

Universidad Autónoma de Nuevo León
Facultad de Ciencias Físico Matemáticas

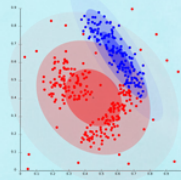
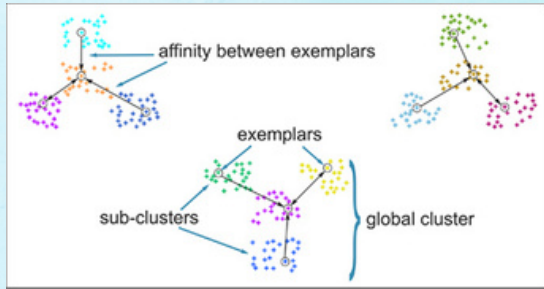
Minería de Datos Grp. 02
Mapa mental II

Alumno: Alfonso Llanos Morales
Matrícula: 1887939

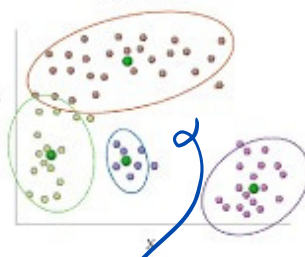
7° Semestre

Clustering

Clustering. Consiste en agrupar puntos de datos y de esta forma crear particiones basándonos en similitudes.

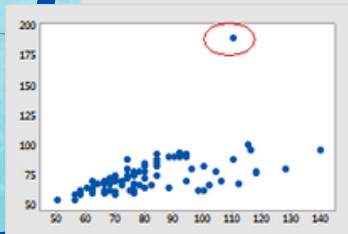
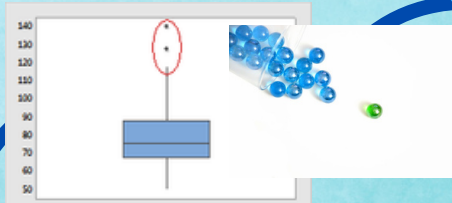
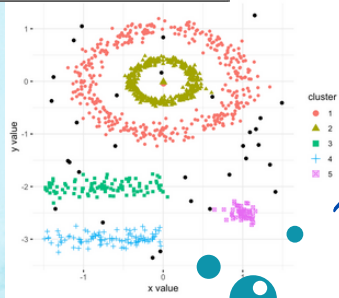


Centroid-Based Clustering



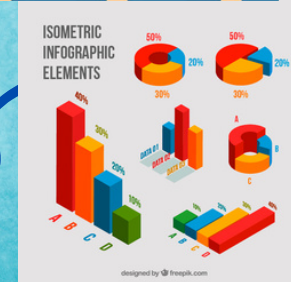
Detección de Outliers.

Los outliers o datos atípicos son los datos que se encuentran alejados del resto de datos, debido a que no siguen el patrón que la mayoría sí. La detección de outliers es fundamental ya que de mantener estos datos dentro de nuestra base de datos pueden ocasionar distorsiones en los parámetros que calculemos.



Detección de Outliers

Descriptivo



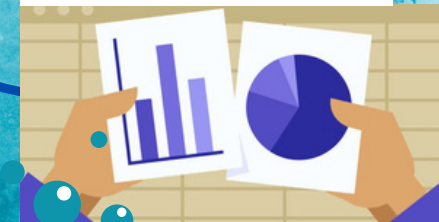
Reglas de Asociación

Reglas de Asociación.

Extraer información por coincidencias con el objetivo de encontrar relaciones dentro de un conjunto de transacciones.



Visualización. Es la representación gráfica de información y datos, proporcionan una manera accesible de ver y comprender tendencias, valores y patrones de los datos.

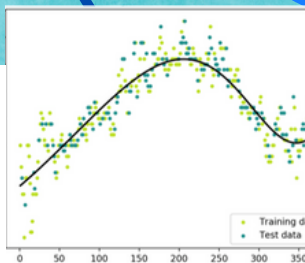
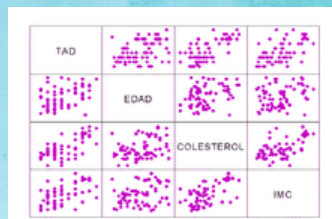


Visualización

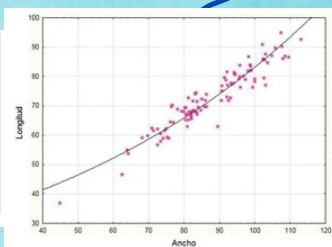
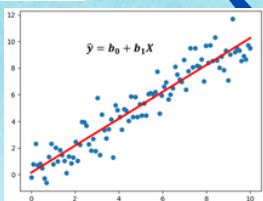
Técnicas de

Minería de Datos

Regresión



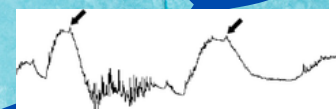
Regresión. Se encarga de analizar el vínculo entre una variable dependiente y una o varias independientes, encontrando una relación matemática.



Clasificación. Es la organización o mapeo de un conjunto de atributos por clase dependiendo de sus características.

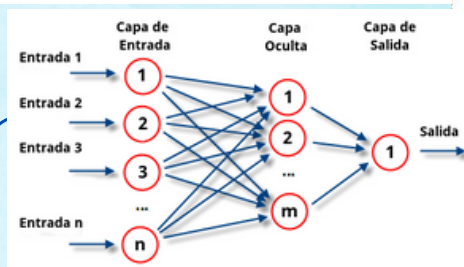
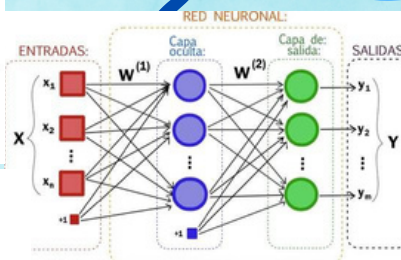
Predictivo

Patrones Secuenciales



Patrones Secuenciales. Clase espacial de dependencia en la que el orden de los acontecimientos es considerado, son eventos que se relacionan con el paso del tiempo. Si sucede el evento x en el instante del tiempo t entonces sucederá el evento y en el instante t+n.

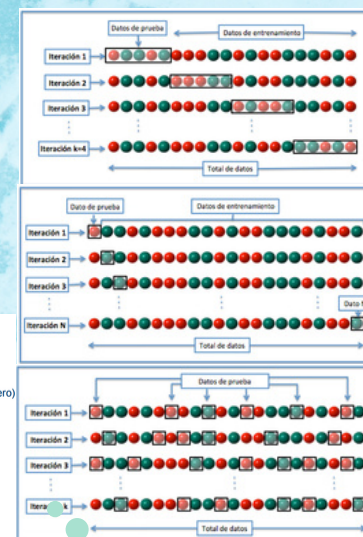
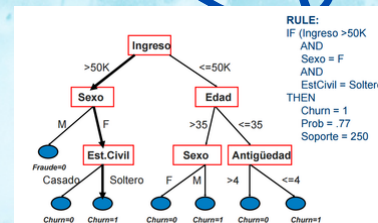
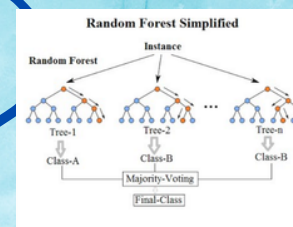
	Reality	
	True	False
True	Correct 😊	Type 1 error False Positive
False	Type 2 error False Negative	Correct 😊



Clasificación



Predicción. Existe una metodología para la partición de los datos a analizar en conjunto de entrenamiento, prueba y validación. Comúnmente se utilizan árboles de decisión y random forest para predecir el resultado de una variable.



Predicción