



Computação Evolutiva para Aprendizado de Máquina Automático (AutoML)

Matéria: Computação Evolucionista

Autor: Matheus Muniz Damasco

Professores: Heder Soares e Hélio José Correa





Introdução

Com o passar dos anos, o machine learning tornou-se cada vez mais complexo, com um número crescente de algoritmos, técnicas de processamento e opções de ajuste de hiperparâmetros. Esse aumento na complexidade torna a construção de modelos não apenas mais demorada e custosa, mas também dependente de profissionais altamente especializados. O AutoML surgiu como uma solução para automatizar essas etapas, democratizando o acesso ao machine learning e aumentando a eficiência no desenvolvimento de modelos.



Motivação para o Problema

Vejo um grande potencial na combinação de AutoML com técnicas de computação evolutiva. O AutoML facilita a criação de modelos de machine learning mas ainda enfrenta desafios em termos de precisão, eficiência e escalabilidade. Os Algoritmos Genéticos com sua habilidade em explorar vastos espaços de soluções, pode acabar complementando o AutoML, tornando-o mais robusto e eficiente. Essa combinação tem o potencial de ampliar o impacto e a acessibilidade do Machine Learning, beneficiando diversos setores.



Descrição do Problema

- **Problema:**
 - Otimização de Hiperparâmetros e Seleção de Características no AutoML usando Algoritmos Genéticos
- **Objetivo:**
 - Aplicar Algoritmos Genéticos para resolver os desafios de otimização dentro do AutoML, especificamente focando na busca por configurações ótimas de hiperparâmetros e na seleção de características relevantes. O objetivo é maximizar a precisão dos modelos de machine learning enquanto minimiza o custo computacional e melhora a escalabilidade.
- **Desafios:**
 - Encontrar uma combinação de hiperparâmetros e um subconjunto de características que acabam maximizando a performance de modelos de machine learning, mantendo a eficiência e escalabilidade do processo.



Sobre o Problema

- **Tipo de Problema**
 - **Multiobjetivo:** O problema envolve a otimização de múltiplos objetivos, como maximizar a precisão dos modelos e minimizar o custo computacional.
 - **Restrições:** Existem restrições de tempo de execução, limite de recursos computacionais e limites no número de iterações ou avaliações do modelo durante o processo de otimização.
- **Variáveis de Projeto**
 - **1. Hiperparâmetros de Modelos:**
 - **Contínuas:** Incluem variáveis como taxa de aprendizado, coeficientes de regularização e parâmetros de decisão (ex: critério de divisão em árvores de decisão).
 - **Discretas:** Podem ser o número de camadas em redes neurais, número de árvores em florestas aleatórias ou até o número de vizinhos em algoritmos kNN.
 - **2. Características Seleccionadas:**
 - **Binárias:** Cada característica no conjunto de dados é representada como uma variável binária, onde 1 indica que a característica foi selecionada para o modelo e 0 que não foi.



Algoritmos

- **Seleção de Características:** Podemos utilizar Algoritmos Genéticos para selecionar automaticamente os conjuntos de características mais relevantes, o que reduz a dimensionalidade dos dados e pode melhorar a performance do modelo, tanto em termos de precisão quanto de eficiência computacional.
- **Ajuste de Hiperparâmetros:** Os Algoritmos Genéticos podem ser utilizados para ajustar hiperparâmetros, como taxa de aprendizado e número de neurônios visando melhorar a performance de modelos complexos.



Referências

[AutoML: A systematic review on automated machine learning with neural architecture search](#)

[Different metaheuristic strategies to solve the feature selection problem](#)