



**UNIVERSIDAD TECNOLÓGICA DE PANAMÁ**  
**FACULTAD DE INGENIERÍA DE SISTEMAS COMPUTACIONALES**  
**DEPARTAMENTO DE SISTEMAS DE INFORMACIÓN, CONTROL Y**  
**EVALUACIÓN DE RECURSOS INFORMÁTICOS**

**Gestión de la Información**

**5003**

**Parcial #1**

Nombre y Apellido del Estudiante: Ernesto Crespo

Cédula: \_8-929-1657

Fecha: 30 de abril de 2024

Instrucciones Generales: Para este examen no se requiere el uso del móvil o celular. El parcial es individual. No se puede consultar ni mirar al compañero.

Lea detenidamente las instrucciones en cada enunciado planteado para poder resolverlo adecuadamente. Se puede consultar internet y apuntes de cada estudiante en su computador.

Total: 100 puntos

1-Parte. De los modelos avanzados de la Gestión de la información y complete la siguiente tabla. 15 puntos.

Modelos	De que se trata	Sector o Caso de Uso / aplicación
Modelo Rowley	Se basa en estos pilares para la gestión de información: el entorno de información, el contexto de información, los sistemas de información y las interfaces persona-información.	Recursos de Información, Cultura Organizacional y Servicios o Sistemas.
Modelo PonJuan	Se centra en la creación de una estructura de datos semántica y jerárquica para organizar y recuperar información de manera eficiente. Utiliza la indexación y la relación de conceptos.	Recursos Humanos, Procesos y Tecnología Usado como apoyo en estos rubros
Modelo Paez Urdaneta	Se basa en la teoría de la semántica cognitiva y propone una estructura de datos que organiza la información en función de relaciones semánticas y contextuales.	Ambiente Social y Cultura Logística y Transporte Servicios y Actividades en Línea

2-Parte. De acuerdo con los datos facilitados para este examen parcial ejecute la sintaxis correspondiente en Jupiter (Python) de al menos los siguientes pasos: 85 puntos

Colocar la sintaxis utilizada en cada espacio de la tabla y adjuntar su notebook.

1. Cargar Librerías (5 puntos) Cargar Datos	# Paso 1 Cargamos las librerias import seaborn as sn #Graficas import pandas as pd #Manejar la data import matplotlib.pyplot as plt #Graficas from sklearn.metrics import confusion_matrix #Matriz de Confusion
--	--

2. Al menos Verificar si hay valores nulos. Analizar los datos. Comente que debe verificar (20 puntos)	<p>Verificare si hay valores nulos y si hay repetidos</p> <pre># Paso 2 Verificamos si hay valores nulos data.dropna(inplace=True) data.shape # Paso 2 Parte 2 - Verificamos datos repetidos print(f'Tamano del Set antes de eliminar repetidos: {data.shape}') data.drop_duplicates(inplace=True) print(f'Tamano del Set despues de eliminar repetidos: {data.shape}')</pre>
3. Están equilibrados os datos ¿Está equilibrada la target? (10 puntos)	<p>Los datos estan desbalanceados, la target esta a 96 contra 204 (relacion)</p> <pre># Paso 3 Equilibrio de Datos - Target num_muertos = data['DEATH_EVENT'].sum() num_noMuertos = 300 - num_muertos #num_noMuertos = data['DEATH_EVENT']==False print("Muertos:",num_muertos) print("No Muertos:",num_noMuertos)</pre>
4. Verificar la correlación y comentar (10 puntos)	<pre># Paso 4 Calcular la correlacion entre variables de time y Death_Event ya que el time es el tiempo de seguimiento del tratamiento y el Death Event nos confirma si murió o no, dentro del tiempo de seguimiento correlation = data["time"].corr(data["DEATH_EVENT"]) print("Correlacion entre el Time y Death Event: ",correlation)</pre>
5. Separar los datos de entrenamiento y prueba. Seleccionar el porcentaje. ¿Qué otro detalle debe tener en cuenta? (10 puntos)	<p>Debemos de tener en cuenta el Random State, jugando con este valor he podido concluir que el randomState de 1 es el que brinda la mayor precisión al modelo</p> <p>Porcentaje 80/20</p>
6. Según el target, ¿Qué tipo de algoritmo utilizará y por qué? Justifique su elección (20 puntos)	<p>Utilizare el algoritmo de Regresión Logística porque es lineal y de entrenamiento rápido, adecuado para problemas de clasificación binarios (0-1)</p> <p>Utilizare RandomForest para buscar la Presicion del Modelo</p>
7. Realice la cross validation y presente la matriz de confusión (10 puntos)	<p>Hecho en el Jupyter</p>

DeathEvent target if == 1 murio en follow up period