

EVALUACIÓN III - 40%

Estudiante	
Fecha de Entrega	10-07-2024
Carrera	Ingeniería Civil Informática Magister en Ciencias de la Computación
Académicos	Natalia Morales Rosales - Cristian Martínez
Curso	Tópicos Avanzados en Manejo de Datos y Aprendizaje Automático

PUNTAJE MÁXIMO	30	PUNTAJE DE CORTE	18	PUNTAJE OBTENIDO		CALIFICACIÓN	
----------------	----	------------------	----	------------------	--	--------------	--

Resultados de Aprendizajes Evaluados:	Desarrollo y presentación de un estudio de Ciencia de Datos aplicado a un dataset de interés
Indicadores de Evaluación:	<ul style="list-style-type: none"> • Aplica técnicas de EDA para tener un mejor conocimiento de la BD de estudio • Aplica técnicas de preprocesamiento de datos para proveer de atributos de entrada de calidad a los modelos desarrollados. • Desarrolla modelos de Aprendizaje No Supervisado de manera de obtener conocimiento de los datos y para su uso en predicción de resultados. Los modelos presentan alta performance. • Desarrolla modelos de Aprendizaje Supervisado de manera de obtener conocimiento de los datos y para su uso en predicción de resultados. Los modelos presentan alta performance. • Expone de manera clara y precisa el estudio y conclusiones llevadas adelante. • Presenta un informe de nivel similar al exigido en un congreso nacional del área.

INSTRUCCIONES
<ul style="list-style-type: none"> • El presente trabajo puede ser desarrollado entre 2 estudiantes. El mismo involucra una presentación del dataset de estudio, informe, defensa oral presencial y entrega de código (notebook y PDF). • La presentación del dataset de estudio es hasta el día miércoles 26 de junio. • La presentación del informe y defensa es el día miércoles 10 de julio a las 18.30¹. El tiempo de exposición es de 20' (desarrollado por un estudiante) o 30' (si es desarrollado por dos estudiantes). Después del límite de tiempo, se descontarán 5 décimas por cada minuto extra. Si la

¹ De ser posible, se iniciará la ronda de presentaciones antes de las 18.30 horas. Esto será consensuado entre los estudiantes y la Cátedra.

presentación es de 2 personas, cada estudiante tiene un tiempo de 15 minutos aproximadamente.

- La plantilla para el informe está disponible en:
<https://www.overleaf.com/1388557611smvgtxcdbkdx#7d9810>
- **La evaluación es de carácter individual.**
- No se aceptarán informes entregados fuera de plazo. Si así fuera, la calificación será mínima.
- Todo el proyecto debe enviarse por correo electrónico a cmartinez@di.unsa.edu.ar y nmoralesr@ucm.cl.

Producto a entregar:

- Dataset de estudio
- Informe en formato PDF (máximo 30 hojas)
- Notebook de lo desarrollado y su respectivo PDF
- Diapositivas de la presentación

En caso de no enviar alguna de las entregas mencionadas (4) se descontará 1 punto.

Restricciones:

1. Se aplicará artículo 67º del reglamento del estudiante, el cual indica que, en caso de sorprender copia parcial o exacta, ya sea entre compañeros o reproducidos de algún medio, lo cual implica un 1,0 para todos los involucrados.
2. Si hubiese algún requerimiento sin contestar, el mismo no aplica puntaje.

DESCRIPCIÓN DE LA ACTIVIDAD Y DE LAS PREGUNTAS A SER RESUELTAS

El informe y su defensa tienen por objetivo abordar, profundizar y aplicar los contenidos de todas las unidades del Curso. El orden de actividades a seguir es:

1. Presentación del dataset: describir la fuente de datos elegida para una primera aproximación a la misma.
2. Resolución: presentar la resolución de cada actividad relevante referida a un proyecto de Ciencia de Datos. Puede incluir texto, tablas y gráficos.
3. Análisis de resultados: realice un análisis correspondiente a los efectos de poder evaluar la interpretación de resultados alcanzados.

Rúbrica

Indicadores	Valores				Obs.
	Destacado (5 puntos)	Competente (4 puntos)	Básico (3 puntos)	En desarrollo (1 punto)	
<p>EDA</p> <ul style="list-style-type: none"> - Utiliza gráficos variados y pertinentes (histogramas, diagramas de dispersión, gráficos de caja, entre otros) con etiquetas claras, títulos y leyendas informativas que mejoran la comprensión. - Realiza un análisis exhaustivo de las variables relevantes, abordando tendencia central, dispersión, relaciones entre variables, y gestionando correctamente valores atípicos y datos faltantes. - Proporciona interpretaciones claras y evidentes de los análisis y visualizaciones, estableciendo conexiones lógicas y precisas. - El análisis está bien estructurado con una clara introducción, desarrollo y conclusiones, usando subtítulos y anotaciones cuando corresponda. 	Cumple con 4 requerimientos	Cumple con 3 requerimientos.	Cumple con 2 requerimientos.	Cumple con 1 requerimiento.	
<p>Aprendizaje no Supervisado</p> <ul style="list-style-type: none"> - Exhibe una selección y aplicación experta de modelos No Supervisados, optimizados 	Cumple con 4 requerimientos	Cumple con 3 requerimientos.	Cumple con 2 requerimientos.	Cumple con 1 requerimiento.	

<p>para extraer conocimiento y mejorar la predicción.</p> <ul style="list-style-type: none"> - Proporciona interpretaciones detalladas y valiosas, mejorando la predicción y la toma de decisiones. - Aplica métodos de evaluación avanzados, alcanzando resultados superiores que resaltan en comparaciones importantes. - Usa estrategias continuas de mejora, adaptando modelos de forma efectiva para optimizar predicciones futuras. 					
<p>Aprendizaje Supervisado</p> <ul style="list-style-type: none"> - Selecciona y aplica modelos de Aprendizaje Supervisado adecuadamente, diseñados específicamente para maximizar la precisión en las predicciones. - Realiza análisis detallado de los resultados, proporcionando interpretaciones que mejoran directamente la calidad de las predicciones y la toma de decisiones estratégicas. - Usa métodos de evaluación actuales, alcanzando resultados destacados que superan los niveles habituales y comparaciones importantes. - Implementa mejoras en los modelos, asegurando adaptaciones y optimizaciones que refinan las predicciones a lo largo del tiempo. 	Cumple con 4 requerimientos	Cumple con 3 requerimientos.	Cumple con 2 requerimientos.	Cumple con 1 requerimiento.	
Informe	Cumple con 4	Cumple con 3	Cumple con 2	Cumple con 1	

<ul style="list-style-type: none"> - Aborda un proyecto de Ciencia de Datos sobre un dataset de complejidad media, relevante y de interés. - Aplica e incluye conveniente y apropiadamente técnicas EDA, pre-procesamiento de datos, de Aprendizaje No Supervisado y Supervisado según corresponda que den un valor agregado al proyecto. - La solución en Python (incluida en la documentación anexa) se ajusta a buenas prácticas de desarrollo. - Los resultados obtenidos por las diferentes técnicas y modelos son competitivos. 	requerimientos	requerimientos.	requerimientos.	requerimiento.	
<p>Presentación oral</p> <ul style="list-style-type: none"> - Explica las técnicas y modelos desarrollados de manera clara y precisa. - Realiza un análisis correcto de los resultados alcanzados. - Responde a las consultas realizadas por la Cátedra. - No sobrepasa el límite de tiempo de la presentación 	Cumple con 4 requerimientos	Cumple con 3 requerimientos.	Cumple con 2 requerimientos.	Cumple con 1 requerimiento.	

<p>Generalidades</p> <ul style="list-style-type: none"> - Solución (informe y exposición) ordenada y legible. - Desarrollo de acuerdo a la solicitud de las instrucciones. - Cumple con los entregables (dataset de estudio, informe, notebook y PDF, presentación) en tiempo y forma. - Buen uso del lenguaje y ortografía. 	Cumple con 4 requerimientos	Cumple con 3 requerimientos.	Cumple con 2 requerimientos.	Cumple con 1 requerimiento.	
---	-----------------------------	------------------------------	------------------------------	-----------------------------	--

Tabla 1: Rúbrica

Escala de Notas

Puntaje	Nota	Puntaje	Nota	Puntaje	Nota	Puntaje	Nota
0.0	1.0	10.0	2.7	20.0	4.5	30.0	7.0
1.0	1.2	11.0	2.8	21.0	4.8		
2.0	1.3	12.0	3.0	22.0	5.0		
3.0	1.5	13.0	3.2	23.0	5.3		
4.0	1.7	14.0	3.3	24.0	5.5		
5.0	1.8	15.0	3.5	25.0	5.8		
6.0	2.0	16.0	3.7	26.0	6.0		
7.0	2.2	17.0	3.8	27.0	6.3		
8.0	2.3	18.0	4.0	28.0	6.5		
9.0	2.5	19.0	4.3	29.0	6.8		

Ilustración 1: Escala de Notas