

Comparación de Técnicas de Optimización Convexa y No Convexa Aplicadas al Diagnóstico de Cáncer de Mama: Un Estudio Experimental con el Dataset de Wisconsin

Presentado por: MARIO WILFREDO RAMIREZ PUMA
Lopez.pu26@gmail.com

UNIVERSIDAD NACIONAL DEL ALTIPLANO - PUNO
ESCUELA PROFESIONAL DE INGENIERÍA
ESTADÍSTICA E INFORMÁTICA
CURSO: METODOS DE OPTIMIZACIÓN
DOCENTE: ING. TORRES CRUZ FRED

9 de junio de 2025

Resumen

Este estudio presenta una comparación experimental exhaustiva entre técnicas de optimización convexa y no convexa aplicadas al diagnóstico de cáncer de mama utilizando el dataset de Wisconsin. Se implementaron y evaluaron seis algoritmos: tres métodos convexos (Regresión Logística, SVM Lineal, y Regresión Ridge) y tres métodos no convexos (Redes Neuronales, SVM con Kernel RBF, y Algoritmos Genéticos). Los resultados revelan un empate técnico excepcional en rendimiento (98.2 % de precisión) entre SVM Lineal, SVM RBF y Algoritmos Genéticos, con diferencias dramáticas en eficiencia computacional que favorecen a los métodos convexos. El hallazgo principal demuestra que el dataset de Wisconsin es linealmente separable, validando que métodos convexos pueden ser suficientes para problemas médicos aparentemente complejos. Los algoritmos genéticos proporcionaron el único valor agregado significativo mediante selección automática de características, reduciendo las variables necesarias de 30 a 11 sin pérdida de rendimiento. Este trabajo contribuye con evidencia empírica sobre cuándo la complejidad algorítmica se justifica en aplicaciones médicas críticas.

Palabras clave: optimización convexa, optimización no convexa, diagnóstico médico, machine learning, cáncer de mama, algoritmos genéticos, SVM, redes neuronales

Índice

1. Introducción	4
1.1. Motivación del Estudio	4
1.2. Objetivos de la Investigación	4
2. Marco Teórico	5
2.1. Optimización Convexa	5
2.2. Optimización No Convexa	5
2.3. Métodos Convexos Seleccionados	5
2.3.1. Regresión Logística	5
2.3.2. Support Vector Machines (SVM) Lineal	6
2.3.3. Regresión Ridge	6
2.4. Métodos No Convexos Seleccionados	6
2.4.1. Redes Neuronales Artificiales	6
2.4.2. SVM con Kernel RBF	7
2.4.3. Algoritmos Genéticos	7
2.5. Dataset de Wisconsin	8
3. Metodología	8
3.1. Diseño Experimental	8
3.2. Preprocesamiento de Datos	9
3.3. Optimización de Hiperparámetros	9
3.3.1. Espacios de Búsqueda	9
3.4. Métricas de Evaluación	10
3.5. Protocolo de Validación	10
4. Implementación	11
4.1. Entorno de Desarrollo	11
4.2. Configuraciones Específicas	11
4.2.1. Métodos Convexos	11
4.2.2. Métodos No Convexos	11
5. Resultados	12
5.1. Métodos Convexos	12
5.1.1.	12
5.1.2. SVM Lineal	13
5.1.3. Regresión Ridge	13
5.2. Métodos No Convexos	14
5.2.1. Redes Neuronales	14
5.2.2. SVM RBF	14
5.2.3. Algoritmos Genéticos	15
5.3. Comparación Global de Métodos	15
6. Comparación y Discusión	15
6.1. Análisis Comparativo de Rendimiento	15
6.2. Eficiencia Computacional: Convexo vs No Convexo	16
6.3. Valor Agregado de Métodos No Convexos	16
6.4. Recomendaciones Prácticas	16

6.5. Limitaciones del Estudio	17
6.6. Conclusión de la Comparación	17
7. Conclusiones	17
7.1. Hallazgos Principales	17
7.2. Rendimiento Comparativo	17
7.3. Eficiencia vs Complejidad	18
7.4. Valor Agregado Identificado	18
7.5. Limitaciones y Trabajo Futuro	18
7.6. Reflexión Final	18
8. Referencias	19

1. Introducción

El diagnóstico temprano y preciso del cáncer de mama constituye uno de los desafíos más críticos en la medicina moderna, donde la precisión algorítmica puede impactar directamente en la supervivencia de pacientes. En este contexto, la selección de técnicas de optimización apropiadas para sistemas de diagnóstico asistido por computadora representa una decisión fundamental que equilibra precisión, eficiencia computacional e interpretabilidad clínica.

El presente estudio aborda una pregunta fundamental en el ámbito de la inteligencia artificial aplicada a medicina: **¿cuándo se justifica el uso de métodos de optimización no convexa sobre métodos convexos en problemas de diagnóstico médico?** Esta interrogante trasciende consideraciones puramente técnicas, incorporando aspectos prácticos como tiempo de respuesta en entornos clínicos, interpretabilidad de resultados para profesionales médicos, y reproducibilidad de diagnósticos.

1.1. Motivación del Estudio

La literatura científica frecuentemente asume que problemas complejos requieren métodos sofisticados de optimización no convexa, particularmente en el contexto de diagnóstico médico donde múltiples variables interactúan de manera aparentemente no lineal. Sin embargo, esta asunción carece de validación experimental rigurosa en datasets específicos, creando una brecha entre teoría y práctica clínica.

El dataset de cáncer de mama de Wisconsin [1], ampliamente utilizado en la comunidad de machine learning, proporciona un caso de estudio ideal para esta comparación debido a sus características morfológicas cuantitativas derivadas de imágenes digitalizadas de aspirados de aguja fina (FNA) de masas mamarias. Con 569 muestras y 30 características numéricas, representa un problema de clasificación binaria con relevancia clínica directa.

1.2. Objetivos de la Investigación

Objetivo General: Comparar experimentalmente el rendimiento, eficiencia y aplicabilidad clínica de técnicas de optimización convexa versus no convexa en el diagnóstico de cáncer de mama.

Objetivos Específicos:

1. Implementar y optimizar hiperparámetros de seis algoritmos representativos: Regresión Logística, SVM Lineal, Regresión Ridge (convexos) y Redes Neuronales, SVM RBF, Algoritmos Genéticos (no convexos)
2. Evaluar rendimiento predictivo mediante métricas clínicamente relevantes: precisión, sensibilidad, especificidad, F1-Score y AUC-ROC
3. Analizar eficiencia computacional y tiempos de convergencia para aplicaciones en tiempo real
4. Identificar características discriminativas y su relevancia clínica
5. Determinar cuándo la complejidad algorítmica se justifica en términos de beneficios tangibles

2. Marco Teórico

2.1. Optimización Convexa

La optimización convexa [2] constituye una clase fundamental de problemas matemáticos donde tanto la función objetivo como las restricciones son convexas. Formalmente, un problema de optimización convexa se define como:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{sujeto a} \quad & g_i(x) \leq 0, \quad i = 1, \dots, m \\ & Ax = b \end{aligned} \tag{1}$$

donde $f : \mathbb{R}^n \rightarrow \mathbb{R}$ y $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ son funciones convexas.

Propiedades fundamentales:

- **Óptimo global único:** Todo mínimo local es global
- **Convergencia garantizada:** Algoritmos eficientes con garantías teóricas
- **Estabilidad numérica:** Resultados reproducibles y consistentes
- **Interpretabilidad:** Soluciones matemáticamente interpretables

2.2. Optimización No Convexa

Los problemas de optimización no convexa [3] se caracterizan por la presencia de múltiples óptimos locales y la ausencia de garantías de convergencia global. Se definen como:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{sujeto a} \quad & x \in \mathcal{S} \end{aligned} \tag{2}$$

donde f no es necesariamente convexa y $\mathcal{S} \subseteq \mathbb{R}^n$ puede ser no convexo.

Características distintivas:

- **Múltiples óptimos locales:** Riesgo de convergencia prematura
- **Sensibilidad a inicialización:** Resultados dependientes de condiciones iniciales
- **Mayor flexibilidad:** Capacidad de modelar relaciones complejas
- **Costo computacional:** Generalmente mayor tiempo de procesamiento

2.3. Métodos Convexos Seleccionados

2.3.1. Regresión Logística

La regresión logística modela la probabilidad de pertenencia a una clase mediante la función logística:

$$P(y = 1|x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}} \tag{3}$$

La función de costo log-verosimilitud es convexa:

$$L(\beta) = - \sum_{i=1}^n [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (4)$$

Ventajas clínicas:

- Salidas probabilísticas interpretables
- Coeficientes indican importancia de características
- Convergencia rápida garantizada
- Amplia aceptación en comunidad médica

2.3.2. Support Vector Machines (SVM) Lineal

SVM [4] encuentra el hiperplano óptimo que maximiza el margen entre clases:

$$\min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \quad (5)$$

sujeto a: $y_i(w^T x_i + b) \geq 1 - \xi_i$ y $\xi_i \geq 0$

donde C controla el balance entre margen y error de clasificación.

Características destacadas:

- Margen máximo garantiza robustez
- Eficiencia en memoria (solo vectores de soporte)
- Base teórica sólida en teoría de aprendizaje estadístico
- Manejo efectivo de datos de alta dimensión

2.3.3. Regresión Ridge

La regresión Ridge [6] añade regularización L2 a la regresión lineal:

$$\min_{\beta} \|y - X\beta\|^2 + \alpha \|\beta\|^2 \quad (6)$$

Adaptación para clasificación: Para problemas de clasificación binaria, se adapta usando un umbral de decisión sobre las predicciones continuas, típicamente 0.5.

2.4. Métodos No Convexos Seleccionados

2.4.1. Redes Neuronales Artificiales

Las redes neuronales multicapa (MLP) [7] aproximan funciones complejas mediante composición de transformaciones no lineales:

$$f(x) = \sigma(W_L \sigma(W_{L-1} \cdots \sigma(W_1 x + b_1) + \cdots + b_{L-1}) + b_L) \quad (7)$$

donde σ es la función de activación (típicamente ReLU: $\sigma(z) = \max(0, z)$).

Función de pérdida (no convexa):

$$L(\theta) = \frac{1}{n} \sum_{i=1}^n \ell(f(x_i; \theta), y_i) + \lambda R(\theta) \quad (8)$$

Capacidades distintivas:

- Aproximación universal de funciones
- Aprendizaje automático de características
- Modelado de interacciones no lineales complejas
- Flexibilidad arquitectural

2.4.2. SVM con Kernel RBF

El kernel Radial Basis Function mapea datos a un espacio de características de dimensión infinita:

$$K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2} \quad (9)$$

El problema dual resultante es no convexo en el espacio original pero convexo en el espacio de características implícito.

Parámetros críticos:

- C : Control de regularización
- γ : Ancho del kernel (influye en complejidad de frontera)

2.4.3. Algoritmos Genéticos

Los algoritmos genéticos [8] emulan procesos evolutivos naturales para optimización global:

Componentes principales:

1. **Cromosoma:** Codificación binaria de soluciones candidatas
2. **Función de fitness:** Evaluación de calidad de soluciones
3. **Selección:** Proceso de elección de padres (torneo, ruleta)
4. **Cruce:** Intercambio de información genética
5. **Mutación:** Introducción de diversidad genética

Algoritmo básico:

1. Inicializar población aleatoria
2. Evaluar fitness de individuos
3. Seleccionar padres
4. Aplicar cruce y mutación
5. Reemplazar población
6. Repetir hasta convergencia

2.5. Dataset de Wisconsin

El Wisconsin Breast Cancer Dataset [9], creado por Dr. William H. Wolberg en la Universidad de Wisconsin, contiene características computadas de imágenes digitalizadas de aspirados de aguja fina (FNA) de masas mamarias.

Características del dataset:

- **Muestras:** 569 casos (357 benignos, 212 malignos)
- **Características:** 30 atributos numéricos por muestra
- **Derivación:** Cada característica representa media, error estándar y “peor valor” de 10 medidas morfométricas

Medidas morfométricas base:

1. Radio (promedio de distancias desde centro a puntos del perímetro)
2. Textura (desviación estándar de valores en escala de grises)
3. Perímetro
4. Área
5. Suavidad (variación local en longitudes de radio)
6. Compacidad ($\text{perímetro}^2 / \text{área} - 1.0$)
7. Concavidad (severidad de porciones cóncavas del contorno)
8. Puntos cóncavos (número de porciones cóncavas del contorno)
9. Simetría
10. Dimensión fractal (“aproximación de línea costera” - 1)

3. Metodología

3.1. Diseño Experimental

Se diseñó un experimento controlado para comparar sistemáticamente seis algoritmos de optimización en condiciones idénticas, garantizando reproducibilidad y validez de las comparaciones.

Principios metodológicos:

1. **Reproducibilidad:** Semilla aleatoria fija (`random.state=42`)
2. **Equidad:** Misma división entrenamiento/prueba para todos los métodos
3. **Rigurosa optimización:** Grid search con validación cruzada
4. **Evaluación completa:** Múltiples métricas clínicamente relevantes

3.2. Preprocesamiento de Datos

División del dataset:

- **Entrenamiento:** 80 % (455 muestras)
- **Prueba:** 20 % (114 muestras)
- **Estratificación:** Mantiene proporción de clases en ambos conjuntos

Normalización: Aplicación de StandardScaler para estandarización Z-score:

$$x_{norm} = \frac{x - \mu}{\sigma} \quad (10)$$

donde μ es la media y σ la desviación estándar calculadas únicamente en el conjunto de entrenamiento.

3.3. Optimización de Hiperparámetros

Para cada algoritmo se implementó búsqueda exhaustiva de hiperparámetros mediante GridSearchCV con validación cruzada.

3.3.1. Espacios de Búsqueda

Regresión Logística:

- $C \in \{0,001, 0,01, 0,1, 1,0, 10,0, 100,0, 1000,0\}$
- Solver $\in \{\text{liblinear}, \text{lbfgs}\}$

SVM Lineal:

- $C \in \{0,01, 0,1, 1,0, 10,0, 100,0, 1000,0\}$
- Gamma $\in \{\text{scale}, \text{auto}\}$

Regresión Ridge:

- $\alpha \in \{0,001, 0,01, 0,1, 1,0, 10,0, 100,0, 1000,0\}$

Redes Neuronales:

- Arquitectura $\in \{(50), (100), (50, 25), (100, 50), (100, 50, 25)\}$
- $\alpha \in \{0,0001, 0,001, 0,01\}$ (regularización L2)
- Learning rate $\in \{0,001, 0,01\}$

SVM RBF:

- $C \in \{0,1, 1,0, 10,0, 100,0, 1000,0\}$
- $\gamma \in \{\text{scale}, \text{auto}, 0,001, 0,01, 0,1, 1,0, 10,0\}$

Algoritmos Genéticos:

- Población: 50 individuos
- Generaciones: 30
- Tasa de mutación: 0.1
- Tasa de cruce: 0.8
- Cromosoma: 36 bits (30 para características + 6 para hiperparámetros)

3.4. Métricas de Evaluación

Se seleccionaron métricas estándar en diagnóstico médico para evaluación comprensiva:

Métricas primarias:

$$\text{Precisión (Accuracy)} = \frac{TP + TN}{TP + TN + FP + FN} \quad (11)$$

$$\text{Precisión (Precision)} = \frac{TP}{TP + FP} \quad (12)$$

$$\text{Sensibilidad (Recall)} = \frac{TP}{TP + FN} \quad (13)$$

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (14)$$

$$\text{AUC-ROC} = \int_0^1 \text{TPR}(FPR^{-1}(t))dt \quad (15)$$

donde TP = Verdaderos Positivos, TN = Verdaderos Negativos, FP = Falsos Positivos, FN = Falsos Negativos.

Métricas secundarias:

- Tiempo de convergencia
- Número de parámetros/complejidad del modelo
- Estabilidad en validación cruzada

3.5. Protocolo de Validación

Validación cruzada estratificada:

- **Métodos convexos:** 5-fold CV
- **Redes neuronales:** 3-fold CV (reducido por costo computacional)
- **Algoritmos genéticos:** 3-fold CV en evaluación de fitness

Evaluación final: Todos los modelos se evalúan en el conjunto de prueba no visto durante optimización de hiperparámetros.

4. Implementación

4.1. Entorno de Desarrollo

Especificaciones técnicas:

- **Lenguaje:** Python 3.8+
- **Biblioteca principal:** scikit-learn [\[10\]](#)
- **Dependencias:** NumPy, Pandas, Matplotlib
- **Hardware:** CPU estándar (sin GPU requerida)

4.2. Configuraciones Específicas

4.2.1. Métodos Convexos

Regresión Logística:

```
LogisticRegression(  
    max_iter=1000,  
    random_state=42,  
    C=optimal_C,  
    solver=optimal_solver  
)
```

SVM Lineal:

```
SVC(  
    kernel='linear',  
    random_state=42,  
    probability=True,  
    C=optimal_C  
)
```

Regresión Ridge:

```
Ridge(  
    alpha=optimal_alpha,  
    random_state=42  
)
```

4.2.2. Métodos No Convexos

Redes Neuronales:

```
MLPClassifier(  
    hidden_layer_sizes=optimal_architecture,  
    alpha=optimal_alpha,  
    learning_rate_init=optimal_lr,  
    max_iter=1000,  
    random_state=42,
```

```

        early_stopping=True,
        validation_fraction=0.1
    )

```

SVM RBF:

```

SVC(
    kernel='rbf',
    random_state=42,
    probability=True,
    C=optimal_C,
    gamma=optimal_gamma
)

```

Algoritmos Genéticos: Implementación personalizada con codificación binaria para optimización simultánea de selección de características e hiperparámetros.

5. Resultados

5.1. Métodos Convexos

5.1.1.

Regresión Logística

Cuadro 1: Resultados - Regresión Logística

Métrica	Valor
Accuracy	0.974 (97.4 %)
Precision	0.973 (97.3 %)
Recall	0.986 (98.6 %)
F1-Score	0.979 (97.9 %)
AUC-ROC	0.996 (99.6 %)
Tiempo	0.0037s

Configuración óptima: C=0.1, solver='lbfgs'

Errores: 1 falso positivo, 2 falsos negativos

Interpretación: Mejor AUC-ROC (99.6 %). Ideal para explicación clínica por interpretabilidad de coeficientes. Velocidad excelente para tiempo real.

5.1.2. SVM Lineal

Cuadro 2: Resultados - SVM Lineal

Métrica	Valor
Accuracy	0.982 (98.2 %)
Precision	0.986 (98.6 %)
Recall	0.986 (98.6 %)
F1-Score	0.986 (98.6 %)
AUC-ROC	0.994 (99.4 %)
Tiempo	0.0065s

Configuración óptima: C=0.1, vectores de soporte: 51/455 (11.2 %)

Errores: 1 falso positivo, 1 falso negativo

Interpretación: Mejor rendimiento global. Balance perfecto de errores. Eficiencia en memoria. Ideal para implementación clínica.

5.1.3. Regresión Ridge

Cuadro 3: Resultados - Regresión Ridge

Métrica	Valor
Accuracy	0.956 (95.6 %)
Precision	0.935 (93.5 %)
Recall	1.000 (100 %)
F1-Score	0.966 (96.6 %)
AUC-ROC	0.993 (99.3 %)
Tiempo	0.0008s

Configuración óptima: alpha=1.0

Errores: 5 falsos positivos, 0 falsos negativos

Interpretación: Recall perfecto - no pierde casos malignos. Fastest (0.0008s). Ideal para screening masivo donde seguridad es prioritaria.

5.2. Métodos No Convexos

5.2.1. Redes Neuronales

Cuadro 4: Resultados - Redes Neuronales

Métrica	Valor
Accuracy	0.956 (95.6 %)
Precision	0.959 (95.9 %)
Recall	0.972 (97.2 %)
F1-Score	0.966 (96.6 %)
AUC-ROC	0.990 (99.0 %)
Tiempo	0.0523s

Configuración óptima: 3 capas (100→50→25), alpha=0.0001, lr=0.01

Errores: 4 falsos positivos, 1 falso negativo

Interpretación: Rendimiento inferior a métodos lineales simples. 9,477 parámetros innecesarios. Complejidad no justificada.

5.2.2. SVM RBF

Cuadro 5: Resultados - SVM RBF

Métrica	Valor
Accuracy	0.982 (98.2 %)
Precision	0.986 (98.6 %)
Recall	0.986 (98.6 %)
F1-Score	0.986 (98.6 %)
AUC-ROC	0.998 (99.8 %)
Tiempo	0.0080s

Configuración óptima: C=10.0, gamma=0.01, vectores: 52/455 (11.4 %)

Errores: 1 falso positivo, 1 falso negativo

Interpretación: Idéntico a SVM Lineal. Gamma=0.01 indica comportamiento cuasi-lineal. Confirma que dataset es linealmente separable.

5.2.3. Algoritmos Genéticos

Cuadro 6: Resultados - Algoritmos Genéticos

Métrica	Valor
Accuracy	0.982 (98.2 %)
Precision	0.986 (98.6 %)
Recall	0.986 (98.6 %)
F1-Score	0.986 (98.6 %)
AUC-ROC	0.995 (99.5 %)
Tiempo	51.35s

Configuración óptima: 50 individuos, 30 generaciones, características: 11/30

Errores: 1 falso positivo, 1 falso negativo

Interpretación: Valor único: reducción dimensional 63.3 %. Solo 11 características necesarias. Mismo rendimiento con menos mediciones. Justificado solo para selección de características.

5.3. Comparación Global de Métodos

Cuadro 7: Ranking Final

Posición	Método	Accuracy	F1-Score	Tiempo
	SVM Lineal	0.982	0.986	0.0065s
	SVM RBF	0.982	0.986	0.0080s
	Algoritmos Genéticos	0.982	0.986	51.35s
4°	Regresión Logística	0.974	0.979	0.0037s
5°	Redes Neuronales	0.956	0.966	0.0523s
5°	Regresión Ridge	0.956	0.966	0.0008s

Hallazgos principales:

- Empate triple en el primer lugar con 98.2 % de precisión
- Los métodos convexos dominaron en eficiencia temporal
- Los algoritmos genéticos proporcionaron el único valor agregado significativo: reducción dimensional
- El dataset no se benefició sustancialmente de modelado no lineal complejo

6. Comparación y Discusión

6.1. Análisis Comparativo de Rendimiento

Los resultados experimentales revelan un **empate técnico excepcional** entre tres métodos que alcanzaron 98.2 % de precisión: SVM Lineal (convexo), SVM RBF (no convexo) y Algoritmos Genéticos (no convexo). Este empate es particularmente significativo

porque demuestra que, para el dataset de Wisconsin, **la optimización convexa es suficiente** para alcanzar el rendimiento óptimo.

Hallazgo principal: El dataset de cáncer de mama de Wisconsin es **fundamentalmente linealmente separable**. La evidencia incluye:

- SVM RBF con gamma óptimo de 0.01 (comportamiento cuasi-lineal)
- Redes neuronales no superaron a métodos lineales a pesar de 9,477 parámetros
- Misma cantidad de vectores de soporte en SVM lineal vs RBF (51 vs 52)

6.2. Eficiencia Computacional: Convexo vs No Convexo

La diferencia temporal es **dramática**:

- **Métodos convexos:** 0.0008s - 0.0065s
- **Métodos no convexos:** 0.0080s - 51.35s

Implicaciones prácticas:

- **Para aplicaciones clínicas en tiempo real:** Métodos convexos son ideales
- **Para investigación offline:** Métodos no convexos pueden aportar insights adicionales
- **Costo-beneficio:** Algoritmos genéticos (51.35s) solo se justifican por la selección de características

6.3. Valor Agregado de Métodos No Convexos

Único beneficio real encontrado: Algoritmos genéticos lograron **reducción dimensional del 63.3 %** (de 30 a 11 características) manteniendo rendimiento óptimo.

Características seleccionadas automáticamente:

- Mean concavity, mean symmetry
- Errores: radius, texture, area, compactness, concave points
- Worst: texture, perimeter, area

Interpretación clínica: Solo 11 de 30 mediciones son necesarias para diagnóstico óptimo, lo que podría **reducir costos** y **simplificar protocolos** clínicos.

6.4. Recomendaciones Prácticas

Para implementación clínica:

1. **Primera opción:** SVM Lineal (balance óptimo velocidad-rendimiento)
2. **Para screening masivo:** Regresión Ridge (recall 100 %)
3. **Para reducir costos:** Algoritmos genéticos para identificar características esenciales

4. **Para investigación:** Regresión logística (interpretabilidad máxima)

Evitar en este contexto:

- Redes neuronales (complejidad innecesaria)
- SVM RBF (sin ventaja sobre lineal)

6.5. Limitaciones del Estudio

1. **Dataset específico:** Resultados pueden no generalizarse a otros tipos de cáncer
2. **Tamaño muestral:** 569 muestras pueden ser insuficientes para métodos complejos
3. **Preprocesamiento:** Estandarización pudo haber favorecido a métodos lineales
4. **Validación:** Split único (aunque reproducible con `random_state=42`)

6.6. Conclusión de la Comparación

Los métodos **convexos** dominaron en este experimento, confirmando que **no siempre más complejidad significa mejor rendimiento**. El valor principal de los métodos no convexos fue la **capacidad de selección automática de características** demostrada por los algoritmos genéticos, sugiriendo un enfoque híbrido: usar métodos evolutivos para selección de características, seguido de métodos convexos para clasificación final.

7. Conclusiones

7.1. Hallazgos Principales

Este estudio experimental comparó seis técnicas de optimización (tres convexas y tres no convexas) aplicadas al diagnóstico de cáncer de mama utilizando el dataset de Wisconsin. Los resultados revelan conclusiones fundamentales sobre la efectividad relativa de métodos convexos versus no convexos en problemas de clasificación médica.

Hallazgo central: Los métodos de optimización convexa (SVM Lineal, Regresión Logística, Regresión Ridge) demostraron ser **suficientes y superiores** para este problema específico, alcanzando rendimientos equivalentes o superiores a métodos no convexos con **eficiencia computacional dramáticamente mayor**.

7.2. Rendimiento Comparativo

El experimento produjo un **empate técnico excepcional** en el primer lugar con 98.2% de precisión entre:

- **SVM Lineal** (convexo): 0.982 accuracy en 0.0065s
- **SVM RBF** (no convexo): 0.982 accuracy en 0.0080s
- **Algoritmos Genéticos** (no convexo): 0.982 accuracy en 51.35s

Este empate demuestra empíricamente que el dataset de Wisconsin es **linealmente separable**, validando la hipótesis de que no todos los problemas complejos requieren métodos de optimización no convexa.

7.3. Eficiencia vs Complejidad

La diferencia en eficiencia computacional fue **dramática y concluyente**:

- **Métodos convexos:** Convergencia en ¡0.01 segundos
- **Métodos no convexos:** Convergencia entre 0.05-51 segundos

Esta diferencia de **hasta 8,000x en tiempo de ejecución** sin mejora en rendimiento constituye evidencia sólida de que la complejidad algorítmica debe justificarse con beneficios tangibles.

7.4. Valor Agregado Identificado

Única ventaja significativa de métodos no convexos: Los algoritmos genéticos lograron **reducción dimensional del 63.3 %**, identificando que solo 11 de 30 características son necesarias para mantener rendimiento óptimo.

7.5. Limitaciones y Trabajo Futuro

Limitaciones reconocidas:

- Estudio limitado a un dataset específico (Wisconsin)
- Resultados pueden no generalizarse a otros tipos de cáncer
- Tamaño muestral moderado (569 casos) puede favorecer métodos simples

Direcciones futuras:

1. **Validación externa:** Replicar en datasets de otros tipos de cáncer
2. **Estudios longitudinales:** Evaluar performance en datos temporales
3. **Métodos híbridos:** Combinar selección evolutiva con clasificación convexa
4. **Implementación clínica:** Validar protocolo reducido de 11 características en práctica real

7.6. Reflexión Final

Este estudio demuestra que, **la complejidad no es sinónimo de superioridad**. Los métodos de optimización convexa, fundamentados en sólidos principios matemáticos y computacionalmente eficientes, pueden ser la solución óptima para problemas clínicos específicos.

La elección entre métodos convexos y no convexos debe basarse en **evidencia experimental rigurosa** más que en asunciones sobre la complejidad aparente del problema. Para el diagnóstico de cáncer de mama mediante análisis morfométrico, los métodos convexos proporcionan la combinación ideal de **precisión, eficiencia e interpretabilidad** requerida en aplicaciones médicas críticas.

Impacto esperado: Estos resultados pueden informar decisiones de implementación en sistemas de diagnóstico asistido por computadora, priorizando métodos que combinen alta precisión con eficiencia computacional y interpretabilidad clínica.

8. Referencias

Referencias

- [1] Wolberg, W. H., Street, W. N., & Mangasarian, O. L. (1995). Breast cancer Wisconsin (diagnostic) data set. *UCI Machine Learning Repository*. University of California, Irvine, School of Information and Computer Sciences.
- [2] Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge University Press.
- [3] Nocedal, J., & Wright, S. J. (2006). *Numerical optimization* (2nd ed.). Springer Science & Business Media.
- [4] Vapnik, V. N. (1995). *The nature of statistical learning theory*. Springer-Verlag.
- [5] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273-297.
- [6] Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55-67.
- [7] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- [8] Holland, J. H. (1992). *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*. MIT Press.
- [9] Street, W. N., Wolberg, W. H., & Mangasarian, O. L. (1993). Nuclear feature extraction for breast tumor diagnosis. *Biomedical Image Processing and Biomedical Visualization*, 1905, 861-870.
- [10] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.
- [11] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.
- [12] Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- [13] Cristianini, N., & Shawe-Taylor, J. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge University Press.
- [14] Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4), 303-314.
- [15] Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861-874.