

AutoML 과 XAI 의 결합 : 기계학습 모델의 자동화와 해석력 향상을 위하여

손민혁^{1*}, 김남훈^{2*}, 이현지^{3*}, 김도연^{4*}
*공동 주저자

¹ 서울과학기술대학교 글로벌융합산업공학과 학부생

² 한국공학대학교 IT 경영학과 학부생

³ 고려대학교(세종) 빅데이터사이언스학부 학부생

⁴ 명지대학교 데이터테크놀로지학과 학부생

shawn225@naver.com, kimnh097@naver.com, 0105lhl@korea.ac.kr, ehdusrla812@naver.com

Combining AutoML and XAI: Automating machine learning models and improving interpretability

Min Hyeok Son^{1*}, Nam Hun Kim^{2*}, Hyeon Ji Lee^{3*}, Do Yeon Kim^{4*}

¹Dept. of Industrial engineering, SeoulTech

²Dept. of IT management, Tech University of Korea

³Dept. of Big Data Science, Korea University Sejong Campus

⁴Dept. of Data Technology, Myungji University

요 약

본 연구는 최근 기계학습 모델의 복잡성 증가와 '블랙 박스'로 인식된 머신러닝 모델의 해석 문제에 주목하였다. 이를 해결하기 위해, AutoML 기술을 사용하여 효율적으로 최적의 모델을 탐색하고, XAI 기법을 도입하여 모델의 예측 과정에 대한 투명성을 확보하려 하였다. XAI 기법을 도입한 방식은 전통적인 방법에 비해 뛰어난 해석력을 제공하며, 사용자가 머신러닝 모델의 예측 근거와 그 타당성을 명확히 이해할 수 있음을 확인하였다.

1. 서론

현대의 기계학습 모델은 복잡도가 급증하며 다양한 알고리즘이 등장하고 있다. 이런 배경 속에서 AutoML 라이브러리가 모든 알고리즘을 포괄하는 것은 어려워지고 있다. 예를 들면, pycaret 은 M/L 에, Auto-Keras 는 딥러닝에 특화되어 있다. 모델의 예측력은 물론, 변수 중요도와 같은 해석력도 중요하다. 특히 딥러닝은 '블랙 박스'로 인식되기 쉽다.

이 문제를 해결하기 위해, 확장 가능한 AutoML 과 XAI 가 결합될 필요가 있다. 본 연구는 AutoML 에 SHAP, LIME 과 같은 XAI 해석 기법을 통합, 사용자가 모델의 효율과 투명성을 동시에 향상시키는 방안을 제시한다. 이 통합 솔루션은 다양한 알고리즘에 적용 가능하며, 신뢰성 있는 의사결정을 지원할 것으로 예상된다.

2. 본론

2.1 관련 연구

AutoML: 'Automated Machine Learning'을 지칭하며, 머신러닝 과정을 자동화하는 기술이다. 데이터 처리부터 모델 최적화까지를 포괄한다. 최근에는 주요 연구

분야로 각광받고 있으며, 많은 기업들이 투자하고 있다.

XAI: 'Explainable Artificial Intelligence'로, 머신러닝 모델의 결정 과정을 해석 가능하게 하는 기술이다. 특히 딥러닝 같은 복잡한 모델의 'Black Box' 성격을 해석하는 데 중요하다. 특히 SHAP[1]은 설명가능한 인공지능(XAI) 방법 중 하나로, 게임이론 기반으로 모델 예측을 설명하는 방식이다. 플레이어의 기여도 분석 개념을 변수의 기여도 평가에 활용한다.

2.2 실험 및 평가

본 연구는 AutoML 기술을 적용하여 최적의 예측 모델을 자동 탐색하였다. AutoML 을 통해 최적의 모델을 선별한 후, XAI 기법을 활용하여 모델의 예측 기여도를 분석하고 시각화하였다. 이를 통해 중요한 피처와 그들의 영향력을 정밀하게 파악하였다.

(a) AutoML 을 통한 최적의 모델 선정

본 연구에서는 AutoML 을 이용해 최적의 예측 모델을 찾았다. 실험 과정에서 오픈소스 라이브러리인 pycaret 이라는 AutoML 툴을 사용하였으며, 사용한 데이터는 'Boston House Price

