adelie_FEMALE
- Chinstrap_FEMALE
- Chinstrap_MALE
- Gentoo_FEMALE
- Gentoo_MALE