**Módulo: ADM1929 - Business understanding: Pensamiento analítico basado en datos. - (A51)**

**Actividad: Reto de aprendizaje 21. Documento sobre aplicación (método)**

**Nombre: Roberto Mora Balderas**

**Asesor: José Carlos Soto Monterrubio**

**Fecha: 01 de Agosto de 2023**

**Objetivo**:

Aplicar el análisis de clúster con la información presentada con la finalidad de realizar un análisis que optimice la modelación de costos.

**Instrucciones**:

1. Aplica el análisis de clúster X-means con los datos proporcionados, y genera un documento explicando paso por paso las etapas para aplicarlo. Incluye los siguientes aspectos:
2. Importación de data.
3. Aplicación en Auto Model.
4. Presentación del número de clúster y la tabla de centroides.
5. Análisis preliminar de los resultados con respecto a los clústeres generados. ¿Qué caracteriza a cada clúster?

**Desarrollo:**

1. Importar los datos, se inicia dando clic en el icono de importación de datos.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Elegir de donde se importará el data set.

Interfaz de usuario gráfica, Aplicación, Sitio web

Descripción generada automáticamente

1. Elegir la ruta donde se cargará el archivo.

Interfaz de usuario gráfica

Descripción generada automáticamente con confianza media

1. Elegir opciones con respecto a la importación y dar siguiente.

Tabla

Descripción generada automáticamente

1. Elegir la dirección del repositorio donde se guardará el data set.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Irse a la pestania automodel y elegir la data con la que se desea trabajar.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Elegir modelo a utilizar, en este caso utilizaremos el modelo de clustering para aplicar X-means.

Interfaz de usuario gráfica, Texto, Aplicación

Descripción generada automáticamente

1. Seleccionar la variable utilizar en el modelo, como recomendación se utilizan solo las que posean un estatus de color verde, en nuestro caso eso no incluye el ProductoID. El estatus nos brinda información de la relevancia y de la correlación entre las variables y el modelo.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

1. Elegir el modelo, en este caso x-means y correr el modelo.

Interfaz de usuario gráfica, Texto, Aplicación, Correo electrónico

Descripción generada automáticamente

1. Generado el modelo podemos observar como nos arroja dos clusters como modelo optimo.

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

Descripción generada automáticamente

1. A continuación, se muestra la tabla de centroides.

Interfaz de usuario gráfica

Descripción generada automáticamente

La caracterización de los clusters puede interpretarse de la siguiente manera:

Cluster 0:

* Demand\_growth: Tiene un valor positivo, lo que indica que este cluster tiende a tener un crecimiento de la demanda.
* Equated margins: Tiene un valor negativo, lo que sugiere que este cluster tiende a tener márgenes más bajos.
* Fixed costs ratio: Tiene un valor negativo, lo que indica que este cluster tiende a tener una relación más baja entre los costos fijos y los costos totales.
* Process costs: Tiene un valor positivo, lo que sugiere que este cluster tiende a tener costos de proceso más altos.
* Rawmatl Costs: Tiene un valor positivo, lo que indica que este cluster tiende a tener mayores costos de materia prima.
* Revenue: Tiene un valor positivo, lo que sugiere que este cluster tiende a generar mayores ingresos.
* RoboticAssy: Tiene un valor negativo, lo que indica que este cluster tiende a utilizar menos ensamblaje robótico.
* Utilization rate: Tiene un valor negativo, lo que sugiere que este cluster tiende a tener una tasa de utilización más baja.

Cluster 1:

* Demand\_growth: Tiene un valor negativo, lo que indica que este cluster tiende a tener un decrecimiento de la demanda.
* Equated margins: Tiene un valor positivo, lo que sugiere que este cluster tiende a tener márgenes más altos.
* Fixed costs ratio: Tiene un valor positivo, lo que indica que este cluster tiende a tener una relación más alta entre los costos fijos y los costos totales.
* Process costs: Tiene un valor negativo, lo que sugiere que este cluster tiende a tener costos de proceso más bajos.
* Rawmatl Costs: Tiene un valor negativo, lo que indica que este cluster tiende a tener menores costos de materia prima.
* Revenue: Tiene un valor negativo, lo que sugiere que este cluster tiende a generar menores ingresos.
* RoboticAssy: Tiene un valor positivo, lo que indica que este cluster tiende a utilizar más ensamblaje robótico.
* Utilization rate: Tiene un valor positivo, lo que sugiere que este cluster tiende a tener una tasa de utilización más alta.

**Referencias:**

* Anáhuac Online. (2019). *Problema de negocio* [Contenido creado para Anáhuac Online].
* Vohra, G. (2018). *Cluster Analysis For Business*. https://bit.ly/3pICgmq
* Whittaker, C. (2019). *7 Innovative Uses of Clustering Algorithms in the Real World*. https://bit.ly/2KVXP0u