

AutoML: advanced tool for mining multivariate plant traits

Mirza Shoab,^{1,2} Neelesh Sharma,^{1,2} Lars Kotthoff,³ Marius Lindauer,⁴ and Surya Kant^{1,2,5,*}

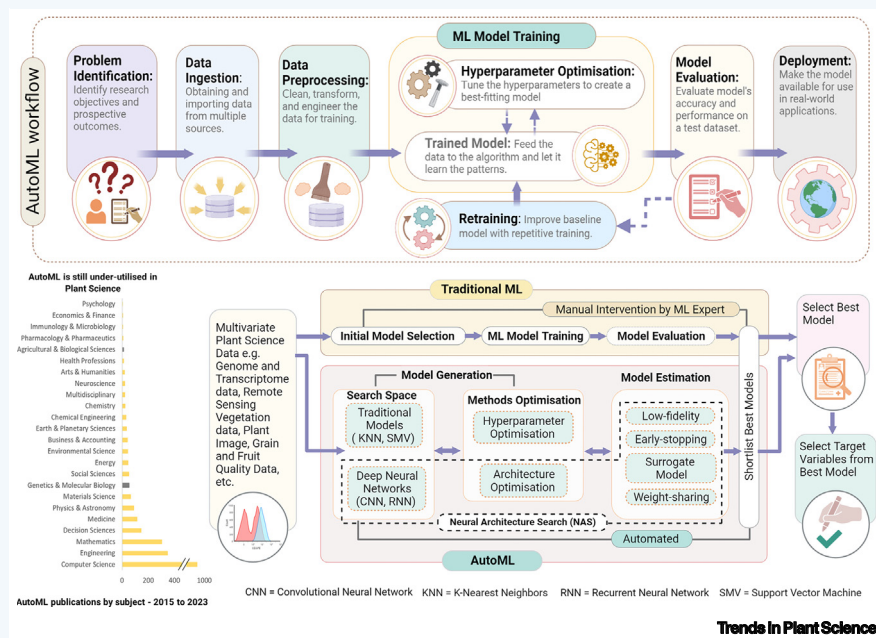
¹Agriculture Victoria, Grains Innovation Park, 110 Natimuk Rd, Horsham, Victoria 3400, Australia

²School of Applied Systems Biology, La Trobe University, Bundoora, Victoria 3083, Australia

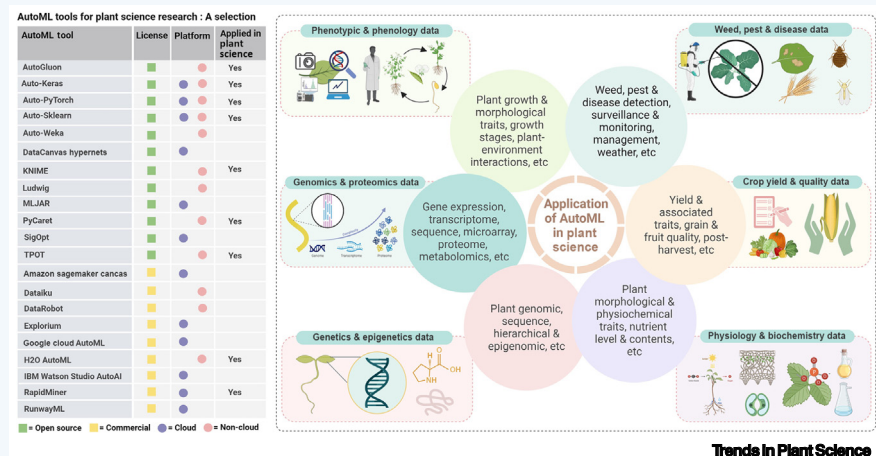
³Department of Electrical Engineering and Computer Science, School of Computing, University of Wyoming, Laramie, WY 82071, USA

⁴Institute of Artificial Intelligence (LUAI), Leibniz University Hannover, Appelstr. 9A, 30167 Hannover, Germany

⁵Agriculture Victoria, AgriBio, Centre for AgriBioscience, 5 Ring Road, Bundoora, Victoria 3083, Australia



Automated machine learning (AutoML) is an automated version of machine learning (ML) that has the potential to become an integral part of plant science research for dealing with large and complex multivariate datasets. AutoML is rapidly advancing and has the potential to revolutionise the development and deployment of ML as it fully automates the ML process, from selecting the appropriate model and optimising its hyperparameters, to handling data preparation in some instances. This automation saves time and effort, enhances model quality, and increases its usability for novice users such as plant scientists.



AutoML excels in extracting meaningful features from diverse datasets in plant science domains, enhancing understanding of precision agriculture, crop breeding, and disease forecasting; estimating abiotic stresses, molecular genetics, and proteomics analysis; and yield prediction. As AutoML tools evolve, they will become even more powerful and user-friendly, accelerating innovation in plant science.

ADVANTAGES:

AutoML automates feature engineering and scalability, making it easier to extract meaningful information from complex multivariate and large biological datasets.

Efficiency is increased by automating repetitive tasks, reducing errors, and automatically selecting the best features and models for improved accuracy.

Reproducibility is improved by generating a detailed report of the model-building and evaluation process.

ML is democratized by providing a drag-and-drop interface that eliminates the need for programming skills, making ML accessible and usable for beginners.

It integrates seamlessly with existing tools and pipelines, streamlining research workflows and simplifying model building.

CHALLENGES:

Interpreting AutoML, particularly deep learning models, can occasionally present challenges in gaining insights from data, like traditional ML. Therefore, consideration of interpretable AutoML tools may be required in some projects.

In some cases, AutoML training and implementation can be challenged by limited or missing data in plant science.

It lacks the same customisation or flexibility as coding-based approaches, necessitating careful evaluation and domain-specific knowledge.

Most AutoML platforms are versatile and can be used for plant science research. Some are already being used in this field, but others still need to be tested, deployed, and evaluated in the specific plant science domain.

*Correspondence: surya.kant@agriculture.vic.gov.au (S. Kant).

Declaration of interests

No interests are declared.

Literature

1. Waring, J. *et al.* (2020) Automated machine learning: review of the state-of-the-art and opportunities for healthcare. *Artif. Intell. Med.* 104, 101822
2. Escalante, H.J. (2021) Automated machine learning—a brief review at the end of the early years. In *Automated Design of Machine Learning and Search Algorithms* (Pillay, N. and Qu, R., eds), pp. 11–28, Springer
3. Hutter, F. *et al.* (2019) *Automated Machine Learning: Methods, Systems, Challenges*, Springer
4. Korot, E. *et al.* (2021) Code-free deep learning for multi-modality medical image classification. *Nat. Mach. Intell.* 3, 288–298
5. He, X. *et al.* (2021) AutoML: a survey of the state-of-the-art. *Knowl. Based Syst.* 212, 106622
6. Nagarajah, T. and Poravi, G. (2019) A review on automated machine learning (AutoML) systems. In *2019 IEEE 5th International Conference for Convergence in Technology (I2CT)*, IEEE
7. Truong, A. *et al.* (2019) Towards automated machine learning: evaluation and comparison of AutoML approaches and tools. In *2019 IEEE 31st International Conference on Tools for Artificial Intelligence (ICTAI)*, IEEE
8. Ferreira, L. *et al.* (2021) A comparison of AutoML tools for machine learning, deep learning and XGBoost. In *2021 International Joint Conference on Neural Networks*, IEEE
9. Doke, A. and Gaikwad, M. (2021) Survey on automated machine learning (AutoML) and meta learning. In *2021 12th International Conference on Computing and Networking Technologies*, IEEE
10. Wever, M. *et al.* (2021) AutoML for multi-label classification: overview and empirical evaluation. *IEEE Trans. Pattern Anal. Mach. Intell.* 43, 3037–3054
11. Das, S. and Cakmak, U.M. (2018) *Hands-on Automated Machine Learning: A Beginner's Guide to Building Automated Machine Learning Systems Using AutoML and Python*, Packt Publishing