

AutoML 프레임워크에서 AI모델 성능 고도화를 위한 기술 전략

서동준*, 정경용**

Technical Strategies for Performance Enhancement of AI Models in AutoML Frameworks

Dongjoon Suh*, Kyungyong Chung**

이 논문은 2025년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. RS-2020-NR049579) 아울러, 본 연구는 2025년 경기대학교 대학원 연구원장학생 장학금 지원에 의하여 수행되었음

요 약

본 논문은 AutoML 프레임워크에서 AI 모델 성능을 향상시키기 위한 기술 전략을 분석한다. Auto-Sklearn, AutoKeras, AutoGluon이라는 세 가지 주요 도구의 구조적 기반, 최적화 접근법, 성능 특성을 비교 평가하였다. 분석 결과, 도메인 특수성 부족, 분절된 인터페이스, 모델 해석 가능성 부족이라는 한계점이 확인되었다. 이를 해결하기 위해 도메인 특화 메타러닝, 뉴로심볼릭 통합 인터페이스라는 두 가지 전략을 제안한다. 이 전략들은 특수 도메인 적용, 다양한 데이터 처리, 규제 산업에서의 투명성 확보라는 과제를 해결한다. 본 접근법은 AutoML을 단순한 하이퍼파라미터 최적화에서 포괄적인 도메인 맞춤형 솔루션으로 발전시켜, 전문가들이 심층적인 AI 지식 없이도 고성능 머신러닝 시스템을 개발할 수 있게 한다.

Abstract

This paper analyzes technical strategies for enhancing AI model performance in AutoML frameworks. We evaluate and compare the architectural foundations, optimization approaches, and performance characteristics of three major tools: Auto-Sklearn, AutoKeras, and AutoGluon. Our analysis identifies key limitations including insufficient domain specificity, fragmented interfaces, and inadequate model interpretability. To address these challenges, we propose two strategies: domain-specific meta-learning, neuro-symbolic integration for unified interfaces. These approaches tackle the challenges of applying AutoML to specialized domains, processing diverse data types, and ensuring transparency in regulated industries. Our framework advances AutoML beyond simple hyperparameter optimization toward comprehensive domain-tailored solutions, enabling experts to develop high-performance machine learning systems without extensive AI knowledge.

Keywords

bayesian optimization, neural architecture search, NAS, multi-layer ensemble, meta-learning, feature engineering

* 경기대학교 컴퓨터과학과 석사과정

- ORCID: <https://orcid.org/0009-0008-5412-382X>

** 경기대학교 AI컴퓨터공학부 정교수(교신저자)

- ORCID: <https://orcid.org/0000-0002-6439-9992>

• Received: Mar. 04, 2025, Revised: Apr. 25, 2025, Accepted: Apr. 28, 2025

• Corresponding Author: Kyungyong Chung

Dept. of Computer Science, Kyonggi University, 154-42 Gwanggyosan-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do 16227, Republic of Korea

Tel.: +82-31-249-9670, Email: fuu7kp@kyonggi.ac.kr

I. 서 론

Kaggle, Dacon, AI Factory와 같은 인공지능 경진 대회에서 AutoML은 회귀, 분류, 이미지 처리, 자연어 분석 등 다양한 문제 해결을 위한 핵심 도구로 자리잡았다. AutoML은 데이터 전처리, 모델 선택, 하이퍼파라미터 최적화, 평가에 이르는 전체 워크플로우를 자동화함으로써 분석가의 생산성을 획기적으로 향상시킨다. 오픈소스 생태계에서는 Auto-Sklearn, AutoKeras, AutoGluon과 같은 도구들이 각기 다른 전략으로 AutoML의 범위를 확장해왔다. 이러한 도구들의 등장은 도메인 전문가들이 복잡한 머신러닝 알고리즘에 대한 심도 있는 이해 없이도 높은 수준의 예측 모델을 구축할 수 있게 한다. Auto-Sklearn은 베이지안 최적화와 메타러닝을 기반으로 하이퍼파라미터 탐색 효율성을 혁신하고[1], AutoKeras는 신경망 구조 탐색을 통해 딥러닝 모델 설계의 자동화를 실현한다[2]. 특히 Auto-Sklearn의 메타러닝 접근방식은 이전 데이터셋들에서 얻은 경험을 활용하여 새로운 문제에 대한 최적화 시간을 대폭 단축시키는 성과를 거두었다. 한편 AutoGluon은 다중 앙상블과 자동 피쳐 엔지니어링을 결합하여 구조화된 데이터 분석에서 뛰어난 성능을 입증한다[3]. 이들 프레임워크는 전통적 알고리즘, 딥러닝, 앙상블 기법을 각각 집중 개선함으로써 AutoML의 다각화된 발전을 이끌고 있다.

그러나 현재 AutoML 프레임워크들은 주로 일반적인 데이터셋에 대한 하이퍼파라미터 최적화에 초점을 맞추고 있어, 특정 도메인의 고유한 특성을 반영하지 못하는 한계가 있다. 또한 데이터 유형에 따라 서로 다른 인터페이스를 사용해야 하는 문제로 인해 복잡한 데이터 환경에서의 활용이 제한적이다. 기존 연구들이 AutoML의 일반적 성능 향상에 중점을 둔 반면, 본 논문은 도메인 특화 접근법과 통합 인터페이스 개발을 통한 AutoML의 실용적 가치 증대에 초점을 맞춘다는 점에서 차별성을 갖는다.

M. A. Zoller 와 M. F. Huber[4]의 연구에서는 다양한 AutoML 도구의 성능을 비교하였으나, 도메인 특화 메타러닝 접근법에 대한 실증적 분석은 미흡하였다. 또한 X. He et al.[5]은 AutoML의 확장성 문

제를 다루었지만, 복합적 데이터 환경에서의 통합 인터페이스 설계 방안에 대한 구체적 해결책을 제시하지 못했다. 본 연구는 이러한 기존 연구의 한계를 극복하고자 도메인 특화 메타러닝과 뉴로심볼릭 통합 인터페이스라는 두 가지 혁신적 전략을 제안하고, 실제 데이터셋을 활용한 정량적 성능 평가를 통해 그 효과성을 검증한다.

본 논문은 세 도구의 기술적 차이와 혁신적 접근법을 비교 분석함으로써 AutoML의 진화 방향과 실용적 가치를 탐구한다. 연구의 주요 목적은 (1) 현재 AutoML 프레임워크들의 기술적 특성과 한계점을 체계적으로 분석하고, (2) 이를 기반으로 도메인 특화 메타러닝과 뉴로심볼릭 통합 인터페이스라는 두 가지 핵심 전략을 제안하며, (3) 실험적 검증을 통해 제안된 전략의 효과성을 정량적으로 입증하는 것이다. 특히 각 도구가 지닌 고유한 장단점을 비교하고, 이를 바탕으로 향후 AutoML 발전 방향에 대한 통찰을 제시하고자 한다. 제안된 전략 중 도메인 특화 메타러닝에 대한 실험적 검증을 통해 의료 및 금융 도메인에서의 성능 향상을 정량적으로 측정함으로써, 해당 접근법의 효과성을 실증적으로 입증하였다.

II. 관련 연구

2.1 AutoML 기술 트렌드

AutoML 기술은 2010년대 초반부터 본격적으로 발전하기 시작하여 알고리즘 접근 방식과 적용 범위 측면에서 다각화된 진화를 거듭해왔다. 초기에는 단순 하이퍼파라미터 최적화에 초점을 맞추었으나, 점차 모델 선택, 피쳐 엔지니어링, 신경망 구조 설계까지 자동화 영역이 확장되었다. 표 1은 주요 AutoML 프레임워크의 발전 과정을 연도별, 기술적 특성별로 정리한 것이다.

AutoML 기술의 진화는 크게 세 가지 경로로 발전해왔다. 첫째, 베이지안 최적화와 메타러닝을 결합한 효율적 탐색 전략이 등장했다. 둘째, 딥러닝 모델의 자동 설계를 위한 신경망 구조 탐색 기법이 개발되었다. 셋째, 다양한 기본 모델을 조합한 앙상블 기반 접근법이 고도화되었다.

표 1. AutoML 프레임워크의 진화
Table 1. Evolution of AutoML frameworks

Framework	Year	Developer/ institution	Algorithms
Auto-WEKA [6]	2013	University of british columbia	Bayesian optimization
Auto-Sklearn [7]	2015	Fraunhofer institute	Bayesian optimization, meta-learning
Auto-Net[8]	2016	Freiburg university	Bayesian optimization
TPOT	2016	UPenn	Genetic programming
ATM[9]	2017	MIT	Bayesian optimization, meta-learning
H2O AutoML [10]	2017	H2O.ai	Stacked ensembles
OBOE[11]	2018	Cornell university	Meta-Learning, collaborative filtering
Google cloud AutoML	2018	Google	NAS (Neural Architecture Search)
AutoKeras[2]	2018	Texas A&M university	NAS, bayesian optimization
NNI	2018	Microsoft	NAS
PyCaret	2020	PyCaret team	Stacked ensembles
AutoGluon[3]	2020	Amazon web services	Stacked ensembles

이러한 기술 발전은 AI 기술을 보다 널리 보급하는 중요한 원동력이 되고 있으며, 프로그래밍에 대한 깊은 지식이 없는 전문가들도 최신 AI 기술을 실무에 적용할 수 있게 해주고 있다. AutoML 기술은 베이지안 최적화 → 메타러닝 → 유전자 알고리즘 → 신경망 탐색 → 앙상블 통합으로 발전해오며 점차 복잡한 문제 영역으로 확장되고 있다. 특히 클라우드 기반 서비스와 오픈소스 프레임워크의 상호작용이 기술 진화를 가속화하는 추세이며, 최근에는 자동화 범위가 데이터 전처리부터 모델 배포까지 전체 MLops 사이클로 확대되고 있다.

2.2 베이지안 최적화와 Auto-Sklearn 활용

AutoML은 다양한 모델과 하이퍼파라미터 조합을 탐색하는 과정에서 두 가지 문제점에 직면한다. 첫 번째 근본적인 문제는 차원의 저주이다. 이는 하이퍼파라미터 최적화 과정에서 발생하는 기하급수적인 탐색 공간의 증가를 의미한다. 하이퍼파라미터의 수가 증가할수록 가능한 조합의 수는 지수적으로 증가하여, 전체 탐색 공간이 방대해진다. 이러한 고차원 공간에서는 전통적인 격자 탐색이나 무작위 탐색 방식으로는 최적의 조합을 찾는 것이 사실상 불가능해진다. 두 번째로 직면하는 문제는 계산 비용의 폭발적 증가이다. 하이퍼파라미터 최적화 과정에서는 각각의 파라미터 조합에 대해 모델을 처음부터 새로 학습시키고 평가해야 한다. 특히 현대의 딥러닝 모델들은 단일 학습에도 상당한 시간이 소요되며, 교차 검증을 수행할 경우 이 비용은 더욱 증가한다. 게다가 앙상블 모델링을 고려할 경우, 개별 모델들의 조합에 대한 평가까지 필요하므로 계산 비용은 더욱 가파르게 상승한다.

이러한 문제를 해결하기 위해 베이지안 최적화는 확률적 추정과 메타 지식 활용을 결합한 전략을 제시한다. 먼저, 차원의 저주로 인해 무작위 탐색이 비효율적인 문제는 메타-러닝으로 접근한다[12]. Auto-Sklearn은 사전에 수집된 140개 이상의 데이터셋에서의 모델 성능 데이터를 분석하여 유사한 데이터셋에 대한 최적의 초기 하이퍼파라미터 후보를 추천함으로써, 탐색 시작점을 가속화한다.

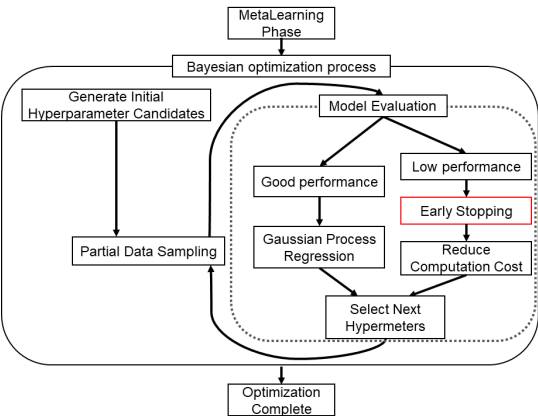


그림 1. Auto-Sklearn 과정
Fig. 1. Auto-Sklearn process

계산 비용 문제는 부분 평가 전략으로 대응한다. auto-sklearn은 전체 데이터가 아닌 부분 샘플링된 데이터로 모델을 빠르게 평가하고, 성능이 낮은 후보는 조기 종료를 적용하여 불필요한 계산을 생략한다. 동시에 대리 모델로 가우시안 프로세스를 사용하여 하이퍼파라미터 조합과 모델 성능 간의 관계를 추정함으로써, 실제 평가 횟수를 최소화하면서도 정확한 탐색이 가능하다. 그림 1은 Auto-Sklearn 과정을 나타낸다.

2.3 AutoKeras와 신경망 구조 탐색

AutoKeras는 딥러닝을 AutoML에 접목한 대표적인 시스템으로, 전통적인 머신러닝에 국한되었던 AutoML의 범위를 확장했다[13]. Auto-sklearn이 SVM, Random Forest와 같은 전통적 알고리즘의 하이퍼파라미터 최적화에 중점을 두었다면, AutoKeras는 합성곱 신경망(CNN), 순환 신경망(RNN)과 같은 딥러닝 아키텍처의 자동 설계까지 그 범위를 넓혔다. AutoKeras는 Keras를 기반으로 구축되어 높은 접근성과 사용 편의성을 제공한다. 사용자는 복잡한 신경망 구조의 설계 없이도 ImageClassifier, TextClassifier 등의 높은 수준의 API를 통해 자동화된 모델 개발이 가능하다. AutoKeras는 Neural Architecture Search를 효율적으로 구현하기 위해 트리 구조 기반의 탐색 전략을 도입했다[14]. 이 트리 구조에서 각 노드는 하나의 신경망 구조를 나타내며, 자식 노드는 부모 노드에서 파생된 변형 구조를 표현한다. 예를 들어, CNN의 경우 컨볼루션 레이어의 필터 수 변경, 풀링 레이어 추가, 스킵 연결 추가 등의 연산이 각각의 가치를 형성한다[15]. 이러한 트리 구조는 모델 구조의 진화 과정을 체계적으로 추적하고, 유망한 구조를 효율적으로 탐색할 수 있게 한다. Network Morphism은 이러한 트리 구조 탐색의 핵심 메커니즘이다. 기존 구조의 성능을 보존하면서 새로운 구조로 변형하는 이 기법은, 각 트리 노드에서 다음 탐색 방향을 결정할 때 중요한 역할을 한다. 이는 무작위 탐색이나 전수 조사 방식보다 훨씬 효율적인 탐색을 가능하게 하며, 특히 깊은 신경망 구조에서도 안정적인 성능 향상을 달성할 수 있게 한다.

그림 2는 AutoKeras 과정을 나타내고, 다음과 같이 작동한다. 초기 CNN 모델에서 시작하여 지속적인 성능 평가와 구조 최적화를 수행한다. 주요 최적화는 세 가지 방향으로 진행된다. Add_Conv_Layer에서는 새로운 합성곱 층을 추가하여 모델의 복잡도를 증가시킨다. Channel Widening에서는 기존 층의 채널 수를 확장하여 모델의 표현력을 강화시킨다. Skip Connection에서는 레이어 간 건너뛰기 연결을 추가하여 그래디언트 흐름을 개선한다. Network Morphism은 이러한 과정에서 핵심적인 역할을 수행하며, 각 변경이 이전 모델의 성능을 유지하면서도 새로운 가능성을 탐색할 수 있게 한다. 이 과정은 최대 반복 횟수에 도달하거나 최적의 모델이 발견될 때까지 반복적으로 수행된다.

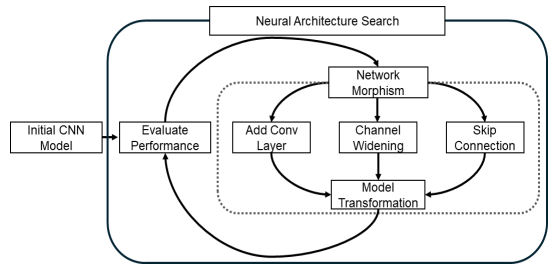


그림 2. AutoKeras 과정
Fig. 2. AutoKeras process

2.4 AutoGluon과 다층 앙상블 전략

AutoGluon은 기존 AutoML 프레임워크의 접근 방식을 확장하여 다층 앙상블과 자동화된 피쳐 엔지니어링을 통해 예측 성능을 극대화하는 전략을 제시한다. Auto-Sklearn이 베이지안 최적화와 메타러닝을 활용한 하이퍼파라미터 탐색에 집중하고, AutoKeras가 신경망 구조 탐색(Neural architecture search)을 통해 딥러닝 모델을 자동화한 것과 달리, AutoGluon은 다양한 기본 모델의 생성과 계층적 조합에 중점을 둔다. 이는 특히 테이블 형식의 구조화된 데이터에서 높은 성능을 발휘하도록 설계되었으며, 복잡한 모델 설계 없이도 강력한 앙상블 효과를 달성한다. AutoGluon의 핵심 프로세스는 다음과 같다. 먼저 데이터 전처리 단계에서는 결측치 처리, 데이터 타입 변환, 피쳐 엔지니어링과 같은 기본적인 전처리 작업을 자동으로 수행한다.

이후 XGBoost, LightGBM, Neural Networks, Random Forest, CatBoost 등 다양한 기본 모델들을 병렬적으로 학습한다. 이러한 기본 모델들의 예측값은 다시 새로운 특성으로 활용되어 상위 계층의 모델 학습에 입력된다. 이 과정이 여러 계층에 걸쳐 반복되면서, 각 모델의 장점을 조합한 강력한 앙상블 모델이 구축된다. AutoGluon의 핵심은 스택킹과 블렌딩을 기반으로 한 다단계 앙상블 전략이다. 첫 단계는 랜덤 포레스트, XGboost, K-최근접 이웃 등 이질적인 특성을 가진 다양한 기본 모델을 병렬로 생성하고, 이들의 예측 결과를 메타 데이터로 활용한다[16]. 두 번째 단계는 이러한 메타 데이터를 입력으로 받는 상위 모델을 학습시켜 개별 모델의 오차를 보정하며, 필요에 따라 이 과정을 여러 계층으로 반복한다. 이때 AutoGluon은 메타러닝 기반의 모델 선택 알고리즘을 적용하여 계산 비용을 관리한다. 예를 들어, 초기 단계에서 검증 성능이 낮은 모델은 후보군에서 조기 제외되며, 상위 계층으로 갈수록 고성능 모델에 집중하는 방식으로 자원을 최적화한다. 또한 AutoGluon은 자동화된 피쳐 엔지니어링을 통해 사용자의 전처리 부담을 줄인다. 범주형 변수의 임베딩, 결측치 보간 등을 자동으로 수행하며, GPU 가속을 활용하여 대규모 데이터 처리 속도를 향상시킨다. 이는 기존 AutoML 시스템이 가진 계산 비용 문제를 완화하는 동시에, 복잡한 신경망 구조보다는 전통적인 머신러닝 모델의 조합에 의존함으로써 효율성을 균형 있게 유지한다. 그림 3은 AutoGluon 과정을 나타내고, 다음과 같이 작동한다.

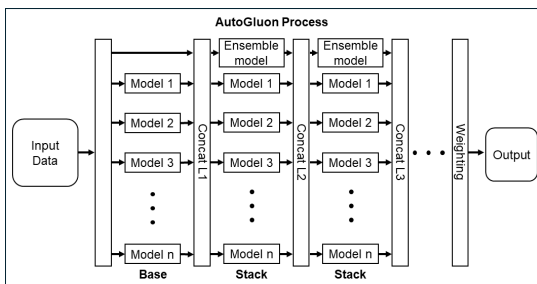


그림 3. AutoGluon 과정
Fig. 3. AutoGluon process

그림 3은 입력 데이터가 여러 계층의 모델을 통과하는 구조로 설계되어 있다. 각 계층은 여러 개의

모델로 구성되며, 이전 계층의 결과가 다음 계층의 입력으로 사용된다. Stack, Ensemble, Weighting 단계를 통해 모델 간 정보가 교환되고 누적되며, 이러한 다층 구조는 개별 모델의 장점을 결합하여 최종 예측 성능을 향상시킨다. 이 접근 방식은 복잡한 데이터 패턴을 효과적으로 포착하면서도 계산 효율성을 유지한다.

III. AutoML 프레임워크의 한계점 분석 및 성능 고도화 방안

3.1 AutoML 도구의 특성 및 한계점 분석

현재 주요 AutoML 도구들은 각각의 고유한 특성과 한계점을 가지고 있다. Auto-Sklearn은 베이지안 최적화와 메타러닝을 결합해 구조화된 데이터에서 빠른 하이퍼파라미터 탐색이 가능하다는 장점이 있지만, 딥러닝 모델을 지원하지 않고 메타 데이터셋에 대한 의존성으로 인해 특정 도메인에서 편향이 발생할 수 있다는 한계가 있다. 또한 복잡한 피쳐 엔지니어링을 자동화하는 기능이 제한적이어서 데이터 전처리에 사용자의 개입이 여전히 필요하다. AutoKeras는 신경망 구조 탐색(NAS)을 통해 이미지, 텍스트와 같은 비정형 데이터 처리 능력을 확장했다는 점에서 주목할 만하지만, GPU 리소스에 대한 의존도가 높아 계산 비용이 큰 문제를 안고 있다. 특히 복잡한 아키텍처 탐색 과정에서 학습 시간이 기하급수적으로 증가하는 경향이 있으며, 최적 모델 선정을 위한 평가 지표가 제한적이어서 특정 비즈니스 목표에 적합한 모델을 찾기가 어렵다.

AutoGluon은 다층 앙상블 전략으로 구조화된 데이터에서 최상위 수준의 예측 성능을 달성하는 강점이 있지만, 복잡한 모델 조합으로 인한 모델 해석력 저하와 고사양 컴퓨팅 인프라 요구라는 트레이드오프가 존재한다. 또한 모델 경량화나 추론 속도 최적화와 같은 실시간 서비스 배포에 필요한 기능이 아직 충분히 발전되지 않았다는 점도 한계로 작용한다.

이러한 AutoML 프레임워크들은 공통적으로 머신러닝의 발전 과정에서 베이지안 최적화, 신경망 구조 탐색, 다층 앙상블과 같은 흐름을 반영하며 하이

퍼파라미터의 미세조정에 집중해왔다. 그러나 실제 산업 현장에서의 활용도를 높이기 위해서는 하이퍼파라미터 최적화를 넘어서는 다양한 기술적 혁신이 필요한 시점이다.

3.2 AutoML 시스템의 한계 및 개선 방향

첫째, 현재의 AutoML 도구들은 전문 도메인 영역에서의 특수성을 충분히 반영하지 못하는 문제가 있다. Auto-Sklearn은 320개의 일반적인 데이터셋에 기반한 메타러닝을 활용하고 있으나, 금융, 의료, 제조업 등 특정 산업 영역의 고유한 데이터 특성과 문제 유형을 충분히 고려하지 못한다. 이로 인해 도메인 전문가들은 여전히 상당한 수준의 수동 개입을 통해 모델을 조정해야 하는 상황이다. 둘째, 동일한 AutoML 프레임워크 내에서도 데이터 유형에 따라 서로 다른 함수와 인터페이스를 사용해야 하는 번거로움이 존재한다. 예를 들어, AutoGluon을 사용할 때 구조화된 데이터는 `autogluon_tabular`, 시계열 데이터는 `autogluon_timeseries`, 다중 모달 데이터는 `autogluon_multimodal` 등 각기 다른 모듈을 별도로 호출해야 한다. 이는 다양한 데이터 유형이 혼합된 실제 비즈니스 문제에서 통합적 접근을 어렵게 만든다.

3.3 성능 고도화를 위한 기술 전략 및 구현 방안

3.3.1 도메인 특화 메타러닝 기술

현재의 일반적인 메타러닝 접근법을 넘어, 산업별로 특화된 메타 데이터베이스를 구축해야 한다. 의료 영상 분석을 위한 전문화된 메타 지식 베이스나, 금융 데이터를 위한 특화된 메타 모델 라이브러리 개발이 가능하다. 이를 통해 특정 도메인의 데이터 특성과 문제 유형에 최적화된 초기 모델 설정과 하이퍼파라미터 추천이 가능해진다.

각 산업 분야별 대표적인 데이터셋과 모델링 경험을 체계적으로 수집하여 도메인 특화 메타 데이터베이스를 구축하는 것이 중요하다. 예를 들어 의료 영상 진단에 최적화된 컨볼루션 신경망 아키텍처와 하이퍼파라미터 조합을 축적하고, 도메인 지식

그래프를 활용한 문맥적 유사도 측정 메커니즘을 개발할 수 있다[17]. 이를 통해 새로운 데이터셋이 입력되었을 때 도메인 특성을 고려한 정확한 유사 데이터셋 매칭이 가능해진다. 또한, 데이터가 부족한 새로운 도메인에서도 효과적으로 모델링할 수 있도록 기존 도메인의 메타 지식을 활용하는 전이 메타러닝 알고리즘을 적용해야 한다. 아울러, 새로운 데이터셋과 모델링 결과가 축적될수록 추천 정확도가 향상되는 지속적 학습 메커니즘을 통합적으로 구현해야 한다.

3.3.2 뉴로심볼릭 통합 인터페이스 기술

현재는 데이터 유형에 따라 서로 다른 함수를 사용해야 하는 분절된 접근 방식이 일반적이다. 이를 개선하기 위해 뉴로심볼릭 기법을 활용한 통합 인터페이스가 필요하다. 뉴로심볼릭 기법은 머신러닝 알고리즘과 논리적 추론을 결합하여, 시스템이 데이터의 특성을 자동으로 인식하고 적절한 처리 파이프라인을 구성할 수 있도록 한다.

특히, 입력 데이터의 구조와 특성을 분석하여 다양한 데이터 유형을 자동으로 식별하고, 이에 맞는 전처리 파이프라인을 구성하는 다중 모달 자동 인식 시스템을 핵심 기술로 구현해야 한다. 또한, 도메인 전문가의 지식을 논리적 규칙으로 표현하고 이를 신경망 학습 과정에 통합하는 심볼릭 규칙과 신경망의 하이브리드 아키텍처가 필요하다.

이와 함께, 사용자가 입력한 데이터셋에 따라 도메인 특화 데이터베이스를 자동으로 연결하는 메커니즘을 구축해야 한다. 더불어, 다양한 데이터 유형과 모델링 목표를 단일 인터페이스로 표현할 수 있는 고수준 쿼리 언어를 개발하여, 복잡한 다중 모달 문제도 직관적이고 일관된 방식으로 정의할 수 있도록 지원해야 한다.

IV. 실험 및 성능평가

4.1 실험 셋업

본 실험의 목적은 메타러닝이 특정 도메인별로

어떠한 성능 향상을 가져오는지 정량적으로 평가하는 데 있다. 이를 위해 OpenML에서 제공하는 공개 데이터셋을 기반으로 의료(Medical), 금융(Finance), 그리고 대조군으로 무작위(Random) 도메인으로 나누어 분석을 진행하였다. 데이터셋 선정 기준은 (1) 해당 도메인의 전형적인 특성을 반영하는지, (2) 충분한 샘플 수를 가지고 있는지, (3) 클래스 불균형 정도가 극단적이지 않은지를 고려하였다.

실험 설계는 각 도메인별로 대표적인 데이터셋을 3개씩 선정하여 일반적인 AutoML 프레임워크인 TPOT(Tree-based Pipeline Optimization Tool)와 도메인 특화 메타러닝을 적용한 TPOT의 성능을 비교하는 방식으로 진행하였다. 도메인 특화 메타러닝은 동일 도메인 내의 다른 데이터셋에서 얻은 최적 파이프라인 정보를 활용하여 새로운 데이터셋에 적용하는 방식으로 구현되었다.

4.2 실험 결과

실험 결과, 도메인 특화 메타러닝의 효과는 도메인에 따라 상이한 패턴을 보였다. 표 2는 각 도메인별 평균 정확도(Accuracy) 및 F1 점수(F1 Score) 변화량을 요약한 것으로, 도메인 특화 메타러닝 적용 전후의 성능 차이를 보여준다.

표 2에서 확인할 수 있듯이, 의료 도메인에서는 도메인 특화 메타러닝 적용 시 평균적으로 정확도가 0.0233, F1 점수가 0.0261 향상되는 긍정적인 효과를 보였다. 특히 'diabetes' 데이터셋에서 정확도가 3.25% 향상되었으며, F1 점수는 3.48% 증가하는 주목할 만한 개선을 보였다. 이는 의료 데이터의 특성이 유사한 데이터셋 간에 지식 전이가 효과적으로 이루어질 수 있음을 시사한다.

표 2. 도메인 특화 메타러닝의 성능 비교 분석 결과
Table 2. Performance comparison analysis of domain-specific meta-learning

Domain	Accuracy (Average)	F1 Score (Average)
Medical	+0.0233	+0.0261
Finance	+0.0156	+0.0142
Random	-0.0104	-0.0092

금융 도메인에서는 'German-Credit-Data' 데이터셋에서 정확도가 1.81%, F1 점수가 1.64% 향상되었고, 'credit' 데이터셋에서도 정확도와 F1 점수 모두 1.50% 개선되었다. 'Give-Me-Some-Credit' 데이터셋에서는 정확도가 소폭 개선되었으나 F1 점수는 약간 감소하였다. 이러한 결과는 금융 데이터에서도 도메인 특화 메타러닝이 일정 수준의 성능 향상을 가져올 수 있음을 보여준다.

반면, 대조군으로 설정한 무작위 도메인에서는 메타러닝 적용 효과가 미미하거나 부정적인 결과를 보였다. 'credit' 데이터셋에서는 정확도가 1.68%, F1 점수가 2.03% 감소하였으며, 'Heart_disease_prediction_20' 데이터셋에서는 변화가 거의 없었다. 이는 서로 연관성이 낮은 데이터셋 간의 메타 지식 전이가 효과적이지 않을 수 있음을 시사한다.

V. 결 론

본 연구에서는 현재 AutoML 프레임워크들의 기술적 특성과 한계점을 분석하고, 이를 기반으로 AI 모델의 성능을 고도화하기 위한 혁신적인 기술 전략을 제시하였다. 도메인 특화 메타러닝과 뉴로심볼릭 통합 인터페이스 두 가지 핵심 전략을 중심으로 구체적인 구현 방안을 제안하고, 도메인 특화 메타러닝에 대한 정량적 검증을 수행하였다.

비록 뉴로심볼릭 통합 인터페이스의 효과성에 대한 실험적 검증은 후속 과제로 남겨두었으나, 기존 문헌을 통해 그 가능성을 간접적으로 확인할 수 있다. 예를 들어, A. D. Garcez 와 L. C. Lamb[18]의 연구에서는 뉴로심볼릭 아키텍처가 문맥 이해 및 의미 기반 추론 작업에서 전통적인 딥러닝 모델 대비 높은 성능을 보였으며, 특히 구조화된 지식을 활용함으로써 데이터가 제한된 상황에서도 학습 효율과 정확도를 향상시킬 수 있음을 보고하였다. 이러한 접근은 단순한 패턴 인식에 머무르지 않고, 논리적 제약과 의미 기반 규칙을 결합함으로써 복잡한 문제 해결에 있어 더 나은 일반화 성능을 기대할 수 있음을 시사한다.

또한, K. Yi et al.[13]은 뉴로심볼릭 AI가 모델 해석 가능성을 크게 향상시킬 수 있으며, 그 과정에

서 전통적인 신경망의 성능을 유지하거나 경우에 따라 개선할 수도 있음을 실험을 통해 입증하였다.

이러한 선행 연구들은 본 논문에서 제시한 도메인 특화 메타러닝 프레임워크에 뉴로심볼릭 추론 모듈을 결합할 경우, 데이터 효율성과 해석 가능성 모두를 향상시킬 수 있는 강력한 방향성을 제시해준다.

실험 결과, 의료 도메인에서는 메타러닝 적용 시 정확도가 평균 2.33%, F1 점수는 2.61% 향상되었으며, 금융 도메인에서도 정확도 1.56%, F1 점수 1.42%의 개선 효과를 확인하였다. 특히 의료 도메인의 'diabetes' 데이터셋에서는 F1 점수가 3.48% 증가하는 주목할 만한 성능 향상을 보였다. 반면, 서로 연관성이 낮은 무작위 도메인에서는 메타러닝 적용 효과가 미미하거나 오히려 부정적인 결과를 보여, 도메인 특화 접근의 중요성을 실증적으로 입증하였다.

본 연구의 주요 기여점은 다음과 같다. 첫째, 현재 AutoML 프레임워크들의 한계점을 체계적으로 분석하여 개선 방향을 명확히 제시하였다. 둘째, 도메인 특화 메타러닝의 효과성을 실험을 통해 정량적으로 검증함으로써, 특정 도메인에 최적화된 모델링 환경 구축의 중요성을 입증하였다. 셋째, 뉴로심볼릭 통합 인터페이스를 통해 다양한 데이터 유형을 자동 인식하고 처리하는 새로운 패러다임을 제안하였다.

향후 연구에서는 뉴로심볼릭 통합 인터페이스에 대한 실제 구현과 정량적 검증을 추가로 수행하고, 다양한 산업 분야의 실제 사례에 적용하여 실용성을 평가할 필요가 있다. 또한 각 전략 간의 시너지 효과를 극대화할 수 있는 통합 프레임워크를 개발하고, 제안된 방법론의 계산 효율성과 확장성을 개선하기 위한 추가 연구가 요구된다. 이러한 노력을 통해 AutoML은 단순한 하이퍼파라미터 최적화 도구를 넘어, 도메인 특화된 종합적 AI 솔루션으로 발전할 수 있을 것이다.

References

- [1] Z. Guo, X. Zhang, H. Mu, W. Heng, Z. Liu, Y. Wei, and J. Sun, "Single Path One-Shot Neural Architecture Search with Uniform Sampling", Proc. European Conference on Computer Vision (ECCV), online, pp. 544-560, Aug. 2020. <https://doi.org/10.48550/arXiv.1904.00420>.
- [2] H. Jin, Q. Song, and X. Hu, "Auto-Keras: An Efficient Neural Architecture Search System," in Proc. 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Anchorage, AK, USA, pp. 1946-1956, Aug. 2019. <https://doi.org/10.1145/3292500.3330648>.
- [3] N. Erickson, J. Mueller, A. Shirikov, H. Zhang, P. Larroy, M. Li, and A. Smola, "AutoGluon-Tabular: Robust and Accurate AutoML for Structured Data", arXiv preprint arXiv:2003.06505, Mar. 2020. <https://doi.org/10.48550/arXiv.2003.06505>.
- [4] M. A. Zoller and M. F. Huber, "Benchmark and Survey of Automated Machine Learning Frameworks", Journal of Artificial Intelligence Research, Vol. 70, No. 1, pp. 409-472, Jan. 2021. <https://doi.org/10.48550/arXiv.1904.12054>.
- [5] X. He, K. Zhao, and X. Chu, "AutoML: A Survey of the State-of-the-Art," Knowledge-Based Systems, Vol. 212, No. 4, pp. 106622-106640, Feb. 2021. <https://doi.org/10.1016/j.knosys.2020.106622>.
- [6] C. Thornton, F. Hutter, H. H. Hoos, and K. Leyton-Brown, "Auto-WEKA: Combined Selection and Hyperparameter Optimization of Classification Algorithms", Proc. 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), Chicago, IL, USA, pp. 847-855, Aug. 2013. <https://doi.org/10.1145/2487575.2487629>.
- [7] M. Feurer, A. Klein, K. Eggenberger, J. T. Springenberg, M. Blum, and F. Hutter, "Efficient and Robust Automated Machine Learning", Proc. Advances in Neural Information Processing Systems 28 (NIPS), Montreal, Canada, pp. 2962-2970, Dec. 2015.
- [8] H. Mendoza, A. Klein, M. Feurer, J. T. Springenberg, and F. Hutter, "Towards

- Automatically-Tuned Neural Networks", Proc. Workshop on Automatic Machine Learning (AutoML), New York, NY, USA, pp. 58-65, Jun. 2016.
- [9] T. Swearingen, W. Drevo, B. Cyphers, A. Cuesta-Infante, K. Veeramachaneni, and I. Arnaldo, "ATM: A Distributed, Collaborative, Scalable Framework for Automated Machine Learning", Proc. IEEE International Conference on Big Data (Big Data), Boston, MA, USA, pp. 151-162, Dec. 2017. <https://doi.org/10.1109/BigData.2017.8257923>.
- [10] E. LeDell and S. Poirier, "H2O AutoML: Scalable Automatic Machine Learning", Proc. AutoML Workshop at ICML, Vienna, Austria, Jul. 2020.
- [11] C. Yang, A. Gittens, and I. Stoica, "OBOE: Collaborative Filtering for AutoML", arXiv preprint arXiv:1808.03233, Aug. 2018. <https://doi.org/10.48550/arXiv.1808.03233>.
- [12] F. Hutter, L. Kotthoff, and J. Vanschoren, "Automated Machine Learning: Methods, Systems, Challenges", Springer, pp. 1-270, Dec. 2019. <https://doi.org/10.1007/978-3-030-05318-5>.
- [13] H. Jin, F. Chollet, Q. Song, and X. Hu, "AutoKeras: An AutoML library for deep learning", Journal of Machine Learning Research, Vol. 24, No. 1, pp. 169-174, Jan. 2023. <https://dl.acm.org/doi/10.5555/3648699.3648705>.
- [14] T. Elsken, J. H. Metzen, and F. Hutter, "Neural architecture search: A survey", Journal of Machine Learning Research, Vol. 20, No. 55, pp. 1-74, Mar. 2019. <https://doi.org/10.48550/arXiv.1808.05377>.
- [15] N. Lee, J. Kim, S. Lee, and W. Kim, "2D-CNN of behavior classification model using EEG spectrogram images", Journal of Artificial Intelligence Convergence Technology, Vol. 4, No. 4, pp. 99-104, Dec. 2024.
- [16] J. C. Kim, W. H. Kim, J. H. Jang, and H. Jung, "Search and recommendation systems with metadata extensions", Journal of Artificial Intelligence Convergence Technology, Vol. 4, No. 2, pp. 55-60, Jun. 2024.
- [17] H. Yoo, H. Jung, C. R. Park, S. Jo, and M. H. Lee, "LSTM-based predicting arrhythmia through heart rate sensors in wearable devices", Journal of Artificial Intelligence Convergence Technology, Vol. 4, No. 2, pp. 61-66, Jun. 2024.
- [18] A. D. Garcez and L. C. Lamb, "Neurosymbolic AI: The 3rd Wave," Artif. Intell. Rev., Vol. 56, No. 11, pp. 12387-12406, Nov. 2023. <https://doi.org/10.1007/s10462-023-10448-w>.
- [19] K. Yi, J. Wu, C. Gan, A. Torralba, P. Kohli, and J. B. Tenenbaum, "Neural-Symbolic VQA: Disentangling reasoning from vision and language understanding", arXiv preprint, arXiv:1810.02338, Oct. 2018. <https://doi.org/10.48550/arXiv.1810.02338>.

저자소개

서 동 준 (Dongjoon Suh)



2023년 2월 : 경기대학교
산업공학과(공학사)
2025년 3월 ~ 현재 : 경기대학교
컴퓨터과학과 석사과정
관심분야 : 데이터마이닝,
컴퓨터비전, 딥러닝, 빅데이터,
인공지능

정 경 용 (Kyungyong Chung)



2000년 2월 : 인하대학교
전자계산공학과(공학사)
2002년 2월 : 인하대학교
전자계산공학과(공학석사)
2005년 8월 : 인하대학교
컴퓨터정보공학부(공학박사)
2006년 3월 ~ 2017년 2월 :
상지대학교 컴퓨터정보공학부 교수
2017년 3월 ~ 현재 : 경기대학교 컴퓨터공학부 교수
관심분야 : 데이터 마이닝, 빅데이터, HCI, 인공지능