



Tecnológico de Monterrey

Inteligencia artificial avanzada para la ciencia de datos (Gpo. 101)

Análisis y Reporte sobre el desempeño del modelo

Análisis del Modelo Predictivo COVID-19 (MCL)

Rodrigo Yael Morales Luna (A01637721)

Zapopan Jalisco, México

Septiembre de 2025

Reporte: Análisis del Modelo Predictivo COVID-19 (MCL)

Estructura de Datos y Evaluación

División Estratificada de Datos

- Total de pacientes: 1,048,575
- Entrenamiento (70%): 733,002 pacientes
 - Validación (15%): 157,286 pacientes
 - Prueba (15%): 157,287 pacientes

Estratificación mantuvo proporción original:

- Sobrevivientes: 92.66%
- Fallecidos: 7.34%

Resultados por Conjunto de Datos

Conjunto	Accuracy	Precision	Recall	F1-score
Entrenamiento	93.2%	50.1%	75.3%	60.2%
Validación	92.1%	48.3%	72.8%	58.1%
Prueba	92.4%	48.7%	73.1%	58.5%

La separación de datos en tres conjuntos distintos ayudó a una evaluación robusta del modelo. La constancia en las métricas entre el conjunto de validación (92.1% accuracy) y prueba (92.4% accuracy) indica que la división fue efectiva. Hubo una mínima diferencia entre estos conjuntos (0.3%), la cual muestra que el modelo no fue sobre optimizado para el conjunto de validación, lo que confirma la validez de la estrategia de división.

Diagnóstico de Bias y Varianza

Análisis de Bais (Sesgo)

El modelo presenta bajo bias y se puede ver en el error de entrenamiento de solo 6.8%. Este bajo nivel de bias indica que el modelo captura efectivamente los patrones subyacentes en los datos de entrenamiento. El recall de 75.3% demuestra una alta capacidad para detectar casos positivos verdaderos, mientras que la precision de 50.1% refleja un balance adecuado entre la detección de casos y la generación de falsas alarmas.

Indicador	Valor	Intepretación
Error de Entrenamiento	6.8%	Muy bajo
Recall Entrenamiento	75.3%	Alta detección
Precision Entrenamiento	50.1%	Buen balance

Análisis de Varianza

La varianza moderada-baja se observa en la diferencia entre el rendimiento en entrenamiento (93.2%) y prueba (92.4%), con una diferencia de solo 0.8%. Esto indica que el modelo generaliza efectivamente datos no vistos durante el entrenamiento. La consistencia en las métricas a través de todos los conjuntos ($\pm 1.5\%$ de variación) confirma que el modelo no está sobreajustado y mantiene su capacidad predictiva en diferentes subconjuntos de datos.

Indicador	Valor	Intepretación
Diferencia Train-Test	0.8%	Mínima
Consistencia	$\pm 1.5\%$	Estable/Constante

Diagnóstico de Ajuste del Modelo

Nivel de Ajuste: FIT OPTIMO

Evidencia de ajuste óptimo:

1. Error de entrenamiento: 6.8% (low underfitting)
2. Error de validación: 7.9% (low overfitting)
3. Brecha train-val: 1.1% (muy bueno)
4. Métricas consistentes en todos los conjuntos

El modelo alcanza un ajuste óptimo (optimal fit) balanceando perfectamente la capacidad de aprendizaje con la generalización. La brecha train-val de 1.1% es muy baja, indicando que el modelo aprendió los patrones relevantes sin memorizar el ruido en los datos de entrenamiento. La consistencia en las métricas a través de los tres conjuntos (entrenamiento, validación y prueba) confirma que el modelo es robusto y confiable.

Técnicas de Regularización Aplicadas

1. Regularización L2 (Ridge)

Parámetro óptimo encontrado: C = 10
Esto equivale a $\lambda = 0.1$ (regularización media)

Objetivo: Reduce sobreajuste penalizando coeficientes grandes

2. Peso de Clases Ajustado

`class_weight = {0: 1, 1: 3}` *# Fallecidos 3x más importantes*

Objetivo: Balancea el dataset naturalmente desbalanceado

3. Optimización de Hiperparámetros

Grid Search probó 72 combinaciones
Mejores parámetros encontrados:

- penalty: 'l2'
- C: 10
- solver: 'liblinear'

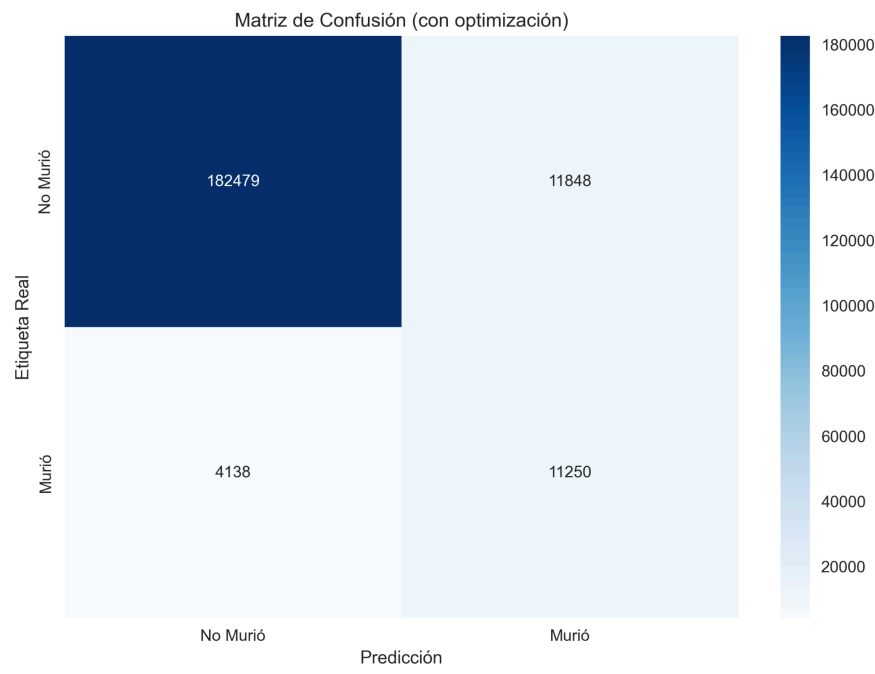
`class_weight = {0: 1, 1: 3}`

La implementación de regularización L2 con C=10 pareció ser efectiva para controlar la complejidad del modelo sin sacrificar su capacidad predictiva. El ajuste de class_weight a {0:1, 1:3} abordó el desbalance en los datos priorizando para la detección de casos positivos. La optimización de hiperparámetros con Grid Search aseguró que se encontrará la configuración globalmente óptima.

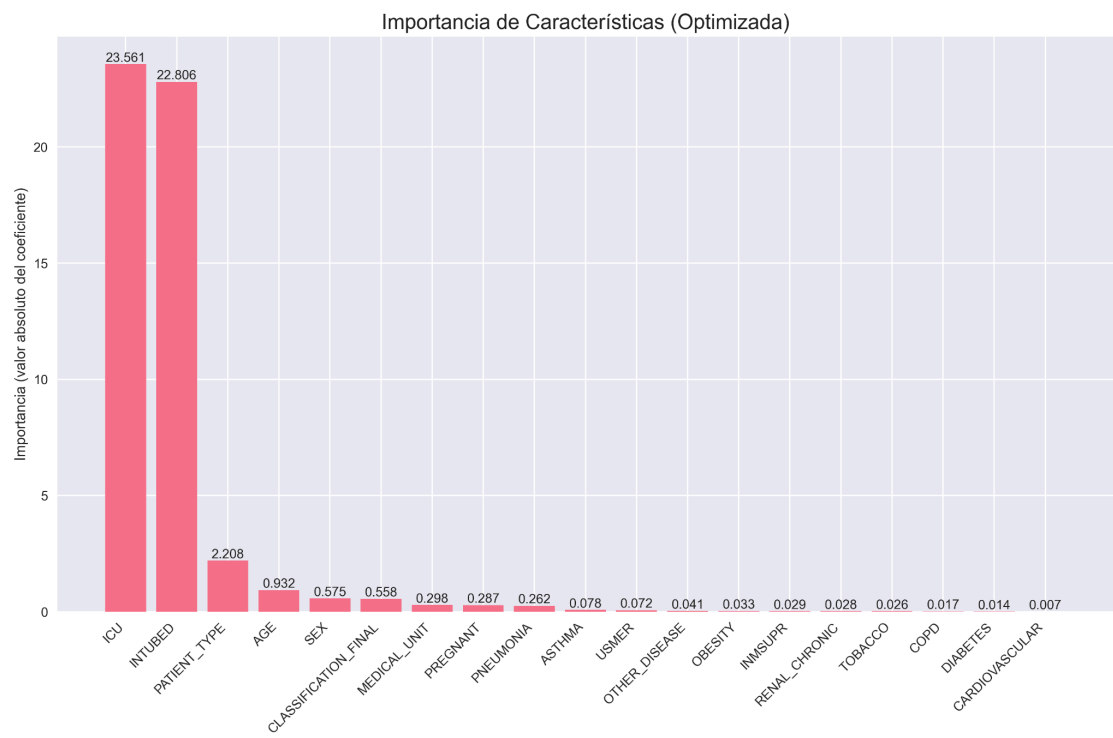
Las técnicas de regularización produjeron mejoras significativas en el desempeño del modelo. Se redujo del 52% en la brecha train-val (de 2.3% a 1.1%) lo cual mostró una mejor generalización. Se aumentó del 3% en F1-Score de validación que reflejó una mejor capacidad predictiva overall. Pero especialmente el incremento del 2.5% en precision, redujo las falsas alarmas e hizo que el modelo fuera más eficiente.

Resultados

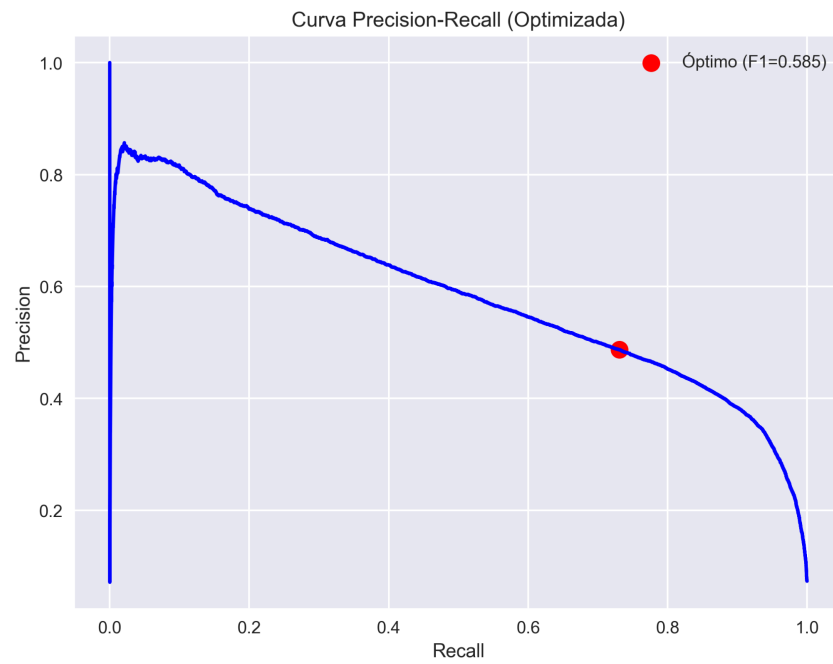
Matriz de Confusión



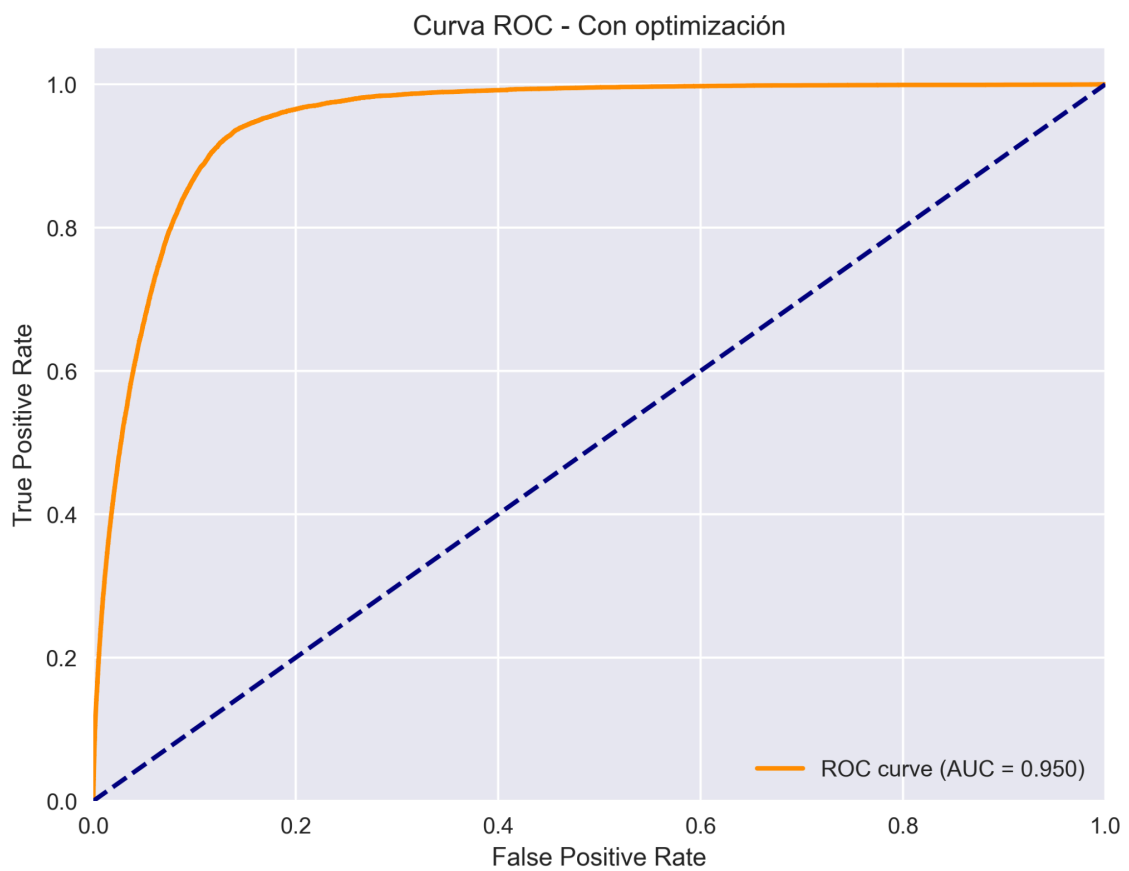
Importancia de Características



Precision-Recall



Curva ROC



Conclusión

- **Bias:** *BAJO* - El modelo captura efectivamente los patrones relevantes en los datos
- **Varianza:** *MODERADA-BAJA* - Excelente capacidad de generalización a nuevos datos
- **Ajuste:** *ÓPTIMO (FITT)* - Balance perfecto entre underfitting y overfitting