

Actividad de Evaluación

Objetivos de Aprendizaje

OA 2 Representar diferente tipo de datos en una variedad de formas que incluya textos, sonidos, imágenes y números.

OA 3 Desarrollar y programar algoritmos para ejecutar procedimientos matemáticos, realizar cálculos y obtener términos definidos por una regla o patrón.

OA g Elaborar representaciones, tanto en forma manual como digital, y justificar cómo una misma información puede ser utilizada según el tipo de representación.

OA i. Buscar, seleccionar, manejar y producir información matemática/cuantitativa confiable a través de la web.

Indicadores de evaluación

- Identifican tipos de datos para elaborar diversas representaciones.
- Utilizan códigos para traducir algoritmos a lenguaje de programación.
- Programan algoritmos que permiten ejecutar procedimientos matemáticos de cálculo.
- Ejecutan programas que incluyen diferentes tipos de datos.
- Programan algoritmos que permiten obtener un término definido al ejecutar procedimientos matemáticos.

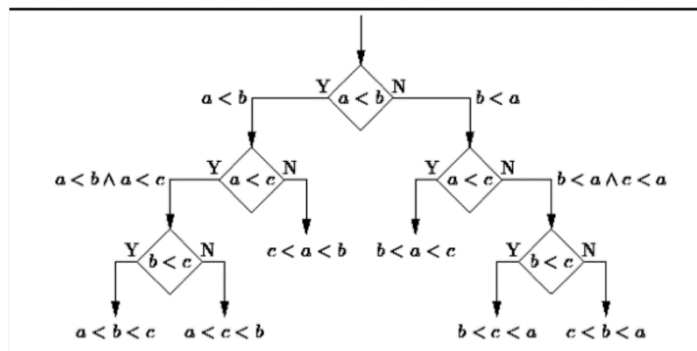
Duración: 6 horas pedagógicas

Se puede usar las siguientes actividades como ejemplos de evaluaciones para la unidad 2, cada una por sí misma o en conjunto. Se sugiere delimitar la evaluación según el contexto y el tiempo disponible.

Primera parte individual

1. Explica el funcionamiento del siguiente diagrama de flujo; recorre el árbol con los siguientes valores:

a	b	c
1	5	-3
0	-2	4
10	3	8



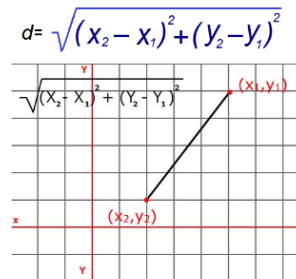
¿Puedes indicar qué hace este diagrama de flujo?

- ¿Para qué sirve?
- Los tres casos de la tabla, ¿permiten recorrer todas las ramas de este diagrama?

2. Construye el algoritmo correspondiente a este diagrama de flujo y pruébalo para los mismos tres casos de la tabla. ¿Obtienes los mismos resultados?

Segunda parte individual

1. Construye un algoritmo con su respectivo diagrama de flujo que permita determinar la distancia entre dos puntos, utilizando la herramienta PSeInt. Pruébalo para, al menos, 3 pares de puntos. ¿Es siempre positiva la distancia entre 2 puntos?
2. Construye un algoritmo en la herramienta PSeInt que permita generalizar este proceso de la distancia euclidiana entre los puntos $P = (p_1, p_2, \dots, p_n)$ y $Q = (q_1, q_2, \dots, q_n)$ del espacio euclidiano n -dimensional; se define como:



$$d_E(P, Q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}.$$

- ¿Cómo se relaciona este algoritmo con el de la pregunta anterior?
- ¿Qué representa esta nueva fórmula? ¿Qué interpretación le podrías dar?
- ¿En qué ejemplo la podrías utilizar?
- Construye una planilla Excel con los resultados correspondientes a estas fórmulas y comprueba los resultados obtenidos con, al menos, tres casos, tanto en la planilla como con tu algoritmo.

Tercera parte grupal

- La siguiente tabla muestra un grupo de 20 personas y sus gustos sobre películas. Abran una planilla Excel, y nomínenla “Ejemplo K-means Estudiantes”.

Persona	Star Wars	Jurassic Park	Toy Story	Avenger
1	4	4	1	2
2	5	4	2	1
3	2	2	4	3
4	2	1	3	4
5	5	3	2	2
6	5	5	1	3
7	3	2	1	2
8	2	1	2	2
9	1	3	3	5
10	1	1	1	3
11	2	4	4	1
12	5	1	5	2
13	3	2	4	4
14	3	3	2	5
15	5	1	1	5
16	5	4	3	3
17	4	5	3	1
18	1	5	2	1
19	1	4	1	2
20	3	3	1	1

- ¿Entre qué rango de valores se mueven las preferencias de las distintas películas?
 - ¿Pueden decir a simple vista qué personas tienen gustos similares de películas a partir de la valoración que le dan a cada una?
 - ¿Qué pasa si tenemos 1 000 personas o 1 000 000 de personas?
 - ¿Qué características tiene este problema, que permite que el algoritmo K-means ayude a resolverlo?
- Este algoritmo ayuda a encontrar grupos similares. Requiere, en primer lugar, definir el número de *k-clusters* o grupos que se implementará, los cuales se llaman centroides.



Abran la planilla electrónica “Ejemplo_K-means_Estudiante”. En este problema, seleccionamos 3 personas al azar que representan los 3 centroides como puntos de partida. En nuestro caso, serán las personas 5, 7 y 17.

En la figura anterior, los centroides quedarían graficados, por ejemplo, así:



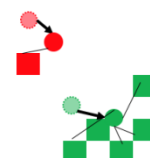
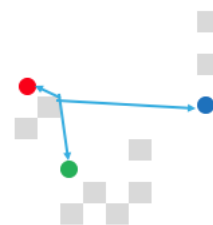
Persona	Star Wars	Jurassic Park	Toy Story	Avenger
1	4	4	1	2
2	5	4	2	1
3	2	2	4	3
4	2	1	3	4
5	5	3	2	2
6	5	5	1	3
7	3	2	1	2
8	2	1	2	2
9	1	3	3	5
10	1	1	1	3
11	2	4	4	1
12	5	1	5	2
13	3	2	4	4
14	3	3	2	5
15	5	1	1	5
16	5	4	3	3
17	4	5	3	1
18	1	5	2	1
19	1	4	1	2
20	3	3	1	1

3. Se calculará la distancia euclidiana entre cada punto y cada uno de los centroides. Observen la planilla: en el rango AB1...AF4, se colocó una tabla con las valoraciones para los 3 *clústers* seleccionados. Además, en las celdas G2, H2 e I2 ya se ha ingresado las fórmulas para calcular la distancia de la primera persona. Observen la fórmula de cada celda y vean las celdas involucradas; pueden hacer doble clic para que resaltar las celdas involucradas en cada fórmula. ¿Qué representa la fórmula de cada celda?
 - a. Calculen las distancias para el resto de las personas, copiando las fórmulas anteriores.
 - b. Observen la fórmula de la celda J2; ¿qué hace? Cópíenla para cada fila de datos.

4. Los valores de la columna J, llamada “Mínimo”, se mueven entre 1, 2 y 3, y representan a cada uno de los *clústers*. En este punto, ya tienen una primera organización de todas las personas. Con la opción de filtro, observen los datos de las personas para cada grupo; es decir, las personas que están agrupadas en 1, luego las en 2 y finalmente las que están en 3.
 - a. Al mirar cada grupo, ¿qué pueden decir sobre las valoraciones para las diferentes películas?
 - b. ¿Se observa alguna tendencia o patrón?

5. Se calculará una nueva media para cada centroide como la media de los puntos del mismo grupo. Observen que, en el rango AB6... AF10, se ingresó una nueva tabla, llamada “Después de la iteración 1”, y para cada película y cada *clúster* se calculó una nueva media.
 - a. ¿Qué representa esta nueva fórmula?
 - b. ¿Qué datos están involucrados?

6. Calculen una nueva iteración para los tres grupos. Para esto, observen las fórmulas de la celda K2 e ingresen las fórmulas correspondientes para L2 y M2. Copien esta fórmula para cada una de las personas (para cada fila). Esto mueve el centro de gravedad de la nueva ubicación para cada *clúster*. Copien la fórmula de la celda de J2 en N2 para calcular el mínimo valor de la iteración de los tres *clústers* para la persona 1. Luego copien N2 en el resto de las filas.
 - a. Observen la nueva tabla que se genera en N12 “Después de la iteración 2” y comparen estos valores con los de la tabla anterior “Después de la iteración 1”. Las tablas, ¿son iguales o distintas?
 - b. Si las tablas de iteración son iguales, pueden terminar; de lo contrario, continúen repitiendo esta actividad desde el punto 13, hasta que la tabla de la iteración que se genere sea igual a la anterior.
 - c. ¿En qué ciclo de la iteración terminó el proceso?



7. Con la opción de filtro, observen los datos de las personas para cada grupo de los valores de la columna Z; es decir, las personas que están agrupadas en 1, luego en 2 y finalmente en 3. Al mirar cada grupo, ¿qué pueden decir sobre las valoraciones de las diferentes películas por parte de las personas del grupo? ¿Se observa alguna tendencia o patrón? Comparen esto con lo obtenido en la primera tabla; es decir, con los grupos generados en la columna J. ¿Cómo son los grupos en cada iteración? ¿Por qué la última es mejor?
8. ¿Qué información pudieron obtener sobre el uso del algoritmo K-means en este ejemplo? Den algunas ideas de dónde podrían usar este algoritmo. ¿Por qué piensan que este algoritmo se usa en el ámbito de ML respecto del concepto de aprendizaje de máquinas?

PAUTA DE EVALUACIÓN

Criterios de evaluación	Niveles de logros		
	Completamente logrado	Se observa aspectos específicos que pueden mejorar	No logrado por ausencia o no se puede entender las instrucciones o la secuencia
Identifican tipos de datos y sus representaciones.			
Representan diferentes tipos de datos al programar algoritmos.			
Ejecutan programas que incluyen diferentes tipos de datos.			
Traducen un algoritmo a programa con códigos proporcionados.			
Programan, de manera individual y colaborativa, algoritmos que permiten ejecutar procedimientos matemáticos de cálculo.			
Programan, de manera individual y colaborativa, algoritmos que permiten obtener un término definido al ejecutar procedimientos matemáticos.			