
ACTIVIDAD PREVIA

- Realizar el cuestionario de este [enlace](#) sobre el artículo Non-negative Matrix Factorization: A Survey.

RESULTADO DE APRENDIZAJE

RdA de la asignatura:

- **RdA 1:** Plantear los conceptos fundamentales del aprendizaje automático, incluyendo los principios básicos, técnicas de preprocesado de datos, métodos de evaluación y ajuste de modelos, destacando su importancia en el análisis y resolución de problemas de datos.

RdA de la clase:

- Comprender qué es el aprendizaje no supervisado y sus aplicaciones generales.
- Diferenciar entre los tipos de agrupamiento: exclusivo, superposición y jerárquico.

INTRODUCCIÓN

Pregunta inicial: Imagina que el Profesor Oak pierde todos los datos sobre los «Tipos» (Fuego, Agua, Planta, etc.). Si solo tenemos sus estadísticas numéricas (Ataque, Defensa, Velocidad), ¿podría un algoritmo redescubrir por sí solo que Charizard y Blastoise pertenecen a diferentes grupos basados en sus características?

DESARROLLO

Actividad 1: Introducción al Aprendizaje No Supervisado

Se utilizará un video introductorio mediante proyección y discusión para familiarizar a los estudiantes con el concepto de aprendizaje no supervisado y sus aplicaciones en problemas de datos sin etiquetas.

¿Cómo lo haremos?

- **Proyección de video:** Se presentará un video educativo que explique qué es el aprendizaje no supervisado.

Enlace al video: [Qué es el Aprendizaje No Supervisado](#).

- **Interacción con ChatGPT:** Después del video, cada estudiante formulará preguntas a ChatGPT relacionadas con los términos o conceptos que no entendieron durante la visualización. Términos sugeridos para investigar: K-Means, Clusterización jerárquica, DBSCAN, Modelos de mezclas gaussianas, Ruido en los datos, Detección de anomalías.
- **Discusión grupal:** Los estudiantes compartirán los términos que investigaron y las respuestas proporcionadas por ChatGPT.

Actividad 2: Definiciones clave

En esta actividad los estudiantes explorarán las definiciones y características de los principales tipos de agrupamiento en el aprendizaje no supervisado mediante clase magistral, exploración con ChatGPT y búsqueda de referencias científicas en Consensus.

¿Cómo lo haremos?

- **Presentación de conceptos:** El docente explicará brevemente las definiciones y ejemplos de los principales tipos de agrupamiento en aprendizaje no supervisado:
 - **Agrupamiento exclusivo:** Cada dato pertenece exclusivamente a un único grupo. Ejemplo: k-means.
 - **Agrupamiento con superposición:** Los datos pueden pertenecer a múltiples grupos con grados de pertenencia. Ejemplo: clustering difuso (*fuzzy clustering*).
 - **Agrupamiento jerárquico:** Los datos son organizados en una estructura de árbol jerárquico. Ejemplo: dendrogramas.
- **Exploración guiada:** Los estudiantes realizarán dos tareas principales en grupos pequeños:
 1. **Preguntar a ChatGPT:** Cada grupo formulará las siguientes preguntas a la herramienta:
 - ¿Cuál es la principal diferencia entre el agrupamiento jerárquico y el agrupamiento exclusivo?
 - ¿Cuándo es preferible usar un modelo probabilístico frente a k-means?
 - ¿Qué limitaciones tienen los métodos de agrupamiento con superposición en aplicaciones reales?
 2. **Búsqueda de referencias:** Usando la plataforma [Consensus](#), los estudiantes buscarán un artículo científico que respalde o amplíe las respuestas obtenidas de ChatGPT, identificando la idea principal del artículo y cómo se relaciona con el tema discutido.
- **Presentación de hallazgos:** Cada grupo presentará un resumen de las respuestas obtenidas de ChatGPT y la referencia científica encontrada en Consensus.

Actividad 3: Efecto de la Normalización en la Clusterización

En esta actividad los estudiantes analizarán mediante exploración de cuaderno de Jupyter cómo las diferentes técnicas de normalización pueden afectar el resultado de los algoritmos de clusterización.

¿Cómo lo haremos?

- **Presentación inicial:** Se presenta un gráfico que muestra cómo diferentes normalizaciones (StandardScaler, MinMaxScaler, Normalizer) afectan la distribución de los datos en un espacio bidimensional.
- **Exploración individual:** Cada estudiante deberá analizar el gráfico y responder preguntas como:
 - ¿Cuál técnica de normalización parece más adecuada para separar los grupos visualmente?
- **Implementación en Python:** Los estudiantes accederán a un cuaderno de Jupyter previamente preparado.

Enlace al cuaderno: [07-Normalizacion-Agrupamiento.ipynb](#).
- **Experimentación:** Agrega otras normalizaciones (por ejemplo, RobustScaler) y observa cómo afectan los resultados.

CIERRE

Verificación de aprendizaje:

1. ¿Cuál es la diferencia principal entre agrupamiento exclusivo y agrupamiento con superposición?
2. ¿Por qué es importante seleccionar una técnica de normalización adecuada antes de aplicar un algoritmo de clusterización?

Pregunta de investigación:

1. ¿Cómo podrían detectarse anomalías en un conjunto de datos usando aprendizaje no supervisado?