
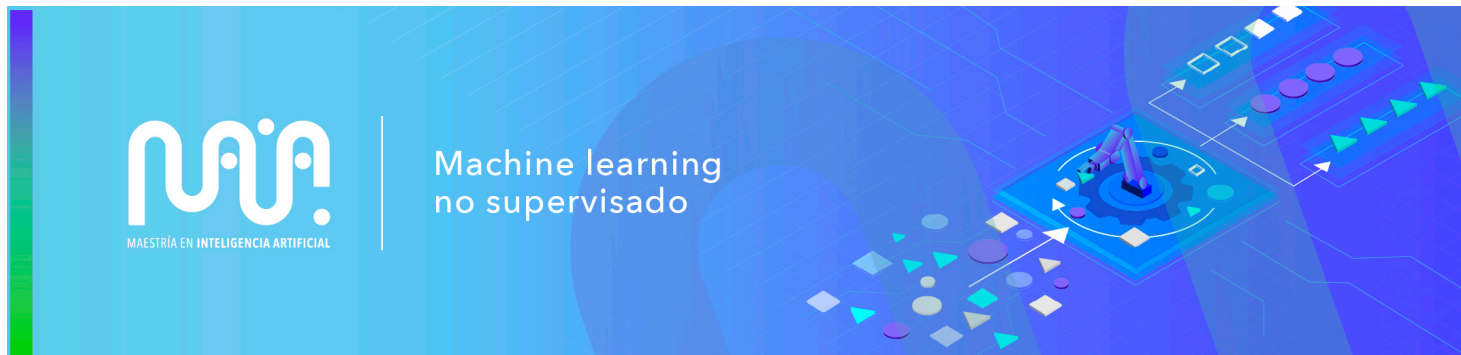


# ¿Cómo evaluar modelos de agrupación?

 This item includes content that hasn't been translated into your preferred language yet.



Para la evaluación de modelos de agrupación podemos utilizar diversos métodos, extrínsecos e intrínsecos. Los primeros utilizan el conocimiento experto de manera implícita, a través de un conjunto de datos etiquetados, los cuales representan la "verdad o realidad fundamental" (*ground truth*) del problema. Las etiquetas indican las clases a la que pertenecen los objetos del conjunto de datos y son utilizadas de manera explícita en métricas que determinan el grado al cual los datos que son de la misma clase están incluidos en el mismo grupo, es decir, el "acuerdo" en las asignaciones de los datos a los grupos y el *ground truth*. Por su parte, los métodos intrínsecos se basan en métricas que cuantifican la compactidad y separabilidad de los grupos. La compactidad está relacionada con la distancia intragrupo (*inter-clúster*) y mide su nivel de cohesión interna, es decir, qué tan cercanos están los datos dentro de un grupo. La separabilidad mide la distancia inter-grupos (*inter-clúster*), es decir, qué tan separados están los grupos entre sí.

Una detallada exposición sobre diversas métricas de evaluación se encuentra disponible en el sitio de scikit-learn bajo la sección "*Clustering performance evaluation*":

<https://scikit-learn.org/stable/modules/clustering.html#clustering-performance-evaluation> .

Este recurso no solo proporciona una descripción exhaustiva de las métricas acompañadas de ejemplos de código, sino que también ofrece recomendaciones sobre los contextos de aplicación donde resultan apropiadas, destacando sus ventajas y desventajas.

**Actividad.** Una vez realices la lectura, construye una tabla que muestre cómo se interpretan estas métricas, con base en los valores que pueden asumir.