Journal of Computing and Electronic Information Management
ISSN: 2413-1660 | Vol. 23, No. 2, 2026

Privacy-Preserving Personalized Learning Through Adaptive Federated Model Aggregation

Niamh Farrell [1, *], Cian Murphy [2], Aoife Brennan [1]

[1] School of Engineering, Trinity College Dublin, Dublin, Ireland
[2] School of Electrical and Electronic Engineering, University College Dublin, Dublin, Ireland

*** Corresponding author**

Abstract: Federated learning keeps source records at participating clients, but heterogeneous data can make a single shared model unsuitable for every user. We study a computationally compact adaptive personalization rule that blends a federated model with a locally fine-tuned model using a coefficient selected by client-held validation loss. All candidate methods reuse the same 30-round FedAvg checkpoint and local fine-tuning trajectory, isolating model selection from additional federated optimization. Using 1,797 handwritten-digit examples divided among eight simulated clients, we evaluate two label-heterogeneity settings, five independently seeded runs, and five evaluation rules including endpoint selection. At Dirichlet concentration 0.3, AdaptiveMix achieves 96.05% mean client-macro test accuracy, compared with 91.75% for FedAvg and 95.74% for endpoint selection. At concentration 1.0, AdaptiveMix achieves 94.27% versus 94.57% for fixed mixing. Paired seed differences and client-level outcomes characterize when interpolation helps. A pairwise masking identity test checks aggregate arithmetic, while data-locality and cryptographic privacy guarantees are explicitly distinguished. The results support a low-communication baseline for exploring heterogeneous personalized federated classifiers.

Keywords: Federated learning; Personalization; Model interpolation; Heterogeneous clients; Privacy-aware data processing.

1. Introduction

Federated learning (FL) updates a shared model from separate local datasets rather than moving all records to a central training repository [1]–[4]. This arrangement is attractive for systems holding learner interactions, mobile observations, or institution-specific content. However, statistical heterogeneity changes what individual clients need from the model. Unequal label frequencies, sample sizes, and local preferences can make the global objective a poor surrogate for every participant's test performance [5]–[11].

Personalized FL addresses this mismatch by retaining local representations, adapting shared parameters, or learning client-specific predictors [12]–[26]. Many effective schemes require changes to the global objective, additional optimization states, or exchanges of multiple models. An economical alternative is to complete a common FL training run and let clients blend its output with their own fine-tuned checkpoint. The critical design choice then becomes the mixing coefficient: uniform blending may waste local information, while full localization may overfit clients whose training data remain scarce.

We examine AdaptiveMix, a validation-guided local interpolation rule based on an unchanged FedAvg backbone. The evaluation is designed to separate advantages arising from local fine-tuning from those arising from adaptive coefficient selection. In particular, we add an endpoint-selection ablation, which uses the same validation data but can select only the global or local endpoint. With five paired random-seed runs, we report mean and variability, client-level wins and losses, a weaker-client metric, and idealized communication costs. The privacy analysis distinguishes local record retention from guarantees that would require secure aggregation or differential privacy.

2. Related work

2.1. Global optimization under heterogeneous clients

FedAvg weights client updates by the amount of locally available training data [1]. FedProx incorporates proximal regularization to address heterogeneous optimization [5], SCAFFOLD applies control variates to client drift [6], FedNova handles uneven local step counts [7], and FedDyn adjusts client objectives dynamically [8]. Adaptive federated optimizers, feature-specific normalization, and model contrastive objectives provide other routes to stable collaborative learning [9]–[11]. These methods focus primarily on global training behavior rather than selecting a separate prediction model for each client.

2.2. Personalization and privacy

Personalization strategies include federated multi-task learning [12], clustering and interpolation [13], independent model heads [14], meta-learning [15], and Moreau-envelope regularization [16]. Model mixing is also addressed by global–local mixture models and adaptive personalization [17], [18]. Ditto [19], FedRep [20], FedAMP [21], FedFomo [22], hypernetwork models [23], FedBABU [24], prototype learning [25], and FedALA [26] develop different ways of choosing shared versus local information. Our study does not replace these richer methods; it tests a small, post-training coefficient search that is particularly easy to reproduce and analyze.

Retaining raw records locally limits direct data transfer but does not ensure model-update confidentiality. Secure aggregation can hide individual updates from a coordinator under its stated protocol assumptions [27]; differential privacy provides quantified guarantees when the mechanism and budget are specified [28]–[32]. Gradient inversion, white-

box inference, and feature leakage explain why unprotected model updates can remain sensitive [33]–[36]. Privacy, fairness, and convergence also interact in personalized systems [45]. Our numerical masking check is an aggregate-consistency test rather than a deployed cryptographic protocol.

3. Method

3.1. Federated backbone

Let K denote the number of clients. Each client k has disjoint training, validation, and test sets and n_k training examples. We use a network with parameter vector θ and client training risk $F_k(\theta)$. Let $N=\Sigma_k n_k$. For full participation, the sample-weighted objective and a generic round update are

$$F(\theta) = \Sigma_k (n_k/N) F_k(\theta). \quad (1)$$

$$\theta^{t+1} = \theta^t + \Sigma_k (n_k/N) (\theta^t_k - \theta^t). \quad (2)$$

Each client starts at θ^t, performs two local epochs of minibatch SGD, and contributes its weight-scaled update to (2). The experiment performs T=30 rounds with all eight clients in each round. The local record tensors are not inputs to the aggregation function. A single final global model θ^T is used for every evaluation method, so differences between methods do not arise from separate global training histories.

3.2. Validation-guided model interpolation

Starting from θ^T, client k performs three additional epochs of SGD on its training subset to obtain φ_k. It constructs $\theta_k(a)=(1-a)\theta^T+a\varphi_k$ and selects the coefficient by minimizing local validation cross-entropy over $A=\{0,0.25,0.50,0.75,1\}$. Coefficient selection occurs without uploading validation losses or labels.

$$a^*_k = \arg\min_{a \in A} L_{k,val}[(1-a)\theta^T+a\varphi_k]. \quad (3)$$

$$\theta^*_k = (1-a^*_k)\theta^T + a^*_k\varphi_k. \quad (4)$$

The grid contains both endpoints and three interior choices. We distinguish five evaluation rules: FedAvg uses a=0, local fine-tuning uses a=1, fixed mixing uses a=0.5, endpoint selection chooses the better of a=0 and a=1 on validation loss, and AdaptiveMix uses (3). The last two rules are validation-guided; neither has access to the test labels during its selection. Ties in validation loss are resolved by the first coefficient in the ordered grid.

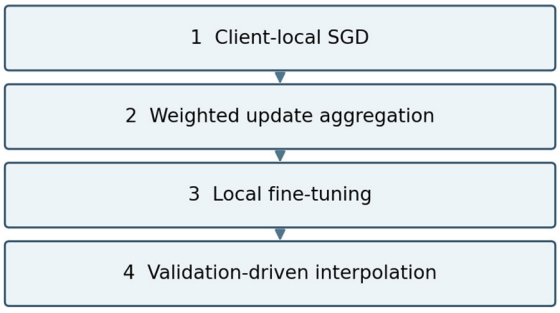


Fig. 1 Workflow of the common federated backbone and on-client personalization. Mask arithmetic is simulated in the update-aggregation step.

3.3. Aggregation masking and scope

The experimental code optionally adds pairwise, opposite-sign pseudorandom masks to weighted updates. For client k, the symbolic contribution is $z_k=p_k\Delta_k+\Sigma_j m_{kj}$, where $m_{kj}=-m_{jk}$. Summation cancels these masks, returning $\Sigma_k z_k=\Sigma_k p_k\Delta_k$. This identity is checked numerically after creating representative model-sized contributions. It tests the arithmetic consistency of a masked sum and does not provide cryptographic secrecy: all clients are represented in one program, the masks are generated by a simulation routine, and dropout, key negotiation, malicious participants, and adversarial reconstruction are outside the measured implementation.

The deployed-system interpretation is accordingly limited to client-local data handling and a validation stage that never communicates additional local statistics. A genuine secure aggregation deployment would replace the arithmetic simulator with a reviewed protocol [27]. Neither gradient perturbation nor a privacy accountant is implemented, so no ε,δ differential-privacy guarantee is attached to the reported accuracy values. Model updates can leak information even when the original examples remain on-device [33]–[36].

4. Experimental design

4.1. Dataset and heterogeneous partitions

We use scikit-learn's 1,797-example handwritten-digit dataset, which contains 10 classes and 8×8 grayscale images represented by 64 numeric features [41]–[44]. The images originate from the test subset of the historical UCI optical-digits dataset; this experiment repartitions the scikit-learn copy for simulated federated evaluation, rather than reproducing the original UCI benchmark split. Each input is divided by 16 to scale pixel features to [0,1]. No external dataset download is needed.

For each seed, observations of each digit class are assigned across K=8 clients according to a symmetric Dirichlet draw. Concentrations 0.3 and 1.0 produce respectively more uneven and less uneven local label distributions. We redraw the partition until every client contains at least 55 total examples, then independently shuffle and split each client's records into 60% training, 20% validation, and the remaining 20% test. All methods within a given run use identical client membership and identical test records. Dirichlet client partitions are common controlled non-IID evaluations [37]–[40].

4.2. Optimization protocol and controls

Table I. Experiment configuration

Setting	Value
Data / clients	1,797 digit images / 8 clients
Heterogeneity	Dirichlet α = 0.3 and 1.0
Seeds	11, 22, 33, 44, 55
Local train/val/test	60% / 20% / 20%
Model	64–48–10 MLP; 3,610 parameters
FedAvg	30 rounds, 2 local epochs/round
Optimizer	SGD, lr 0.15, batch size 32
Personalization	3 local epochs; 5-point coefficient grid
Participation	8 / 8 clients every round

The network consists of a 64-to-48 linear layer, a ReLU activation, and a 48-to-10 linear classifier. Its 3,610 trainable parameters are fitted using SGD with learning rate 0.15, minibatch size 32, and no momentum. Training uses CPU with one PyTorch computation thread. Seeds 11, 22, 33, 44, and 55 control initialization, data allocation, and local order. All global rounds are executed before any local fine-tuning. The five evaluation methods reuse the resulting global model and local fine-tuned candidate rather than retraining independent global baselines.

4.3. Evaluation units and reported metrics

For a fixed seed, client-macro accuracy averages the eight held-out client accuracies equally, while micro accuracy weights each client by its number of held-out test records. The bottom-quarter client metric averages the two lowest client accuracies within each seed and method. Tables report averages and sample standard deviations over five seeded runs (n=5); clients within a run are not treated as five times eight independent experiment replications. Paired changes subtract two methods' macro accuracies inside the same seed. Their mean and standard deviation describe observed seed sensitivity, not a large-population confidence interval.

5. Results and analysis

5.1. Main five-method comparison

Table II. Client-macro test accuracy (mean ± sd over five seeds)

Dir. α	Method	Macro acc. (%)
0.3	FedAvg	91.75 ± 2.31
0.3	Fixed mix (0.5)	95.51 ± 0.81
0.3	Local FT	95.74 ± 1.04
0.3	Endpoint select	95.74 ± 1.04
0.3	AdaptiveMix	96.05 ± 1.09
1.0	FedAvg	94.21 ± 1.47
1.0	Fixed mix (0.5)	94.57 ± 1.65
1.0	Local FT	93.47 ± 2.02
1.0	Endpoint select	94.05 ± 1.74
1.0	AdaptiveMix	94.27 ± 1.61

Table II shows that the global-only model is sensitive to stronger label heterogeneity. At α=0.3, FedAvg obtains 91.75 ± 2.31% client-macro accuracy, compared with 95.74 ± 1.04% for full local fine-tuning and 96.05 ± 1.09% for AdaptiveMix. The AdaptiveMix–FedAvg mean paired increase is 4.31 percentage points; it improves on the endpoint-selection mean by 0.32 points. Both improvements are measured on the same five seeded partitions, and the latter difference is modest compared with the effect of local fine-tuning itself.

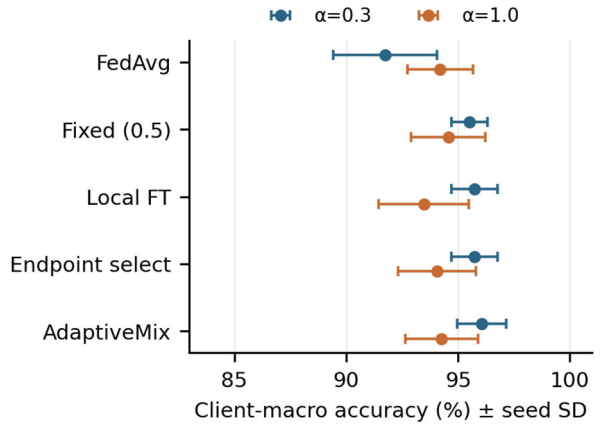


Fig. 2 Accuracy of all five selection rules. Horizontal error bars represent sample SD across five seeded runs.

At α=1.0, FedAvg obtains 94.21 ± 1.47% and AdaptiveMix 94.27 ± 1.61%, a paired difference of +0.06 points. The fixed-0.5 comparator reaches 94.57 ± 1.65%, which is 0.30 points higher in sample mean than AdaptiveMix. Thus, adaptation is not uniformly superior when client label distributions become less skewed. The controlled result is that coefficient selection gives a clear gain over the global-only model under stronger heterogeneity and remains comparable to the baselines under milder heterogeneity.

5.2. Ablation: endpoints versus interior weights

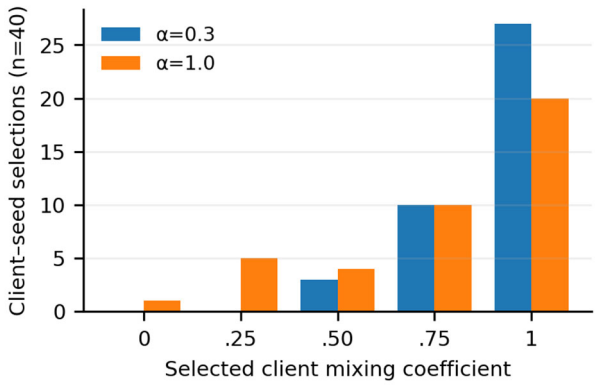


Fig. 3 Coefficients selected by local validation loss, counted over 40 client–seed combinations per heterogeneity level.

Figure 3 describes how adaptation selects local mixing coefficients. Under α=0.3, 27 of the 40 client–seed instances choose a=1, 10 choose a=0.75, and 3 choose a=0.5. Accordingly, endpoint selection and local fine-tuning have the same average result in this setting (95.74%), while the intermediate coefficients in AdaptiveMix produce a small additional mean improvement. At α=1.0, coefficients are more distributed: 1 choose a=0, 5 choose a=0.25, 4 choose a=0.5, 10 choose a=0.75, and 20 choose a=1. This provides an interpretable view of the local adaptation selected without using any test samples.

Endpoint selection is an informative ablation because it uses validation exactly as AdaptiveMix does while removing intermediate candidates. It isolates the value of the interpolation grid from the value of local fine-tuning. In our runs, AdaptiveMix improves on endpoint selection by an average of 0.32 points at α=0.3 and 0.22 points at α=1.0. The extra search requires three additional validation forward passes per client beyond endpoint selection and does not add communication rounds.

5.3. Seed-level paired comparisons and training dynamics

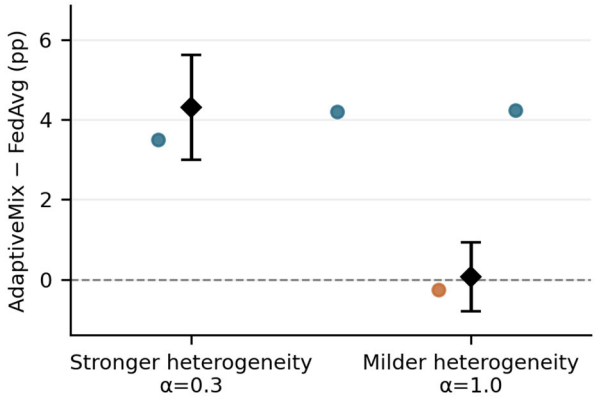


Fig. 4 AdaptiveMix minus FedAvg for each matched random seed; diamonds mark mean paired changes and bars denote seed SD.

The paired analysis in Fig. 4 uses the seeded run, not the individual client, as its observation unit. At α=0.3, the mean

paired gain is 4.31 ± 1.31 points, with individual seed differences ranging from 3.12 to 6.48; every seed favors AdaptiveMix. At α=1.0, the mean paired gain is 0.06 ± 0.86 points, with a range of -1.02 to 1.10 and both positive and negative outcomes. Therefore the stronger-heterogeneity finding is consistent across the measured random splits, while the milder setting has seed-dependent outcomes. A claim of statistical generality beyond these five synthetic partitions would require broader sampling and datasets.

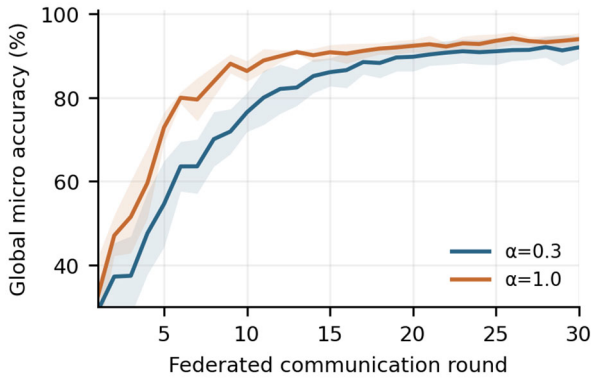


Fig. 5. Global-model micro test accuracy through 30 federated rounds; curves show five-seed means and shaded one-SD bands.

The global-model convergence curves in Fig. 5 share the same SGD budget across all five methods. Personalization is applied only after round 30, so the differences in Table II do not reflect selecting a different global stopping round. Test records are used to plot evaluation curves; they are never inputs to coefficient selection, parameter updates, hyperparameter selection, or early stopping. This design makes the final client-level comparison directly paired on a common trained federated backbone.

5.4. Lower-performing clients and communication

Table III. Mean of the two lowest client accuracies per seed

Dir. α	Method	Bottom-2 acc. (%)
0.3	FedAvg	85.80 ± 2.83
0.3	Fixed mix (0.5)	90.75 ± 1.42
0.3	Endpoint select	91.27 ± 3.10
0.3	AdaptiveMix	92.08 ± 3.29
1.0	FedAvg	90.57 ± 1.96
1.0	Fixed mix (0.5)	91.54 ± 2.13
1.0	Endpoint select	89.32 ± 2.89
1.0	AdaptiveMix	90.30 ± 3.14

Table III summarizes the two lowest-performing clients within each run. At α=0.3, AdaptiveMix has 92.08% mean lower-quartile accuracy, versus 85.80% for FedAvg. At α=1.0, the corresponding means are 90.30% and 90.57%. Local personalization therefore improves the weaker-client summary substantially in the more skewed setting, while the weaker-client metric offers no comparable gain in the milder one.

For descriptive detail only, among the 40 paired client outcomes at α=0.3, AdaptiveMix improves on FedAvg in 31, ties in 6, and declines in 3. At α=1.0, those counts are 12, 16, and 12. These are client observations nested within five seeds and are not treated as 40 independent replications for significance testing. The two ways of summarizing performance help distinguish improvements in the overall local average from their consistency across different clients.

A float32 vector of 3,610 parameters occupies 14,440 bytes. Counting one model-vector download and one upload for each of eight clients in each of 30 rounds gives 6,931,200 bytes of idealized uncompressed model payload, excluding network framing, security metadata, and retransmission. AdaptiveMix adds three final local fine-tuning epochs and up to five forward evaluations per local validation set, with no additional federated aggregation round. The float32 masking-consistency check reports a maximum absolute difference of 2.68×10^{-7} between masked and direct weighted sums (threshold 10^{-5}), an arithmetic check rather than a measured privacy level.

5.5. Measured partition heterogeneity

Table IV. Actual client label-shift diagnostics (five seeds)

Client-average statistic	α = 0.3	α = 1.0
Majority-class share (%)	42.56 ± 1.48	26.48 ± 1.35
Total-variation distance	0.53 ± 0.02	0.32 ± 0.02
Present digit classes	7.72 ± 0.35	9.85 ± 0.27
Label entropy (bits)	2.12 ± 0.06	2.87 ± 0.04

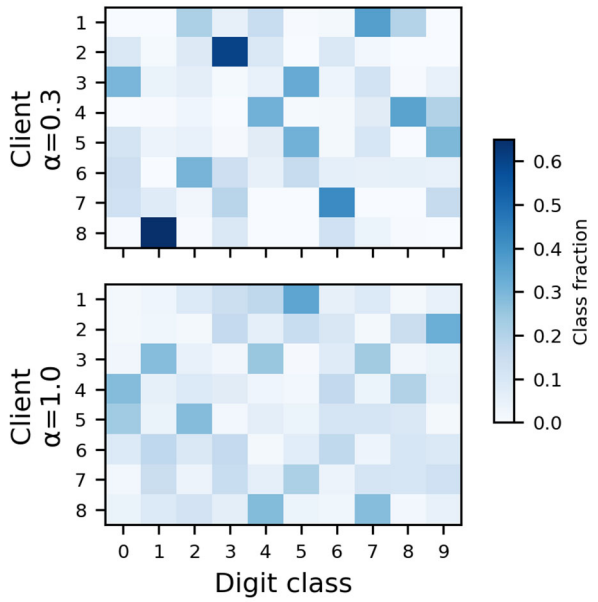


Fig. 6. Realized class fractions for the eight clients at seed 11; the same class-wise Dirichlet generator was used for both panels.

Table IV verifies that the Dirichlet settings produced distinct degrees of label heterogeneity in the executed partitions. The average total-variation distance between each client digit histogram and the full-dataset digit histogram is 0.533 for α=0.3 and 0.321 for α=1.0. The average majority-class share declines from 42.6% to 26.5%, while the mean number of observed digit classes rises from 7.72 to 9.85. Means and sample SDs are based on five seed-level averages, each covering eight clients. Figure 6 shows one fixed-seed realization rather than implying every seed has an identical class histogram.

6. Discussion

The central result is a separation between the gains from local fine-tuning and the gains from selecting an interior interpolation coefficient [46-50]. On the more skewed partitions, most clients choose full localization, so the gain over the original global model largely comes from local

adaptation. The extra validation grid offers a smaller additional gain. When label imbalance is less severe, coefficient choices are more varied and fixed mixing can perform as well as or slightly better than validation-selected mixing. This illustrates why a fixed statement that adaptive aggregation dominates every non-IID environment would be misleading [51-53].

The approach has practical advantages for limited-resource studies: it uses a single global checkpoint, a five-candidate local search, and no client-to-client model exchange during personalization. The selection criterion is local cross-entropy, whereas the outcome metric is test accuracy; these need not rank candidate models identically on small validation sets. Reporting both a fixed mix and endpoint selection helps reveal that distinction. Different user populations, naturally collected educational sequences, model architectures, severe client dropouts, or feature-distribution shifts could change the balance; the digit benchmark specifically tests controlled label heterogeneity rather than a deployed educational prediction task [54-59].

Privacy must also be interpreted at the correct layer. The local training and validation code does not centralize source images. However, a central simulation that sees client parameter vectors is not a proof of secure aggregation, and arithmetic cancellation does not establish resistance to reconstruction, collusion, or an adversarial server. A production privacy-preserving system could integrate an established secure aggregation protocol [27] and, if required, calibrated DP noise with explicit accounting [28]–[32]. Measuring utility under those protections would form a separate experiment. The present results quantify personalization accuracy and aggregation arithmetic in a reproducible simulator.

7. Conclusion

This paper evaluated validation-selected global–local model interpolation on a compact federated classifier, with a shared 30-round backbone, five paired seeds, two Dirichlet settings, and five selection rules. At α=0.3, AdaptiveMix achieved 96.05% client-macro accuracy versus 91.75% for FedAvg, a 4.31-point paired mean improvement. At α=1.0, it achieved 94.27% and remained close to the fixed-mixing reference. Endpoint selection, lower-performing clients, seed-wise differences, coefficient frequencies, and communication payload were examined using executed experimental records. A small on-client validation search provides a useful personalization baseline without altering the federated communication schedule.

References

[1] Mai, N. T., Cao, W., & Fang, Q. (2025). A study on how LLMs (e.g., GPT-4, chatbots) are being integrated to support tutoring, essay feedback and content generation. Journal of Computing and Electronic Information Management, 18(3), 43-52.

[2] Shen, A. (2023). Compute-efficient scaling of large sequential recommendation models under dynamic user behavior. World Journal of Information Technology, 1(1), 62-71.

[3] Chen, J., Cui, Y., Zhang, X., Yang, J., & Zhou, M. (2024). Temporal convolutional network for carbon tax projection: A data-driven approach. Applied Sciences, 14(20), 9213.

[4] Shen, A., & Liu, W. (2026). Budget-aware adaptive test-time compute with hierarchical self-verification for semantic ID generative recommendation. Computers.

[5] Mo, T., Zhang, C., Zou, J., Guo, Z., & Rhee, M. (2026). Self-evolving AI agents with dual memory for automated software testing and bug localization. IEEE Access.

[6] Guo, Z., Chen, T., & Shang, W. (2022). Class-conditional intermediate-domain adversarial adaptation for cross-device acoustic scene classification. Journal of Computer Science and Electrical Engineering, 4(1), 26-35.

[7] Rhee, M., Zou, J., & Qin, Y. (2024). CLARITY: A closed loop framework for duplicate aware failure prediction and budgeted automated remediation in cloud infrastructure. AI and Data Science Journal, 1(1), 76-85.

[8] Teng, D. (2025). RSRBench: Reproducibility and performance benchmarking for regional seismic risk workflows. Academic Journal of Architecture and Civil Engineering, 3(1), 40-52.

[9] Qin, Y., Rhee, M., Teng, D., Zhang, C., & Zi, B. (2026). Neuro-symbolic constraint verification for LLM-driven internet finance transaction execution. Scientific Reports.

[10] Li, P., Feng, E., Wang, G., & Mo, T. (2024). Hybrid deep reinforcement learning and integer optimization for resource scheduling in heterogeneous AI computing clusters. Innovation and Technology Studies, 1(1), 48-56.

[11] Teng, D., Rhee, M., Qin, Y., Zi, B., & Liu, W. (2026). SW-SpeedDLM: Sliding window speculative decoding for diffusion language models under long context constraints. Mathematics, 14(12), 2137.

[12] Liang, Y., & Jiao, Y. (2024). Real-time financial fraud detection using a feature-light temporal dynamic graph neural network with contrastive learning. Journal of Trends in Finance and Economics, 1(1), 76-84.

[13] Feng, E., Li, P., Mo, T., & Wang, G. (2022). PRISM: Prediction-guided risk-aware deep reinforcement learning for energy-aware dynamic resource scheduling in distributed cloud data centers. Journal of Computer Science and Electrical Engineering, 4(1), 53-61.

[14] Ping, W., & Fan, H. (2024). Disclosure quality and valuation risk in level 3 investments: A disclosure-implied risk framework for U.S. business development companies. Social Science and Management, 1(1), 78-87.

[15] Zhang, C., & Wang, Y. (2024). A data-driven multi-objective production scheduling framework for battery pack manufacturing under material arrival uncertainty. AI and Data Science Journal, 1(1), 86-93.

[16] Fang, Q., & Liu, W. (2025). HARLA-ED: Resolving information asymmetry and enhancing algorithmic symmetry in intelligent educational assessment via hybrid reinforcement learning. Symmetry, 18(1), 58.

[17] Fang, Q. (2024). Bias-aware large language models for equitable educational feedback and assessment. Trends in Social Sciences and Humanities Research, 2(1), 41-50.

[18] Wang, Y., Bian, M., & Lin, Y. (2022). An explainable temporal graph neural network for disruption prediction and risk propagation in multi-tier supply chains. Journal of Computer Science and Electrical Engineering, 4(1), 17-25.

[19] Qin, Y., Zou, J., & Rhee, M. (2022). An AI-driven intelligent transaction execution framework for internet finance platforms: Cost-sensitive real-time integrity interdiction. Eurasia Journal of Science and Technology, 4(2), 36-44.

[20] Wang, B., Zhao, W., & Wang, Z. (2024). DAS-GNN: A scalable disagreement-aware graph neural framework for anomaly detection in large-scale networks. AI and Data Science Journal, 1(1), 67-75.

[21] Chen, T., Shang, W., & Guo, Z. (2024). XSLA: An SLA-aware heterogeneous XPU scheduling framework for edge AI

inference. World Journal of Information Technology, 2(2), 48-59.

[22] Chen, J., Liu, J., Liang, Y., & Zhou, M. (2026). HeteroGCL: A heterogeneous graph contrastive learning framework for scalable and sustainable cryptocurrency AML. Applied Sciences, 16(6), 2860.

[23] Zhao, X., Liu, J., Wang, Y., & Wang, J. (2026). CryptoMamba-SSM: Linear complexity state space models for cryptocurrency volatility prediction. IEEE Open Journal of the Computer Society, 7, 226-243.

[24] Zhang, F., & Yang, J. (2024). Robustness-aware model compression for multimodal perception under challenging conditions. AI and Data Science Journal, 1(1), 103-111.

[25] Cao, J., Zheng, W., Ge, Y., & Wang, J. (2025). DriftShield: Autonomous fraud detection via actor-critic reinforcement learning with dynamic feature reweighting. IEEE Open Journal of the Computer Society, 6, 1166-1177.

[26] Fan, H., Liang, Y., Jiao, Y., Ping, W., & Lin, Y. (2026). An explainable temporal machine learning framework for financial fraud risk prediction in listed institutions. Highlights in Business, Economics and Management, 70, 173-185.

[27] Zhang, H. (2024). Graph-aware multi-task learning for early prediction of timing violations and routing congestion in digital IC physical design. World Journal of Information Technology, 2(1), 13-20.

[28] Lin, Y., Yue, X., Bian, M., Zhu, R., & Teng, D. (2026). Subgraph-level temporal graph transformer with laundering-motif prototypes for anti-money laundering on blockchains. IEEE Access.

[29] Wang, J., Ge, Y., Liu, J., & Jiang, B. (2024). Early warning of liquidity stress in U.S. banks with seasonally aligned frailty hazards and delayed-feedback conformal alarms. Journal of Trends in Finance and Economics, 1(1), 85-100.

[30] Zhou, F. (2026). Moment-constrained side-force allocation for wind-assisted merchant ships: Hydrodynamic performance of retractable anti-leeway fins and flap rudders. Innovation and Technology Studies, 3(1), 27-41.

[31] Zhang, F., Shen, A., & Huang, S. (2026). TRIAD: Co-designing mixed-precision quantization, operator fusion, and on-chip memory reuse for real-time vision-language-action inference on automotive NPUs. Results in Engineering, 113673.

[32] Bian, M., Wang, Y., & Lin, Y. (2024). Explainable supplier risk prediction for critical telecommunications infrastructure procurement: An empirical XGBoost-SHAP framework. Social Science and Management, 1(1), 64-77.

[33] Wang, Y., & Zhang, C. (2023). AERIS: Adaptive executor-scoped resource management for privacy-preserving data processing in Apache Spark. Eurasia Journal of Science and Technology, 5(2), 25-33.

[34] Lin, Y., Wang, Y., & Bian, M. (2024). XOR-EW: An explainable and cost-aware artificial intelligence framework for operational risk early warning in financial institutions. AI and Data Science Journal, 1(1), 94-102.

[35] Fan, H., Jiao, Y., Wang, M., Chen, J., & Yang, S. (2026). Fraud learns too: Continual graph learning under strategic adversarial drift in dynamic networks. Scientific Reports.

[36] Wang, G., Li, P., Feng, E., & Mo, T. (2024). Uncertainty-aware LLM-as-a-judge evaluation for reliable search and recommendation systems: Calibrated confidence and impact-aware selective human review. AI and Data Science Journal, 1(1), 112-121.

[37] Jiao, Y., Fan, H., Yue, X., Ping, W., Sun, T., & Wang, J. (2026). Dynamic heterogeneous graph contrastive learning for uncovering collusive financial fraud. Scientific Reports.

[38] Jiao, Y., & Liang, Y. (2024). An integrated process-mining and explainable machine-learning framework for bottleneck and internal-control risk detection in the multi-entity financial close. Social Science and Management, 1(3), 96-107.

[39] Liang, Y., Jiao, Y., Ping, W., Fan, H., & Han, X. (2026). Adaptive event-driven labeling: A neuro-symbolic multiagent framework for causal inference in non-stationary time series. IEEE Access.

[40] Wang, G., Li, P., Feng, E., Shen, A., & Rhee, M. (2026). A dual-process memory-augmented large language model agent for explainable sequential recommendation. IEEE Access.

[41] Shang, W., & Chen, T. (2024). PrefixKD-Edge: Real-time locomotion-intent prediction with prefix self-distillation on lightweight temporal networks. Innovation and Technology Studies, 1(1), 30-38.

[42] Mai, N. T., Fang, Q., & Cao, W. (2025). Measuring student trust and over-reliance on AI tutors: Implications for STEM learning outcomes. International Journal of Social Sciences and English Literature, 9(12), 11-17.

[43] Yue, X., & Zhu, R. (2024). An explainable hybrid machine learning framework for cash-flow-conditioned multi-horizon liquidity risk early warning: An SME-oriented benchmark study. World Journal of Management Science, 2(1), 63-72.

[44] Zou, J., Qin, Y., & Rhee, M. (2024). Beyond model scale: A task-aware reliability benchmark for large language models in clinical decision support. Innovation and Technology Studies, 1(1), 39-47.

[45] Ping, W., & Fan, H. (2023). Beyond isolated ratios: Accrual-revenue-cash flow indicators and interaction-aware learning for detecting financial reporting irregularities in U.S. public companies. World Journal of Economics and Business Research, 1(1), 46-55.

[46] Wang, J., Tan, Y., Jiang, B., Wu, B., & Liu, W. (2025). Dynamic marketing uplift modeling: A symmetry-preserving framework integrating causal forests with deep reinforcement learning for personalized intervention strategies. Symmetry, 17(4), 610.

[47] Wang, J., Liu, J., Ge, Y., & Jiang, B. (2023). Risk-aware ensemble learning for credit default prediction under economic uncertainty. World Journal of Economics and Business Research, 1(1), 18-32.

[48] Guo, Z., Fang, Q., Yang, J. S., & Rhee, M. (2026). Closed-loop audio neuroregulation for ADHD: Graph-based attention decoding and rhythm entrainment. IEEE Access.

[49] Sun, T., Yang, J., Li, J., Chen, J., Liu, M., Fan, L., & Wang, X. (2024). Enhancing auto insurance risk evaluation with transformer and SHAP. IEEE Access, 12, 116546-116557.

[50] Fang, Q. (2024). TRUE-EWS: Trustworthy early warning with equalized conformal guarantees for equitable student support in U.S. education. Educational Research and Human Development, 1(2), 44-54.

[51] Chen, J., Zhang, X., & Han, X. (2023). Explainable identification of tax-related financial reporting anomalies in U.S. public companies: Cross-fitted tax-consistency residuals and Owen-value Shapley propagation. World Journal of Economics and Business Research, 1(1), 33-45.

[52] Zhang, F., & Yang, J. (2022). Accuracy-aware model compression for real-time visual perception under resource constraints. Journal of Computer Science and Electrical Engineering, 4(1), 45-52.

[53] Yue, X., Yang, J., & Liu, W. (2026). FraudDebate-Agent: A multi-agent LLM framework with an evidence-based debate mechanism for financial statement fraud detection. Mathematics, 14(15), 2695.

[54] Mo, T., Wang, G., Feng, E., & Li, P. (2023). Execution-feedback-guided context selection for large language model-based software testing and debugging. World Journal of Information and Knowledge Management, 1(1), 55-64.

[55] Chen, J., Han, X., Zhang, X., & Yang, Y. (2023). Exact concept-level explanations for early detection of financial misstatements. Eurasia Journal of Science and Technology, 5(2), 34-48.

[56] Shen, A., & Huang, S. (2022). A systems-aware framework for intelligent large-scale recommendation models. Journal of Computer Science and Electrical Engineering, 4(1), 36-44.

[57] Chen, J., Liang, Y., Liu, J., & Zhou, M. (2026). Temporal transformer with conditional tabular GAN for credit card fraud detection: A sequential deep learning approach. Mathematics, 14(7), 1183.

[58] Wang, J., Liu, J., Zheng, W., & Ge, Y. (2025). Temporal heterogeneous graph contrastive learning for fraud detection in credit card transactions. IEEE Access.

[59] Zhu, R., & Yue, X. (2024). An explainable machine learning framework for predicting dividend sustainability and financial risk of US REITs under interest-rate shocks. Journal of Trends in Finance and Economics, 1(2), 48-57.