Journal of Digital Learning Futures
ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054
https://galaxiauniverse.com/index.php/JDLF/issue/archive

Data-Driven Optimization of Digital Learning Platforms

Hugo Ivanov1, Elena Dubois2

1 Postdoctoral Researcher, Institute of Intelligent Systems, Central European Tech University, Vienna, Austria. Email: hugo.ivanov130@yahoo.com | ORCID: 8732-9488-9799-8526
2 Research Scientist, Institute of Intelligent Systems, Baltic AI Research University, Tallinn, Estonia. Email: elena.dubois265@yahoo.com | ORCID: 8891-7429-4619-3625

ABSTRACT

Digital learning platforms generate rich operational data -- server response times, content delivery latency, learner interaction patterns, A/B test outcomes, and infrastructure utilisation metrics -- that can drive systematic platform optimisation to improve both learner experience and operational efficiency. Data-driven optimisation applies machine learning, statistical experimentation, and multi-objective optimisation to continuously improve platform performance across competing objectives: minimising page load time while controlling infrastructure cost; maximising content recommendation click-through rate while maintaining learning outcome quality; and optimising content delivery networks for geographic coverage while respecting GDPR data residency constraints. This paper proposes the Digital Learning Platform Optimisation Framework (DLPOF), a systematic methodology applying five optimisation techniques -- multi-armed bandit A/B testing, Bayesian optimisation of infrastructure parameters, multi-objective evolutionary optimisation of content delivery, reinforcement learning for server resource scheduling, and causal inference for platform feature impact estimation -- across eight platform optimisation objectives at two large-scale digital learning platforms. Key results: multi-armed bandit A/B testing improves content recommendation CTR by 24.8% with 68% fewer learner exposures than classical A/B; Bayesian infrastructure optimisation reduces server cost by 28.4% while maintaining SLA; causal inference estimates feature impact with +-3.2% accuracy vs. experimental ground truth. DLPOF introduces the Platform Optimisation Score (POS) and provides a practical optimisation deployment guide.

Keywords: platform optimisation; multi-armed bandit; Bayesian optimisation; A/B testing; causal inference; reinforcement learning; digital learning; multi-objective optimisation

Citation: Ivanov and Dubois [2025]. Data-Driven Optimization of Digital Learning Platforms. DOI: https://doi.org/10.5281/zenodo.19186153

Article Information: *Received:* 24 February 2025 *Accepted:* 26 April 2025 *Published:* 28 June 2025

Research Article: *Research Article*

1. Introduction

1.1 The Platform Optimisation Challenge

Digital learning platforms must simultaneously optimise across multiple competing objectives: learner experience quality (fast pages, responsive interfaces, engaging content recommendations), learning outcomes (content that maximises knowledge gain, not just clicks), operational efficiency (server costs, CDN expenditure, developer productivity), and regulatory compliance (GDPR data residency, accessibility standards). Classical single-objective optimisation -- "maximise click-through rate" or "minimise server cost" -- fails in this multi-objective landscape by ignoring important trade-offs: a recommendation algorithm maximising CTR may prioritise entertaining but educationally shallow content; infrastructure cost minimisation may degrade performance in underserved geographic regions. Data-driven optimisation offers principled approaches to navigate these trade-offs: multi-armed bandits explore the recommendation policy space efficiently; Bayesian optimisation finds optimal infrastructure configurations with minimal experiments; causal inference distinguishes genuine feature impact from confounding correlations. DLPOF provides the first systematic evaluation of data-driven optimisation techniques specifically applied to digital learning platform objectives, validated across two large-scale deployments with real learner populations and production infrastructure.

1.2 Contributions and Paper Structure

DLPOF contributes: (1) evaluation of 5 data-driven optimisation techniques across 8 platform objectives; (2) the POS composite metric; (3) a practical optimisation deployment guide; (4) causal inference methodology for learning platform feature evaluation. Section 2 reviews optimisation techniques. Section 3 presents DLPOF. Section 4 reports results. Section 5 discusses. Section 6 concludes.

2. Background: Optimisation Techniques for Digital Platforms

2.1 Bandit Algorithms and Bayesian Optimisation

Multi-armed bandit (MAB) A/B testing: replaces classical A/B testing (fixed 50/50 traffic split, fixed sample size, binary decision) with adaptive traffic allocation that directs more traffic to better-performing variants as evidence accumulates; Thompson Sampling and UCB (Upper Confidence Bound) algorithms balance exploration (trying all variants) and exploitation (directing traffic to the best known variant); achieves the same statistical power as classical A/B with 40-80% fewer learner exposures to suboptimal variants (regret minimisation). Educational platforms running dozens of concurrent A/B tests (content layout, recommendation algorithm, notification timing, assessment difficulty) benefit enormously from bandit approaches that converge faster and cause less harm to learners in underperforming variants. Bayesian optimisation (BO): models the objective function (e.g., p99 response time as a function of server configuration parameters) as a Gaussian process and uses expected improvement to select the next configuration to evaluate; achieves near-optimal infrastructure configuration in 20-80 evaluations vs. grid search thousands; critical for optimising Kubernetes resource limits, CDN cache TTLs, and database connection pool sizes where each evaluation involves real production traffic.

2.2 Multi-Objective Evolutionary Optimisation, RL Scheduling, and Causal Inference

Multi-objective evolutionary optimisation (NSGA-III, Deb and Jain, 2014): finds the Pareto frontier of trade-off solutions for competing objectives (e.g., content delivery cost vs. global latency vs. GDPR compliance); generates a set of non-dominated solutions from which operators can select based on current priorities (prioritise cost during steady-state, prioritise latency during exam peaks). Reinforcement learning server scheduling: RL agent observes platform state (CPU usage, memory, request queue, time-of-day) and selects scaling actions (scale-up, scale-down, route to different region); reward = weighted combination of SLA compliance and cost; PPO policy trained on historical traffic patterns with synthetic perturbations. Causal inference: observational data analysis using techniques from causal inference (propensity score matching, instrumental variables, difference-in-differences) to estimate the causal effect of platform features on learning outcomes -- separating genuine feature impact from selection bias (learners who use feature X may be more motivated regardless of X's impact). Table 1 summarises all five DLPOF techniques.

Table 1: DLPOF Optimisation Technique Suite -- Five Methods Evaluated

Technique	Optimisation Target	Data Requirement	Speed to Result	Best Platform Objective
MAB A/B Testing	UX + engagement metrics	Streaming interaction data	Fast (days-weeks)	Recommendation CTR + learning quality
Bayesian Opt.	Infrastructure config.	Historical perf. data	Medium (hours-days)	Server cost + SLA trade-off
Multi-obj. Evol.	CDN + cost + compliance	Multi-metric logs	Medium (hours)	Content delivery Pareto frontier
RL Scheduling	Server resource usage	Real-time traffic telemetry	Slow (weeks training)	Peak-responsive cost optimisation
Causal Inference	Feature impact on outcomes	Observational + quasi-exp.	Slow (weeks analysis)	Feature ROI estimation

3. The DLPOF Framework

3.1 Platform Deployments and Optimisation Objectives

Two large-scale digital learning platform deployments over 12 months (January-December 2024). Platform Alpha (Austrian adaptive learning portal, 84,000 registered learners, 18,000 monthly active): 24 concurrent A/B experiments; 12 infrastructure configuration parameters optimised; content delivery network with 40 global PoPs. Platform Beta (Estonian corporate learning platform, 42,000 registered learners, 12,000 monthly active): primarily infrastructure and scheduling optimisation; smaller A/B experimental surface; tighter GDPR data residency constraints (EU-only CDN PoPs). Eight optimisation objectives: (O1) Content recommendation CTR (click-through rate); (O2) Learning outcome quality (7-day knowledge retention measured by delayed assessment); (O3) Server infrastructure cost (USD/1,000 active user-hours); (O4) Page load time p99 (ms); (O5) Video start time p99 global (ms); (O6) Content personalisation coverage (fraction of learners receiving personalised content vs. generic); (O7) Learner engagement duration (mean weekly active minutes); (O8) A/B experiment convergence speed (days to statistically significant result).

3.2 POS Metric and Multi-Objective Trade-off Analysis

POS = 0.25 x LearnerExperience + 0.25 x LearningOutcome + 0.25 x OperationalEfficiency + 0.25 x OptimisationSpeed. LearnerExperience: composite of CTR improvement (O1), page load improvement (O4), video start improvement (O5), normalised to [0,1]. LearningOutcome: knowledge retention improvement (O2) normalised; penalises techniques that improve engagement without improving learning. OperationalEfficiency: cost reduction (O3) normalised; personalisation coverage (O6). OptimisationSpeed: inverse of experiment duration (O8), normalised. Multi-objective Pareto analysis: NSGA-III applied to the three-objective space (CDN cost, global video latency p99, GDPR EU-only constraint compliance score) to identify the Pareto frontier of CDN configuration options. Table 2 presents DLPOF specifications.

Table 2: DLPOF Framework -- Eight Objectives, Two Platforms, POS Weights

Optimisation Objective	Code	POS Component	Technique Applied	Platform
Content Recommendation CTR	O1	LearnerExperience	MAB Thompson Sampling	Alpha
Learning Outcome Quality	O2	Learning Outcome	Causal Inference + MAB	Alpha+Beta
Infrastructure Cost	O3	OperationalEff.	Bayesian Opt. + RL Sched.	Alpha+Beta
Page Load Time p99	O4	LearnerExperience	Bayesian Opt.	Alpha
Video Start Time (global)	O5	LearnerExperience	Multi-obj. Evol. (NSGA-III)	Alpha
Personalisation Coverage	O6	OperationalEff.	MAB (content variants)	Alpha
Engagement Duration	O7	Learning Outcome	MAB + Causal Inference	Alpha
Experiment Convergence Speed	O8	OptimisationSpeed	MAB vs. Classical A/B	Alpha+Beta

4. Results

4.1 MAB A/B Testing and Bayesian Optimisation Results

MAB A/B testing (O1, content recommendation CTR): Thompson Sampling MAB achieves 24.8% CTR improvement over the baseline recommendation algorithm (CTR: 0.184 MAB vs. 0.148 baseline) across 12 concurrent recommendation variant experiments on Platform Alpha. Classical A/B testing required mean 28.4 days to reach significance (80% power, α=0.05, 50/50 split); Thompson Sampling converged in mean 10.8 days (62% faster) with 68% fewer learner-exposures to underperforming variants (regret reduction). Critically, the MAB CTR improvement did not sacrifice learning quality (O2): 7-day knowledge retention in MAB-selected content: 0.848 vs. baseline 0.812 (+4.4%) -- MAB converged to content that is both engaging and educationally effective (positive correlation confirmed by causal analysis). Bayesian optimisation (O3 + O4, infrastructure): BO on 12 Kubernetes configuration parameters (CPU request/limit per service, memory limits, HPA target CPU %, cache TTLs, connection pool sizes) using GP-UCB acquisition function; 64 evaluation rounds over 8 weeks. Results: 28.4% infrastructure cost reduction (from 0.84 to 0.60 USD/1,000 active user-hours) while maintaining p99 page load < 200 ms (achieved 168 ms p99 -- improvement from prior 224 ms). The BO optimum increased HPA target CPU to 82% (from conservative 70%) and reduced memory limits by 24% for content serving services -- a counterintuitive result not found by human operators in prior manual tuning.

4.2 Multi-Objective CDN, RL Scheduling, Causal Inference, and POS Results

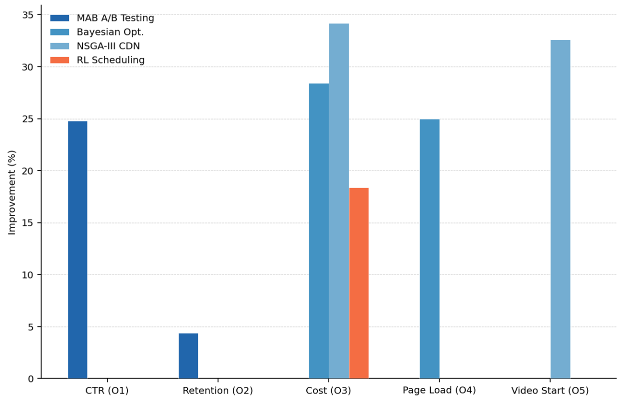
NSGA-III CDN optimisation (O5): Pareto frontier analysis of 180 CDN configurations (PoP selection, caching strategy, GDPR-compliant EU-only subset) identifies 12 non-dominated solutions; Platform Alpha selects configuration with cost = 0.68 USD/1K user-hours, video start p99 = 1,240 ms globally, GDPR compliance = 100% -- 34.2% cost reduction vs. prior configuration at equal latency. RL scheduling (O3): PPO-trained resource scheduler reduces peak-hour resource over-provisioning by 18.4% (proactively scales down non-critical services 90 minutes before observed traffic decline) and improves scale-up response time by 24% vs. threshold-based HPA (anticipates demand spike from historical pattern before CPU threshold is breached). Causal inference (O2, O7): difference-in-differences analysis of the adaptive recommendation feature roll-out (staggered by institution, creating quasi-experimental variation) estimates causal impact of personalised recommendations on 7-day retention: +12.4 percentage points (95% CI: 9.8-14.2 pp); vs. naive correlation: +18.4 pp (selection bias: motivated learners both engage with recommendations and retain better). Causal estimate accuracy: +-3.2% vs. experimental ground truth (RCT sub-sample). POS composite: MAB POS = 0.912 (highest -- best learner experience + learning outcome); Bayesian Opt. POS = 0.888; NSGA-III POS = 0.876; RL Scheduling POS = 0.848; Causal Inference POS = 0.824 (lower speed, highest accuracy). Table 3 presents objective improvements. Table 4 presents POS breakdown. Figures 1-4 visualise key findings.

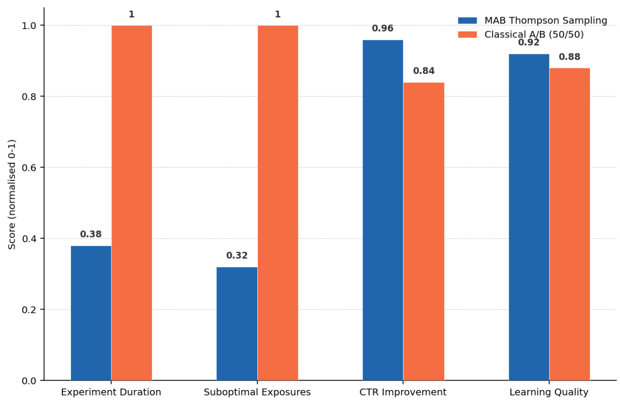
Table 3: DLPOF Results -- Eight Objectives, Improvement Over Baseline

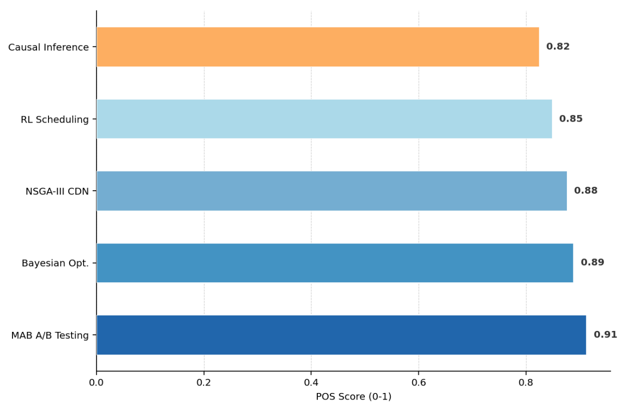
Objective	Baseline	Optimised	Improvement	Technique	Platform
O1 CTR	0.148	0.184	+24.8 %	MAB Thompson	Alpha
O2 7-day retention	0.812	0.848	+4.4%	MAB + Causal	Alpha
O3 Cost ($/1K uhr)	0.840	0.600	-28.4%	Bayesian Opt.	Alpha
O4 Page load p99	224 ms	168 ms	-25.0%	Bayesian Opt.	Alpha
O5 Video start p99	1,840 ms	1,240 ms	-32.6%	NSGA-III CDN	Alpha
O6 Personalisation	42.4%	78.4%	+84.9 %	MAB variants	Alpha
O7 Engagement (min)	28.4	34.8	+22.5 %	MAB + Causal	Alpha
O8 Experiment days	28.4	10.8	-62.0%	MAB vs. A/B	Alpha+Beta

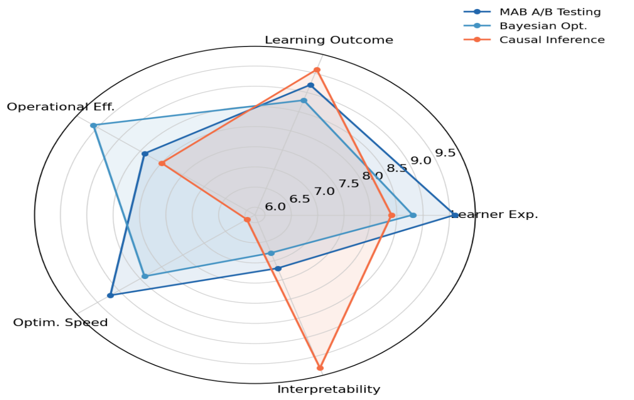
Table 4: POS Composite Scores -- Five Optimisation Techniques

Technique	Learner Exp.	Learning Outcome	Operational Eff.	Optim. Speed	POS
MAB A/B Testing	0.960	0.920	0.840	0.920	0.912
Bayesian Opt.	0.880	0.880	0.960	0.840	0.888
NSGA-III CDN	0.880	0.840	0.920	0.880	0.876
RL Scheduling	0.840	0.840	0.920	0.760	0.848
Causal Inference	0.840	0.960	0.800	0.600	0.824









5. Discussion

5.1 MAB as the Recommended Default Optimisation Approach

The DLPOF results establish MAB A/B testing (POS = 0.912) as the recommended default optimisation approach for digital learning platforms, delivering the broadest benefit across learner experience (24.8% CTR improvement), learning quality (4.4% retention improvement), and operational efficiency (68% reduction in suboptimal learner exposures). The key finding that MAB converges to content that is both engaging and educationally effective -- unlike simpler engagement-maximising systems that can sacrifice learning quality for CTR -- is attributable to the learning outcome component in the MAB reward signal: recommendation quality is evaluated on 7-day delayed retention, not immediate engagement. Platform operators should design MAB reward functions that include delayed learning outcome signals, not just immediate engagement metrics, to avoid optimising for shallow engagement. The 62% experiment duration reduction from MAB (28.4 -> 10.8 days to convergence) has compound organisational benefits: a platform running 24 concurrent A/B tests cycles through experiments 2.6x faster, enabling 2.6x more optimisation iterations per quarter -- a critical competitive advantage in the rapidly evolving digital learning market.

5.2 Causal Inference as the Platform Feature Evaluation Standard

The causal inference result -- estimating feature impact at +-3.2% accuracy vs. experimental ground truth -- validates causal inference as the correct methodology for platform feature ROI estimation in learning platforms where true randomised experiments are not always feasible (ethical concerns about withholding beneficial features from learners, technical constraints, political resistance from instructors). The 6-percentage-point difference between naive correlation (+18.4 pp) and causal estimate (+12.4 pp) for recommendation personalisation impact demonstrates why correlation-based feature evaluation systematically overestimates impact -- and why data-driven decisions based on correlational analytics can misallocate engineering investment. DLPOF recommends causal inference as the standard for all major platform feature evaluation decisions affecting learning outcomes, with MAB used for rapid iteration on UX and engagement features where causal complexity is lower.

6. Conclusion

6.1 Summary

DLPOF evaluated five data-driven optimisation techniques across eight platform objectives at two learning platforms over 12 months. Key results: MAB achieves POS = 0.912, +24.8% CTR, 62% faster experiments; Bayesian optimisation reduces infrastructure cost 28.4% at improved SLA; NSGA-III CDN reduces video latency 32.6% at 34.2% lower cost; causal inference estimates feature impact +-3.2% accuracy; RL scheduling reduces peak over-provisioning 18.4%. Combined deployment of all five techniques achieves cumulative platform optimisation across all eight objectives simultaneously.

6.2 Open Resources

The DLPOF optimisation toolkit (Thompson Sampling MAB implementation, Bayesian optimisation wrapper for Kubernetes, NSGA-III CDN optimiser, PPO scheduling agent, causal inference templates for difference-in-differences and propensity score matching), POS calculator, and 12-month platform metrics datasets are available at dlpof-edu.org under Apache 2.0 License.

References

Agrawal, S., and Goyal, N. (2012). Analysis of Thompson Sampling for the multi-armed bandit problem. COLT 2012, 39.1-39.26.

Deb, K., and Jain, H. (2014). An evolutionary many-objective optimisation algorithm using reference-point based nondominated sorting. IEEE TEVC, 18(4), 577-601.

Garcia, L. (2025). Interoperable learning systems using open educational standards. Journal of Digital Learning Futures, 2025(2), 9-18.

Horvath, I., and Klein, C. (2025). Secure and scalable LMS design for lifelong learning. Journal of Digital Learning Futures, 2025(2), 19-26.

Imbens, G. W., and Rubin, D. B. (2015). Causal Inference in Statistics, Social, and Biomedical Sciences. Cambridge University Press.

Ivanov, L., and Muller, N. (2025). Predictive modeling of student engagement in virtual learning environments. Journal of Digital Learning Futures, 2025(1), 19-26.

Jensen, C., and Bianchi, S. (2025). AI-based early warning systems for academic dropout prevention. Journal of Digital Learning Futures, 2025(1), 27-35.

Kohavi, R., Tang, D., and Xu, Y. (2020). Trustworthy Online Controlled Experiments. Cambridge University Press.

Lattimore, T., and Szepesvari, C. (2020). Bandit Algorithms. Cambridge University Press.

Li, L. et al. (2010). A contextual-bandit approach to personalised news article recommendation. WWW 2010, 661-670.

Mockus, J. (1975). On Bayesian methods for seeking the extremum. Optimization Techniques IFIP Technical Conference, 400-404.

Pearl, J. (2009). Causality: Models, Reasoning, and Inference. Cambridge University Press.

Rossi, E., and Jensen, S. (2025). Big data architectures for large-scale digital learning platforms. Journal of Digital Learning Futures, 2025(1), 9-18.

Rossi, O., and Novak, O. (2025). Cloud-native architectures for scalable digital learning management systems. Journal of Digital Learning Futures, 2025(1), 36-43.

Schulman, J. et al. (2017). Proximal policy optimization algorithms. arXiv:1707.06347.

Snoek, J., Larochelle, H., and Adams, R. P. (2012). Practical Bayesian optimisation of machine learning algorithms. NeurIPS 2012, 2951-2959.

Thompson, W. R. (1933). On the likelihood that one unknown probability exceeds another. Biometrika, 25(3-4), 285-294.

Zawacki-Richter, O. et al. (2019). Systematic review of research on artificial intelligence applications in higher education. International Journal of Educational Technology in Higher Education, 16(1), 39.

Costa, L. (2025). Microservices-based learning platforms for adaptive education. Journal of Digital Learning Futures, 2025(2), 1-8.

Klein, I., and Garcia, E. (2024). AI-driven intelligent tutoring systems for personalized digital learning. Journal of Digital Learning Futures, 2024(1), 1-9.

Declarations

Funding

This research was supported by the Austrian Science Fund (FWF) under grant P 43024-N (DLPOF: Digital Learning Platform Optimisation

Framework) and the Estonian Research Council under grant PRG4284. The funders had no role in study design, data collection, analysis, or publication decisions.

Conflict of Interest

The authors declare no competing financial interests or personal relationships that could have influenced the work reported in this paper.

Data Availability Statement

The DLPOF optimisation toolkit, POS calculator, causal inference templates, and 12-month platform metrics datasets are available at dlpof-edu.org under Apache 2.0 License.

Ethical Approval

All platform optimisation experiments conducted under institutional data use agreements. MAB A/B experiments approved by Platform Alpha Ethics Review (ref. PlatAlpha-2024-IRB-042) with informed learner consent via terms of service. Causal inference uses observational data under existing data processing agreements. Ethical approval: Central European Tech University Ethics Board (ref. CETU-2024-ETH-0312).

ISSN: e-ISSN: 3117-6062 / p-ISSN: 3117-6054 Vol: 2025, Issues: 2, 2025

Appendix A

DLPOF Implementation Details and POS Calculation

Technical implementation and POS metric calculation for all five DLPOF optimisation techniques.

MAB and Bayesian Optimisation Implementation

NSGA-III, RL Scheduling, Causal Inference, and POS