

AUTOMATED MACHINE LEARNING YORDAMIDA AVTOMATIK MODEL TANLASH

*Farg'ona davlat texnika universiteti
Axborot texnologiyalar va telekommunikatsiya
fakulteti Kompyuter injineri: At-servis yo'nalishi
681-23 guruh talabasi
Topivoldiyev Abbosbek
abbosbektovoldiyev99@gmail.com*

Annotatsiya: Ushbu maqola AutoML texnologiyalaridan foydalanib avtomatik model tanlash jarayonini tahlil qiladi. An'anaviy mashinaviy o'qitishda ko'p vaqt oladigan algoritm tanlash va giperparametrlarni sozlash AutoML yordamida avtomatlashtiriladi. Tadqiqotda AutoML tizimlarining ishlash prinsiplari, optimallashtirish usullari va platformalarning samaradorligi ko'rib chiqildi. Natijalar AutoML model ishlab chiqishni soddalashtirishi va yuqori aniqlikdagi modellarni tez yaratishga imkon berishini ko'rsatdi.

Kalit so'zlar: AutoML, model tanlash, MO, giperparametr sozlash, meta-o'qitish, Bayesian optimallashtirish, NAS, xususiyatlar muhandisligi, model baholash, oldindan qayta ishlash, SI.

Annotation: This article analyzes the automated model selection process using AutoML technologies. In traditional machine learning, algorithm selection and hyperparameter tuning are time-consuming tasks, while AutoML automates these steps to choose optimal models efficiently. The study reviews the principles, optimization techniques, and effectiveness of modern AutoML platforms. Results show that AutoML simplifies model development and enables fast creation of high-accuracy models.

Keywords: AutoML, model selection, ML, hyperparameter tuning, meta-learning, Bayesian optimization, NAS, feature engineering, model evaluation, preprocessing, AI.

Аннотация: В данной статье анализируется процесс автоматического выбора модели с использованием технологий AutoML. В традиционном машинном обучении выбор алгоритма и настройка гиперпараметров требуют много времени, тогда как AutoML автоматизирует эти этапы и позволяет эффективно находить оптимальные модели. В работе рассматриваются принципы AutoML, методы оптимизации и эффективность современных платформ. Результаты показывают, что AutoML упрощает разработку моделей и ускоряет создание высокоточных решений.

Ключевые слова: AutoML, выбор модели, МО, настройка гиперпараметров, мета-обучение, Bayesian-оптимизация, NAS, инженерия признаков, оценка модели, предобработка, ИИ.

Kirish. Sun'iy intellekt va mashinaviy o'qitishning tez rivojlanishi real sektor va sanoatda ma'lumotlarga asoslangan qarorlarni muhimlashtirmoqda. Samarali model yaratish uchun algoritm tanlash, ma'lumotlarni tayyorlash va giperparametrlarni sozlash ko'p vaqt talab qiladi va ekspert tajribasiga tayanadi. Shu sababli AutoML texnologiyasi ahamiyat kasb etadi. AutoML mashinaviy o'qitish pipeline'ini avtomatlashtirib, ma'lumotga mos algoritmlarni solishtiradi, parametrlar makonini izlaydi va eng yaxshi modelni tanlaydi. Bu jarayonda Bayesian optimization, grid search, evolyutsion izlash va meta-learning qo'llanadi, natijada model samaradorligi va umumlashma qobiliyati yaxshilanadi. Maqola AutoML'ning nazariy asoslari, amaliy bosqichlari, mashhur platformalari va real loyihalardagi natijalarini yoritadi hamda afzalliklari, cheklovlari va rivojlanish istiqbollarini muhokama qiladi.

Asosiy qism.

1. AutoML texnologiyasining ilmiy asoslari

AutoML mashinaviy o'qitish jarayonini optimallashtirish va avtomatlashtirishga qaratilgan texnologiyadir. Uning nazariy asosini quyidagi yo'nalishlar tashkil qiladi:

1.1. Model tanlash nazariyasi

Model tanlash statistik yondashuvlarda AIC, BIC kabi kriteriyalar bilan, ML'da esa cross-validation orqali baholanadi. AutoML esa bu jarayonni turli algoritmlar ustida avtomatik ravishda olib boradi.

1.2. Giperparametrlarni qidirish nazariyasi

Giperparametrlar makoni (Hyperparameter Space) ko'pincha millionlab kombinatsiyalarga ega bo'ladi. Uni izlash NP-murakkab masaladir. AutoML bu izlashni quyidagi matematik asoslar orqali hal qiladi:

ehtimollikka asoslangan modellar (Gaussian Process)

sirt bo'yicha optimallashtirish (Surrogate Models)

evolyutsion optimallashtirish

Bu yondashuvlar AutoML'ga murakkab konfiguratsiyalarni samarali topishga yordam beradi.

2. Avtomatik model tanlash jarayoni — chuqur tahlil

AutoML'da model tanlash pipeline ko'p bosqichli jarayon bo'lib, quyidagilardan iborat:

2.1. Ma'lumotlarni avtomatik tahlil qilish

Ma'lumot turlari: kategorik, sonli, matnli, rasm

Disbalansni aniqlash

Korrelyatsiya darajasi

Yo'qolgan qiymatlar diagnostikasi

Outlierlarni aniqlash

AutoML ushbu ma'lumotlarga mos algoritmlar toifasini belgilaydi.

2.2. Avtomatik xususiyatlar muhandisligi

AutoML ko'pincha quyidagi funktsiyalarni avtomatlashtiradi:

Kategorik xususiyatlarni kodlash (One-Hot, Target Encoding)

Scaling / Normalization

Polynomial features generatsiyasi

Automatic feature selection (L1 shrinkage, mutual information)

Bu jarayon odatda model sifatining 40–60% ini belgilaydi.

2.3. Algoritmlar orasidan optimalini tanlash

AutoML binar, multiklass, regressiya va klasterlash vazifalariga qarab mos algoritmnı sinab ko'radi. Masalan:

Kichik dataset uchun: Logistic Regression, Decision Trees

O'rta o'lchamli dataset uchun: Random Forest, XGBoost

Juda murakkab ma'lumotlar uchun: Neural Networks, NAS orqali yaratilgan tarmoqlar

Tanlov cross-validation orqali tasdiqlanadi.

2.4. Giperparametrlarni chuqur optimallashtirish

Giperparametr optimallashtirishda AutoML quyidagi mexanizmlardan foydalanadi:

- **Bayesian Optimization**

Qidiruvni oldingi natijalardan o'rganib boradi. Tez va aniq yo'l.

- **Genetik algoritmlar**

Tabiatdagi evolyutsiya prinsipiga asoslanadi:

mutatsiya

krossover

tanlanish

- **Hyperband**

Vaqtni oshirmasdan optimal parametr topishga yordam beradi.

- **Reinforcement Learning asosidagi qidiruv**

Asosan NAS'larda qo'llanadi.

3. Meta-learningning AutoMLdagi o'rni

Meta-learning AutoML tizimlariga "tajriba banki" vazifasini bajaradi. Bu jarayon quyidagicha ishlaydi:

1. Minglab ML muammolarining tarixi saqlanadi.

2. Har bir datasetning morfologik xususiyatlari (feature statistics) o'rganiladi.

3. O'xshash datasetlar uchun samarali bo'lgan algoritmlar tavsiya qilinadi.

Natija: izlash va tanlash jarayoni 3–10 baravarga tezlashadi.

4. Neural Architecture Search (NAS) — AutoMLning eng ilg'or yo'nalishi

NAS avtomatlashtirilgan neyron tarmoq arxitekturasini yaratish texnologiyasidir.

4.1. NAS turlari

RL-NAS — mukofotga asoslangan arxitektura tanlash

Evo-NAS — evolyutsiyaga asoslangan qidiruv

Gradient-based NAS — Differentiable Architecture Search (DARTS)

4.2. NAS afzalliklari

Qo'lda yaratilgan arxitekturalardan 15–30% yuqori aniqlik

Layerlar soni, turi, bog'lanishi avtomatik topiladi

Turli vazifalar uchun maxsus model ishlab chiqiladi

5. Zamonaviy AutoML platformalari — chuqur solishtirish

Jadval qvolasan

Platforma	Afzalliklar	Cheklovlar
Google AutoML	Matn, rasm, video bilan yaxshi ishlaydi; yuqori aniqlik	Narxi yuqori
Auto-sklearn	Meta-learning, Bayesian Optimization	Katta datasetlarda sekin
H2O AutoML	Juda tez, ko'p algoritmlar	NAS yo'q
TPOT	Evolyutsion pipeline optimallashtirish	Hisoblash talabi yuqori
AutoKeras	Kuchli NAS	Ma'lumot oldindan tayyorligi talab qilinadi

6. AutoMLning amaliy afzalliklari

70–80% vaqt tejash

Ekspert bo'lmagan foydalanuvchi ham kuchli model yarata oladi

Ko'plab sinovlar avtomatik o'tkaziladi

Tizim har doim eng yaxshi model kombinatsiyasini topishga harakat qiladi

7. Cheklovlar va muammolar

Qidiruv jarayoni juda ko'p resurs talab qiladi

Modelning qanday tanlanganini tushuntirish qiyin ("black box")

Shovqinli yoki sifatsiz ma'lumotlarda unumdorlik pasayadi

Xulosa

AutoML yordamida avtomatik model tanlash jarayoni zamonaviy mashinaviy o'qitishning eng muhim va istiqbolli yo'nalishlaridan biri bo'lib, u murakkab hisoblangan modellash jarayonlarini sezilarli darajada soddalashtirishga hamda optimallashtirishga xizmat qiladi. Tadqiqot davomida ko'rsatildiki, AutoML platformalari algoritm tanlash, giperparametrlarni izlash, xususiyatlar muhandisligi va modelni baholash kabi jarayonlarni integratsiyalashgan holda amalga oshirish orqali

nafaqat ekspert bilimiga bo'lgan ehtiyojni kamaytiradi, balki model aniqligini oshiruvchi optimal konfiguratsiyalarni ham samarali ravishda topa oladi.

Ilmiy tahlillar shuni ko'rsatadiki, Bayesian Optimization, meta-learning, evolyutsion qidiruv va Neural Architecture Search kabi ilg'or texnologiyalarning AutoML jarayoniga tatbiqi model tanlashning aniqligi va barqarorligini sezilarli oshiradi. Bu esa AutoML'ni nafaqat eksperimental tadqiqotlar uchun, balki real sektor, sanoat ishlab chiqarishi, tibbiy tashxislash, moliyaviy prognozlash kabi amaliy yo'nalishlarda ham jiddiy qo'llash imkonini yaratadi.

Biroq AutoML'ning hisoblash resurslariga yuqori talab qo'yishi, ba'zi jarayonlarning "qora quti" sifatida talqin qilinishi va ma'lumot sifati past bo'lgan holatlarda unumdorlikning pasayishi kabi cheklovlar mavjudligi ham ilmiy asosda qayd etildi. Ushbu cheklovlar AutoML sohasining kelajakdagi tadqiqotlarini yanada chuqurlashtirish, algoritmik shaffoflikni oshirish, izlash jarayonini tezlashtirish va resurs sarfini kamaytirishga qaratilgan yangi yechimlarni ishlab chiqishni talab qiladi.

Umuman olganda, olib borilgan tahlil va kuzatuvlar AutoML yordamida avtomatik model tanlash texnologiyasi mashinaviy o'qitishning kelajakdagi rivojida muhim o'rin tutishini, ilmiy va amaliy jarayonlarda samaradorlikni sezilarli oshirishini hamda zamonaviy intellektual tizimlarning ajralmas qismi bo'lib borayotganini yaqqol ko'rsatmoqda.

Foydalanilgan adabiyotlar

1. Hutter, F., Kotthoff, L., & Vanschoren, J. (2019). *Automated Machine Learning: Methods, Systems, Challenges*. Springer.
2. Feurer, M., & Hutter, F. (2019). Hyperparameter Optimization. In *Automated Machine Learning* (pp. 3–33). Springer.
3. Olson, R. S., & Moore, J. H. (2016). TPOT: A Tree-Based Pipeline Optimization Tool for Automating Machine Learning. In *Proceedings of the 2016 Python in Science Conference*.
4. He, X., Zhao, K., & Chu, X. (2021). *AutoML: A Survey of the State-of-the-Art Knowledge-Based Systems*.
5. Thornton, C., Hutter, F., Hoos, H. H., & Leyton-Brown, K. (2013). *Auto-WEKA: Combined Selection and Hyperparameter Optimization of Classification Algorithms*. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.