

## Indicaciones para el Proyecto Final 2023-1

Grupo: 3CV11

No. Equipo: \_\_\_\_\_

Nombre de los integrantes del equipo:

- 1) \_\_\_\_\_
- 2) \_\_\_\_\_
- 3) \_\_\_\_\_

### INDICACIONES GENERALES

- a) Se organizarán en equipo de 2 ó 3 alumnos
- b) Los alumnos deben presentar dos revisiones previas del proyecto
- c) Agregar una portada de presentación del proyecto y en cada una de las diversas técnicas de análisis de datos
- d) Elaborar un índice de contenido
- e) Anexar al final de cada apartado las referencias bibliográficas
- f) Presentarán en la clase y generarán un reporte
- g) El proyecto final lo debe desarrollar en Python y Knime.
- h) Puede localizar el conjunto de datos a utilizar en los sitios Kaggle, UCI Machine Learning o algún otro repositorio que contengan conjuntos de datos.

### DESCRIPCIÓN DEL PROYECTO

#### 1. RECURSOS.

Cada equipo creará un archivo empaquetado con el nombre del tema (ProyectoFinal\_EquipoNúmero, por ejemplo: ProyectoFinal\_Equipo2.rar) en donde colocarán los recursos del proyecto: archivos de datos, reporte y lo que considere necesario. Colocará este archivo .zip o .rar en la asignación.

#### 2. ESTRUCTURA DEL REPORTE:

El reporte debe incluir los siguientes aspectos:

##### A. **CONJUNTO DE DATOS.**

Identificar un conjunto de datos con diversos tipos de datos (categóricos y numéricos), al que le pueda aplicar técnicas de tratamiento de datos, de un tamaño y estructura similar al de museos. En este apartado debe agregar los siguientes datos:

Nombre del conjunto de datos:	
Fecha de publicación:	

Autor(es):	
Objetivo del conjunto de datos:	
Cantidad de atributos:	
Cantidad de registros:	
Enlace de acceso:	

**B. DICCIONARIO DE DATOS.**

Diccionario de datos: número de atributo, nombre, tipo, significado y dominio.

No. Atrib.	Nombre	Tipo	Significado	Dominio

**C. PROPUESTA DE TÉCNICAS DE MINERÍA DE DATOS:**

Plantear los objetivos de las técnicas de minería de datos a aplicar en su conjunto de datos.

Técnica	Objetivo	Atributos
Árboles (Tipo)		
Bagging		
Boosting		
Agrupamiento (#)		
Regresión lineal (&)		
Reglas de asociación (*)		
Análisis de componentes principales		

(\*) En Reglas de Asociación, puede crear un catálogo de oferta de productos, integrando imágenes. No debe ser el que realizó en el proyecto de tema.

(#) Agregar las métricas de la silueta para analizar los elementos de los grupos creados.

(&) En la regresión lineal agregar gráfica de residuales, prueba de normalidad y prueba de significancia estadística.

**D. ASPECTOS A CONSIDERAR**

La estructura de la investigación es la siguiente (reporte):  
En el reporte: Crear una portada para cada técnica, mencionando el nombre que corresponda.

Para cada una de las técnicas de minería de datos debe agregar los aspectos indicados a continuación.

**i. Introducción.**

- Desarrollar el marco teórico del tema
- Agregar una breve descripción del estudio que se realizó en esta sección, así como los resultados encontrados.
- Intención de la aplicación de la técnica
- Justificación de la técnica a aplicar

- ii.** Diccionario de datos. Fracción del conjunto de datos que utilizó en cada técnica, considerando: nombre de la variable, tipo de dato (nominal, ordinal, numérico discreto o continuo), dominio de valores.
- iii.** Resultados. Medidas de estimación y valoración de la obtención y resultados encontrados.
  - a.** Diagrama generado.
  - b.** Medidas obtenidas
  - c.** Descripción de las características de los resultados generados.
  - d.** Tipo de muestra que utilizó para prueba y entrenamiento.
- iv.** Análisis de los resultados. Describir los resultados encontrados proporcionando la explicación de los mismos. Así como las conclusiones que obtenga. Las métricas de evaluación deben corresponder a la técnica aplicada.
- v.** Anexo. Indicar el nombre del archivo que contiene el código correspondiente, colocado en el archivo empaquetado.