International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-15 Issue-3, February 2026

Comparative Study of Automated Machine Learning Services on Cloud Platforms

Duckki Lee

Abstract*: Automated machine learning (AutoML) has emerged as a practical approach to facilitate the adoption of machine learning by automating model development tasks, including preprocessing, model selection, and hyperparameter optimisation, thereby reducing reliance on specialised expertise. Recently, major cloud providers have integrated AutoML into their platforms to offer end-to-end machine-learning pipelines as managed services. However, the practical implications of cloud-based AutoML systems, particularly their system and operational aspects, remain insufficiently explored. This paper presents an empirical, system-oriented analysis of AutoML services provided by Microsoft Azure, Amazon Web Services, and Google Cloud Platform. Using representative regression and binary classification tasks, the predictive performance, evaluation metrics, and feature-importance results produced by each platform are compared. The study also examines how platform-level design choices influence usability, reproducibility, and lifecycle management. The results demonstrate that cloud AutoML platforms deliver high-performing models that operate without manual intervention, whereas differences among providers primarily reflect architectural and operational abstractions rather than algorithmic limitations. These findings suggest that cloud AutoML should be understood as an integrated system that combines automated modelling and MLOps capabilities and offers a viable pathway toward production-ready machine learning under real-world constraints.

***Keywords*: AutoML, MlaaS, Cloud Computing, Performance Evaluation**

Nomenclature:
AutoML: Automated Machine Learning
AWS: Amazon Web Services
GCP: Google Cloud Platform
RMSE: Root Mean-Squared Error
MAE: Mean Absolute Error

I. INTRODUCTION

Machine learning has become a core technology for extracting practical knowledge from rapidly growing data volumes across diverse domains, including computer vision, natural language processing, healthcare, finance, and industrial systems [1], [2]. As the availability of data and computational resources increases, machine-learning models have demonstrated remarkable performance on tasks ranging from image and speech recognition to complex decision-making and generative modelling. Consequently, organizations across various industries are actively attempting to incorporate machine learning into their business and operational processes. Despite these advances, the practical adoption of machine learning in real-world environments remains challenging. Prior studies have consistently reported that the performance and reliability of machine-learning systems are highly sensitive to data quality, feature engineering, model selection, and hyperparameter optimization [3]–[5]. Modern machine-learning pipelines require extensive manual effort by skilled data scientists to preprocess heterogeneous datasets, select appropriate algorithms, iteratively tune hyperparameters, and maintain reproducibility across heterogeneous environments [6], [7]. These requirements significantly increase development time and cost, while limiting the accessibility of machine-learning technologies to non-expert users.

Recent research has highlighted that, in addition to model-level challenges, many failures in production machine-learning systems stem not from algorithmic limitations but from system-level issues, such as latent technical debt, poor pipeline management, and lack of lifecycle governance [8], [9]. As machine-learning systems evolve from experimental prototypes to long-running services, problems related to versioning, reproducibility, monitoring, and collaboration have become increasingly critical. Empirical studies have shown that even state-of-the-art machine-learning platforms often struggle to provide reproducible out-of-the-box results, particularly when models are retrained or deployed across heterogeneous environments [10].

Automated machine learning (AutoML) has emerged as a promising approach to address these challenges by automating key stages of machine-learning pipelines, including data preprocessing, feature selection, model search, hyperparameter optimisation, and evaluation [5], [11]. By reducing reliance on expert intervention, AutoML aims to democratize machine learning and enable practitioners without deep expertise in statistics or optimization to build competitive models. Early AutoML research primarily focused on algorithmic efficiency and search strategies; however, recent studies have emphasized that AutoML should be understood as a system-level solution rather than a standalone optimization technique.

In this context, cloud-based AutoML services are a significant development in the AutoML paradigm. By integrating AutoML capabilities into cloud platforms, these services combine automated model development with scalable infrastructure, a standardised environment, collaborative workflows, and built-in

Manuscript received on 31 January 2026 | First Revised Manuscript received on 08 February 2026 | Second Revised Manuscript received on 12 February 2026 | Manuscript Accepted on 15 February 2026 | Manuscript published on 28 February 2026.
*Correspondence Author(s)
Duckki Lee*, Associate Professor, Department of Smart Software, Yonam Institute of Technology, Jinju-si (Gyeongsangnam-do), Korea Republic of. Email ID: dlee@yc.ac.kr, ORCID ID: 0000-0002-1711-8094

deployment pipelines. Cloud AutoML systems are often offered as Artificial Intelligence-as-a-Service (AIaaS) or Machine Learning-as-a-Service (MLaaS), providing users with end-to-end machine learning pipelines via web-based interfaces and managed services [12].

From a systems engineering perspective, cloud AutoML can be viewed as a foundational framework for adopting MLOps, offering automated solutions for managing model lifecycles, tracking experiments, managing artefacts, and maintaining continuous deployment [13]–[14]. Unlike traditional machine-learning workflows, in which reproducibility and governance must be manually engineered, cloud platforms increasingly embed best practices for production readiness directly into their AutoML pipelines. This integration addresses several long-standing challenges identified in prior studies [8], [9], [15], including the accumulation of technical debt, the lack of standardised evaluation criteria, and fragmented toolchains.

Although several comparative studies have examined AutoML frameworks in specific application domains [16], [17], empirical analyses that focus on cloud-native AutoML platforms as integrated systems rather than isolated modelling tools are lacking. Understanding how different cloud providers operationalise AutoML through design choices in optimisation metrics, feature-importance reporting, and pipeline automation would provide valuable insights for researchers and practitioners seeking to adopt AutoML in real-world environments.

Therefore, this study investigated the AutoML services provided by major cloud platforms, Microsoft Azure, Amazon Web Services (AWS), and Google Cloud Platform (GCP), from a practical and system-oriented perspective. Using representative regression and binary classification tasks, we compared model performance, evaluation metrics, and feature-importance results across the platforms. Rather than proposing new algorithms, our goal was to elucidate how cloud AutoML systems translate theoretical automation concepts into deployable machine-learning solutions and identify their strengths and limitations in applied settings.

The contributions of this paper are summarized as follows:

- We provide a system-level overview of cloud AutoML as an integrated solution that combines automated modelling, scalable infrastructure, and MLOps capabilities.
- The leading cloud AutoML platforms are empirically compared using standardized datasets and evaluation metrics.
- The implications of platform-level design choices on model performance, reproducibility, and practical usability are discussed.

Through this analysis, we aim to offer guidance to organizations and researchers considering cloud AutoML as a viable approach for deploying machine-learning systems under real-world constraints.

II. RELATED WORK

A. AutoML Foundations

AutoML was originally proposed to reduce the reliance on expert knowledge in the design and optimization of machine-learning models. Early AutoML research focused on automating algorithm selection, feature engineering, and hyperparameter optimization to improve model performance while reducing human intervention. Comprehensive overviews of AutoML methods and challenges [5], [11] classified AutoML approaches into those for pipeline composition, optimization strategies, and evaluation mechanisms.

Moreover, as noted in surveys [4], [5], AutoML is not a single technique, but a collection of methods aimed at automating different stages of the machine-learning workflow, including data preprocessing, model selection, and validation. These studies highlight that the effectiveness of AutoML depends not only on search algorithms but also on the surrounding system architecture, which determines scalability, robustness, and usability.

B. AutoML and Hyperparameter Optimization

Hyperparameter optimization remains a central component of AutoML systems because model performance is highly sensitive to hyperparameter configurations. Because no closed-form solution exists for selecting optimal hyperparameters, most AutoML frameworks rely on iterative search strategies, such as grid search, random search, Bayesian optimisation, and multifidelity optimisation [6], [7].

A recent study [6] demonstrated that massively parallel hyperparameter tuning can significantly reduce optimization time but introduces new challenges related to resource allocation and experiment management. Consequently, effective AutoML systems must balance optimization efficiency with computational cost, particularly in large-scale or cloud-based environments. These considerations have motivated the increasing integration of AutoML with scalable infrastructure and automated experiment tracking on cloud platforms.

C. Production Machine Learning, Technical Debt, and Reproducibility

As machine-learning systems transition from research prototypes to production services, system-level concerns have become as critical as model accuracy. Sculley et al. introduced the concept of hidden technical debt in machine-learning systems, arguing that data dependencies, configuration complexity, and pipeline fragility can lead to long-term maintenance challenges [8]. Subsequent studies proposed structured evaluation frameworks, such as the ML Test Score, to assess production readiness and reduce technical debt through systematic testing and governance [9].

Reproducibility has emerged as a major concern in machine-learning applications. Empirical studies have shown that many machine-learning platforms do not guarantee reproducible results without additional manual configuration, particularly when pipelines involve multiple tools and environments [10].

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Furthermore, the management of machine-learning artefacts, such as datasets, trained models, and evaluation results, has been identified as a key challenge in the lifecycle of machine-learning systems [15].

These studies collectively suggest that the successful deployment of machine learning requires not only accurate models but also a robust infrastructure for lifecycle management, reproducibility, and collaboration.

D. Cloud AutoML Platforms and Comparative Studies

Cloud computing has played a crucial role in addressing the scalability and operational challenges of AutoML by providing managed infrastructure and a standardised environment. Cloud AutoML platforms integrate automated model development with scalable computing resources, enabling users to perform complex optimisation tasks without managing the underlying hardware.

The convergence between AutoML and MLOps has grown [12]–[14]. Automated pipelines support continuous training, evaluation, deployment, and monitoring of models, and cloud AutoML can be viewed as an operationalization of AutoML principles within mature MLOps ecosystems.

Several comparative studies have evaluated AutoML platforms in specific application domains. For example, Choi et al. conducted a comparative analysis of AutoML platforms for anthropometry-based typology analysis. They demonstrated that platform-level design choices influence both performance metrics and the interpretation of feature importance [17]. However, many studies have focused on algorithmic outcomes rather than on cloud AutoML platforms as integrated systems encompassing automation, reproducibility, and lifecycle management.

This study extends prior work by providing an empirical comparison of major cloud AutoML platforms from a systems-oriented perspective, emphasizing practical performance metrics and usability considerations in real-world machine-learning tasks.

III. AUTOML IN CLOUD SERVICES

Cloud AutoML extends the concept of automation from simple model optimization to the entire machine-learning lifecycle. Whereas traditional AutoML focuses on automating tasks such as data preprocessing, model selection, and hyperparameter tuning, cloud AutoML platforms integrate these capabilities with scalable infrastructure, standardized execution environments, and operational pipelines.

From a systems perspective, cloud AutoML can be regarded as a managed abstraction layer that shields users from the complexity of machine-learning system engineering. Instead of manually configuring libraries, managing dependencies, and orchestrating experiments, users interact with AutoML services through web-based interfaces or APIs. In contrast, the underlying platform handles execution, scaling, and resource allocation.

A. End-to-End Pipeline Automation

Cloud AutoML encapsulates the entire machine-learning lifecycle, from data ingestion to development, within a managed framework. Once a dataset has been uploaded and a target variable has been specified, the platform automatically performs data validation, preprocessing, feature transformation, model training, and evaluation. This pipeline-level automation addresses several sources of technical debt identified in production machine-learning systems, such as ad hoc preprocessing scripts and undocumented configuration dependencies [8], [9].

Standardising pipeline execution enables cloud platforms to reduce variability across experiments and to facilitate reproducibility, a persistent challenge in applied machine learning [10].

B. Scalability and Resource Management

Cloud AutoML platforms leverage elastic cloud infrastructure to scale computation in response to workload demand dynamically. Hyperparameter optimization and model search, which are computationally expensive in traditional environments, can be executed in parallel across distributed resources without requiring users to manage clusters explicitly.

This scalability enables efficient exploration of large search spaces while maintaining reasonable execution time and cost, consistent with prior research on massively parallel hyperparameter optimisation [6]. Moreover, the pay-as-you-go pricing model allows organizations to control costs by allocating resources only when needed.

C. Standardized Environments and Reproducibility

A major barrier to collaboration in machine-learning projects is environmental inconsistency. Differences in library versions, hardware configurations, and execution settings often result in irreproducible results. Cloud AutoML platforms mitigate this issue by providing prebuilt and version-controlled environments, ensuring that experiments are executed under consistent conditions.

Such standardisation addresses concerns regarding reproducibility and lifecycle artefact management in machine-learning systems [10], [15]. As a result, cloud AutoML both simplifies development and improves governance and traceability.

D. Integration with MLOps Practices

Successful machine-learning deployment requires robust MLOps practices, including continuous integration, model versioning, monitoring, and lifecycle management [12]–[14]. Cloud AutoML platforms increasingly embed these practices into their workflows by supporting automated pipelines.

Through features such as experiment tracking, model registry, and deployment automation, cloud AutoML serves as a practical gateway to MLOps for organizations that lack dedicated machine-learning operations teams. This integration transforms AutoML from a convenient modelling tool to a production-ready system component.

IV. COMPARISON OF AUTOML IN CLOUD SERVICES

In this section, we evaluate the system-oriented performance of AutoML services provided by Microsoft Azure Machine Learning, AWS SageMaker,

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

and GCP Vertex AI. The study focuses on representative regression and binary classification tasks to compare the automated modelling capabilities of these tasks.

A. Experimental Setup and Dataset Description

To ensure a rigorous and reproducible evaluation, we established a protocol focused on "out-of-the-box" performance using the managed web interfaces of each provider. The following criteria governed the experiments:

i. *Managed Infrastructure:* We utilized the fully managed automation modes, where the platforms automatically orchestrate the underlying compute infrastructure and end-to-end pipelines without manual user intervention.

ii. *Operational Protocol:* All configurations were kept at their default values, with user input strictly limited to defining the target variable for each task.

iii. *Regression Dataset:* The US Health Insurance dataset [18] consists of 1,338 individual records used to predict insurance charges based on personal attributes.

iv. *Classification Dataset:* The Bank Marketing dataset [19] comprises 32,950 records from a Portuguese bank's marketing campaign to determine term deposit subscriptions.

B. Evaluation Metrics

As recommended for systematic AutoML benchmarking, the following metrics were defined and selected before the experiments to ensure a consistent comparative analysis:

i. *Regression Metrics:* Performance was measured using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). Notably, RMSE was prioritised as a consistent benchmark because GCP Vertex AI primarily reports RMSE as its default metric for regression.

ii. *Classification Metrics:* For binary classification, the Area Under the Receiver Operating Characteristic Curve (AUC) was used as the primary metric for optimization, as it is a standard accuracy metric for such models [4], [20]. To provide a balanced assessment, four additional metrics derived from the confusion matrix were employed:

- **Accuracy:** The proportion of correct predictions to the total number of predictions.
- **Precision:** The proportion of correct positive predictions to all positive predictions.
- **Recall:** The proportion of correct positive predictions to the actual number of positive instances.
- **F1-Score:** The harmonic average of precision and recall, providing a performance measure without bias toward either metric.

C. Performance Comparison Result

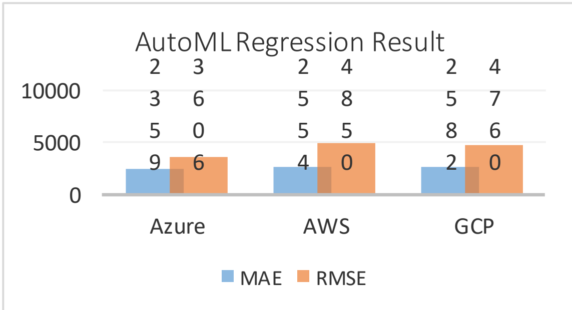
Each cloud provider provides AutoML services under the names Azure Machine Learning, AWS SageMaker, and GCP Vertex AI. The results of regression on the US Health Insurance are reported in Table I and Fig. 1, and the results of binary classification on the Bank Marketing dataset are summarized in Table II and Fig. 2. In the tables, only performance indicators provided by all platforms are reported. For example, GCP only provides the root mean-squared error (RMSE) for regression, instead of the more commonly used mean-squared error; therefore, only the RMSE was reported in Table I. For the binary classification in Table II, the most frequently used performance indicators (AUC, accuracy, precision, recall, and F1-Score) are listed. The top 5 most important features of the models are also presented.

Table I: Results of Regression on the US Health Insurance Dataset

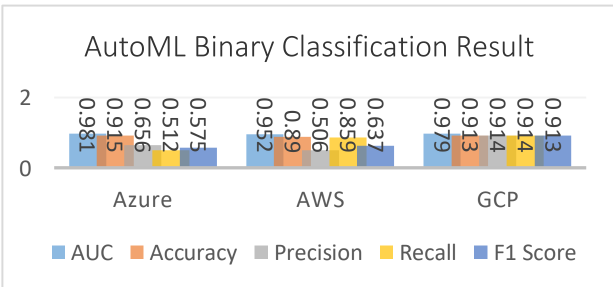
Metric	Platform		
	Azure Machine Learning	AWS SageMaker	GCP Vertex AI
MAE	2,359	2,554	2,582
RMSE	3,606	4,850	4,760
Feature Importance (Top 5)	1. Smoker 2. BMI 3. Age 4. Children 5. Region	1. Smoker 2. Age 3. BMI 4. Children 5. Region	1. Smoker 2. Age 3. BMI 4. Children 5. Region

Table II: Results of Binary Classification on the Bank Marketing Dataset?

Metric	Platform		
	Azure Machine Learning	AWS SageMaker	GCP Vertex AI
AUC	0.981	0.952	0.979
Accuracy	0.915	0.89	0.913
Precision	0.656	0.506	0.914
Recall	0.512	0.859	0.914
F1-Score	0.575	0.637	0.913
Feature Importance (Top 5)	1. duration 2. nr.employed 3. emp.var.rate 4. euribor3m 5. cons.conf.idx	1. duration 2. nr.employed 3. emp.var.rate 4. euribor3m 5. month	1. duration 2. nr.employed 3. euribor3m 4. month 5. cons.price.idx



[Fig.1: AutoML Regression Results]



[Fig.2: AutoML Binary Classification Result

i. *Regression Analysis:* The MAE and RMSE results reveal that Azure performed the best, while AWS

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

and GCP achieved comparable performance levels. Feature importance rankings were consistent across platforms, with 'Smoker' identified as the most critical factor.

ii. *Classification Analysis:* While all models exhibited similar Accuracy, the F1-Score showed significant variation across platforms. Notably, Azure achieved the highest overall accuracy, whereas GCP demonstrated the superior F1-Score (0.913), indicating a highly balanced predictive performance between precision and recall.

V. DISCUSSION

The experimental results presented in this study demonstrate that cloud AutoML platforms can produce competitive models for both regression and binary classification tasks without manual intervention. However, we should be careful not to treat these performance differences as a simple contest of which provider's algorithm is better than the others. Instead, these variations are more likely to reflect how each company has designed its overall automated system. Since each platform follows its own rules for cleaning data, selecting features, or optimising model settings, the results we observe are influenced by the overall design choices rather than by the underlying mathematics alone.

A clear example of this is the difference in how evaluation outputs are handled. For instance, the fact that GCP Vertex AI defaults to RMSE for regression, whereas other platforms might emphasise different metrics, shows how platform-level abstractions can shape a user's understanding of their model's quality. While these simplified interfaces make machine learning more accessible to non-experts, they can also hide the specific methodological differences between providers. This underscores the need to view AutoML platforms as integrated systems that manage the entire model lifecycle, rather than as standalone optimisation tools.

Despite these internal differences, it is encouraging that the feature-importance rankings were highly consistent across all three services. All platforms successfully identified the same primary drivers for our datasets, such as 'Smoker' for the insurance task and 'Duration' for the bank marketing task. This level of reliability suggests that cloud-native AutoML is sufficiently robust to capture the most important data patterns, making it a useful tool for rapid analysis in real-world settings. These observations also support the idea that the success of machine learning in production is as much about system-level robustness and MLOps as it is about raw accuracy.

A. Study Limitations and Practical Challenges

It is important to acknowledge the specific constraints of this study when interpreting these findings. First, our analysis was limited to three major cloud providers and to two primary datasets. While these are common business problems, they do not represent the full range of data complexities or the specific challenges found in other industrial fields.

Second, we encountered practical technical hurdles that underscore the rigidity of these managed services. For example, during our initial setup, we found that certain platforms impose stringent data requirements; we observed that GCP Vertex AI could not process datasets with fewer than 1,000 records. This type of system constraint forced us to limit our final analysis to larger datasets, indicating that cloud AutoML is not always a flexible solution across all data scales. Finally, the "black-box" nature of these services remains a barrier for advanced users seeking full transparency. Since we cannot observe how the internal architectures were selected, there remains a need for more open benchmarking protocols and improved interpretability in future research.

VI. CONCLUSION AND FUTURE WORKS

This study examined cloud AutoML platforms from both practical and system-oriented perspectives. Rather than proposing new learning algorithms, this study focused on how major cloud providers operationalise AutoML as an integrated solution that combines automated modelling, scalable infrastructure, and lifecycle management. Empirical comparisons across regression and binary classification tasks demonstrate that cloud AutoML platforms can deliver competitive predictive performance without requiring extensive manual intervention or expert-level configuration.

The findings suggest that the primary value of cloud AutoML lies not only in model optimization but also in its ability to reduce system complexity and the technical debt associated with traditional machine-learning pipelines. By standardizing data preprocessing, model training, and evaluation workflows, cloud AutoML platforms address long-standing challenges related to reproducibility, collaboration, and operational readiness. These characteristics align with recent research guidance that emphasises the importance of system robustness and MLOps integration in real-world machine-learning deployments.

From a comparative standpoint, the differences observed among the cloud providers reflect platform-specific design choices rather than the fundamental limitations of AutoML. Variations in performance metrics, optimisation strategies, and feature-importance reporting indicate that cloud AutoML should be understood as a managed abstraction layer, where ease of use and operational efficiency are prioritised alongside predictive accuracy. Consequently, direct performance comparisons should be interpreted in the context of each platform's architectural and operational assumptions.

The implications of this work extend to organisations seeking to adopt machine learning under practical constraints, such as limited expertise, resource constraints, or operational maturity. Cloud AutoML offers a viable entry point for deploying machine-learning systems that adhere to emerging best practices in MLOps, thereby enabling broader adoption while mitigating common risks associated with production systems.

Future work should explore the standardization of evaluation protocols and benchmarking methodologies for cloud AutoML platforms to improve cross-platform comparability. Further investigation into transparency, controllability, and lifecycle governance within AutoML systems would support the

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

development of more robust and interpretable machine-learning services. As cloud AutoML continues to evolve, its role as a key component of production-ready machine-learning infrastructure is expected to grow.

DECLARATION STATEMENT

After aggregating input from all authors, I must verify the accuracy of the following information as the article's author.

- **Conflicts of Interest/ Competing Interests:** Based on my understanding, this article has no conflicts of interest.
- **Funding Support:** Yes, I have received financial assistance for this article. **The funding assistance for this article is as follows:** The Yonam Institute of Technology provided the funding assistance for this article through its internal research fund.
- **Ethical Approval and Consent to Participate:** The content of this article does not necessitate ethical approval or consent to participate with supporting documentation.
- **Data Access Statement and Material Availability:** The adequate resources of this article are publicly accessible.
- **Author's Contributions:** The author solely contributed to this research and is fully responsible for the content of this article.

REFERENCES

1. E. Dritsas and M. Trigka, "Exploring the Intersection of Machine Learning and Big Data: A Survey," *MAKE*, vol. 7, no. 1, p. 13, Feb. 2025, DOI: https://doi.org/10.3390/make7010013
2. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 770-778, 2016, DOI: https://doi.org/10.1109/CVPR.2016.90
3. R. Elshawi, M. Maher, and S. Sakr, "Automated machine learning: State-of-the-art and open challenges," arXiv preprint arXiv:1906.02287, 2019. [Online]. Available: https://arxiv.org/abs/1906.02287
4. M. Zöller and M. F. Huber, "Benchmark and Survey of Automated Machine Learning Frameworks," Journal of Artificial Intelligence Research, vol. 70, pp. 409-472, 2021, DOI: https://doi.org/10.1613/jair.1.11854
5. X. He, K. Zhao, and X. Chu, "AutoML: A survey of the state-of-the-art," Knowledge-Based Systems, vol. 212, Art. no. 106622, 2021, DOI: https://doi.org/10.1016/j.knosys.2020.106622
6. L. Li et al., "Massively parallel hyperparameter tuning," arXiv preprint arXiv:1810.05934, 2018. [Online]. Available: https://arxiv.org/abs/1810.05934
7. Microsoft Azure, "Tune hyperparameters for your model," 2024. [Online]. Available: https://learn.microsoft.com/en-us/azure/machine-learning/how-to-tune-hyperparameters
8. G. A. A. Santos et al., "Technical Debt in Machine Learning: A Systematic Mapping Study," Journal of Systems and Software, vol. 184, art. no. 111130, 2022, DOI: https://doi.org/10.1016/j.jss.2021.111130
9. E. Breck et al., "The ML test score: A rubric for ML production readiness and technical debt reduction," in Proc. IEEE Int. Conf. Big Data, pp. 1123-1132, 2017, DOI: https://doi.org/10.1109/BigData.2017.8258038
10. E. Gundersen et al., "Do machine learning platforms provide out-of-the-box reproducibility?" Future Generation Computer Systems, vol. 126, pp. 34-47, 2022, DOI: https://doi.org/10.1016/j.future.2021.07.017
11. F. Hutter et al., Automated Machine Learning: Methods, Systems, Challenges. Cham, Switzerland: Springer, 2019, DOI: https://doi.org/10.1007/978-3-030-05318-5
12. F. Bayram and B. S. Ahmed, "Towards Trustworthy Machine Learning in Production: An Overview of the Robustness in MLOps Approach," ACM Computing Surveys, vol. 57, no. 5, 2025, DOI: https://doi.org/10.1145/3708497
13. M. Zarour, H. Alzabut, and K. T. Al-Sarayreh, "MLOps best practices, challenges and maturity models: A systematic literature review," Information and Software Technology, vol. 183, art. no. 107733, 2025, DOI: https://doi.org/10.1016/j.infsof.2024.107733
14. M. Steidl, "The pipeline for the continuous development of AI models: Practice and research challenges," Journal of Systems and Software, vol. 199, art. no. 111615, 2023, DOI: https://doi.org/10.1016/j.jss.2023.111615
15. S. Schlegel and K.-U. Sattler, "Management of machine learning lifecycle artefacts: A survey," ACM SIGMOD Record, vol. 51, no. 4, pp. 28-39, 2022, DOI: https://doi.org/10.1145/3579051.3579056
16. M. Xhepa and N. Sree, "Machine Learning Model Computation in AWS and Azure," Frankfurt University of Applied Sciences, Frankfurt, Germany, 2022. [Online]. Available: https://sites.google.com/view/ai-as-a-service/comparison
17. W. Choi, T. Choi, and S. Heo, "A comparative study of automated machine learning platforms for exercise anthropometry-based typology analysis," Bioengineering, vol. 10, no. 8, Art. no. 891, 2023, DOI: https://doi.org/10.3390/bioengineering10080891
18. US Health Insurance Dataset, 2022. [Online]. Available: https://www.kaggle.com/datasets/teertha/ushealthinsurancedataset
19. Bank Marketing Dataset, 2022. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Bank+Marketing
20. AWS Autopilot Model Insight, 2023. [Online]. Available: https://docs.aws.amazon.com/sagemaker/latest/dg/autopilot-model-insights.html

AUTHOR'S PROFILE

Duckki Lee is currently an Associate Professor in the Department of Smart Software, Yonam Institute of Technology. His research interests include automated machine learning, mobile computing, smart systems, and Internet of Things.