

Predicción de quiebra de empresas mediante técnicas de Machine Learning

Autor:

Ing. Gaspar Acevedo Zain

Director:

Título y Nombre del director (pertenencia)

${\rm \acute{I}ndice}$

1. Descripción técnica-conceptual del proyecto a realizar.	•	•	•	 •	•	•	•	•	•	•	•	5
2. Identificación y análisis de los interesados												6
3. Propósito del proyecto												6
4. Alcance del proyecto												6
5. Supuestos del proyecto												7
6. Product Backlog								•				7
7. Criterios de aceptación de historias de usuario												11
8. Fases de CRISP-DM												12
9. Desglose del trabajo en tareas					•							13
10. Planificación de Sprints												20
11. Diagrama de Gantt (sprints)					•						. :	33
12. Normativa y cumplimiento de datos (gobernanza)											. :	38
13. Gestión de riesgos								•			. ;	38
14. Sprint Review											. ;	39
15. Sprint Retrospective				 _		_		_		_		40



Registros de cambios

Revisión	Detalles de los cambios realizados	Fecha
0	Creación del documento	24 de junio de 2025
1	Se completa hasta el punto 5 inclusive	6 de Julio de 2025
2	Se completa hasta el punto 9 inclusive	15 de Julio de 2025
3	Se completa hasta el punto 12 inclusive	29 de Julio de 2025



Acta de constitución del proyecto

Buenos Aires, 24 de junio de 2025

Por medio de la presente se acuerda con el Ing. Gaspar Acevedo Zain que su Trabajo Final de la Carrera de Especialización en Inteligencia Artificial se titulará "Predicción de quiebra de empresas mediante técnicas de Machine Learning" y consistirá en el desarrollo de una herramienta basada en Machine Learning que permitirá predecir si una empresa puede entrar en quiebra o no. El trabajo tendrá un presupuesto preliminar estimado de 604 horas y un costo estimado de \$XXX, con fecha de inicio el 24 de junio de 2025 y fecha de presentación pública el a definir.

Se adjunta a esta acta la planificación inicial.

Dr. Ing. Ariel Lutenberg Director posgrado FIUBA Nombre del cliente Empresa del cliente

Título y Nombre del director Director del Trabajo Final



1. Descripción técnica-conceptual del proyecto a realizar

Este proyecto consiste en un emprendimiento personal cuyo objetivo es utilizar técnicas de aprendizaje de máquina para detectar si una empresa puede entrar en quiebra o no. Este tipo de análisis puede resultar de gran interés y utilidad para distintos actores del mercado financiero, tales como bancos, compañías aseguradoras, fondos de inversión o consultoras especializadas en riesgo crediticio. Por ello, estos se considerarán como potenciales clientes.

Para llevarlo a cabo, se utilizará un dataset publicado por el Taiwan Economic Journal, que contiene información financiera de empresas del mercado de Taiwán entre los años 1999 y 2009. Al ser estos datos públicos, hoy en día existen soluciones que exploran esta temática. Algunas de ellas hacen uso de modelos de machine learning tales como SVM y XGBoost, junto con algunas técnicas de preprocesamiento de datos como Smote y de búsqueda de hiperparámetros como Random Search.

Con el fin de diferenciarse de estas soluciones, se propone implementar el marco de trabajo basado en *MLFlow* definido en la figura 1. Se detalla una serie de etapas cuyas salidas se refinarán durante distintas iteraciones. Esto permitirá a los usuarios finales trabajar en un entorno seguro, robusto, y reproducible.

El proyecto se encuentra en la etapa de planificación. El desarrollo e implementación se realizará en distintas etapas. Se comenzará con un análisis exploratorio de datos, que nos permitirá conocer mejor al dataset en cuestión. Luego, se realizarán iteraciones sobre las siguientes etapas:

- Preprocesamiento de datos: en la primer iteración se implementarán técnicas de tratamiento de nulos y desbalance de clases. En las siguientes iteraciones, se estudiarán técnicas de extracción e ingeniería de features.
- Entrenamiento de modelos: se implementará un modelo distinto en cada iteración. Los modelos a explorar son regresión logística, SVM y XGBoost. También, se explorará la optimización de hiperparámetros mediante búsqueda bayesiana.
- Evaluación y refinamiento: en esta etapa se evaluará al modelo entrenado en la etapa anterior. Se generarán métricas que permitirán compararlo con resultados obtenidos en otras iteraciones.

La innovación de este proyecto radica en el uso del marco de trabajo definido en la figura 1. Éste proporciona un ambiente productivo, reproducible y escalable, en donde se podrán analizar diversas técnicas de aprendizaje de máquina para detectar si una empresa puede entrar en quiebra o no.



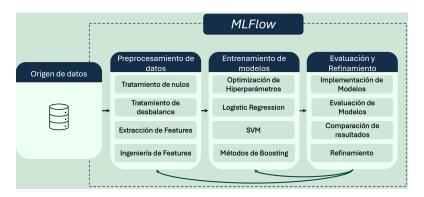


Figura 1. Diagrama en bloques del sistema.

2. Identificación y análisis de los interesados

Rol	Nombre y Apellido	Organización	Puesto
Responsable	Ing. Gaspar Acevedo	FIUBA	Alumno
	Zain		
Orientador	Título y Nombre del	pertenencia	Director del Trabajo Final
	director		
Cliente	Actores del mercado	-	-
	financiero		
Usuario final	Trabajadores de clien-	-	-
	tes		

- Orientador: podrán ayudar en la recomendación y evaluación de técnicas a explorar en las diferentes etapas del proyecto.
- Cliente: si bien es un proyecto personal, se considerarán como potenciales clientes a distintos actores del mercado financiero, tales como bancos, compañías aseguradoras, fondos de inversión o consultoras especializadas en riesgo crediticio.
- Usuario final: analistas de riesgos, ejecutivo de créditos, entre otros integrantes que trabajan para los potenciales clientes.

3. Propósito del proyecto

Predecir si una empresa puede entrar en quiebra o no, al explorar técnicas de *machine learning* en un marco de trabajo productivo, reproducible y escalable.

4. Alcance del proyecto

El alcance del proyecto incluye:

■ Análisis exploratorio de datos: se analizarán las distintas variables presentes en el *dataset* de estudio, con el fin de conocer sus características y poder tomar decisiones con base en ellas.



- Preprocesamiento de datos: se realizarán técnicas de tratamiento de datos faltantes, selección y/o extracción de variables, como así también de ingeniería de features.
- Implementación de modelos de machine learning: se estudiarán diversos modelos de aprendizaje de máquina sobre los datos procesados, tales como logistic regression, SVM y XGBoost. Además, se optimizarán los hiperparámetros de estos modelos mediante búsqueda bayesiana.
- Evaluación y comparación de modelos: se obtendrán métricas relacionadas a los modelos explorados, con el fin de poder determinar cuál de ellos realiza una mejor predicción.
- Implementación de un entorno basado en *MLFlow*: este entornó facilitará la realización, la reproducibilidad y la escalabilidad de las distintas etapas de trabajo que se realizarán en este proyecto. Este será de caracter local, es decir, no se implementará en una plataforma de *cloud computing*.

No se incluye:

- El despliegue del entorno de trabajo en una plataforma de *cloud computing*, tales como *Azure*, *AWS*, entre otros.
- El análisis de otros datasets distintos al propuesto.

5. Supuestos del proyecto

Para el desarrollo del presente proyecto se supone que:

- Supuesto 1: el *dataset* de estudio presenta datos fiables, y no tiene restricciones en cuanto a licencias de uso.
- Supuesto 2: una *laptop* como equipo de trabajo es más que suficiente para realizar el preprocesamiento y entrenamiento de los modelos de aprendizaje automático.
- Supuesto 3: el entorno de *MLFlow* podrá desarrollarse en etapas futuras del proyecto, posteriores a la exploración de los modelos de aprendizaje automático.
- Supuesto 4: el entorno de *MLFlow* podrá desplegarse de manera local, sin necesidad de recurrir a plataforma de *cloud computing*, tales como *Azure*, *AWS*, entre otros.
- Supuesto 5: se disponen de al menos 15 horas semanales para realizar el proyecto.

6. Product Backlog

Roles

- Ingeniero del proyecto: es quien se encarga del análisis, diseño, desarrollo y despliegue del proyecto.
- *Usuario final*: es quien consulta y analiza las predicciones de los modelos explorados en el proyecto.



Criterios de ponderación de historias de usuario

Esto son los criterios que se utilizan para ponderar a las historias de usuario mediante *Story Points*:

- Dificultad: representa la cantidad de trabajo estimado que requiere la historia de usuario para realizarse.
- Complejidad: representa la dificultad de realizar la historia de usuario a nivel técnico.
- Incertidumbre: representa el riesgo asociado a la historia de usuario.

Cada criterio tiene asociado las ponderaciones baja, media y alta, que se detallan en el cuadro 1. Los Story Points de una historia de usuario quedan definidos por la suma de los valores de estas ponderaciones redondeada hacia el número superior más próximo en la serie de Fibonacci.

Criterio\Ponderación	Baja	Media	Alta
Dificultad	1	3	5
Complejidad	1	3	5
Incertidumbre	1	5	8

Cuadro 1. Tabla de ponderaciones de historia de usuario.

Épicas

• Épica 1 - Análisis y procesamiento de datos

- HU1 Análisis exploratorio
 - Como ingeniero del proyecto, quiero realizar un análisis exploratorio de datos para conocer las distribuciones, formas y otras particularidades de las variables del dataset con el que se trabajará.
 - Ponderación
 - ♦ Dificultad: media 3 Story Points
 - ♦ Complejidad: baja 1 Story Points
 - ♦ Incertidumbre: baja 1 Story Points
 - ♦ Suma: 5
 - ♦ Total: 5 Story Points
- HU2 Procesamiento de datos faltantes y datos atípicos
 - o Como ingeniero del proyecto, quiero realizar un procesamiento de datos faltantes y de datos atípicos con el fin de asegurar la calidad del dataset.
 - $\circ\,$ Ponderación
 - ♦ Dificultad: media 3 Story Points
 - ♦ Complejidad: media 3 Story Points
 - ♦ Incertidumbre: baja 1 Story Points
 - ♦ Suma: 7
 - ♦ Total: 8 Story Points



• HU3 - Feature Engineering

- o Como ingeniero del proyecto, quiero implementar *Feature Engineering* con el fin de crear nuevos atributos en el dataset.
- Ponderación
 - ♦ Dificultad: media 3 Story Points
 - ♦ Complejidad: media 3 Story Points
 - ♦ Incertidumbre: baja 1 Story Points
 - ♦ Suma: 7
 - ♦ Total: 8 Story Points

• Épica 2 - Implementación y comparación de modelos

- HU4 Implementación de modelos de Machine Learning
 - o Como ingeniero del proyecto, quiero implementar los modelos de *Machine Learning* de *Logistic Regression*, *SVM* y *XGBoost* que permitan predecir si una empresa entra en quiebra o no.
 - Ponderación
 - \diamond Dificultad: media 3 Story Points
 - ♦ Complejidad: media 3 Story Points
 - \diamond Incertidumbre: media 5 Story Points
 - ♦ Suma: 11
 - ♦ Total: 13 Story Points
- HU5 Optimización de hiperparámetros
 - o Como ingeniero del proyecto, quiero implementar técnicas de optimización de hiperparámetros y aplicarlas a los modelos de *Machine Learning* implementados.
 - Ponderación
 - ♦ Dificultad: media 3 Story Points
 - ♦ Complejidad: media 3 Story Points
 - \diamond Incertidumbre: media 5 Story Points
 - ♦ Suma: 11
 - ♦ Total: 13 Story Points
- HU6 Métricas de modelos
 - \circ Como ingeniero del proyecto, quiero calcular las métricas de AUC-ROC y F1-score en cada modelo de Machine Learning implementado y comparar sus resultados.
 - Ponderación
 - $\diamond\,$ Dificultad: media 3 $Story\,Points$
 - ♦ Complejidad: media 3 Story Points
 - ♦ Incertidumbre: baja 1 Story Points
 - ♦ Suma: 7
 - ♦ Total: 8 Story Points



• Épica 3 - Despliegue en entorno *MLFlow*

- \bullet HU7 Despliegue en MLFlow
 - Como ingeniero del proyecto, quiero desplegar un entorno local de MLFLow en donde se repliquen los pasos de procesamiento de datos e implementación y comparación de modelos.
 - o Ponderación
 - ♦ Dificultad: alta 5 Story Points
 - \diamond Complejidad: media 3 Story Points
 - ♦ Incertidumbre: media 5 Story Points
 - ♦ Suma: 13
 - ♦ Total: 13 Story Points
- \bullet HU8 API para entorno MLFlow
 - \circ Como ingeniero del proyecto, quiero exponer el entorno de MLFlow mediante una API para facilitar el acceso y su utilización.
 - Ponderación
 - \diamond Dificultad: baja 1 Story Points
 - ♦ Complejidad: media 3 Story Points
 - ♦ Incertidumbre: baja 1 Story Points
 - ♦ Suma: 5
 - \diamond Total: 5 Story Points

• Épica 4 - Gestión de calidad del código fuente

- HU9 Implementación de buenas prácticas
 - Como ingeniero del proyecto, quiero asegurar que el código siga las buenas prácticas y estándares de la industria.
 - Ponderación
 - \diamond Dificultad: media 3 Story Points
 - ♦ Complejidad: baja 1 Story Points
 - ♦ Incertidumbre: baja 1 Story Points
 - ♦ Suma: 5
 - ♦ Total: 5 Story Points
- HU10 Documentación
 - o Como ingeniero del proyecto, quiero documentar todos los pasos realizados durante el proyecto.
 - o Ponderación
 - ♦ Dificultad: baja 1 Story Points
 - ♦ Complejidad: baja 1 Story Points
 - ♦ Incertidumbre: baja 1 Story Points
 - ♦ Suma: 3
 - ♦ Total: 3 Story Points



• HU11 - Validación de API de MLFlow

- Como usuario final, quiero consultar los resultados y comparaciones de los modelos mediante la API del entorno de MLFlow, para poder analizarlos.
- Ponderación
 - ♦ Dificultad: media 3 Story Points
 - ♦ Complejidad: baja 1 Story Points
 - ♦ Incertidumbre: baja 1 Story Points
 - ♦ Suma: 5
 - ♦ Total: 5 Story Points

7. Criterios de aceptación de historias de usuario

Épica 1 - Análisis y procesamiento de datos

- Criterios de aceptación HU1 Análisis exploratorio
 - o Se estudia la presencia de datos atípicos y de datos faltantes para cada variable.
 - o Se grafican las distribuciones de las variables del dataset.
 - o Se realiza un estudio de correlaciones entre variables numéricas.
 - o Se documentan los hallazgos del análisis de cada variable.
- Criterios de aceptación HU2 Procesamiento de datos faltantes y datos atípicos
 - o Se realiza una imputación de datos faltantes a las variables del dataset.
 - o Se justifican los métodos de imputación utilizados.
 - o Se ajustan los datos atípicos de las variables del dataset.
 - $\circ\,$ Se justifican los métodos de ajuste utilizados.
 - Se justifican los casos en donde se decide no imputar ni ajustar.
- Criterios de aceptación HU3 Feature Engineering
 - o Se crean nuevas variables en el dataset a partir de las existentes.
 - Se estudia el impacto por separado de estas variables en los modelos generados, a partir de sus métricas.
 - o Se justifica la inclusión o no en el modelo de cada variable generada.

Épica 2 - Implementación y comparación de modelos

- Criterios de aceptación HU4 Implementación de modelos de Machine Learning
 - Se implementan distintos modelos de Machine Learning.
 - o Se justifica el uso de cada uno de los modelos implementados.
 - Se persisten los modelos generados en GitHub, para futuros análisis y comparaciones.
- Criterios de aceptación HU5 Optimización de hiperparámetros
 - o Se seleccionan los hiperparámetros de cada modelo a optimizar.
 - o Se define el rango sobre el que se optimizará cada hiperparámetro.
 - Se realiza una búsqueda del valor óptimo de los hiperparámetros en los rangos definidos.
 - o Se justifican las decisiones tomadas en cada paso.



- Criterios de aceptación HU6 Métricas de modelos
 - o Se definen las métricas de análisis para cada modelo.
 - o Se justifica la selección de cada métrica para cada modelo.
 - o Se obtienen las métricas de análisis de cada modelo.
 - Se comparan los distintos modelos mediante las métricas definidas.

• Épica 3 - Despliegue en entorno MLFlow

- Criterios de aceptación HU7 Despliegue en MLFlow
 - \circ Se crea un entorno MLFlow local desde cero
 - o Se configura el paso correspondiente al análisis de datos en el entorno.
 - o Se replican las técnicas exploradas de análisis de datos en el paso correspondiente.
 - o Se configura el paso de entrenamiento de modelos en el entorno.
 - Se replican las técnicas exploradas de entrenamiento de modelos en el paso correspondiente.
 - o Se configura el paso de evaluación de modelos en el entorno.
 - o Se replican las técnicas exploradas de evaluación de modelos en el paso correspondiente.
- Criterios de aceptación HU8 API para entorno MLFlow
 - Se exponen los resultados de los modelos explorados en el entorno de MLFlow mediante una API.
 - \circ Se exponen las comparaciones de los modelos explorados en el entorno de MLFlow mediante una API.

• Épica 4 - Gestión de calidad del código fuente

- Criterios de aceptación HU9 Implementación de buenas prácticas
 - o Se implementan buenas prácticas de código Python en el proyecto.
- Criterios de aceptación HU10 Documentación
 - o Se documentan todos los pasos realizados durante el desarrollo del proyecto.
 - o Se valida que cada paso realizado esté correctamente justificado.
- Criterios de aceptación HU11 Validación de API de MLFlow
 - $\circ\,$ Se valida el acceso a los resultados de los modelos mediante la API del entorno MLFlow.
 - $\circ\,$ Se valida el acceso a la comparación de los modelos mediante la API del entorno MLFlow.

8. Fases de CRISP-DM

1. Comprensión del negocio:

- Objetivo: predecir si una empresa va a entrar en quiebra o no.
- Impacto: ayudar en la toma de decisiones a empresas especializadas en finanzas, en inversiones, en prestación de seguros, entre otras, permitiéndoles saber si una empresa sobre la que se quiere invertir o a la que se le quiere otorgar un préstamo puede entrar en quiebra o no.
- *Métricas:* se predice correctamente si la empresa quiebra o no.



2. Comprensión de los datos

- Tipos de datos: datos tabulares.
- Fuente de datos: datos publicados por el Taiwan Economic Journal.
- Cantidad de datos: 6819 registros con 96 columnas.

3. Preparación de los datos

- Transformaciones
 - Análisis y ajuste de datos atípicos.
 - Análisis y ajuste de datos faltantes.
 - Creación de nuevas variables al combinar las variables existentes.
 - Normalización de datos.
- Características clave
 - Indicador de si la empresa entró en quiebra o variable target.
 - Distintas métricas del desempeño de la empresa a nivel económico y contable.

4. Modelado

- Tipo de problema: clasificación.
- Arquitecturas posibles: modelos de clasificación como Logistic Regression, Support Vector Machines y XGBoost.

5. Evaluación del modelo

■ F1-score y AUC-ROC.

6. Despliegue del modelo

• Despliegue local usando *MLFlow*.

9. Desglose del trabajo en tareas

Historia de usuario	Tarea técnica	Estimación	Prioridad
HU1 - Análisis exploratorio	Identificar variables categóricas	4 h	Media
HU1 - Análisis exploratorio	Identificar variables numéricas	4 h	Media
HU1 - Análisis exploratorio	Graficar la distribución de las	6 h	Media
	variables numéricas		
HU1 - Análisis exploratorio	Realizar análisis de correlacio-	4 h	Media
	nes entre variables numéricas		
HU1 - Análisis exploratorio	Documentar pasos y decisiones	3 h	Media
	tomadas		
HU2 - Procesamiento de datos	Investigar técnicas de balanceo	6 h	Media
faltantes y datos atípicos	de clases para algoritmos de		
	clasificación		
HU2 - Procesamiento de datos	Implementar técnicas de balan-	5 h	Media
faltantes y datos atípicos	ceo de clases para algoritmos de		
	clasificación		
HU2 - Procesamiento de datos	Separar dataset en train y test	3 h	Media
faltantes y datos atípicos			



Historia de usuario	Tarea técnica	Estimación	Prioridad
HU2 - Procesamiento de datos	Identificar variables con datos	4 h	Alta
faltantes y datos atípicos	faltantes		
HU2 - Procesamiento de datos	Analizar causas de datos faltan-	6 h	Alta
faltantes y datos atípicos	tes		
HU2 - Procesamiento de datos	Corregir datos faltantes	8 h	Alta
faltantes y datos atípicos			
HU2 - Procesamiento de datos	Identificar datos con valores	6 h	Alta
faltantes y datos atípicos	atípicos		
HU2 - Procesamiento de datos	Analizar causas de datos atípi-	8 h	Alta
faltantes y datos atípicos	cos		
HU2 - Procesamiento de datos	Graficar variables que presentan	5 h	Media
faltantes y datos atípicos	de datos atípicos		
HU2 - Procesamiento de datos	Corregir datos atípicos	8 h	Alta
faltantes y datos atípicos			
HU2 - Procesamiento de datos	Documentar pasos y decisiones	5 h	Media
faltantes y datos atípicos	tomadas		
HU3 - Feature Engineering	Identificar variables menos im-	5 h	Alta
	portantes para eliminarlas		
HU3 - Feature Engineering	Implementar técnicas de elimi-	5 h	Media
	nación de features		
HU3 - Feature Engineering	Crear nuevas variables mediante	7 h	Alta
	combinaciones lineales de varia-		
	bles existentes		
HU3 - Feature Engineering	Investigar otras técnicas de	5 h	Media
	creación de variables		
HU3 - Feature Engineering	Aplicar otras técnicas de crea-	8 h	Alta
	ción de variables		
HU3 - Feature Engineering	Evaluar nuevas variables en	5 h	Alta
	modelos		
HU3 - Feature Engineering	Documentar pasos y decisiones	3 h	Baja
	tomadas		



Historia de usuario	Tarea técnica	Estimación	Prioridad
HU4 - Implementación de mo-	Implementar código de vali-	4 h	Media
delos de Machine Learning	dación cruzada para <i>Logistic</i>		
	Regression		
HU4 - Implementación de mo-	Implementar modelo Logistic	6 h	Alta
delos de Machine Learning	Regression, sin considerar featu-		
	re engineering	4.1	2.5.11
HU4 - Implementación de mo-	Evaluar modelo Logistic Re-	4 h	Media
delos de Machine Learning	gression, sin considerar feature		
IIII4 Implementación de ma	engineering	6 h	Alta
HU4 - Implementación de modelos de Machine Learning	Implementar modelo Logistic Regression, considerando featu-	0 11	Ana
delos de <i>Machine Learning</i>	re engineering		
HU4 - Implementación de mo-	Evaluar modelo Logistic Re-	4 h	Media
delos de Machine Learning	gression, considerando feature	4 11	Wiedia
delos de Machine Bearning	engineering		
HU4 - Implementación de mo-	Implementar código de valida-	4 h	Media
delos de Machine Learning	ción cruzada para SVM		
HU4 - Implementación de mo-	Implementar modelo SVM, sin	6 h	Alta
delos de Machine Learning	considerar feature engineering		
HU4 - Implementación de mo-	Evaluar modelo SVM, sin consi-	4 h	Media
delos de Machine Learning	derar feature engineering		
HU4 - Implementación de mo-	Implementar modelo SVM, con-	6 h	Alta
delos de Machine Learning	siderando feature engineering		
HU4 - Implementación de mo-	Evaluar modelo SVM, conside-	4 h	Media
delos de Machine Learning	rando feature engineering		
HU4 - Implementación de mo-	Implementar código de valida-	4 h	Media
delos de Machine Learning	ción cruzada para XGBoost		
HU4 - Implementación de mo-	Implementar modelo XGBoost,	8 h	Alta
delos de Machine Learning	sin considerar feature enginee-		
	ring		
HU4 - Implementación de mo-	Evaluar modelo XGBoost, sin	6 h	Media
delos de Machine Learning	considerar feature engineering		
HU4 - Implementación de mo-	Implementar modelo XGBoost,	8 h	Alta
delos de Machine Learning	considerando feature enginee-		
IIII4 Il	ring	C l	N/ - J: -
HU4 - Implementación de mo-	Evaluar modelo XGBoost, con-	6 h	Media
delos de Machine Learning	siderando feature engineering Persistir modelos en GitHub	4 h	Madia
HU4 - Implementación de modelos de Machine Learning	r ersistir modelos en Gumu	4 h	Media
HU4 - Implementación de mo-	Documentar pasos y decisiones	5 h	Media
delos de Machine Learning	tomadas	0 11	ivicula
HU5 - Optimización de hiper-	Identificar hiperparámetros y	5 h	Media
parámetros	rangos de Logistic Regression		Titodia
HU5 - Optimización de hiper-	Optimizar hiperparámetros de	6 h	Media
parámetros	Logistic Regression, sin conside-		1.10414
	rar feature engineering		
	ل.٠٠٠٠٠ ل	I.	



HU5 - Optimización de hiperparámetros se en Logistic Regression, sin considerar feature engineering en	Historia de usuario	Tarea técnica	Estimación	Prioridad
Regression, sin considerar feature engineering Feature enginee	_		4 h	Media
Feature engineering	parámetros	_		
HU5 - Optimización de hiperparámetros optimos, sin considerar feature engineering HU5 - Optimización de hiperparámetros HU5 - Optimiza		,		
parámetros gression con hiperparámetros óptimos, sin considerar feature engineering considerando feature engineering limplementar hiperparámetros de Logistic Regression, considerando feature engineering limplementar hiperparámetros más óptimos en Logistic Regression, considerando feature engineering limplementar hiperparámetros más óptimos en Logistic Regression, considerando feature engineering limplementar hiperparámetros optimios, considerando feature engineering limplementar hiperparámetros optimios, considerando feature engineering limplementar hiperparámetros optimios, considerando feature engineering limplementar hiperparámetros de SVM, sin considerar feature engineering limplementar hiperparámetros más óptimos en SVM, sin considerar feature engineering limplementar hiperparámetros de SVM con hiperparámetros optimización de hiperparámetros en SVM, sin considerar feature engineering limplementar hiperparámetros de SVM, considerando feature engineering limplementar hiperparámetros de A hiperparámetros optimos, considerando feature engineering limplementar hiperparámetros de SVM, considerando feature engineering li	HILE O .:		4.1	N
foptimos, sin considerar feature engineering HU5 - Optimización de hiperparámetros HU5 - Optimización de hiperparámetros de SVM conhiperparámetros HU5 - Optimización de hiperparámetros de SVM conhiperparámetros de SVM conhi			4 n	Media
HU5 - Optimización de hiperparámetros de Logistic Regression, considerando feature engineering HU5 - Optimización de hiperparámetros de SVM con hiperparámetros HU5 - Optimización de hiperparámetros de SVM con hiperparámetros HU5 - Optimización de hiperparámetros de SVM, considerar feature engineering HU5 - Optimización de hiperparámetros de SVM, considerando feature engineering HU5 - Optimización de hiperparámetros de SVM, considerando feature engineering HU5 - Optimización de hiperparámetros de SVM, considerando feature engineering HU5 - Optimización de hiperparámetros de SVM, considerando feature engineering HU5 - Optimización de hiperparámetros de SVM, considerando feature engineering HU5 - Optimización de hiperparámetros de SVM con hiperp	parametros	_ = =		
HU5 - Optimización de hiperparámetros de Logistic Regression, considerando feature engineering HU5 - Optimización de hiperparámetros de feature engineering HU5 - Optimización de hiperparámetros optimos, considerando feature engineering HU5 - Optimización de hiperparámetros optimos, considerando feature engineering HU5 - Optimización de hiperparámetros optimos, considerando feature engineering HU5 - Optimización de hiperparámetros optimos, considerando feature engineering HU5 - Optimización de hiperparámetros optimos, considerando feature engineering HU5 - Optimización de hiperparámetros optimos, considerando feature engineering HU5 - Optimización de hiperparámetros optimos, considerando feature engineering HU5 - Optimización de hiperparámetros optimos, co				
Deparametros Logistic Regression, considerando feature engineering HU5 - Optimización de hiperparámetros Evaluar modelo de Logistic Regression con hiperparámetros optimos, considerando feature engineering HU5 - Optimización de hiperparámetros HU5 - Optimización de hiperparámetros HU5 - Optimización de hiperparámetros SVM, sin considerar feature engineering HU5 - Optimización de hiperparámetros HU5 - Optimización de hiperparámetros HU5 - Optimización de hiperparámetros Evaluar modelo de SVM SVM, sin considerar feature engineering HU5 - Optimización de hiperparámetros HU5 - Optimización de hiperparámetros Evaluar modelo de SVM con hiperparámetros HU5 - Optimización de hiperparámetros Optimizar hiperparámetros HU5 - Optimización de hiperparámetros Optimizar hiperparámetros HU5 - Optimización de hiperparámetros Implementar hiperparámetros Optimización de hiperparámetros Optimizar hiperparámetros Optimización de hiperparámetros Optimizar hiperparámetros Optimización de hiperparámetros Optimizar hiperparámetros Optimizar hiperparámetros Optimización Optimizar hiperparámetros Optimización Optimizar hiperparámetros Optimización Optimizar hiperparámetros Optimizar hiperparámetros Optimización Optimizar hiperparámetros Optimización Optimizar hiperparámetros Optimización Optimizació	HII5 - Ontimización de hiper-		6 h	Media
do feature engineering HU5 - Optimización de hiperparámetros más óptimos en Logistic Regression, considerando feature engineering HU5 - Optimización de hiperparámetros más óptimos, considerando feature engineering HU5 - Optimización de hiperparámetros HU5 -	_		0 11	Media
HU5 - Optimización de hiperparámetros más óptimos en Logistic Regression, considerando feature engineering HU5 - Optimización de hiperparámetros óptimos, considerando feature engineering HU5 - Optimización de hiperparámetros óptimos, considerando feature engineering HU5 - Optimización de hiperparámetros optimización de hiperparámetros HU5 - Optimización de hiperparámetros optimos, sin considerar feature engineering HU5 - Optimización de hiperparámetros optimos, sin considerar feature engineering HU5 - Optimización de hiperparámetros optimos, sin considerar feature engineering HU5 - Optimización de hiperparámetros optimos, sin considerando feature engineering HU5 - Optimización de hiperparámetros optimos, en SVM, considerando feature engineering HU5 - Optimización de hiperparámetros optimos, considerando feature engineering HU5 - Optimización de hiperparámetros optimos, considerando feature engineering HU5 - Optimización de hiperparámetros optimos, considerando feature engineering HU5 - Optimización de hiperparámetros optimos, considerando feature engineering HU5 - Optimización de hiperparámetros optimos, considerando feature engineering HU5 - Optimización de hiperparámetros optimos, considerando feature engineering HU5 - Optimización de hiperparámetros optimos, considerando feature engineering HU5 - Optimización de hiperparámetros optimos, considerando feature engineering HU5 - Optimización de hiperparámetros de XGBoost, sin considerar feature engineering HU5 - Optimización de hiperparámetros de XGBoost, sin considerar feature engineering HU5 - Optimización de hiperparámetros de XGBoost, sin considerar feature engineering	parametros	,		
más óptimos en Logistic Regression, considerando feature engineering HU5 - Optimización de hiper-parámetros optimos, considerando feature engineering HU5 - Optimización de hiper-parámetros optimos, considerando feature engineering HU5 - Optimización de hiper-parámetros de SVM, sin considerar feature engineering HU5 - Optimización de hiper-parámetros de SVM, sin considerar feature engineering HU5 - Optimización de hiper-parámetros de SVM, sin considerar feature engineering HU5 - Optimización de hiper-parámetros de SVM con hiperparámetros HU5 - Optimización de hiper-parámetros optimos, sin considerar feature engineering HU5 - Optimización de hiper-parámetros optimos, sin considerar feature engineering HU5 - Optimización de hiper-parámetros optimos, sin considerando feature engineering HU5 - Optimización de hiper-parámetros optimos en SVM, considerando feature engineering HU5 - Optimización de hiper-parámetros optimos en SVM, considerando feature engineering HU5 - Optimización de hiper-parámetros optimos, considerando feature engineering HU5 - Optimización de hiper-parámetros optimos, considerando feature engineering HU5 - Optimización de hiper-parámetros optimos, considerando feature engineering HU5 - Optimización de hiper-parámetros optimos, considerando feature engineering HU5 - Optimización de hiper-parámetros optimos, considerando feature engineering HU5 - Optimización de hiper-parámetros optimos, considerando feature engineering HU5 - Optimización de hiper-parámetros de XGBoost HU5 - Optimización de hiper-parámetros de XGBoost HU5 - Optimización de hiper-parámetros de XGBoost HU5 - Optimización de hiper-parámetros de XGBoost, sin considerar feature engineering HU5 - Optimización de hiper-parámetros de XGBoost, sin considerar feature engineering HU5 - Optimización de hiper-parámetros de XGBoost, sin considerar feature engineering	HU5 - Ontimización de hiper-		4 h	Media
Regression, considerando feature engineering HU5 - Optimización de hiperparámetros Brushar modelo de Logistic Respression con hiperparámetros Optimización de hiperparámetros y rangos de SVM HU5 - Optimización de hiperparámetros HU5 - Optimización de hiperparámetros HU5 - Optimización de hiperparámetros Brushar modelo de SVM HU5 - Optimización de hiperparámetros HU5 - Optimización de hiperparámetros Brushametros Brushametro			1 11	Wiedia
HU5 - Optimización de hiperparámetros optimos, considerando feature engineering HU5 - Optimización de hiperparámetros optimos, considerando feature engineering HU5 - Optimización de hiperparámetros engineering	parametros	1 5		
HU5 - Optimización de hiperparámetros soptimos, considerar feature engineering HU5 - Optimización de hiperparámetros y rangos de SVM HU5 - Optimización de hiperparámetros y rangos de SVM HU5 - Optimización de hiperparámetros y rangos de SVM, sin considerar feature engineering HU5 - Optimización de hiperparámetros y rangos de SVM, sin considerar feature engineering HU5 - Optimización de hiperparámetros y rangos de SVM, sin considerar feature engineering HU5 - Optimización de hiperparámetros y rangos de SVM, sin considerar feature engineering HU5 - Optimización de hiperparámetros optimos, sin considerar feature engineering HU5 - Optimización de hiperparámetros de SVM, considerando feature engineering HU5 - Optimización de hiperparámetros de SVM, considerando feature engineering HU5 - Optimización de hiperparámetros de SVM, considerando feature engineering HU5 - Optimización de hiperparámetros de SVM con hiperparámetros de rengineering HU5 - Optimización de hiperparámetros de SVM con hiperparámetros de rengineering HU5 - Optimización de hiperparámetros de SVM con hiperparámetros optimos, considerando feature engineering HU5 - Optimización de hiperparámetros de rangos de XGBoost HU5 - Optimización de hiperparámetros de Nedia rangos de XGBoost HU5 - Optimización de hiperparámetros de Nedia rangos de XGBoost, sin considerar feature engineering HU5 - Optimización de hiperparámetros de Nedia rangos de XGBoost, sin considerar feature engineering HU5 - Optimización de hiperparámetros de Nedia parámetros HU5 - Optimización de hiperparámetros de Nedia parámetros de Nedia pará				
parámetros gression con hiperparámetros óptimos, considerando feature engineering	HU5 - Optimización de hiper-		4 h	Media
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	_	_		
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$				
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$		engineering		
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	HU5 - Optimización de hiper-	Identificar hiperparámetros y	5 h	Media
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	parámetros	rangos de SVM		
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	HU5 - Optimización de hiper-	Optimizar hiperparámetros de	6 h	Media
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	parámetros	SVM, sin considerar feature		
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$				
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	HU5 - Optimización de hiper-		4 h	Media
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	parámetros			
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$				
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$			4 h	Media
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	parámetros			
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$				1.5
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	_		6 h	Media
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	parametros			
parámetros más óptimos en SVM, considerando feature engineering HU5 - Optimización de hiper-parámetros óptimos, considerando feature engineering HU5 - Optimización de hiper-parámetros óptimos, considerando feature engineering HU5 - Optimización de hiper-parámetros y rangos de XGBoost HU5 - Optimización de hiper-parámetros de XGBoost, sin considerar feature engineering HU5 - Optimización de hiper-parámetros de XGBoost, sin considerar feature engineering HU5 - Optimización de hiper-parámetros de xGBoost, sin considerar feature engineering HU5 - Optimización de hiper-parámetros de xGBoost, sin más óptimos en XGBoost, sin		0	4.1	N. 6 11
considerando feature engineering HU5 - Optimización de hiper- parámetros			4 h	Media
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	parametros	,		
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$		· ·		
parámetros hiperparámetros óptimos, considerando feature engineering HU5 - Optimización de hiperparámetros y rangos de $XGBoost$ HU5 - Optimización de hiperparámetros de $XGBoost$ HU5 - Optimización de hiperparámetros de $XGBoost$, sin considerar feature engineering HU5 - Optimización de hiperparámetros de $XGBoost$, sin considerar feature engineering HU5 - Optimización de hiperparámetros $XGBoost$, sin $XGBOOST$	HII5 - Ontimización de hiper-		4 h	Media
derando feature engineering HU5 - Optimización de hiperparámetros y rangos de XGBoost HU5 - Optimización de hiperparámetros de Alta parámetros HU5 - Optimización de hiperparámetros de Alta y Alta			4 11	Media
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	parametros			
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	HU5 - Optimización de hiper-		7 h	Media
$\begin{array}{c ccccccccccccccccccccccccccccccccccc$	-			2.20010
parámetros $XGBoost$, sin considerar $feature$ $engineering$ HU5 - Optimización de hiper- Implementar hiperparámetros 5 h Media parámetros más óptimos en $XGBoost$, sin	-	_	8 h	Alta
engineeringHU5 - Optimización de hiper- parámetrosImplementar hiperparámetros más óptimos en XGBoost, sin5 hMedia	-			
HU5 - Optimización de hiper- Implementar hiperparámetros 5 h Media parámetros más óptimos en XGBoost, sin	•	,		
parámetros más óptimos en XGBoost, sin	HU5 - Optimización de hiper-	<u> </u>	5 h	Media
	_			
considerar feature engineering		considerar feature engineering		



Historia de usuario	Tarea técnica	Estimación	Prioridad
HU5 - Optimización de hiper-	Evaluar modelo de XGBoost	7 h	Media
parámetros	con hiperparámetros óptimos,		
	sin considerar feature enginee-		
	ring	0.1	A 1.
HU5 - Optimización de hiper-	Optimizar hiperparámetros de	8 h	Alta
parámetros	XGBoost, considerando feature engineering		
HU5 - Optimización de hiper-	Implementar hiperparámetros	5 h	Media
parámetros	más óptimos en XGBoost,	0 11	Media
parametros	considerando feature		
	engineering		
HU5 - Optimización de hiper-	Evaluar modelo de XGBoost	7 h	Media
parámetros	con hiperparámetros óptimos,		
	considerando feature enginee-		
	$\mid ring \mid$		
HU5 - Optimización de hiper-	Documentar pasos y decisiones	5 h	Media
parámetros	tomadas		
HU6 - Métricas de modelos	Obtener métricas de F1-score	3 h	Media
	para Logistic Regression, sin		
	considerar feature engineering		
HU6 - Métricas de modelos	Obtener métricas de AUC-ROC	5 h	Media
	para Logistic Regression, sin		
	considerar feature engineering, y		
HU6 - Métricas de modelos	graficar Obtener métricas de F1-score	3 h	Media
HU6 - Metricas de modelos		o n	Media
	para Logistic Regression, considerando feature engineering		
HU6 - Métricas de modelos	Obtener métricas de AUC-ROC	5 h	Media
1100 Worldoo de modelos	para Logistic Regression, consi-	0 11	Wicaia
	derando feature engineering, y		
	graficar		
HU6 - Métricas de modelos	Obtener métricas de F1-score	3 h	Media
	para SVM, sin considerar featu-		
	re engineering		
HU6 - Métricas de modelos	Obtener métricas de AUC-ROC	5 h	Media
	para SVM, sin considerar featu-		
	re engineering, y graficar		
HU6 - Métricas de modelos	Obtener métricas de F1-score	3 h	Media
	para SVM , considerando $feature$		
HU6 - Métricas de modelos	engineering	F 1.	М- 1:-
noo - Metricas de modelos	Obtener métricas de AUC-ROC para SVM, considerando feature	5 h	Media
	engineering, y graficar		
HU6 - Métricas de modelos	Obtener métricas de F1-score	3 h	Media
1100 Montons de moderos	para XGBoost, sin considerar	0 11	wicula
	feature engineering		
HU6 - Métricas de modelos	Obtener métricas de AUC-ROC	5 h	Media
	para $XGBoost$, sin considerar	_	
	feature engineering, y graficar		
	0 0,00	1	



Historia de usuario	Tarea técnica	Estimación	Prioridad
HU6 - Métricas de modelos	Obtener métricas de F1-score	3 h	Media
	para XGBoost, considerando		
	feature engineering		
HU6 - Métricas de modelos	Obtener métricas de AUC-ROC	5 h	Media
	para XGBoost, considerando		
IIII Maria	feature engineering, y graficar	0.1	N. f. 1.
HU6 - Métricas de modelos	Comparar métricas de distintos modelos	3 h	Media
HU6 - Métricas de modelos	Documentar pasos y decisiones	5 h	Media
1100 - Metricas de modelos	tomadas	9 11	Media
HU7 - Despliegue en MLFlow	Investigar buenas prácticas para	5 h	Alta
Tro Pesphegue en M21 vou	despliegues de <i>MLFlow</i>	0 11	71100
HU7 - Despliegue en MLFlow	Crear entorno local para des-	7 h	Alta
1 10 11	pliegue MLFlow		
HU7 - Despliegue en MLFlow	Replicar técnicas de análisis de	8 h	Alta
	datos en entorno MLFlow		
$\mathrm{HU7}$ - Despliegue en MLFlow	Replicar técnicas de entrena-	8 h	Alta
	miento de modelos de <i>Logistic</i>		
	Regression en entorno $MLFlow$		
$\mathrm{HU7}$ - Despliegue en MLFlow	Replicar técnicas de entrena-	8 h	Alta
	miento de modelos de SVM en		
	entorno MLFlow		
$\mathrm{HU7}$ - Despliegue en MLFlow	Replicar técnicas de entrena-	8 h	Alta
	miento de modelos de XGBoost		
	en entorno MLFlow		
HU7 - Despliegue en <i>MLFlow</i>	Replicar técnicas de evaluación	8 h	Alta
11112 D 1: 141 D1	de modelos en entorno MLFlow	4.1	3.5.11
$\mathrm{HU7}$ - Despliegue en MLFlow	Ejecutar localmente el entorno	4 h	Media
IIII7 Dl: MI El	MLFlow	4 h	Media
$\mathrm{HU7}$ - Despliegue en MLFlow	Validar ejecución local del entorno <i>MLFlow</i>	4 n	Media
HU7 - Despliegue en <i>MLFlow</i>	Documentar pasos y decisiones	6 h	Media
1107 - Desphegue en MBP tow	tomadas MLFlow	0 11	Wiedia
HU8 - API para entorno ML-	Investigar como exponer un	5 h	Media
Flow	entorno MLFlow mediante API	0 11	Wicala
HU8 - API para entorno ML-	Exponer resultados de modelos	8 h	Alta
Flow	explorados en entorno <i>MLFlow</i>		
	mediante API		
HU8 - API para entorno ML-	Exponer comparación de mode-	8 h	Alta
Flow	los en entorno <i>MLFlow</i> median-		
	te API		
HU8 - API para entorno ML -	Documentar pasos y decisiones	3 h	Baja
Flow	tomadas		
HU9 - Implementación de bue-	Investigar buenas prácticas en	4 h	Baja
nas prácticas	código Python		
HU9 - Implementación de bue-	Aplicar buenas prácticas en	8 h	Media
nas prácticas	código Python		



Historia de usuario	Tarea técnica	Estimación	Prioridad
HU10 - Documentación	Asegurar que cada decisión to-	6 h	Media
	mada haya sido justificada y		
	documentada		
HU10 - Documentación	Asegurar ortografía y formato	8 h	Media
	en documentación		
HU11 - Validación de API de	Validar acceso a modelos explo-	8 h	Media
MLFlow	rados mediante API de entorno		
	MLFlow		
HU11 - Validación de API de	Validar acceso a comparación	8 h	Media
MLFlow	de modelos mediante API de		
	entorno MLFlow		
$oxed{HU11}$ - Validación de API de	Crear documentación sobre el	5 h	Media
MLFlow	uso de API de entorno $MLFlow$		
Planificación del proyecto y con-	Planificación del proyecto	8 h	Alta
fección de informes de avance			
(opcional)			
Planificación del proyecto y con-	Informe de avance - Secciones 1	6 h	Media
fección de informes de avance	a 5 inclusive		
(opcional)			
Planificación del proyecto y con-	Informe de avance - Secciones 6	5 h	Media
fección de informes de avance	a 9 inclusive		
(opcional)			7.5.11
Planificación del proyecto y con-	Informe de avance - Secciones 10	4 h	Media
fección de informes de avance	a 12 inclusive		
(opcional)		4.1	3.6.11
Planificación del proyecto y con-	Informe de avance - Secciones 13	4 h	Media
fección de informes de avance	a 15 inclusive		
(opcional)		F 1	3.6.11
Planificación del proyecto y con-	Informe de avance - Correciones	5 h	Media
fección de informes de avance	generales		
(opcional)	D 1 '' 1 '' 1	C 1	N. 1.
Redacción de memoria (opcio-	Redacción de sección sobre pro-	6 h	Media
nal)	cesamiento de datos	2 1-	M - 1: -
Redacción de memoria (opcio-	Redacción de sección sobre Fea-	3 h	Media
nal) Redacción de memoria (opcio-	ture Engineering Redacción de sección sobre im-	7 h	Media
		/ 11	Media
nal) Redacción de memoria (opcio-	plementación de modelos Redacción de sección sobre op-	6 h	Media
nal)	timización de hiperparámetros	0 11	Media
Redacción de memoria (opcio-	Redacción de sección sobre <i>ML</i> -	5 h	Media
nal)	Flow	0 11	ivicula
Redacción de memoria (opcio-	Correcciones generales	8 h	Media
nal)	Correctiones generales	0 11	ivicula
Preparación de presentación fi-	Confección de presentación Po-	6 h	Alta
nal (opcional)	werPoint	0 11	And
Preparación de presentación fi-	Confección de video demostra-	8 h	Alta
nal (opcional)	ción	0 11	And
nor (operonar)	01011	<u> </u>	



$10.\ Planificación de Sprints$

Sprint	HU o fase	Tarea	Horas/SP	Responsable	% Completado
Sprint 0	Planificación	Planificación	8 h / 5 SP	Alumno	80 %
	del proyecto	del proyecto			
	y confección				
	de informes de				
Sprint 0	avance Planificación	Informe de	6 h / 3 SP	Alumno	100 %
Sprine 0	del proyecto	avance -		Alumno	100 70
	y confección	Secciones 1			
	de informes de	a 5 inclusive			
	avance				
Sprint 0	Planificación	Informe de	5 h / 3 SP	Alumno	100 %
_	del proyecto	avance -	,		
	y confección	Secciones 6			
	de informes de	a 9 inclusive			
	avance				
Sprint 0	Planificación	Informe de	4 h / 2 SP	Alumno	100 %
	del proyecto	avance -			
	y confección	Secciones 10 a			
	de informes de	12 inclusive			
Comint 0	avance Planificación	Informa	4 b / 9 CD	A 1	0.07
Sprint 0	del proyecto	Informe de avance -	4 h / 2 SP	Alumno	0 %
	y confección	Secciones 13 a			
	de informes de	15 inclusive			
	avance	10 1110101011			
Sprint 0	Planificación	Informe de	5 h / 3 SP	Alumno	0 %
	del proyecto	avance -	,		
	y confección	Correciones			
	de informes de	generales			
	avance				
Sprint 1	HU1 - Análisis	Identificar	4 h / 2 SP	Alumno	0 %
	exploratorio	variables			
0	TTTT1 A /11 ·	categóricas	4.1. / 0.CD	A 1	0.04
Sprint 1	HU1 - Análisis	Identificar	4 h / 2 SP	Alumno	0 %
	exploratorio	variables numéricas			
Sprint 1	HU1 - Análisis	Graficar la	6 h / 3 SP	Alumno	0 %
Sprint 1	exploratorio	distribución de		Alumno	0 70
	CAPIOTATOTIO	las variables			
		$num\'ericas$			
Sprint 1	HU1 - Análisis	Realizar análi-	4 h / 2 SP	Alumno	0 %
	exploratorio	sis de correla-			
	_	ciones entre va-			
		riables numéri-			
		cas			



Sprint	HU o fase	Tarea	Horas/SP	Responsable	% Completado
Sprint 1	HU1 - Análisis exploratorio	Documentar pasos y decisiones tomadas	3 h / 2 SP	Alumno	0 %
Sprint 1	HU2 - Procesa- miento de datos faltantes y da- tos atípicos	Investigar técnicas de balanceo de clases para algoritmos de clasificación	6 h / 3 SP	Alumno	0 %
Sprint 1	HU2 - Procesa- miento de datos faltantes y da- tos atípicos	Implementar técnicas de balanceo de clases para algoritmos de clasificación	5 h / 3 SP	Alumno	0 %
Sprint 1	HU2 - Procesa- miento de datos faltantes y da- tos atípicos	Separar dataset en train y test	3 h / 2 SP	Alumno	0 %
Sprint 1	HU2 - Procesa- miento de datos faltantes y da- tos atípicos	Identificar variables con datos faltantes	4 h / 2 SP	Alumno	0 %
Sprint 1	HU2 - Procesa- miento de datos faltantes y da- tos atípicos	Analizar causas de datos faltan- tes	6 h / 3 SP	Alumno	0 %
Sprint 1	HU2 - Procesa- miento de datos faltantes y da- tos atípicos	Corregir datos faltantes	8 h / 5 SP	Alumno	0 %
Sprint 1	HU2 - Procesa- miento de datos faltantes y da- tos atípicos	Identificar datos con valores atípicos	6 h / 3 SP	Alumno	0 %
Sprint 2	HU2 - Procesa- miento de datos faltantes y da- tos atípicos	Analizar causas de datos atípi- cos	8 h / 5 SP	Alumno	0 %
Sprint 2	HU2 - Procesa- miento de datos faltantes y da- tos atípicos	Graficar variables que presentan de datos atípicos	5 h / 3 SP	Alumno	0 %
Sprint 2	HU2 - Procesa- miento de datos faltantes y da- tos atípicos	Corregir datos atípicos	8 h / 5 SP	Alumno	0 %



Sprint	HU o fase	Tarea	Horas/SP	Responsable	% Completado
Sprint 2	HU2 - Procesa-	Documentar	5 h / 3 SP	Alumno	0 %
	miento de datos	pasos y			
	faltantes y da-	decisiones			
Q	tos atípicos	tomadas			. ~
Sprint 2	HU3 - Feature	Identificar	5 h / 3 SP	Alumno	0%
	Engineering	variables menos			
		importantes			
C 0	IIII9 Et	para eliminarlas	F 1 / 2 CD	A 1	0.07
Sprint 2	HU3 - Feature Engineering	Implementar técnicas de	5 h / 3 SP	Alumno	0 %
	Ендінеетіну 	eliminación de			
		features			
Sprint 2	HU3 - Feature	Crear nuevas	7 h / 5 SP	Alumno	0 %
Sprine 2	Engineering	variables	111/051	Midillio	0 70
		mediante			
		combinaciones			
		lineales de			
		variables			
		existentes			
Sprint 2	HU3 - Feature	Investigar otras	5 h / 3 SP	Alumno	0 %
	Engineering	técnicas de	·		
		creación de			
		variables			
Sprint 2	HU3 - Feature	Aplicar otras	8 h / 5 SP	Alumno	0 %
	Engineering	técnicas de			
		creación de			
G 1 1 2	TITIO E	variables	7 1 / 2 CD	4.1	0.04
Sprint 2	HU3 - Feature	Evaluar nuevas	5 h / 3 SP	Alumno	0%
	Engineering	variables en mo-			
Sprint 2	HU3 - Feature	delos Documentar	3 h / 2 SP	Alumno	0 %
Spriit 2	Engineering			Aluillio	0 70
	Dilgineering	pasos y decisiones			
		tomadas			
Sprint 3	HU4 - Im-	Implementar	4 h / 2 SP	Alumno	0 %
Spring	plementación	código de	1.11 / 2.51		0 70
	de modelos	validación			
	de Machine	cruzada			
	Learning	para Logistic			
		Regression			
Sprint 3	HU4 - Im-	Implementar	6 h / 3 SP	Alumno	0 %
	plementación	modelo Logistic			
	de modelos	Regression, sin			
	de <i>Machine</i>	considerar			
	Learning	feature			
		engineering			



Sprint	HU o fase	Tarea	Horas/SP	Responsable	% Completado
Sprint 3	HU4 - Im-	Evaluar	4 h / 2 SP	Alumno	0 %
	plementación	modelo <i>Logistic</i>			
	de modelos	$Regression, \sin$			
	de Machine	considerar			
	Learning	feature			
G	TTT 1	engineering	0.1 / 0.CD	4.1	0.04
Sprint 3	HU4 - Im-	Implementar	6 h / 3 SP	Alumno	0 %
	plementación	modelo <i>Logistic</i>			
	de modelos de <i>Machine</i>	Regression, considerando			
	$egin{array}{cccc} de & Machine \\ Learning \end{array}$	feature			
	Learning	engineering			
Sprint 3	HU4 - Im-	Evaluar mode-	4 h / 2 SP	Alumno	0 %
Sprine 5	plementación	lo Logistic Re-	411/201	Mullino	0 70
	de modelos	gression, consi-			
	$\begin{array}{ccc} \operatorname{de} & \operatorname{Modeles} \\ \operatorname{de} & \operatorname{Machine} \end{array}$	derando feature			
	Learning	engineering			
Sprint 3	HU4 - Im-	Implementar	4 h / 2 SP	Alumno	0 %
	plementación	código de	,		
	de modelos	validación			
	de Machine	cruzada para			
	Learning	SVM			
Sprint 3	HU4 - Im-	Implementar	6 h / 3 SP	Alumno	0 %
	plementación	modelo SVM,			
	de modelos	sin considerar			
	de Machine	feature			
G : + 9	Learning	engineering	4.1 / 9.CD	A 1	0.07
Sprint 3	HU4 - Im-	Evaluar modelo SVM , sin	4 h / 3 SP	Alumno	0 %
	plementación de modelos	SVM , \sin considerar			
	de = Machine	feature			
	$oxed{Learning}$	engineering			
Sprint 3	HU4 - Im-	Implementar	6 h / 3 SP	Alumno	0 %
r.	plementación	modelo SVM ,	, , , , , ,		,,,
	de modelos	considerando			
	de <i>Machine</i>	feature			
	Learning	engineering			
Sprint 3	HU4 - Im-	Evaluar mode-	4 h / 2 SP	Alumno	0 %
	plementación	lo SVM, consi-			
	de modelos	derando feature			
	de Machine	engineering			
	Learning				
Sprint 3	HU4 - Im-	Implementar	4 h / 2 SP	Alumno	0 %
	plementación	código de			
	de modelos	validación			
	de Machine	cruzada para			
	Learning	XGBoost			



Sprint	HU o fase	Tarea	Horas/SP	Responsable	% Completado
Sprint 3	HU4 - Im-	Implementar	8 h / 5 SP	Alumno	0 %
	plementación	modelo			
	de modelos	XGBoost, sin			
	de Machine	considerar			
	Learning	feature			
G : 4 4	TITTA	engineering	6.1 / 9.CD	A 1	0.07
Sprint 4	HU4 - Im-	Evaluar modelo	6 h / 3 SP	Alumno	0 %
	plementación de modelos	$XGBoost, \sin$			
	de modelos de <i>Machine</i>	considerar $feature$			
	$oxed{Learning}$	jeature engineering			
Sprint 4	HU4 - Im-	Implementar	8 h / 5 SP	Alumno	0 %
Sprine 4	plementación	modelo		Alumno	0 70
	de modelos	XGBoost,			
	de $Machine$	considerando			
	Learning	feature			
		engineering			
Sprint 4	HU4 - Im-	Evaluar modelo	6 h / 3 SP	Alumno	0 %
	plementación	XGBoost, consi-	,		
	de modelos	derando feature			
	de <i>Machine</i>	engineering			
	Learning				
Sprint 4	HU4 - Im-	Persistir mode-	4 h / 2 SP	Alumno	0 %
	plementación	los en $GitHub$			
	de modelos				
	de Machine				
	Learning	D .	* 1 / 2 CD	A 1	0.04
Sprint 4	HU4 - Im-	Documentar	5 h / 3 SP	Alumno	0 %
	plementación	pasos y			
	de modelos	decisiones			
	de Machine	tomadas			
Sprint 4	Learning HU5 - Optimi-	Identificar	5 h / 3 SP	Alumno	0 %
Spring 4	zación de hiper-	hiperparáme-	011/001	Tiuiiiio	0 70
	parámetros	tros y rangos			
	r saturation	$de \qquad Logistic$			
		Regression			
Sprint 4	HU5 - Optimi-	Optimizar hi-	6 h / 3 SP	Alumno	0 %
	zación de hiper-	perparámetros	,		
	parámetros	de Logistic			
		$Regression$, \sin			
		considerar			
		feature			
		engineering			



Sprint	HU o fase	Tarea	Horas/SP	Responsable	% Completado
Sprint 4	HU5 - Optimi-	Implementar	4 h / 2 SP	Alumno	0 %
	zación de hiper-	hiperparáme-			
	parámetros	tros más			
		óptimos			
		en Logistic			
		Regression, sin considerar			
		feature			
		feature engineering			
Sprint 4	HU5 - Optimi-	Evaluar modelo	4 h / 2 SP	Alumno	0 %
Spriit 4	zación de hiper-	de Logistic	111/251	THUIIIIO	0 70
	parámetros	Regression con			
	r	hiperparáme-			
		tros óptimos,			
		sin considerar			
		feature			
		engineering			
Sprint 4	HU5 - Optimi-	Optimizar hi-	6 h / 3 SP	Alumno	0 %
	zación de hiper-	perparámetros			
	parámetros	de <i>Logistic</i>			
		Regression,			
		considerando			
		$ig \ feature\ engineering$			
Sprint 4	HU5 - Optimi-	Implementar	4 h / 2 SP	Alumno	0 %
Spriit 4	zación de hiper-	hiperparáme-	411/251	Midillio	0 70
	parámetros	tros más			
	r	óptimos			
		en Logistic			
		Regression,			
		considerando			
		feature			
		engineering			
Sprint 4	HU5 - Optimi-	Evaluar modelo	4 h / 2 SP	Alumno	0 %
	zación de hiper-	de Logistic			
	parámetros	Regression con			
		hiperparáme-			
		tros óptimos, considerando			
		feature			
		feature engineering			
Sprint 5	HU5 - Optimi-	Identificar hi-	5 h / 3 SP	Alumno	0 %
Spriii	zación de hiper-	perparámetros		1114111110	
	parámetros	y rangos de			
	_	SVM			



Sprint	HU o fase	Tarea	Horas/SP	Responsable	% Completado
Sprint 5	HU5 - Optimización de hiperparámetros	Optimizar hiperparámetros de SVM, sin considerar feature engineering	6 h / 3 SP	Alumno	0 %
Sprint 5	HU5 - Optimización de hiperparámetros	Implementar hiperparámetros más óptimos en SVM, sin considerar feature engineering	4 h / 2 SP	Alumno	0%
Sprint 5	HU5 - Optimización de hiperparámetros	Evaluar modelo de SVM con hiperparámetros óptimos, sin considerar feature engineering	4 h / 2 SP	Alumno	0 %
Sprint 5	HU5 - Optimización de hiperparámetros	Optimizar hiperparáme- tros de SVM, considerando feature engineering	6 h / 3 SP	Alumno	0 %
Sprint 5	HU5 - Optimización de hiperparámetros	Implementar hiperparámetros más óptimos en SVM, considerando feature engineering	4 h / 2 SP	Alumno	0 %
Sprint 5	HU5 - Optimización de hiperparámetros	Evaluar modelo de SVM con hiperparámetros óptimos, considerando feature engineering	4 h / 2 SP	Alumno	0 %
Sprint 5	HU5 - Optimización de hiperparámetros	Identificar hiperparámetros y rangos de XGBoost	7 h / 5 SP	Alumno	0 %



Sprint	HU o fase	Tarea	Horas/SP	Responsable	% Completado
Sprint 5	HU5 - Optimi-	Optimizar hi-	8 h / 5 SP	Alumno	0 %
	zación de hiper-	perparámetros			
	parámetros	de XGBoost,			
		sin considerar			
		feature			
		engineering	~ 1 / 2 GD		0.04
Sprint 5	HU5 - Optimi-	Implementar	5 h / 3 SP	Alumno	0 %
	zación de hiper-	hiperparáme-			
	parámetros	tros más			
		óptimos en			
		XGBoost, sin considerar			
		feature			
		engineering			
Sprint 5	HU5 - Optimi-	Evaluar modelo	7 h / 5 SP	Alumno	0 %
Spriii 0	zación de hiper-	de XGBoost	1 11 / 0 01	THUIIIII	070
	parámetros	con hiper-			
	P	parámetros			
		óptimos, sin			
		considerar			
		feature			
		engineering			
Sprint 6	HU5 - Optimi-	Optimizar hi-	8 h / 5 SP	Alumno	0 %
	zación de hiper-	perparámetros			
	parámetros	de XGBoost,			
		considerando			
		feature			
		engineering			
Sprint 6	HU5 - Optimi-	Implementar	5 h / 3 SP	Alumno	0 %
	zación de hiper-	hiperparáme-			
	parámetros	tros más			
		óptimos en			
		XGBoost, considerando			
		feature			
		$\left \begin{array}{c} feature \\ engineering \end{array} \right $			
Sprint 6	HU5 - Optimi-	Evaluar modelo	7 h / 5 SP	Alumno	0 %
Spriii	zación de hiper-	de XGBoost	111/001	1110111110	
	parámetros	con hiper-			
	•	parámetros			
		óptimos,			
		considerando			
		feature			
		engineering			
Sprint 6	HU5 - Optimi-	Documentar	5 h / 3 SP	Alumno	0 %
	zación de hiper-	pasos y			
	parámetros	decisiones			
		tomadas			



Sprint	HU o fase	Tarea	Horas/SP	Responsable	% Completado
Sprint 6	HU6 - Métricas	Obtener	3 h / 2 SP	Alumno	0 %
	de modelos	métricas	·		
		$de extit{F1-score}$			
		para <i>Logistic</i>			
		Regression, sin			
		considerar			
		feature			
		engineering			
Sprint 6	HU6 - Métricas	Obtener	5 h / 3 SP	Alumno	0%
	de modelos	métricas de			
		AUC- ROC			
		para <i>Logistic</i>			
		$Regression, \sin$			
		considerar			
		feature			
		engineering,			
		y graficar			
Sprint 6	HU6 - Métricas	Obtener	3 h / 2 SP	Alumno	0%
	de modelos	métricas			
		$de extit{F1-score}$			
		para Logistic			
		Regression,			
		considerando			
		feature			
		engineering			. ~
Sprint 6	HU6 - Métricas	Obtener	5 h / 3 SP	Alumno	0%
	de modelos	métricas de			
		AUC-ROC			
		para Logistic			
		Regression,			
		considerando			
		feature			
		engineering,			
Conint C	HU6 - Métricas	y graficar Obtener	2 h / 9 CD	A 1	0 %
Sprint 6			3 h / 2 SP	Alumno	U %
	de modelos	l <u> </u>			
		SVM , \sin considerar			
		feature			
		engineering			
Sprint 6	HU6 - Métricas	Obtener	5 h / 3 SP	Alumno	0 %
Shim 0	de modelos	métricas de		Munno	0 70
		AUC-ROC			
		para SVM , sin			
		considerar			
		feature			
		engineering,			
		y graficar			
	l .	V 0	l		



Sprint	HU o fase	Tarea	Horas/SP	Responsable	% Completado
Sprint 6	HU6 - Métricas	Obtener métri-	3 h / 2 SP	Alumno	0 %
	de modelos	cas de F1-score			
		para SVM, con-			
		siderando featu-			
		re engineering			
Sprint 6	HU6 - Métricas	Obtener	5 h / 3 SP	Alumno	0 %
	de modelos	métricas de			
		AUC-ROC			
		para SVM,			
		considerando			
		$\mid feature \mid$			
		engineering,			
Conint 6	HU6 - Métricas	y graficar Obtener	3 h / 2 SP	A 1	0 %
Sprint 6	de modelos		311/231	Alumno	0 70
	de modelos				
		F1-score para $ XGBoost, $ sin			
		considerar			
		feature			
		$\left \begin{array}{c} feature \\ engineering \end{array}\right $			
Sprint 7	HU6 - Métricas	Obtener	5 h / 3 SP	Alumno	0 %
Sprine (de modelos	métricas de			0 70
		AUC-ROC			
		para XGBoost,			
		sin considerar			
		feature			
		engineering,			
		y graficar			
Sprint 7	HU6 - Métricas	Obtener	3 h / 2 SP	Alumno	0 %
	de modelos	métricas de			
		F1-score para			
		XGBoost,			
		considerando			
		feature			
G =	TITIO DEC.	engineering	K 1 / 2 CD	A 1	0.64
Sprint 7	HU6 - Métricas	Obtener	5 h / 3 SP	Alumno	0 %
	de modelos	métricas de			
		AUC-ROC			
		para XGBoost, considerando			
		feature			
		$\begin{array}{c c} feature \\ engineering, \end{array}$			
		y graficar			
Sprint 7	HU6 - Métricas	Comparar	3 h / 2 SP	Alumno	0 %
- r	de modelos	métricas			
		de distintos			
		modelos			
	<u> </u>	l	I		



Sprint	HU o fase	Tarea	Horas/SP	Responsable	% Completado
Sprint 7	HU6 - Métricas	Documentar	5 h / 3 SP	Alumno	0 %
	de modelos	pasos y decisiones			
		tomadas			
Sprint 7	HU7 - Desplie-	Investigar	5 h / 3 SP	Alumno	0 %
opini .	gue en <i>MLFlow</i>	buenas			0,0
		prácticas para			
		despliegues de			
Q		MLFlow	-1 / - CD		2.04
Sprint 7	HU7 - Desplie-	Crear entorno local para	7 h / 5 SP	Alumno	0 %
	gue en <i>MLFlow</i>	local para despliegue			
		MLFlow			
Sprint 7	HU7 - Desplie-	Replicar técni-	8 h / 5 SP	Alumno	0 %
	gue en <i>MLFlow</i>	cas de análisis	,		
		de datos en en-			
	11115 D	torno MLFlow	0.1 / 5 00		2.04
Sprint 7	HU7 - Desplie-	Replicar	8 h / 5 SP	Alumno	0 %
	gue en <i>MLFlow</i>	técnicas de entrenamiento			
		de modelos			
		$de \qquad Logistic$			
		Regression			
		en entorno			
~		MLFlow			. ~
Sprint 7	HU7 - Desplie-	Replicar técni-	8 h / 5 SP	Alumno	0 %
	gue en <i>MLFlow</i>	cas de entrena- miento de mo-			
		delos de SVM			
		en entorno ML -			
		Flow			
Sprint 8	HU7 - Desplie-	Replicar técni-	8 h / 5 SP	Alumno	0 %
	gue en <i>MLFlow</i>	cas de entrena-			
		miento de mo-			
		$\begin{array}{cccccccccccccccccccccccccccccccccccc$			
		torno MLFlow			
Sprint 8	HU7 - Desplie-	Replicar técni-	8 h / 5 SP	Alumno	0 %
	gue en <i>MLFlow</i>	cas de evalua-	,		
		ción de modelos			
		en entorno <i>ML</i> -			
Cm-i1 O		Flow	4 h / 0 CD	Λ1	0.07
Sprint 8	HU7 - Desplie- gue en <i>MLFlow</i>	Ejecutar localmente	4 h / 2 SP	Alumno	0 %
	Suc on MDI tow	el entorno			
		MLFlow			
Sprint 8	HU7 - Desplie-	Validar	4 h / 2 SP	Alumno	0 %
	gue en <i>MLFlow</i>	ejecución local			
		del entorno			
		MLFlow			



Sprint	HU o fase	Tarea	Horas/SP	Responsable	% Completado
Sprint 8	HU7 - Despliegue en <i>MLFlow</i>	Documentar pasos y decisiones tomadas MLFlow	6 h / 3 SP	Alumno	0 %
Sprint 8	HU8 - API para entorno ML-Flow	Investigar como exponer un entorno <i>MLFlow</i> mediante <i>API</i>	5 h / 3 SP	Alumno	0 %
Sprint 8	HU8 - API para entorno ML-Flow	Exponer resultados de modelos explorados en entorno MLFlow mediante API	8 h / 5 SP	Alumno	0 %
Sprint 8	HU8 - API para entorno ML-Flow	$\begin{array}{ll} \text{Exponer} \\ \text{comparación} \\ \text{de} & \text{modelos} \\ \text{en} & \text{entorno} \\ \\ \textit{MLFlow} \\ \text{mediante } \textit{API} \end{array}$	8 h / 5 SP	Alumno	0 %
Sprint 8	HU8 - API para entorno ML-Flow	Documentar pasos y decisiones tomadas	3 h / 2 SP	Alumno	0 %
Sprint 8	HU9 - Implementación de buenas prácticas	Investigar bue- nas prácticas en código <i>Python</i>	4 h / 2 SP	Alumno	0 %
Sprint 9	HU9 - Implementación de buenas prácticas	Aplicar buenas prácticas en código <i>Python</i>	8 h / 5 SP	Alumno	0 %
Sprint 9	HU10 - Documentación	Asegurar que cada decisión tomada haya sido justificada y documentada	6 h / 3 SP	Alumno	0 %
Sprint 9	HU10 - Documentación	Asegurar orto- grafía y formato en documenta- ción	8 h / 5 SP	Alumno	0 %
Sprint 9	HU11 - Validación de <i>API</i> de <i>MLFlow</i>	Validar acceso a modelos explorados mediante API de entorno MLFlow	8 h / 5 SP	Alumno	0%



Sprint	HU o fase	Tarea	Horas/SP	Responsable	% Completado
Sprint 9	HU11 - Valida-	Validar acceso	8 h / 5 SP	Alumno	0 %
	ción de <i>API</i> de	a comparación			
	MLFlow	$ \text{de} \text{modelos} \\ \text{mediante} API $			
		de entorno			
		MLFlow			
Sprint 9	HU11 - Valida-	Crear	5 h / 3 SP	Alumno	0 %
	ción de API de	documentación			
	MLFlow	sobre el uso de			
		API de entorno $MLFlow$			
Sprint 9	Redacción de	Redacción de	6 h / 3 SP	Alumno	0 %
Sprine 5	memoria	sección sobre		Trumino	0 70
		procesamiento			
		$de \ datos$			
Sprint 9	Redacción de	Redacción	3 h / 2 SP	Alumno	0 %
	memoria	de sección			
		$\begin{array}{ccc} { m sobre} & {\it Feature} \\ {\it Engineering} \end{array}$			
Sprint 9	Redacción de	Redacción de	7 h / 5 SP	Alumno	0 %
Spriit 5	memoria	sección sobre		THUIIIIO	0 70
		$implementaci\'{o}n$			
		$de\ modelos$			
Sprint 10	Redacción de	Redacción de	6 h / 3 SP	Alumno	0 %
	memoria	sección sobre			
		optimización de hiperparáme-			
		tros			
Sprint 10	Redacción de	Redacción de	5 h / 3 SP	Alumno	0 %
	memoria	sección sobre	,		
		MLFlow			
Sprint 10	Redacción de	Correcciones	8 h / 5 SP	Alumno	0 %
Conint 10	memoria Propagaión do	generales Confocción do	6 h / 9 CD	Alumas	0 %
Sprint 10	Preparación de presentación fi-	Confección de presentación	6 h / 3 SP	Alumno	U %
	nal	PowerPoint			
Sprint 10	Preparación de	Confección	8 h / 5 SP	Alumno	0 %
	presentación fi-	de video	,		
	nal	demostración			
Sprint 10	Ajustes finales	Realizar correc-	6 h / 3 SP	Alumno	0 %
		ciones a código			
		de análisis de datos			
Sprint 10	Ajustes finales	Realizar correc-	6 h / 3 SP	Alumno	0 %
- F	J 11110100	ciones a código			
		de entrenamien-			
		to de modelos			



Sprint	HU o fase	Tarea	Horas/SP	Responsable	% Completado
Sprint 10	Ajustes finales	Realizar correc-	6 h / 3 SP	Alumno	0 %
		ciones a códi-			
		go de evalua-			
		ción de modelos			
Sprint 10	Ajustes finales	Realizar	6 h / 3 SP	Alumno	0 %
		pruebas finales			
Sprint 10	Ajustes finales	Realizar ajustes	6 h / 3 SP	Alumno	0 %
		en memoria fi-			
		nal			

11. Diagrama de Gantt (sprints)

En el cuadro 4 se muestra el resumen de los sprints del diagrama Gantt.

En el cuadro 5 se especifíca qué referencia cada color utilizado en el diagrama Gantt.

El diagrama Gantt se puede observar en las figuras 2, 3, 4 y 5.

Sprint	Cantidad de Horas	Fecha Inicio	Fecha Fin
0	32 h	24 de junio de 2025	4 de julio de 2025
1	58 h	7 de julio de 2025	25 de julio de 2025
2	60 h	28 de julio de 2025	15 de agosto de 2025
3	64 h	18 de agosto de 2025	5 de septiembre de 2025
4	62 h	8 de septiembre de 2025	26 de septiembre de 2025
5	60 h	29 de septiembre de 2025	17 de octubre de 2025
6	57 h	20 de octubre de 2025	7 de noviembre de 2025
7	60 h	10 de noviembre de 2025	28 de noviembre de 2025
8	58 h	1 de diciembre de 2025	19 de diciembre de 2025
9	60 h	2 de marzo de 2026	20 de marzo de 2026
10	59 h	23 de marzo de 2026	10 de abril de 2026

Cuadro 4. Resumen de sprints del diagrama Gantt.

Color	Descripción		
Azul	Sprint		
Violeta	Tarea técnica		
Verde	Tarea no técnica		
Amarillo	Hito		

Cuadro 5. Referencias de colores utilizados en el diagrama Gantt.

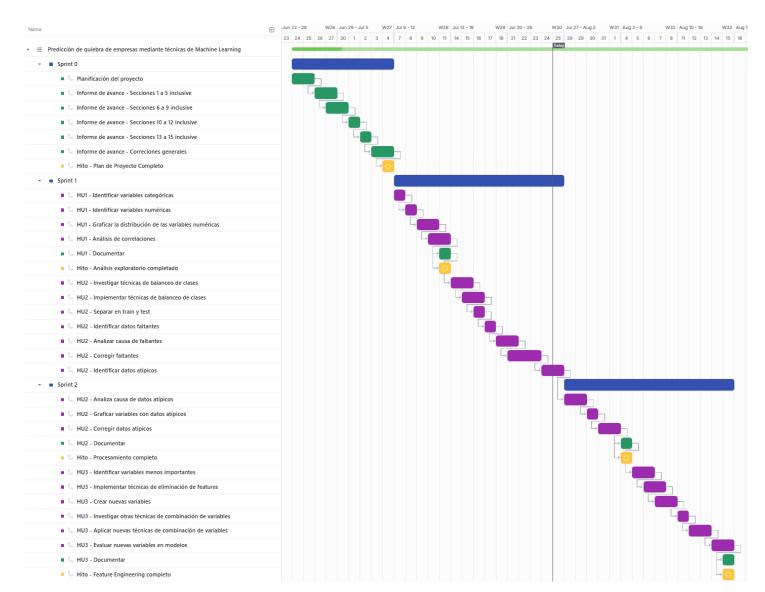


Figura 2. Diagrama de Gantt - Sprints 0, 1 y 2

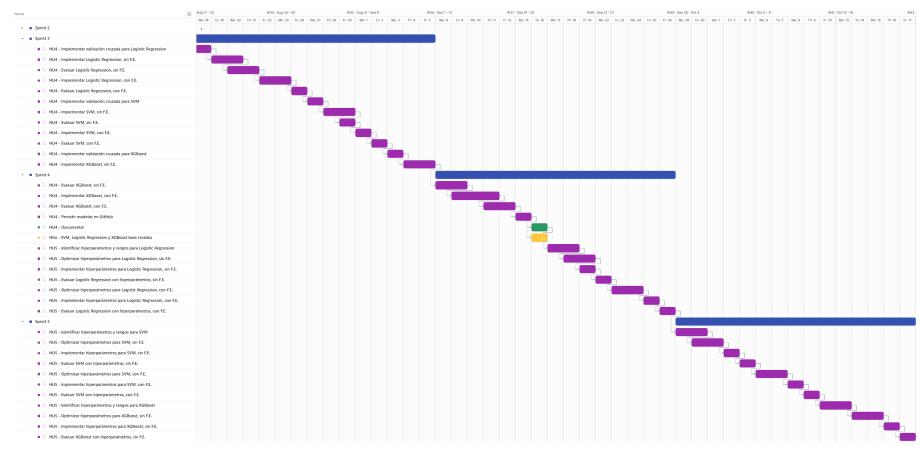


Figura 3. Diagrama de Gantt - Sprints 3, 4 y 5

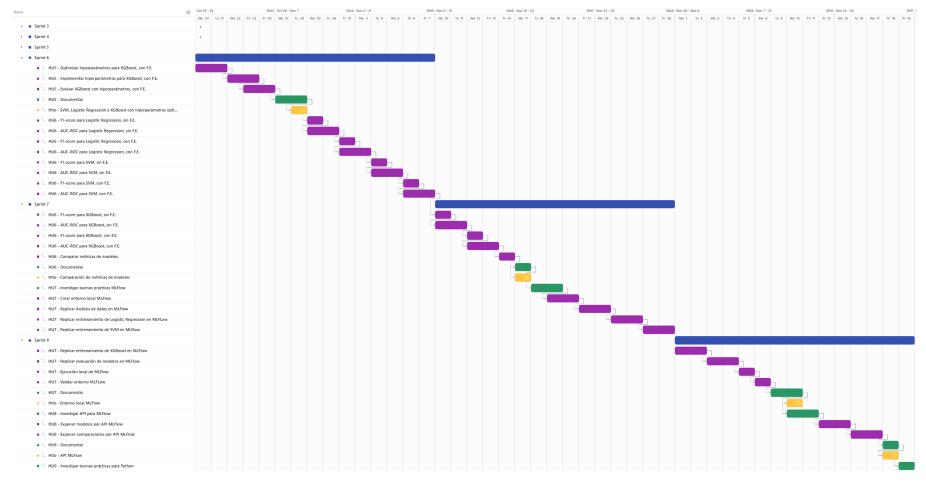


Figura 4. Diagrama de Gantt - Sprints 6, 7 y 8

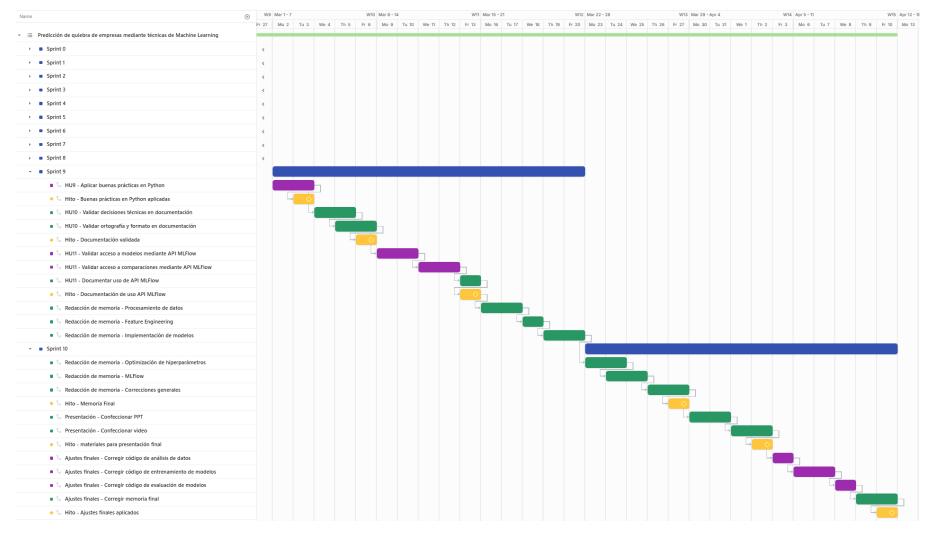


Figura 5. Diagrama de Gantt - Sprints 9 y 10



12. Normativa y cumplimiento de datos (gobernanza)

El proyecto utilizará el dataset de Taiwanese Bankruptcy Prediction.

El dataset está publicado bajo la licencia Creative Commons Attribution 4.0 International, la cual permite la copia, distribución, exhibición y ejecución de los datos siempre y cuando se dé crédito al autor y/o publicador, siendo en este caso el repositorio de Machine Learning de UC Irvine.

La información que presenta el dataset fue recolectada y publicada por el Taiwan Economic Journal. Tal como se menciona en la sección de financial data de su sitio web, todos los datos financieros que ellos presentan se obtienen de:

- Informes auditados por contadores públicos certificados.
- Datos mensuales sobre ingresos proporcionados por empresas que cotizan en la bolsa de Taiwán

Como dato adicional, en el *dataset* no se mencionan los nombres de las empresas ni datos similares, solo presenta información financiera.

Por todo lo mencionado anteriormente podemos garantizar que no hay inconvenientes al utilizar el dataset de Taiwanese Bankruptcy Prediction durante el desarrollo y publicación del proyecto. Solamente hay que dar crédito a su publicador (repositorio de Machine Learning de UC Irvine).

13. Gestión de riesgos

a) Identificación de los riesgos (al menos cinco) y estimación de sus consecuencias:

Riesgo 1: detallar el riesgo (riesgo es algo que si ocurre altera los planes previstos de forma negativa)

- Severidad (S): mientras más severo, más alto es el número (usar números del 1 al 10). Justificar el motivo por el cual se asigna determinado número de severidad (S).
- Probabilidad de ocurrencia (O): mientras más probable, más alto es el número (usar del 1 al 10).

Justificar el motivo por el cual se asigna determinado número de (O).

Riesgo 2:

- Severidad (S): X.
 Justificación...
- Ocurrencia (O): Y. Justificación...

Riesgo 3:



- Severidad (S): X. Justificación...
- Ocurrencia (O): Y. Justificación...
- b) Tabla de gestión de riesgos: (El RPN se calcula como RPN=SxO)

Riesgo	S	О	RPN	S*	O*	RPN*

Criterio adoptado:

Se tomarán medidas de mitigación en los riesgos cuyos números de RPN sean mayores a...

Nota: los valores marcados con (*) en la tabla corresponden luego de haber aplicado la mitigación.

c) Plan de mitigación de los riesgos que originalmente excedían el RPN máximo establecido:

Riesgo 1: plan de mitigación (si por el RPN fuera necesario elaborar un plan de mitigación). Nueva asignación de S y O, con su respectiva justificación:

- Severidad (S*): mientras más severo, más alto es el número (usar números del 1 al 10). Justificar el motivo por el cual se asigna determinado número de severidad (S).
- Probabilidad de ocurrencia (O*): mientras más probable, más alto es el número (usar del 1 al 10). Justificar el motivo por el cual se asigna determinado número de (O).

Riesgo 2: plan de mitigación (si por el RPN fuera necesario elaborar un plan de mitigación).

Riesgo 3: plan de mitigación (si por el RPN fuera necesario elaborar un plan de mitigación).

14. Sprint Review

La revisión de sprint (*Sprint Review*) es una práctica fundamental en metodologías ágiles. Consiste en revisar y evaluar lo que se ha completado al finalizar un sprint. En esta instancia, se presentan los avances y se verifica si las funcionalidades cumplen con los criterios de aceptación establecidos. También se identifican entregables parciales y se consideran ajustes si es necesario.

Aunque el proyecto aún se encuentre en etapa de planificación, esta sección permite proyectar cómo se evaluarán las funcionalidades más importantes del backlog. Esta mirada anticipada favorece la planificación enfocada en valor y permite reflexionar sobre posibles obstáculos.

Objetivo: anticipar cómo se evaluará el avance del proyecto a medida que se desarrollen las funcionalidades, utilizando como base al menos cuatro historias de usuario del *Product Backlog*.



Seleccionar al menos 4 HU del Product Backlog. Para cada una, completar la siguiente tabla de revisión proyectada:

Formato sugerido:

HU seleccionada	Tareas asociadas	Entregable esperado	¿Cómo sabrás que está cumplida?	Observaciones o riesgos	
HU1	Tarea 1	- Módulo funcional	Cumple criterios de aceptación	Falta validar con	
	Tarea 2		definidos	er tutor	
HU3	Tarea 1	Reporte generado	Exportación	Requiere datos	
	Tarea 2	Teperre generado	disponible y clara	reales	
HU5	Tarea 1	Panel de gestión	Roles diferenciados	Riesgo en	
1100	Tarea 2	Taner de gestion	operativos	integración	
HU7	Tarea 1	Informe	PDF con gráficos	Puede faltar	
	Tarea 2	trimestral	y evolución	tiempo para ajustes	

15. Sprint Retrospective

La retrospectiva de sprint es una práctica orientada a la mejora continua. Al finalizar un sprint, el equipo (o el alumno, si trabaja de forma individual) reflexiona sobre lo que funcionó bien, lo que puede mejorarse y qué acciones concretas pueden implementarse para trabajar mejor en el futuro.

Durante la cursada se propuso el uso de la **Estrella de la Retrospectiva**, que organiza la reflexión en torno a cinco ejes:

- ¿Qué hacer más?
- ¿Qué hacer menos?
- ¿Qué mantener?
- ¿Qué empezar a hacer?
- ¿Qué dejar de hacer?

Aun en una etapa temprana, esta herramienta permite que el alumno planifique su forma de trabajar, identifique anticipadamente posibles dificultades y diseñe estrategias de organización personal.

Objetivo: reflexionar sobre las condiciones iniciales del proyecto, identificando fortalezas, posibles dificultades y estrategias de mejora, incluso antes del inicio del desarrollo.

Completar la siguiente tabla tomando como referencia los cinco ejes de la Estrella de la Retrospectiva (Starfish o estrella de mar). Esta instancia te ayudará a definir buenas prácticas



desde el inicio y prepararte para enfrentar el trabajo de forma organizada y flexible. Se deberá completar la tabla al menos para 3 sprints técnicos y 1 no técnico.

Formato sugerido:

Sprint tipo y N°	¿Qué hacer más?	¿Qué hacer menos?	¿Qué mantener?	¿Qué empezar a hacer?	¿Qué dejar de hacer?
Sprint técnico - 1	Validaciones continuas con el alumno	Cambios sin versión registrada	Pruebas con datos simulados	Documentar cambios propuestos	Ajustes sin análisis de impacto
Sprint técnico - 2	Verificar configuraciones en múltiples escenarios	Modificar parámetros sin guardar historial	Perfiles reutilizables	Usar logs para configuración	Repetir pruebas manuales innecesarias
Sprint técnico - 8	Comparar correlaciones con casos previos	Cambiar parámetros sin justificar	Revisión cruzada de métricas	Anotar configuraciones usadas	Trabajar sin respaldo de datos
Sprint no técnico - 12 (por ej.: "De- fensa")	Ensayos orales con feedback	Cambiar contenidos en la memoria	Material visual claro	Dividir la presentación por bloques	Agregar gráficos difíciles de explicar