APRENDIZADO DE MÁQUINA AUTOMATIZADO NA PREPARAÇÃO E MANIPULAÇÃO DE CONJUNTOS DE TREINAMENTO

Ellen Leal dos Santos

Prof. Dr. Luciano Antonio Digiampietri

Escola de Artes, Ciências e Humanidades da USP

ellenleal@usp.br ; digiampietri@usp.br

Objetivos

Este projeto objetiva explorar a Automatização do Aprendizado de Máquina (AutoML) na criação e manipulação de conjuntos de treinamento focando no balanceamento de dados para melhorar o desempenho dos modelos, considerando medidas tradicionais, como acurácia e F1, e de justiça algorítmica (ASHOKAN e HAAS, 2021). Assim, este projeto de AutoML visa automatizar etapas da Mineração de Dados e contribuir para as áreas de Inteligência Artificial e Justiça Algorítmica.

Métodos e Procedimentos

A metodologia deste trabalho consistiu, inicialmente, de uma revisão bibliográfica visando a estudar artigos científicos relacionados à AutoML focando na construção de conjuntos de treinamento e engenharia de características. Também foram pesquisadas métricas relacionadas à justiça.
Com base nessa revisão, foram selecionados nove algoritmos de balanceamento de dados (*Oversampling*, *Undersampling*, *SMOTE*, *ADASYN*, *SMOTETomek*, *SMOTEENN*, *Tomek Links*, *BorderlineSMOTE*, *KMeansSMOTE*) e 19 medidas de injustiça (que avaliam, principalmente, se é uniformidade na classificação considerando as classes majoritárias e minoritárias ou se a justiça considerando um ou mais atributos sensíveis) para serem testadas na primeira etapa deste projeto (ASHOKAN e HAAS, 2021; BISWAS e RAJAN, 2020; MANDHALA et al., 2022). Foram selecionados quatro conjuntos de dados disponibilizados publicamente frequentemente utilizados na literatura sobre justiça algorítmica e oito algoritmos de classificação (Árvore de Decisão, Floresta Aleatória, Gaussian Naive Bayes, KNN, Multilayer perceptron, Regressão Logística, SVM com kernel linear e SVM com kernel polinomial). Foram utilizadas as implementações dos classificadores e algoritmos de balanceamento de dados disponíveis nas bibliotecas *sklearn* e *imblearn* do Python. As métricas relacionadas à justiça foram implementadas com base na literatura.
Cada conjunto de dados foi separado em 70% para treinamento e 30% para teste. Para cada conjunto de dados e cada classificador foram produzidos dez modelos (um treinado com o conjunto de treinamento original e os outros nove aplicando cada um dos algoritmos de balanceamento). Para cada teste realizado foram calculadas 21 métricas (acurácia, F1 macro e 19 medidas relacionadas à justiça).

Resultados

Foi observado que a combinação de classificador e estratégia de balanceamento teve efeitos variados de acordo com o conjunto de dados, não sendo possível encontrar uma combinação que seja a mais adequada para todos os conjuntos de dados.

Considerando o conjunto *Adult Income*, utilizado para prever a renda anual de uma pessoa com base em algumas características, observa-se que a não utilização de balanceamento produziu os melhores resultados de acurácia para três dos oito modelos testados. O uso do algoritmo de balanceamento *KMeansSMOTE* obteve as melhores acurácias para três modelos e do *Tomek Links* para dois. Em termos de medida F1 macro, o treinamento sem balanceamento atingiu o melhor valor para apenas um modelo. Destaca-se o balanceamento usando *SMOTEENN* que atingiu os melhores resultados em três. Em termos de métricas de justiça, o balanceamento usando *Tomek Links* fez com que 16 das 19 métricas ligadas à justiça tivessem seus valores melhorados em relação ao treinamento sem balanceamento. Destacam-se, em seguida, os algoritmos *KMeansSMOTE* (melhoria em 14 métricas) e *SMOTEENN* (melhoria em 10 métricas).

Os resultados para o conjunto *Breast Cancer* foram bastante diferentes. Trata-se de um conjunto simplificado com informações para tentar prever a ocorrência de câncer de mama. A estratégia de balanceamento *ADASYN* atingiu a melhor acurácia para todos os modelos, empatando com outras estratégias para alguns classificadores. Em termos de medida F1 macro, essa estratégia obteve o melhor desempenho para sete dos oito modelos. Considerando as medidas de justiça, duas abordagens de balanceamento (*ADASYN* e *SMOTEENN*) obtiveram resultados semelhantes, melhorando o valor de sete métricas em relação ao treinamento sem balanceamento dos dados.

O conjunto *Ricci* é relacionado à promoção de bombeiros em sua carreira. Para este modelo, no geral, as melhores métricas de justiça foram atingidas sem o balanceamento do conjunto de treinamento. Porém, o balanceamento usando *BorderlineSMOTE* permitiu um aumento de acurácia, na média, superior a 5% e de medida F1 macro superior a 13,9%.

O conjunto *German Credit* é relacionado à avaliação do risco para a concessão de crédito. Os melhores resultados de acurácia consideraram o conjunto sem balanceamento. Já a medida F1 macro teve leve melhora (entre 3% e 4% utilizando os algoritmos *SMOTE*, *ADASYN* e *SMOTETomek*). Em termos de justiça, destacam-se as estratégias *SMOTEENN* (melhora em 12 métricas em relação ao treinamento original) e *KMeansSMOTE* (melhora em 11 métricas).

Conclusões

Algoritmos de aprendizado de máquina são cada vez mais usados para diferentes finalidades. Além da variedade de algoritmos disponíveis, diferentes formas de parametrização, balanceamento e seleção de atributos tornam a tarefa de treinamento adequado destes modelos cada vez mais desafiadora. Neste sentido, este projeto está produzindo ferramentas que combinam abordagens para o treinamento de modelos de forma a identificar, para cada conjunto de dados, o treinamento mais adequado considerando diferentes métricas.

Referências

ASHOKAN, A.; HAAS, C. Fairness metrics and bias mitigation strategies for rating predictions. Information Processing and Management, v. 58, n. 5, 2021.

BISWAS, S.; RAJAN, H. Do the machine learning models on a crowdsourced platform exhibit bias? an empirical study on model fairness. In: Proceedings of the 28th ACM
Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering. New York, NY, USA: Association for Computing
Machinery, 2020, p. 642–653.

MANDHALA, V. N.; BHATTACHARYYA, D.; MIDHUNCHAKKARAVARTHY, D.; KIM, H.-J. Mitigating bias by optimizing the variance between privileged and deprived data using post processing method. Rev. D Intell. Artif., International Information and Engineering Technology Association, v. 36, n. 1, p. 87–91, fev. 2022.