

ISW-911 Minería de Datos

Proyecto III

Supuestos de Negocio:

- Como consultor de la empresa, Grupo Innova S.A, usted fue contratado para implementar algoritmos de Machine learning en la solución de problemas para mejorar el desempeño de la empresa.

Supuestos de Proyecto:

- El tiempo máximo, por grupo, para la defensa del proyecto es de 20 minutos. En dicha presentación deberá explicar el proceso EDA y procesos de transformación realizados para procesar los datos.
- Los grupos ya fueron creados con anterioridad, no se permitirán cambios durante la elaboración de este proyecto.
- Aunque el proyecto es grupal, la evaluación será de forma individual, predominando la participación y el conocimiento de los conceptos de forma equitativa por parte de los integrantes.
- Para desarrollar el ejercicio debe utilizar Python o R.

Objetivo del proyecto:

Comprender los conceptos aprendidos en el curso de minería de datos acerca de las técnicas de machine learning, aplicador a la predicción y clasificación de datos, aplicando un ejercicio de índole investigativo y práctico en la resolución de los enunciados presentados a continuación.

Enunciado I:

- a. Utilice los conjuntos de datos resultados de cualquiera de los ejercicios 1,2 o 3, estos dataset deben de contar con todo el proceso EDA, además de la limpieza y normalización respectiva, si el dataset no presenta este proceso de depuración previo el ejercicio se calificará con 0.
- b. Analice el dataset y determine si existe y cuál es la variable dependiente en el conjunto de datos.
- c. Analice los datos y la composición de la información e investigue y aplique un algoritmo apropiado para realizar un ejercicio de clasificación y una predicción puede utilizar alguno de los visto en clase (Cluster Jerárquico, K-Means, regresión logística).
- d. Documente todos los pasos realizados e interprete adecuadamente los resultados, finalmente expóngale a la junta directiva de la empresa.