

Instrucciones

Seleccione una base de datos que le acomode para desarrollar el siguiente trabajo.

De acuerdo a la base de datos seleccionada, debe elaborar un informe en grupos de **a lo más 3 integrantes**, donde desarrolle los puntos que se detallarán en la siguiente sección (Preguntas). **Este informe debe ser entregado a más tardar el viernes 6 de octubre del presente año, a las 13:00 hrs.**

La estructura general del informe debe considerar el siguiente esquema:

- Portada: Título del trabajo e integrantes que lo desarrollan. **(1 plana)**
- Presentación de la base de datos utilizada, descripción del problema y variables involucradas (destaque el tipo de variables que serán usadas). **(1 plana)**
- Desarrollo del problema planteado: desarrollo a través de las preguntas que se establecerán en la siguiente sección. **(mínimo 2 planas - máximo 10 planas)**
- Conclusiones y/u observaciones relevantes. **(1 plana)**
- Referencias bibliográficas utilizadas. **(1 plana)**
- Código empleado (en el software que le acomode). **(Extensión de planas, libre)**

Se prohíbe el uso de las bases de datos:

- iris: <https://archive.ics.uci.edu/ml/datasets/Iris>,
- wine: <https://archive.ics.uci.edu/ml/datasets/Wine>,
- breast cancer: <https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+28Diagnostic29>

Preguntas

1. Investigue sobre el método gráfico “Caras de Chernof” y represente gráficamente un conjunto de variables mediante las caras de Chernof, ¿qué representa cada cara?, ¿qué representan los rasgos de cada cara (identifique)?
2. Técnicas de reducción de dimensionalidad.
 - (a) Investigue y defina el método de “t-SNE” como técnica de reducción de dimensionalidad.

- (b) Describa sus ventajas y desventajas respecto del análisis de componentes principales.
- (c) Aplique un análisis de componentes principales y el método t-SNE y compare sus resultados obtenidos. Sea detallada/o al momento de describir los hallazgos obtenidos.

3. Técnicas de clustering

- (a) Investigue y defina sobre “Gaussian Mixture” para clasificación de observaciones.
- (b) Establezca una regla de discriminante lineal para su base de datos seleccionada. Indique si tiene información previa de las probabilidades de pertenencia a los determinados grupos (probabilidades a priori), de ser así úselas en su regla de clasificación. Grafique sus resultados (ambas reglas de clasificación).
- (c) Use Gaussian Mixture en la base de datos seleccionada. Compare sus resultados con los obtenidos en el ítem anterior. Sea detallada/o al momento de describir los hallazgos obtenidos.
- (d) Simule **10 nuevas observaciones**. De acuerdo a la regla de clasificación mediante discriminante lineal y Gaussian Mixture, clasifique estas 10 nuevas observaciones. Grafique sus resultados y compárelos.

Rúbrica de Evaluación

Integrante(s): _____

Aspecto a evaluar	1	2	3	4	5	6	7
1.- Presenta el tema de forma clara y convincente							
2.- Pregunta 1							
3.- Pregunta 2a							
4.- Pregunta 2b							
5.- Pregunta 2c							
6.- Pregunta 3a							
7.- Pregunta 3b							
8.- Pregunta 3c							
9.- Pregunta 3d							
10.- Las interpretaciones son correctas							
11.- Los contenidos expuestos son los pertinentes en relación al tema							
12.- Entrega la bibliografía utilizada							
13.- Se apoya en imágenes para explicar lo pertinente							
14.- Se esmera por la calidad de su presentación							
15.- Cumple con los plazos establecidos							

Observaciones: