

Evaluación Parcial 4

Modelos de Clasificación

Sigla	Nombre Asignatura	Tiempo Asignado	% Ponderación
FMY0100	Fundamentos de Machine Learning	2 semanas	25%

01. Agente evaluativo

<input checked="" type="checkbox"/>	Heteroevaluación	<input type="checkbox"/>	Coevaluación	<input type="checkbox"/>	Autoevaluación
-------------------------------------	------------------	--------------------------	--------------	--------------------------	----------------

02. Tabla de Especificaciones

Resultado de Aprendizaje	Indicador de Logro (IL)	Indicador de Evaluación (IE)*	Ponderación Indicador Logro	Ponderación Indicador de Evaluación
RA3 Crea modelos de Machine Learning de aprendizaje supervisado para problemas de regresión considerando el tipo de tarea y la naturaleza de los datos	IL3.1 Comprende las bases de Machine Learning, sus metodologías, herramientas y técnicas asociadas a las diferentes formas de aprendizaje automático.	3.1.1 Sigue a la estructura del proyecto basándose en la metodología CRISP-DM utilizando el formato dispuesto para la evaluación.	30%	10%
		3.1.2 Identifica oportunidades de mejora en el proyecto entre entregas, moviéndose entre fases de la metodología con el fin de profundizar de mejor manera en la problemática.		10%
		3.1.3 Utiliza el formato de Jupyter Notebook, aprovechando cuadros de código, markdown, títulos e índices cuando corresponde.		10%
	IL4.1 Comprende los problemas de clasificación y cómo obtener una	4.1.1 Reconoce las características que tiene un problema de clasificación,	20%	10%

RA4 Crea modelos de Machine Learning de aprendizaje supervisado o no supervisado para problemas de clasificación basándose en el tipo de tarea considerando la naturaleza de los datos y el caso planteado.	predicción mediante el uso de herramientas de Python.	eligiendo un target categórico para el entrenamiento del modelo de ML. 4.1.2 Utiliza las librerías de Scikit learn, pandas y numpy para el desarrollo del modelo de ML en la tarea de clasificación.		
	IL4.2 Aplica un modelo de clasificación a un problema de aprendizaje supervisado o no supervisado para obtener una predicción de nuevos datos.	4.2.1 Realiza la división del dataset en algún porcentaje recomendado por las buenas prácticas de la industria para entrenamiento y prueba.	20%	10%
		4.2.2 Entrena al menos 3 alternativas de modelos usando algoritmos de clasificación y seleccionando cual es el que se adapta mejor a los datos dada la naturaleza del algoritmo implementado.		10%
	IL4.3 Evalúa el modelo de clasificación aplicado al problema de aprendizaje supervisado o no supervisado, a partir del conocimiento de distintas métricas.	4.3.1 Utiliza las métricas y las interpreta correctamente basándose en el tipo de tarea desarrollada, considerando las métricas derivadas de la matriz de confusión.	20%	10%
		4.3.2 Justifica cual es el mejor modelo entrenado de acuerdo con el resultado obtenido, considerando el problema de clasificación con las métricas correspondientes.		10%
	IL4.4 Implementa modelos de clasificación dentro de un entorno interactivo mediante el uso de Python en la etapa de despliegue de solución.	4.1.1 Aplica el modelo predictivo desarrollado en un entorno interactivo como un formulario, permitiendo el ingreso de un nuevo registro a predecir.	10%	10%
	Total		100%	100%

03. Instrucciones para el/la estudiante

Esta es una evaluación sumativa que corresponde a la Evaluación Parcial 4. Consiste en una entrega parcial del caso semestral entregado, que equivale al **25%** de ponderación sobre la nota final de la asignatura.

En esta entrega debes considerar lo necesario basado en la metodología CRISP-DM a las 4° y 5° fase: Modeling (tarea de clasificación), Evaluation y Deployment.

Recuerda que estas fases no son completamente secuenciales, y que puedes iterar y volver a pasos anteriores, hasta tener objetivos definidos del negocio basado en los datos.

El **tiempo** para desarrollar esta evaluación es de 2 semanas y se realiza de manera **grupal** de acuerdo con los equipos definidos.

La evaluación consiste en:

- **Modeling:** Implementar en el formato habilitado, al menos 3 modelos predictivos de Machine Learning en una tarea de clasificación, analizar si los pasos realizados en etapas anteriores fueron suficientes para desarrollar el objetivo, identificar oportunidades de mejora en fases anteriores, y aplicarlas con el objetivo de tener mejores resultados en la etapa de evaluation.
- **Evaluation:** Analizar los modelos desarrollados basándose en métricas de desempeño para tareas de regresión (al menos accuracy, precisión, ROC, F1 score y confusion matrix). Justificar cual es el mejor modelo desarrollado, con qué combinación de features y por qué.
- **Deployment:** Implementar el modelo desarrollado en algún entorno que permita el ingreso de nuevos registros y que el modelo de acuerdo con el proceso de entrenamiento sea capaz de entregar predicciones.

Pauta de Evaluación

Pauta tipo: Escala de valoración

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.
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 del Indicador de Evaluación
	Muy buen desempeño 100%	Desempeño aceptable 60%	Desempeño incipiente 30%	Desempeño no logrado 0%	
Sigue a la estructura del proyecto basándose en la metodología CRISP-DM utilizando el formato dispuesto para la evaluación.	10 pts.	6,0 pts.	3,0 pts.	0 pts.	10%
Identifica oportunidades de mejora en el proyecto entre entregas, moviéndose entre fases de la metodología con el fin de profundizar de mejor manera en la problemática.	10 pts.	6,0 pts.	3,0 pts.	0 pts.	10%
Utiliza el formato de Jupyter Notebook, aprovechando cuadros de código, markdown, títulos e índices cuando corresponde.	10 pts.	6,0 pts.	3,0 pts.	0 pts.	10%
Reconoce las características que tiene un problema de clasificación, eligiendo un target categórico para el entrenamiento del modelo de ML.	10 pts.	6,0 pts.	3,0 pts.	0 pts.	10%
Utiliza las librerías de Scikit learn, pandas y numpy para el desarrollo del modelo de ML en la tarea de clasificación.	10 pts.	6,0 pts.	3,0 pts.	0 pts.	10%

Realiza la división del dataset en algún porcentaje recomendado por las buenas prácticas de la industria para entrenamiento y prueba.	10 pts.	6,0 pts.	3,0 pts.	0 pts.	10%
Entrena al menos 3 alternativas de modelos usando algoritmos de clasificación y seleccionando cual es el que se adapta mejor a los datos dada la naturaleza del algoritmo implementado.	10 pts.	6,0 pts.	3,0 pts.	0 pts.	10%
Utiliza las métricas y las interpreta correctamente basándose en el tipo de tarea desarrollada, considerando las métricas derivadas de la matriz de confusión.	10 pts.	6,0 pts.	3,0 pts.	0 pts.	10%
Justifica cual es el mejor modelo entrenado de acuerdo con el resultado obtenido, considerando el problema de clasificación con las métricas correspondientes.	10 pts.	6,0 pts.	3,0 pts.	0 pts.	10%
Aplica el modelo predictivo desarrollado en un entorno interactivo como un formulario, permitiendo el ingreso de un nuevo registro a predecir.	10 pts.	6,0 pts.	3,0 pts.	0 pts.	10%
Total					100%