

Evaluación Final Transversal

Encargo grupal y Presentación individual - Estudiante

Sigla	Nombre Asignatura	Tiempo Asignado	% Ponderación
MLY0100	MACHINE LEARNING	4 horas	40%

1. Situación Evaluativa 1:

	Ejecución práctica
--	--------------------

X	Entrega de encargo
---	--------------------

	Presentación
--	--------------

2. Situación Evaluativa 2:

	Ejecución práctica
--	--------------------

	Entrega de encargo
--	--------------------

X	Presentación
---	--------------

3. Instrucciones generales Evaluación Final Transversal

Descripción

- El tiempo para desarrollar esta evaluación es durante el semestre, por lo que desde la primera experiencia de aprendizaje deberá integrar elementos para rendir el examen. Luego deberá realizar la entrega del encargo en la semana 17, el cual debe ser presentado junto a su grupo en la semana 18.
- Para este examen, deben mantenerse los mismos grupos de estudiantes que trabajaron durante el semestre, con un máximo de 3 integrantes, quienes deberán exponer el desarrollo del caso, mostrar los insight asociados al contexto abordando CRISP-DM en todas sus fases, incorporando el feedback que entregó su docente durante las entregas parciales.
- La distribución de los porcentajes en esta evaluación es la siguiente:

Evaluación	Tipo de situación evaluativa	Distribución de porcentajes en el ET
Evaluación Transversal (ET) Final	Entrega por encargo grupal	40%
	Presentación individual	60%
Total		100%

- La evaluación consiste en:
 - Informe técnico realizado en Jupyter Notebook con Python sobre todas las fases de la metodología más el feedback entregado por su docente.
 - Presentación del caso (10 minutos de duración, más 5 minutos de preguntas)
- El examen final no es solo una evaluación, sino una celebración de la excelencia y el esfuerzo constante que han demostrado a lo largo de su aprendizaje. Es una oportunidad para mostrar su autonomía y la habilidad de enfrentar desafíos profesionales con confianza. Cada uno de ustedes ha recorrido un camino único de crecimiento y este examen es el momento de brillar con todo lo que han adquirido. ¡Adelante, demuestren lo que son capaces de lograr!

Instrucciones específicas de la Evaluación:

Situación evaluativa 1: Entrega por encargo grupal

En el informe, formato Jupyter Notebook, los estudiantes deben:

1. Identificar dentro del notebook todas las fases de la metodología CRISP-DM.
2. Identificar valores nulos y valores atípicos dentro del conjunto de datos.
3. Aplicar tratamiento de los valores nulos y atípicos identificados.
4. Aplicar técnicas de transformación de los datos para desarrollar los modelos predictivos.
5. Aplicar algoritmos de regresión para construir modelos predictivos.
6. Seleccionar el mejor modelo de regresión en base a las métricas de cada modelo de regresión.
7. Aplicar algoritmos de clasificación para construir modelos predictivos.
8. Aplicar técnicas de balance de clases para mejorar la generalización en el desempeño de los modelos de clasificación.
9. Analizar las métricas obtenidas de los modelos de clasificación.
10. Aplicar algoritmos de aprendizaje no supervisado.
11. Utilizar técnicas para la selección del número óptimo de clusters.
12. Relacionar los resultados obtenidos en el número de clusters con la naturaleza de los datos
13. El notebook debe incluir celdas markdown donde se incluya la justificación de los pasos realizados y de las técnicas aplicadas durante el proceso

Situación evaluativa 2: Presentación individual

Durante la presentación cada estudiante debe:

1. Identificar dentro de los datos valores atípicos y valores nulos
2. Reconocer las diferencias entre una tarea de clasificación de una de regresión
3. Analizar las métricas obtenidas de los modelos de clasificación
4. Relacionar los resultados obtenidos en el número de clusters con la naturaleza de los datos
5. Responder a las preguntas realizadas por el docente.

Los materiales, herramientas o insumos que se requieren para realizar esta evaluación son: Taller de alto cómputo

5. Pauta de Evaluación

Categoría	% logro	Descripción niveles de logro
Muy buen desempeño	100%	Demuestra un desempeño destacado, evidenciando el logro de todos los aspectos evaluados en el indicador.
Buen desempeño	80%	Demuestra un alto desempeño del indicador, presentando pequeñas omisiones, dificultades y/o errores.
Desempeño aceptable	60%	Demuestra un desempeño competente, evidenciando el logro de los elementos básicos del indicador, pero con omisiones, dificultades o errores.
Desempeño incipiente	30%	Presenta importantes omisiones, dificultades o errores en el desempeño, que no permiten evidenciar los elementos básicos del logro del indicador, por lo que no puede ser considerado competente.
Desempeño no logrado	0%	Presenta ausencia o incorrecto desempeño.

Indicador de Evaluación	Categorías de Respuesta					Ponderación Indicador de Evaluación
	Muy buen desempeño 100%	Buen desempeño 80%	Desempeño aceptable 60%	Desempeño incipiente 30%	Desempeño no logrado 0%	
SITUACIÓN EVALUATIVA 1: ENCARGO. INFORME CON EVALUACIÓN GRUPAL.						
1. Utiliza dentro del formato Jupyter Notebook y durante el proceso las fases de la metodología CRISP-DM como un estándar para abordar el proyecto.	Utiliza dentro del notebook las 5 fases de la metodología	Utiliza dentro del notebook sólo 4 fases de la metodología.	Utiliza dentro del notebook sólo 3 fases de la metodología.	Utiliza dentro del notebook sólo 2 fases de la metodología	Utiliza dentro del notebook sólo una fase de la metodología	5%
2. Selecciona el mejor modelo de regresión basándose en los resultados de las métricas y en el desempeño de cada uno de estos.	Selecciona el mejor modelo de regresión considerando al menos r2, score y MSE.	Selecciona el mejor modelo de regresión considerando al menos 2 métricas de desempeño.	Selecciona el mejor modelo de regresión considerando solo una métrica de desempeño.	Selecciona el mejor modelo utilizando métricas erróneas, no asociadas a la tarea de regresión.	No selecciona el mejor modelo, o no utiliza métricas de desempeño.	8%

3. Justifica en Markdown dentro del notebook los pasos realizados y cada técnica utilizada durante el proceso.	Justifica en Markdown todas las transformaciones aplicadas en cada fase de la metodología.	Justifica en Markdown más del 80% y menos de la totalidad de las transformaciones aplicadas en las fases de la metodología.	Justifica en Markdown entre el 60% y el 79% de la totalidad de las transformaciones aplicadas en las fases de la metodología.	Justifica en Markdown entre el 30% y el 59% de la totalidad de las transformaciones aplicadas en las fases de la metodología.	Justifica en Markdown menos del 30% de la totalidad de las transformaciones aplicadas en las fases de la metodología.	5%
--	--	---	---	---	---	----

4. Identifica dentro de los datos, valores atípicos y missing values, aplicando tratamientos para estos valores de acuerdo con los estándares de la industria.	Identifica dentro de los datos, valores atípicos y missing values, aplicando tratamientos para estos valores de acuerdo con los estándares de la industria.	Identifica dentro de los datos sólo valores atípicos o missing values aplicando tratamientos para estos valores de acuerdo con los estándares de la industria.	Identifica dentro de los datos y valores atípicos o missing values SIN aplicar tratamientos para estos valores de acuerdo con los estándares de la industria.	Identifica dentro de los datos o valores atípicos o missing values SIN aplicar tratamientos para estos valores de acuerdo con los estándares de la industria.	No Identifica ni aplica tratamientos para valores atípicos ni para missing values.	8%
5. Realiza limpieza y preparación de datos mediante el uso de lenguaje Python sobre Jupyter notebook de acuerdo con las buenas prácticas de la industria.	Realiza la totalidad de las rutinas de limpieza y preparación de datos mediante el uso de lenguaje Python sobre Jupyter notebook de acuerdo con las buenas prácticas de la industria.	Realiza entre el 80% y el 99% de las rutinas de limpieza y preparación de datos mediante el uso de lenguaje Python sobre Jupyter notebook de acuerdo con las buenas prácticas de la industria.	Realiza entre el 60% y el 79% de las rutinas de limpieza y preparación de datos mediante el uso de lenguaje Python sobre Jupyter notebook de acuerdo con las buenas prácticas de la industria.	Realiza entre el 30% y el 59% de las rutinas de limpieza y preparación de datos mediante el uso de lenguaje Python sobre Jupyter notebook de acuerdo con las buenas prácticas de la industria.	Realiza menos del 30% de las rutinas de limpieza y preparación de datos mediante el uso de lenguaje Python sobre Jupyter notebook de acuerdo con las buenas prácticas de la industria.	8%

6. Reconoce las diferencias entre una tarea de clasificación de una regresión, considerando target continuo para regresión y discreto para clasificación .	Reconoce un target continuo para regresión y uno discreto para clasificación.	NO APLICA	Reconoce un target correcto sólo para regresión o para clasificación.	NO APLICA	No reconoce correctamente el target para tareas de regresión ni para clasificación.	8%
--	---	-----------	---	-----------	---	----

7. Transforma los atributos necesarios para desarrollar un buen modelo de clasificación.	Transforma, correctamente, la totalidad de los atributos necesarios para desarrollar un buen modelo de clasificación.	Transforma, correctamente, entre el 80% y el 99% de los atributos necesarios para desarrollar un buen modelo de clasificación.	Transforma, correctamente, entre el 60% y el 79% de los atributos necesarios para desarrollar un buen modelo de clasificación.	Transforma, correctamente, entre el 30% y el 59% de los atributos necesarios para desarrollar un buen modelo de clasificación.	Transforma, correctamente, menos del 30% de los atributos necesarios para desarrollar un buen modelo de clasificación.	10%
8. Aplica algoritmos de clasificación para construir modelos supervisados según los requerimientos del caso.	Aplica, correctamente, al menos 2 algoritmos de clasificación para construir modelos supervisados según los requerimientos del caso.	NO APLICA	Aplica, correctamente, sólo un algoritmo de clasificación para construir modelos supervisados según los requerimientos del caso.	NO APLICA	No aplica algoritmos de clasificación para construir modelos supervisados según los requerimientos del caso.	8%
9. Aplica técnicas de balance de clases para mejorar la generalización en el desempeño de los modelos de clasificación.	Aplica, correctamente, técnicas de balance de clases para mejorar la generalización en el desempeño de los modelos de clasificación.		Aplica, de forma parcial, técnicas de balance de clases para mejorar la generalización en el desempeño de los modelos de clasificación.		No aplica técnicas de balance de clases para mejorar la generalización en el desempeño de los modelos de clasificación.	8%

10. Analiza las métricas obtenidas de los modelos de clasificación desarrollados , seleccionando el más adecuado dependiendo del caso.	Analiza, correctamente, la totalidad de las métricas obtenidas de los modelos de clasificación desarrollados, seleccionando el más adecuado dependiendo del caso.	Analiza, correctamente, entre el 80% y el 99% de las métricas obtenidas de los modelos de clasificación desarrollados, seleccionando el más adecuado dependiendo del caso.	Analiza, correctamente, entre el 60% y el 79% de las métricas obtenidas de los modelos de clasificación desarrollados, seleccionando el más adecuado dependiendo del caso.	Analiza, correctamente, entre el 30% y el 59% de las métricas obtenidas de los modelos de clasificación desarrollados, seleccionando el más adecuado dependiendo del caso.	Analiza, correctamente, menos del 30% de las métricas obtenidas de los modelos de clasificación desarrollados, seleccionando el más adecuado dependiendo del caso.	8%
11. Construye modelos de aprendizaje no supervisado utilizando algoritmos de segmentación según corresponda.	Construye, correctamente, al menos 2 modelos de aprendizaje no supervisado utilizando algoritmos de segmentación según corresponda.		Construye, correctamente, sólo un modelo de aprendizaje no supervisado utilizando algoritmos de segmentación según corresponda.		No construye modelos de aprendizaje no supervisado utilizando algoritmos de segmentación según corresponda.	8%
12. Utiliza técnicas como Elbow o Silhouette que ayudan a seleccionar la cantidad óptima de clusters	Utiliza, correctamente, técnicas como Elbow o Silhouette que ayudan a seleccionar la cantidad óptima de clusters		Utiliza, de forma parcialmente correcta, técnicas como Elbow o Silhouette que ayudan a seleccionar la cantidad óptima de clusters		No utiliza técnicas como Elbow o Silhouette que ayudan a seleccionar la cantidad óptima de clusters	8%

13. Relaciona los resultados obtenidos en número de clusters con la naturaleza de los datos, métricas de evaluación de aprendizaje no supervisado y el contexto	Relaciona, correctamente, los resultados obtenidos en número de clusters con la naturaleza de los datos, métricas de evaluación de aprendizaje no supervisado y el contexto		Relaciona, de forma parcialmente correcta, los resultados obtenidos en número de clusters con la naturaleza de los datos, métricas de evaluación de aprendizaje no supervisado y el contexto		No relaciona los resultados obtenidos en número de clusters con la naturaleza de los datos, métricas de evaluación de aprendizaje no supervisado y el contexto	8%
Total, Situación evaluativa 1						100%