

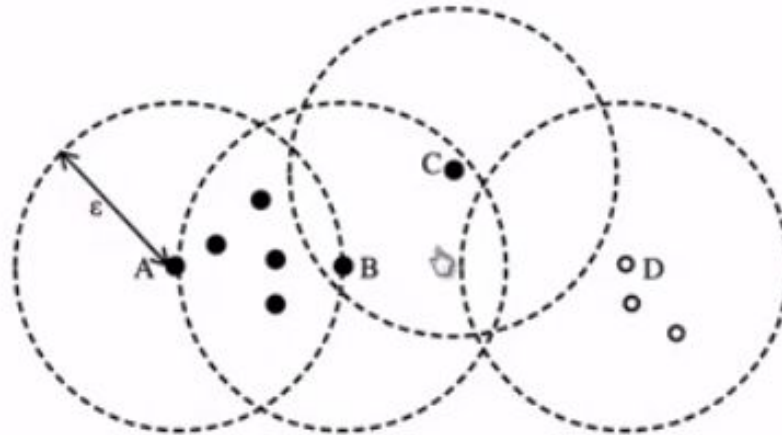
Semana 17 – Algoritmos de machine learning IV

Temas para el día de hoy

- Gaussian Mixture Model (GMM)
- Agglomerative Hierarchical Clustering
- Curso en coursera de google sobre transformación digital e IA.

Recordamos la sesión anterior

¿Defina el algoritmo de DBSCAN?



En la modalidad virtual participamos a través del chat; para ello usamos la opción “Levantar la mano”.

Logro de la semana



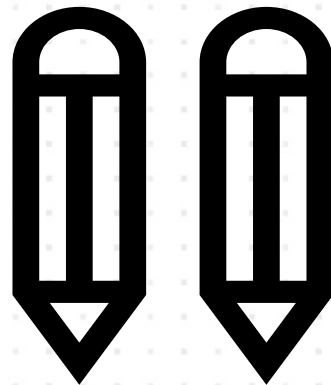
Al finalizar la sesión, el estudiante identifica los principales algoritmos de machine learning y aplica los conceptos en ejercicios.

Actividad inicial

Actividad inicial

Respondamos la siguiente pregunta:

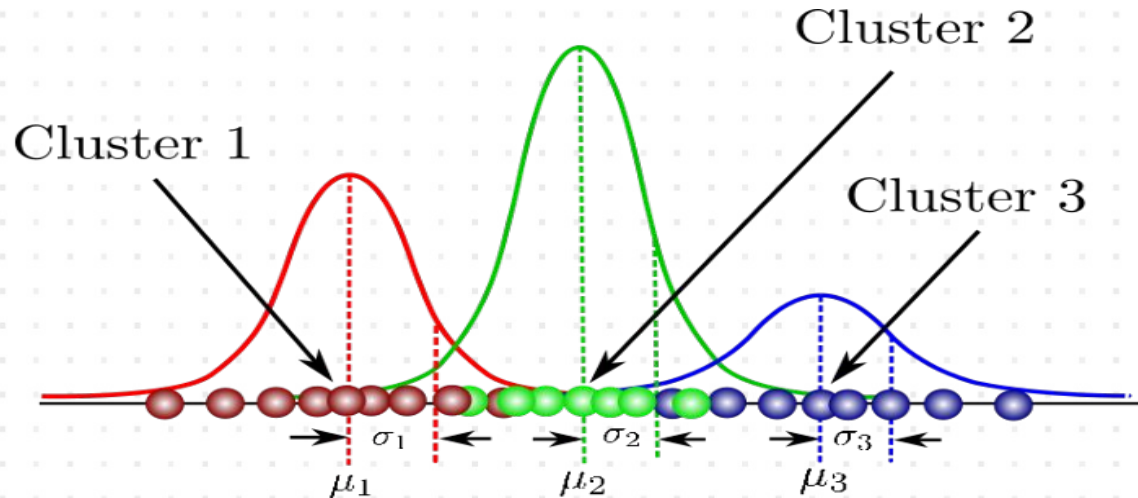
¿Qué es Qwiklabs?



Algoritmos de machine learning

Gaussian Mixture Model (GMM)

Es un modelo probabilístico en el que se considera que las observaciones siguen una distribución probabilística formada por la combinación de múltiples **distribuciones normales**. El algoritmo GMM considera cada distribución gaussiana como un grupo.



Una distribución Gaussiana es lo que también conocemos como distribución Normal. Ya sabes, ese concepto tan difundido de una curva en forma de campana con la media y la mediana como punto central.

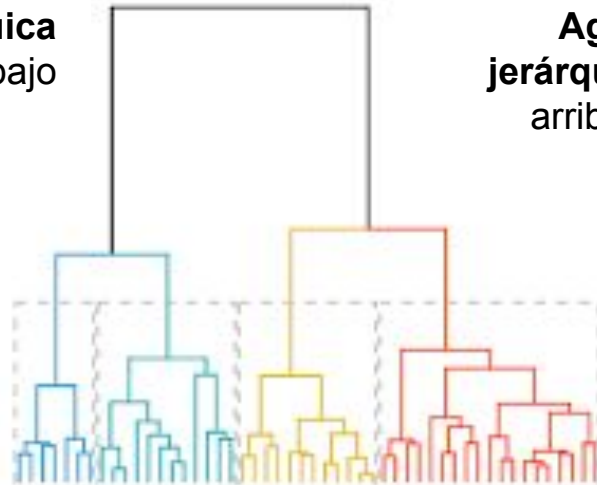
Gaussian Mixture Model (GMM)

El modelo GMM tiene que iniciar usando un parámetro aleatorio y construirlo a partir de eso(**init_params**). Una vez que se elige ese parámetro inicial, el algoritmo comienza una serie de cálculos tratando de encontrar los mejores pesos y medios necesarios para hacer converger las Distribuciones Gaussianas estimadas para cumplir con las que tienen los parámetros iniciales. En otras palabras, elegir bajo qué distribución caerá cada punto de datos.

Agglomerative Hierarchical Clustering

El agrupamiento jerárquico es un algoritmo de aprendizaje no supervisado que se basa en el agrupamiento de datos según el orden jerárquico.

**Agrupación jerárquica
aglomerativa** (de abajo
hacia arriba)



**Agrupamiento
jerárquico divisivo** (de
arriba hacia abajo)



Agglomerative Hierarchical Clustering

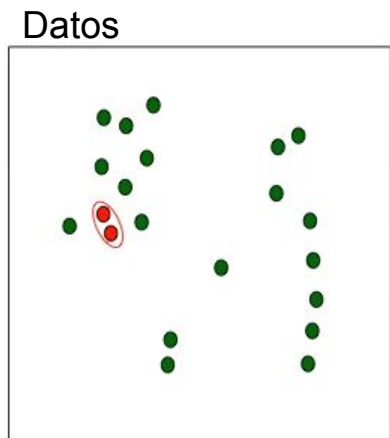
El siguiente parámetro es el **covariance_type**, donde puede elegir entre 'completo', 'atado', 'diagonal' y 'esférico'. Recuerde que K-Means agrupa puntos usando solo la forma circular.

- **Completo:** los componentes pueden adoptar independientemente cualquier posición y forma.
- **Atados:** significa que tienen la misma forma, pero la forma puede ser cualquier cosa.
- **Diagonal:** los ejes de contorno están orientados a lo largo de los ejes de coordenadas, pero de lo contrario, las excentricidades pueden variar entre los componentes.
- **Esférica** es una situación "diagonal" con contornos circulares (esférica en dimensiones más altas, de ahí el nombre).

Agglomerative Hierarchical Clustering

Primera iteración

Construcción de secuencia de grupos



Dendrograma

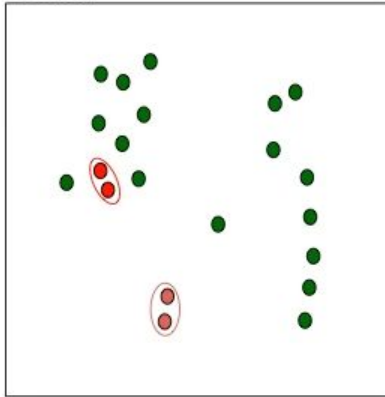


La altura de la unión
indica disimilitud

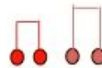
Agglomerative Hierarchical Clustering

Segunda iteración

Datos



Dendrograma

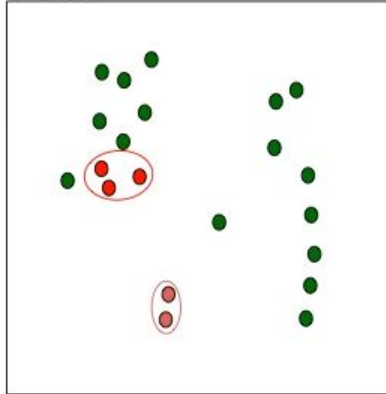


La altura de la unión
indica disimilitud

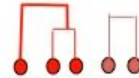
Agglomerative Hierarchical Clustering

Tercera iteración

Datos



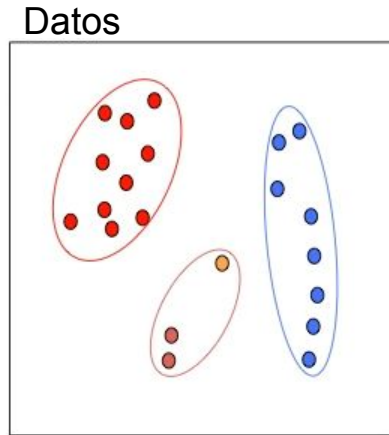
Dendrograma



La altura de la unión
indica disimilitud

Agglomerative Hierarchical Clustering

Cuarta iteración



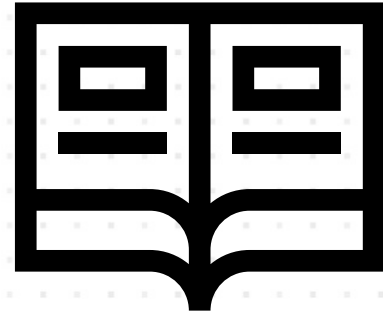
Dendrograma

A dendrogram illustrating the hierarchical clustering process. The horizontal axis represents the distance between clusters, and the vertical axis represents the hierarchy. The dendrogram shows three main branches: a purple branch on the left, a blue branch in the middle, and a red branch on the right. The purple branch connects three orange dots. The blue branch connects eight blue dots. The red branch connects eight red dots. The dendrogram is drawn with black lines, and the clusters are color-coded to match the data points in the scatter plot.

Pongamos en práctica lo aprendido

Resolvamos las actividades propuestas

Implementar el algoritmo de agrupación jerárquica y fundamentar los pasos para su correcto funcionamiento.

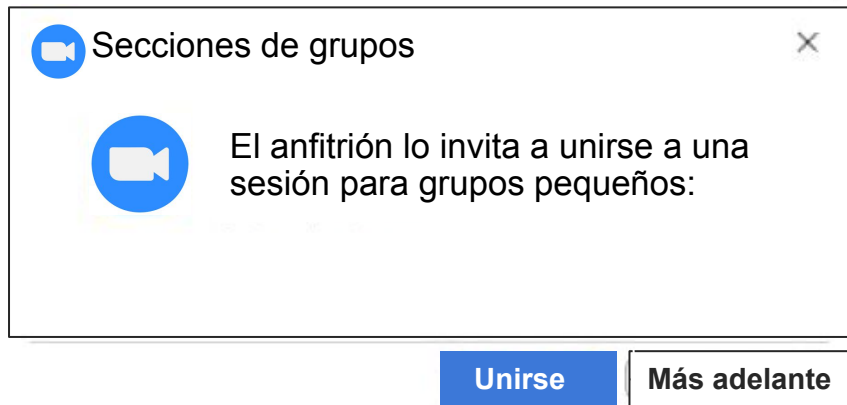


Equipos de Zoom

¿Listos para compartir en equipos de Zoom?



En tu pantalla aparecerá el nombre unirte a una sesión de grupos pequeños. Dale clic en **“Unirse”**.



Conclusiones

Conclusiones



¿Qué aprendiste en esta sesión? Te invitamos a compartir tus conclusiones en clase.

Conclusiones



- Gaussian Mixture Model (GMM), es un modelo probabilístico en el que se considera que las observaciones siguen una distribución probabilística formada por la combinación de múltiples distribuciones normales.
- Agglomerative Hierarchical Clustering (el agrupamiento jerárquico), es un algoritmo de aprendizaje no supervisado que se basa en el agrupamiento de datos según el orden jerárquico.

¿Preguntas?