

Relatório de Análise de Resultados

Régressão Logística e Algoritmos Genéticos

Tech Challenger 2 - Fase 2

Gerado em: 13/01/2026 21:28:16

Relatório Técnico de Otimização de Modelo de Régressão Logística com Algoritmos Genéticos para Diagnóstico de Diabetes

1. Resumo Executivo

A aplicação de Algoritmos Genéticos (AG) para a otimização do modelo de Régressão Logística trouxe melhorias significativas no desempenho diagnóstico de diabetes. O modelo otimizado apresentou um aumento no recall de 0,50 para 0,83, indicando uma capacidade melhorada de identificar corretamente pacientes com diabetes. Houve também uma redução nos falsos negativos, melhorando a confiabilidade clínica do modelo.

2. Análise da Evolução do Algoritmo Genético

O gráfico de evolução da função de aptidão demonstra um rápido aumento inicial da aptidão, estabilizando a partir da 20^a geração. Isto indica que as melhores soluções foram descobertas rapidamente, sugerindo que as características relevantes para a classificação foram identificadas precocemente. A estabilização posterior sugere uma convergência para uma solução quase ótima.

3. Análise Comparativa dos Modelos

Modelo Original

- **Recall:** 0.50
- **Falsos Positivos:** 20
- **Falsos Negativos:** 27

Modelo Otimizado

- **Recall:** 0.83
- **Falsos Positivos:** 29
- **Falsos Negativos:** 9

A otimização através de AG resultou em uma melhora substancial no recall, com um leve aumento nos falsos positivos. No entanto, a significativa redução em falsos negativos é crítica no contexto de diagnóstico médico, onde a identificação de pacientes com diabetes é prioritária.

4. Análise das Matrizes de Confusão

As matrizes de confusão destacam mudanças nos padrões de classificação:

- **Modelo Original:** Tinha maior dificuldade em identificar corretamente os casos de diabetes (27 falsos negativos).
- **Modelo Otimizado:** Mostrou-se mais eficiente na detecção de casos positivos, reduzindo os falsos negativos para 9. A leve elevação nos falsos positivos (de 20 para 29) é um trade-off aceitável devido à importância da detecção correta de condições verdadeiras.

5. Análise das Métricas de Performance

Recall (Taxa de Verdadeiros Positivos)

- Melhoria de 33 pontos percentuais demonstra eficácia do AG na otimização.

Falsos Positivos

- Aumento marginal implica maior vigilância, mas dentro de limites aceitáveis.

Falsos Negativos

- Redução de 27 para 9 é crucial para diagnósticos corretos, minimizando o erro em não identificar pacientes que de fato possuem diabetes.

6. Conclusões e Insights

A aplicação de Algoritmos Genéticos provou ser uma técnica eficiente para aprimorar modelos de Regressão Logística no diagnóstico de saúde. A significativa redução em falsos negativos é especialmente valiosa. Este experimento demonstra o potencial dos AG para ajustar parâmetros de modelos complexos de forma eficaz sem depender exclusivamente de métodos de otimização tradicionais.

7. Impacto Prático

Os resultados têm ampla aplicação prática, especialmente em ambientes clínicos, proporcionando diagnósticos mais confiáveis e rápidos para diabetes. A melhora no recall assegura que menos pacientes passem despercebidos, o que é crucial no manejo de uma condição médica crônica. A utilização de AG para otimização pode ser expandida para outros modelos e condições médicas, ampliando o impacto desta abordagem em sistemas de saúde.

Esta análise confirma a relevância e eficácia de otimizar modelos de aprendizado de máquina utilizando técnicas evolutivas, melhorando a precisão diagnóstica de forma significativa e confiável.

Top 5 Configurações do Algoritmo Genético

Rank 1: Fitness=833.3, Recall=83.33%, FP=29, FN=9, Acc=75.32%, C=45.347, Solver=liblinear, IQR=1.67, Features=5

Rank 2: Fitness=833.3, Recall=83.33%, FP=29, FN=9, Acc=75.32%, C=28.406, Solver=saga, IQR=1.67, Features=5

Rank 3: Fitness=833.3, Recall=83.33%, FP=29, FN=9, Acc=75.32%, C=42.678, Solver=lbfsgs, IQR=1.67, Features=6

Rank 4: Fitness=833.3, Recall=83.33%, FP=29, FN=9, Acc=75.32%, C=38.185, Solver=liblinear, IQR=1.67, Features=5

Rank 5: Fitness=833.3, Recall=83.33%, FP=29, FN=9, Acc=75.32%, C=5.861, Solver=lbfsgs, IQR=1.67, Features=5