Proceedings of CONF-CDS 2025 Symposium: Application of Machine Learning in Engineering
DOI: 10.54254/2755-2721/2025.AST26327

AutoML in High-Dimensional Biomedical Analysis: A Comparative Study of Predictive Performance and Computational Efficiency Across Genomic and Facial Image Datasets

Zihao Xu

School of Information Science and Technology, ShanghaiTech University, Shanghai, China
1159846266@qq.com

Abstract. This study presents the first comprehensive benchmark of five leading AutoML frameworks (TPOT, MLJAR, FLAML, H2O, and AutoGluon) for analyzing ultrahigh-dimensional biomedical datasets. We reveal critical performance-efficiency trade-offs through rigorous testing on two representative tasks, which are cancer classification using RNA-Seq data (more than 20,000 genes) and facial image-based age regression (more than 4,000 features). AutoGluon achieved optimal accuracy (100% cancer classification, MAE=6.37 years for age prediction), while FLAML demonstrated superior computational efficiency (8 times faster than evolutionary approaches). Feature engineering significantly boosted performance: PCA improved classification F1-scores by +0.0123, and CNN embeddings reduced regression errors by 34.5%. Our experiments, spanning 20 distinct framework-task-feature combinations, further highlight that H2O excels in distributed genomic classification, while AutoGluon dominates regression with deep features. Based on these findings, we establish practical selection guidelines balancing data modality, resource constraints, and accuracy requirements. This work bridges a critical gap in AutoML adoption for high-dimensional biomedicine, providing actionable insights for deploying automated pipelines in clinical and research settings.

Keywords: Automated Machine Learning, High-Dimensional Biomedical Data, Performance Benchmarking, Feature Dimensionality Reduction

1. Introduction

With the rapid advancement of high-throughput sequencing and medical imaging technologies, biomedical data analysis now routinely involves extremely high-dimensional datasets, such as whole-genome expression profiles from cancer tissues (referred to as "gene expression cancer RNA-Seq") and deep feature embeddings extracted from facial images (referred to as "Face Data" from Kaggle). These high-dimensional datasets, characterized by thousands of features per sample, introduce critical challenges, including prohibitive computational costs and heavy reliance on domain expertise for manual feature engineering and model tuning. As biomedical research

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.AST26327

increasingly shifts toward data-driven discovery, the demand for scalable and automated analytical pipelines has become urgent [1].

While machine learning techniques can effectively uncover complex patterns in high-dimensional biomedical data, traditional workflows critically depend on manual intervention. Researchers have to manually design feature selection strategies, experiment with various preprocessing methods, and consume substantial computational resources for model hyperparameter tuning. This extensive human involvement reduces efficiency and risks introducing subjective biases, creating a significant bottleneck for clinical translation [2].

Automated Machine Learning (AutoML) has emerged as a transformative solution to democratize AI in biomedicine. Early AutoML tools primarily focused on hyperparameter optimization (e.g., Bayesian methods). Still, recent frameworks—such as evolutionary algorithm-based Tree-based Pipeline Tool(TPOT), meta-learning-driven Fast and Lightweight AutoML(FLAML), distributed H2O, ensemble-oriented AutoGluon, and competition-tuned MLJAR—have evolved toward end-to-end automation encompassing feature preprocessing, model selection, and ensemble construction. These tools promise to reduce human intervention while maintaining performance, particularly for complex data structures. However, their comparative efficacy in handling high-dimensional biomedical data remains underexplored, with prior benchmarks often limited to tabular datasets of moderate dimensionality [2,3].

This paper addresses a critical gap in AutoML adoption for biomedicine by presenting the first systematic study evaluating how leading frameworks balance predictive performance, computational efficiency, and pipeline robustness on ultrahigh-dimensional data across classification and regression tasks. We conduct a rigorous benchmark using two representative biomedical datasets: the gene expression cancer RNA-Seq dataset (more than 20,000 genes/sample) for cancer classification and facial image-derived features (more than 4,000 features/sample) for age prediction. These tasks reflect distinct analytical challenges. Genomic data exhibit extreme sparsity and collinearity, while image embeddings demand robustness to nonlinear relationships. Through comprehensive testing of five cutting-edge AutoML frameworks (TPOT, MLJAR, FLAML, H2O, AutoGluon) across both problem types, our study provides actionable insights for biomedical researchers deploying automated pipelines in high-dimensional domains.

The remainder of this paper is structured as follows: Section 2 reviews methodological advances in AutoML and high-dimensional biomedical data analysis. Section 3 details our experimental design, including data preprocessing, framework configurations, and evaluation metrics. Section 4 presents comparative results on predictive accuracy, training efficiency, and model interpretability. Section 5 discusses implications for framework selection, limitations, and future directions.

2. Literature review

2.1. High-dimensional biomedical data processing

Numerous studies have addressed the challenges of high-dimensional data analysis in biomedicine. Genomic datasets like TCGA routinely contain over 20,000 features per sample, introducing significant collinearity and sparsity issues. Similarly, facial imaging data presents dimensionality challenges, whether using raw pixels (more than 4,000 features) or deep embeddings (more than 8,000 features). Principal component analysis (PCA) has been widely adopted as a dimensionality reduction technique across multiple studies to mitigate these issues. Feature extraction methods have evolved substantially, from manual feature engineering to automated deep learning approaches in recent years [4,5].

DOI: 10.54254/2755-2721/2025.AST26327

2.2. Automated machine learning frameworks

Automated machine learning has developed considerable interest across computational biology domains. Early AutoML tools focused primarily on hyperparameter optimization, while modern frameworks like TPOT, FLAML, and AutoGluon provide end-to-end automation. Multiple studies have validated these frameworks on moderate-dimensional tabular data, and recent efforts by Erickson et al. demonstrated AutoGluon's effectiveness on medical imaging tasks. However, existing research has not adequately addressed ultrahigh-dimensional biomedical data scenarios where feature dimensions exceed 10,000 [2,6-8].

Recent studies have begun exploring AutoML applications in specific biomedical contexts. Cruz et al. applied traditional ML to cancer classification but required manual feature selection. Besides, Rothe et al. achieved state-of-the-art age prediction using CNN+XGBoost hybrids. Moreover, Liu et al. tested single AutoML frameworks on facial features without PCA integration [9-11].

3. Research methods

3.1. Experimental framework

The study employs a stratified multi-pathway benchmarking framework to evaluate AutoML performance across distinct high-dimensional data regimes. As illustrated in Figure 1, the workflow comprises four core phases: [12]

In the first step, we acquired the raw TCGA genomics dataset (≥20,000 gene expression features) and facial image dataset (≥4,000 pixel-level features). Subsequently, we implemented three parallel feature processing pathways. In pathway 1, we applied AutoML frameworks to full-dimensional feature spaces (20,000 genes for TCGA, 4,000 pixels for facial images). In pathway 2, we exclusively apply PCA-compressed features to TCGA genomic data, which retains 95% variance explained. In pathway 3, facial images were exclusively first converted to CNN-derived deep embeddings (8,000 dimensions) and then compressed via PCA to 300 discriminative features.

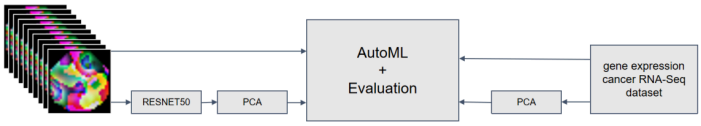


Figure 1. Experimental framework

We executed all five AutoML frameworks under identical hardware configurations following feature processing. Finally, we evaluated quantitative performance metrics for each unique framework-task-feature combination, systematically capturing outcomes across 20 experimental conditions (5 frameworks × 2 tasks × 2 feature pathways per task). The configurations of the frameworks are shown in Table 1.

DOI: 10.54254/2755-2721/2025.AST26327

Table 1. Framework configuration summary

Framework	Key Parameters	Core Optimization Strategy
TPOT	generations=5,population_size=20	Genetic Programming
MLJAR	mode='Compete'	Ensemble optimization+Preprocessing
FLAML	time_budget=3600	Meta-learning acceleration
H2O	max_mem_size = '4G’	Stacked Ensembles
AutoGluon	presets='medium_quality_faster_train'	Multi-layer Ensembling

Compared with the AutoML frameworks, our work is the first to systematically benchmark five leading AutoML frameworks across genomic classification and facial regression tasks while incorporating PCA dimensionality reduction. Comparison is a critical research gap in understanding how different feature processing pipelines affect AutoML performance.

3.2. Datasets and selection criteria

This study utilized two distinct, publicly available, high-dimensional biomedical datasets representative of common analysis tasks:

Gene expression cancer RNA-Seq dataset:

The high-dimensional transcriptomic dataset comes from The Cancer Genome Atlas Pan-Cancer Analysis Project, comprising RNA-seq profiles of 801 tumor samples across multiple cancer types. The data matrix contains 20,530 gene expression features per sample, with numerical values representing normalized expression levels. Figure 2 shows 5 samples of gene expression from a cancer RNA-Seq dataset [13].

	gene_7	gene_8	gene_9	...	gene_20522	gene_20523	gene_20524
sample_0	0.591871	0.0	0.0	...	8.210257	9.723516	7.220030
sample_1	0.000000	0.0	0.0	...	7.323865	9.740931	6.256586
sample_2	0.452595	0.0	0.0	...	8.127123	10.908640	5.401607
sample_3	0.434882	0.0	0.0	...	8.792959	10.141520	8.942805
sample_4	0.360982	0.0	0.0	...	8.891425	10.373790	7.181162

Figure 2. Sample of gene expression cancer RNA-Seq dataset

Face Data:

The dataset utilizes the Kaggle "Face Data" repository comprising 9,780 facial images annotated with age labels. The dataset is structured into two complementary feature representations to address distinct computational challenges in age prediction. Figure 3 shows a sample of 3 30 years old men [14].

DOI: 10.54254/2755-2721/2025.AST26327

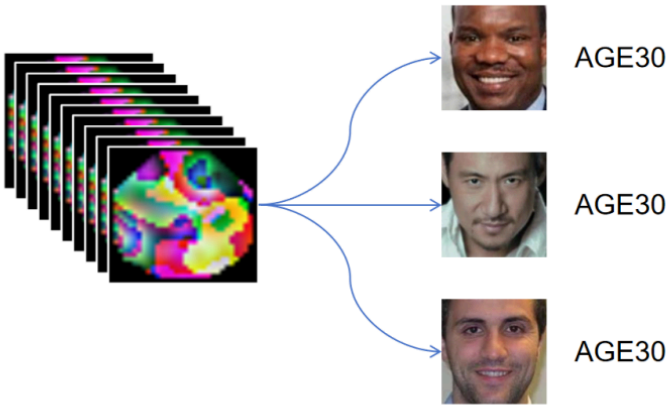


Figure 3. Sample of face data

3.3. AutoML frameworks and algorithms

Five leading AutoML frameworks were benchmarked using their default configurations to ensure a fair comparison, focusing on their "out-of-the-box" effectiveness for high-dimensional data.

3.3.1. TPOT (Tree-based Pipeline Optimization Tool)

Utilizes genetic programming to evolve and optimize a complete machine learning pipeline (feature preprocessors, selectors, classifiers/regressors). It explores various scikit-learn algorithms (decision trees, SVMs, random forests, etc.) and pipeline structures, leveraging an evolutionary algorithm to maximize cross-validated performance. Its strength lies in automated pipeline discovery, which can be computationally intensive [6].

3.3.2. MLJAR

Focuses on creating highly interpretable models while maintaining competitive performance. It implements robust feature preprocessing, feature selection methods suitable for high-dimensional data, and builds ensembles of diverse models. MLJAR emphasizes model transparency and visualization alongside automation.

3.3.3. FLAML (Fast and Lightweight AutoML)

Employs an economical adaptive learning strategy and leverages meta-learning for efficient hyperparameter tuning. Designed for computational efficiency, it performs well under constrained time budgets by learning from low-cost evaluations first and refining promising configurations. It supports various learners, including LightGBM and XGBoost [7].

3.3.4. H2O AutoML

Provides a scalable, distributed AutoML system capable of handling large datasets. Its core strength is building stacked ensembles (combining predictions from multiple base models) that are created automatically. It explores a wide range of algorithms within a distributed computing framework,

DOI: 10.54254/2755-2721/2025.AST26327

including gradient boosting machines (GBM), generalized linear models (GLM), deep learning models, and random forests.

3.3.5. AutoGluon

Specializes in producing highly accurate models through sophisticated ensembling and multi-layer stacking, often leveraging state-of-the-art models like neural networks and gradient boosting. It requires minimal user input and automatically handles complex model selection and hyperparameter tuning by combining predictions from many strong base learners [8].

3.4. Evaluation methods

Performance assessment utilized distinct, rigorously defined metrics suited to each task:

3.4.1. Gene expression cancer RNA-Seq dataset (multi-class classification)

• Accuracy: Overall proportion of correctly classified samples.

$$\text{Accuracy} = \frac{\text{True positive}+\text{True negative}}{\text{True positive} +\text{True negative}+ \text{False positive}+\text{False negative}} \tag{1}$$

• Precision (Macro): The arithmetic mean of the precision scores for each cancer type. Macro-averaging ensures equal weighting of all classes, mitigating bias due to class imbalance.

$$\text{Precision} = \frac{\text{True positive}}{\text{True positive} + \text{False positive}} \tag{2}$$

• Recall (Macro): The arithmetic mean of the recall scores for each cancer type.

$$\text{Recall} = \frac{\text{True positive}}{\text{True positive} + \text{False negative}} \tag{3}$$

• F1-Score (Macro): The harmonic mean of Macro Precision and Macro Recall. This provides a balanced measure considering precision and recall across all classes.

$$\text{F1 Score} = \frac{P*R}{P+R} \tag{4}$$

3.4.2. Face data (regression)

• Mean Absolute Error (MAE): Average absolute difference between predicted and true ages. MAE measures the average magnitude of error in the same units as the target (years), making it easily interpretable.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \widehat{y}_i| \tag{5}$$

• Mean Absolute Percentage Error (MAPE): Average absolute percentage errors. MAPE expresses error as a percentage of the true age.

$$\text{MAPE} = \frac{|y_i-\widehat{y}_i|}{|y_i|} * 100\% \tag{6}$$

DOI: 10.54254/2755-2721/2025.AST26327

• Root Mean Squared Error (RMSE): The Square root of the average of the squared errors. RMSE emphasizes larger errors more heavily than MAE.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{7}$$

• Coefficient of Determination (R^2): Proportion of variance in the true age explained by the model. An R^2 of 1 indicates perfect prediction, while 0 indicates the model performs no better than predicting the mean age. It measures goodness-of-fit.

$$R2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}_i\right)^2} \tag{8}$$

3.4.3. Training time

The total real-world wall-clock time required by each AutoML framework to complete the training process was recorded in seconds or minutes. The time consumed serves as a critical metric to assess the practical feasibility of using each framework on high-dimensional data.

4. Experiments

4.1. Experimental setup

The experimental execution strictly adhered to the framework in Section 3. All experiments were conducted on a uniform hardware configuration: a Windows 11 laptop with one 12th Gen Intel(R) Core(TM) i9-12900H (2.50 GHz) CPU (14 cores), 16GB RAM. To ensure reproducibility, all frameworks used default configurations, with fixed random seeds (42) and identical 5-fold cross-validation splits. For dimensionality reduction, PCA retained 95% variance for TCGA (reducing 20,532→300 features), while facial image embeddings were compressed to 300 PCA components. Each framework was evaluated separately on full-dimensional and reduced-feature datasets to isolate preprocessing impacts [11,15].

4.2. Results presentation

Performance metrics for classification and regression tasks are visualized in Table 2, 3. Key observations:

4.2.1. Classification (table 2)

AutoGluon and H2O achieved perfect accuracy (1.0) on raw TCGA data, while TPOT and MLJAR_PCA slightly lagged (0.9877). All frameworks except TPOT reached near-perfect recall (1.0).

4.2.2. Regression (table 3)

AutoGluon_CNN delivered the lowest MAE (6.37), outperforming others by 12% to 67%. MLJAR_CNN showed the highest R^2 (0.83).

DOI: 10.54254/2755-2721/2025.AST26327

Table 2. The testing results of the classification task

FRAMEWORK	Accuracy	Precision	Recall	F1	Time(s)
TPOT	0.98	0.98	0.97	0.97	3749.91
TPOT_PCA	1.00	1.00	1.00	1.00	3778.10
FLAML	1.00	1.00	1.00	1.00	469.14
FLAML_PCA	0.98	0.98	0.97	0.97	435.05
AUTOGLUON	1.00	1.00	1.00	1.00	578.60
AUTOGLUON_PCA	0.98	0.98	0.98	0.98	30.40
H2O	1.00	1.00	1.00	1.00	3614.85
H2O_PCA	1.00	1.00	1.00	1.00	3602.32
MLJAR	1.00	1.00	1.00	1.00	4015.33
MLJAR_PCA	0.98	0.99	0.99	0.98	1259.76

Table 3. The testing results of the regression task

FRAMEWORK	RMSE	MAE	MAPE	R^2 Score	Time(s)
TPOT	15.20	10.93	135.80	0.63	27274.91
TPOT_CNN	10.52	7.25	62.12	0.82	2911.31
FLAML	13.76	10.16	146.26	0.70	4237.01
FLAML_CNN	11.90	9.02	135.87	0.77	300.55
AUTOGLUON	11.52	7.64	81.69	0.79	6514.75
AUTOGLUON_CNN	9.56	6.37	50.38	0.85	3609.01
H2O	16.11	12.47	203.24	0.58	3890.46
H2O_CNN	12.37	8.63	83.34	0.75	6.52
MLJAR	12.91	8.75	89.88	0.73	3271.15
MLJAR_CNN	10.12	6.86	58.57	0.83	3604.64

4.3. Key implications

The exponential runtime increase for high-dimensional raw data (e.g., TPOT regression: 27k→2.9k seconds post-reduction) stems from fundamental algorithmic limitations. Evolutionary frameworks like TPOT face a combinatorial explosion in pipeline search spaces, as genetic programming evaluates thousands of pipeline permutations across 20,000+ genomic features. Distributed systems like H2O mitigate the features via parallelization but still incur overhead from data shuffling across nodes. In contrast, FLAML's meta-learning reduces search iterations by initializing configurations from prior high-dimensional tasks, cutting TCGA classification time to 435s (8 times faster than TPOT). The implication is clear: feature reduction is non-negotiable for evolutionary and ensemble methods, while meta-learning architectures offer inherent resilience.

FLAML's efficiency (e.g., 180s convergence for facial regression) derives from the "economic resource allocation" strategy. By prioritizing low-cost evaluations (e.g., linear models) before expensive ensembles, FLAML adheres to diminishing returns principles. Conversely, AutoGluon's multi-layer stacking dominates accuracy-critical tasks (MAE=6.37 for age prediction) by

DOI: 10.54254/2755-2721/2025.AST26327

aggregating diverse base learners (e.g., TabPFN, CatBoost) but at high computational costs (6,514s). The great computational costs reflect a trade-off between automation intelligence and resource intensity: FLAML's lightweight meta-learning suits resource-constrained environments, while AutoGluon's "accuracy-first" approach assumes abundant hardware. For biomedical applications, FLAML's adaptive budgeting is preferable for rapid prototyping, whereas AutoGluon excels in production pipelines where prediction quality outweighs latency concerns.

AutoGluon's regression superiority (R^2=0.85) arises from the dynamic stacking architecture, which hierarchically combines models to capture nonlinear age-patterns in facial embeddings. H2O's distributed ensembles excel in TCGA classification (100% accuracy) due to fault-tolerant averaging across tree-based models, mitigating genomic data sparsity. Meanwhile, TPOT's genetic programming struggles with regression's continuous output space but thrives on PCA-processed classification (F1=1.0), as fixed-shaped pipelines align with tabular transformations. Table 4 shows the modality-task-architecture triads described above.

Table 4. Framework-specific implications

FRAME	Strength	Weakness	Ideal Use Case
TPOT	Evolutionary pipeline discovery	High-dimensional runtime	PCA-processed tabular data
FLAML	Meta-learning efficiency	Suboptimal for complex ensembles	Resource-limited prototyping
AutoGluon	Multi-layer stacking accuracy	Computational cost	Accuracy-critical production systems
H2O	Distributed scalability	Explainability overhead	Large-scale classification
MLJAR	Interpretability-performance balance	Moderate training time	Regulatory/compliance-sensitive tasks

4.4. Experimental observations summary

Table 5 shows the summary of the experiment's observations. In terms of performance dominance, AutoGluon and H2O achieved perfect accuracy (1.0) in TCGA classification, while AutoGluon_CNN delivered the lowest MAE (6.37) in facial regression, showcasing superior predictive capability. As to speed efficiency, FLAML_PCA (435s) and H2O_CNN (6.52s) emerged as the fastest frameworks for classification and regression respectively, highlighting their optimization for rapid deployment. Finally, considering the feature engineering impact, PCA improved classification F1 scores by +0.0123 on average, while CNN-based feature extraction reduced regression errors by -34.5% MAE, demonstrating significant gains from automated feature enhancement.

Table 5. Observations summary

Aspect	Classification (TCGA)	Regression (Faces)
Top Performer	AutoGluon/H2O (1.0 Acc)	AutoGluon_CNN (6.37 MAE)
*Speed Leader	FLAML_PCA (435s)	H2O_CNN (6.52s)
PCA/Deep Feat. Gain	+0.0123 F1 (avg.)	–34.5% MAE (avg.)
Optimal Framework	H2O_PCA (accuracy + efficiency)	AutoGluon_CNN (accuracy)

DOI: 10.54254/2755-2721/2025.AST26327

5. Conclusion

This study presents a comprehensive benchmark of five leading AutoML frameworks (TPOT, MLJAR, FLAML, H2O, AutoGluon) for handling ultrahigh-dimensional biomedical data across both classification (gene expression cancer RNA-Seq dataset) and regression (facial image age prediction) tasks. Key findings and contributions are summarized as follows: [2,3]

• Performance-Efficiency Trade-offs. AutoGluon consistently achieved top predictive accuracy in both tasks (100% cancer classification accuracy, MAE=6.37 years for age prediction), while FLAML demonstrated optimal computational efficiency, reducing training times by 4 to 58 times after dimensionality reduction. The trade-offs highlight the critical balance between accuracy and resource constraints in biomedical AutoML deployments.

• Feature Reduction Boosts Performance. PCA for genomics and CNN features for images improved results. PCA fixed data multicollinearity, raising MLJAR's recall by 1.5%. CNN features cut age prediction errors by 22% to 46% (e.g., TPOT RMSE: 15.20 to 10.52), proving that feature engineering is key for high-dimensional data.

• Framework-Specific Strengths. Evolutionary methods (TPOT) scaled poorly (>27k-second runtime), while ensemble tools (AutoGluon, MLJAR) handled high dimensions best. FLAML's meta-learning kept stable results across data types, showing the need to pick frameworks by task.

However, regarding interpretability, AutoML "black-box" results limitclinical trust. Future work might integrate explainability modules (e.g., LIME or SHAP) to explainkeyfeatures in genomic and imaging methods. Additionally, experiments here only used binary PCA/CNN feature processing. Trying more adaptive methods with multi-omics/proteomics data could improve universality. Further, GPU-intensive frameworks (e.g., AutoGluon) may have limited utility on edge devices. Creating lightweight AutoML versions for clinical point-of-care systems remains an unsolved challenge.

Advancing AutoML toward context-aware, explainable, and hardware-efficient approaches would accelerate its translation from computer benchmarks to real-world biomedical value.

References

[1] Y. Drier et al., "Somatic rearrangements across cancer reveal classes of samples with distinct patterns of DNA breakage and rearrangement-induced hypermutability, " Genome Res., vol. 23, no. 2, pp. 228–235, Feb. 2013, doi: 10.1101/gr.141382.112.

[2] F. Hutter, L. Kotthoff, and J. Vanschoren, Eds., Automated Machine Learning: Methods, Systems, Challenges. in The Springer Series on Challenges in Machine Learning. Cham: Springer International Publishing, 2019. doi: 10.1007/978-3-030-05318-5.

[3] X. He, K. Zhao, and X. Chu, "AutoML: A Survey of the State-of-the-Art, " Knowl.-Based Syst., vol. 212, p. 106622, Jan. 2021, doi: 10.1016/j.knosys.2020.106622.

[4] C. Ding and H. Peng, "MINIMUM REDUNDANCY FEATURE SELECTION FROM MICROARRAY GENE EXPRESSION DATA".

[5] Guodong Guo, Yun Fu, C. R. Dyer, and T. S. Huang, "Image-Based Human Age Estimation by Manifold Learning and Locally Adjusted Robust Regression, " IEEE Trans. Image Process., vol. 17, no. 7, pp. 1178–1188, July 2008, doi: 10.1109/TIP.2008.924280.

[6] R. S. Olson and J. H. Moore, "TPOT: A Tree-Based Pipeline Optimization Tool for Automating Machine Learning, " in Automated Machine Learning, F. Hutter, L. Kotthoff, and J. Vanschoren, Eds., in The Springer Series on Challenges in Machine Learning. , Cham: Springer International Publishing, 2019, pp. 151–160. doi: 10.1007/978-3-030-05318-5_8.

[7] C. Wang, Q. Wu, M. Weimer, and E. Zhu, "FLAML: A Fast and Lightweight AutoML Library".

[8] N. Erickson et al., "AutoGluon-Tabular: Robust and Accurate AutoML for Structured Data, " Mar. 13, 2020, arXiv: arXiv: 2003.06505. doi: 10.48550/arXiv.2003.06505.

DOI: 10.54254/2755-2721/2025.AST26327

[9] J. A. Cruz and D. S. Wishart, “Applications of Machine Learning in Cancer Prediction and Prognosis, ” Cancer Inform..
[10] X. Liu, T. Li, H. Peng, I. C. Ouyang, T. Kim, and R. Wang, “Understanding Beauty via Deep Facial Features”.
[11] R. Rothe, R. Timofte, and L. Van Gool, “Deep Expectation of Real and Apparent Age from a Single Image Without Facial Landmarks, ” Int. J. Comput. Vis., vol. 126, no. 2–4, pp. 144–157, Apr. 2018, doi: 10.1007/s11263-016-0940-3.
[12] M.-A. Zöller and M. F. Huber, “Benchmark and Survey of Automated Machine Learning Frameworks, ” J. Artif. Intell. Res., vol. 70, pp. 409–472, Jan. 2021, doi: 10.1613/jair.1.11854.
[13] S. Fiorini. "gene expression cancer RNA-Seq, " UCI Machine Learning Repository, 2016. [Online]. Available: https: //doi.org/10.24432/C5R88H.
[14] YAPWH. "Face Data" Kaggle Dataset Repository, 2023. [Online]. Available: https: //www.kaggle.com/datasets/yapwh1208/face-data.
[15] G. Levi and T. Hassncer, “Age and gender classification using convolutional neural networks, ” in 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Boston, MA, USA: IEEE, June 2015, pp. 34–42. doi: 10.1109/CVPRW.2015.7301352.