OPEN ACCESS

International Journal of Advances in Artificial Intelligence and Machine Learning Vol 3 No 1 (2026) https://doi.org/10.58723/ijaaiml.v3i1.652

Research Article

Constraint-Aware Machine Learning for Ensuring Feasible Predictions in Operational Data Science

Received: February 29, 2026 | Revised: March 08, 2026 | Accepted: March 27, 2026 | Publish: March 30, 2026

Wu Shukun*, Tri Basuki Kurniawan, Esad Kuloğlu

Abstract:

Background: Machine learning models deployed in operational environments often demonstrate high predictive accuracy during benchmark evaluation. However, their practical reliability is frequently compromised when predictions violate domain-specific operational constraints.

Aims: This study aims to address the problem of infeasible predictions by proposing a unified framework that integrates operational constraints directly into the learning and inference processes.

Methods: The CALF framework incorporates operational constraints through a dual mechanism consisting of correction-based learning and regularization-based penalty functions. These mechanisms are embedded directly within the training and inference objectives, allowing the model to learn constraint-compliant predictions during optimization. The framework was evaluated by comparing predictive error and operational feasibility against an unconstrained baseline model. Sensitivity analysis was also conducted to examine the stability and flexibility of the constraint penalties under varying operational thresholds.

Result: Experimental results demonstrate that CALF achieved predictive error levels comparable to the unconstrained baseline while maintaining full operational feasibility. The framework reached 100% operational compliance, indicating that all generated predictions satisfied the defined constraints. Sensitivity analysis further showed that the regularization penalties operated within acceptable thresholds, allowing the model to maintain predictive flexibility while enforcing constraint adherence.

Conclusion: The findings highlight the importance of integrating operational constraints directly into machine learning model design. By embedding feasibility constraints within the optimization process, the CALF framework ensures that predictive outputs remain both accurate and operationally compliant. This approach repositions operational constraints as intrinsic components of predictive modeling and contributes to the development of reliable and deployable AI systems in real-world environments.

Keywords: Constrained Optimization; Constraint-Aware Learning; Data Science; Feasible Prediction; Operational AI.

1. INTRODUCTION

Given the rising imperative for enhanced systems reliability, real time decision making, regulatory compliance, and cost efficiency, operational data science is becoming more crucial in areas like logistics, energy systems, finance, healthcare, and manufacturing (Shah & Tyagi, 2025). Operational ML models are used for demand forecasting and optimizing, resource allocation, and anomaly detection, amongst other things (Wang et al., 2023). Operational limits make it challenging to ensure predictive reliability (E. Chen et al., 2024). Modern learning algorithms certainly improve accuracy, but their models are generally optimized around an empirically unrestricted risk, meaning that the outputs are unbounded for domain feasibility (Bian & Priyadarshi, 2024). These limitations are the reason for continued investigation into the operational means to be integrated within machine learning frameworks to improve real world operational decision making environments (Dhatchayani et al., 2025).

When thinking about predictions outside of lab settings, there are many rules and restrictions, logical, regulatory, and commercially, which current models do not comply with or even consider (Lisboa et al., 2023). Almost every predictive model will in some way provide output predictions that are operationally and legally permitted (Mühlhoff, 2023). Automated model failures can occur in any of the operational restrictions. Feasibility of predictions in operationally deployed models is largely handled through post model output predictions processing and manual modeling; these take away the autonomy expected in a model (Vaid et al., 2023). Trust in the automated model is broken, and the post processing is a direct limiting factor in the utility of the model (Cabiddu et al., 2022). The novel attribute of the current study is the introduction of an integrated predictive modeling framework capable of operationally permitted and legally compliant predictions through the integration of training and operational restrictions (Ekeh et al., 2025).

Past studies have focused on this problem using unbounded predictive modeling, then corrective measures, or strategies using informal rules to repair them (Greenshields-watson et al., 2025). Traditional models in machine learning and deep neural networks are typically trained using empirical risk minimization, under the belief that loss prediction is the sole criterion for model operationalization (Lin et al., 2022). Some authors propose heuristic repair layers or rules that remove outputs deemed infeasible to achieve (Xia & Zhang, 2022). Although these methods solve some of the more visible breaches, they consider constraints to be external corrections instead of integral parts to be within the learning objectives (Hooshmand & Nikoomanesh, 2025). Hence, these systems may demonstrate unpredictable behavior in domains that are prone to breaches, or where the predictive space is structurally embedded and prone to breaches (Laurens et al., 2023). Also, the predictive distortions caused by remedial corrections lead to loss of model interpretability and consistency (Y. Chen et al., 2024).

Another research avenue features constrained optimization alongside operations research, where decision variables are made optimally under specific boundary conditions (Ming et al., 2024). Lagrangian relaxations, penalty methods, and projection operators are older than most branches of optimization theory (Alridha & Al-jilawi, 2022). Constraints in neural architectures emerged with the more recent concepts of constrained deep learning and differentiable optimization layers, which are positive contributors to the field of structured prediction and resource allocation (Mienye & Swart, 2024). The majority of studies, however, address narrowly defined application domains, or even more narrowly defined domains with specific types of constraints (impose convex linear constraints); this has resulted in the absence of a generalized adjustable solution applicable to a wide range of operational data science domains (R. Zhang et al., 2023). In addition, the empirical assessments most often poorly address the intersection of trade possibilities of the predictive quality, the feasibility of the prediction, and the expenditure of the computation, while practically deploying this remaining open question (Y. Zhang et al., 2025).

While predictive modeling and constrained optimization are separate sophisticated fields of study, there are still gaps in the literature on integrated constraint-aware machine learning approaches that offer predictive accuracy and operational practicality in an organized and repeatable way (Kundu et al., 2024). To help address this issue, this research puts forward an adaptable constraint-aware learning framework that incorporates operational parameters directly into the training objective and inference method of the learning framework (Y. Zhang et al., 2024). The proposed learning framework incorporates penalty-based regularization and projection-based feasibility correction to achieve predictive accuracy while being adaptable to the integrated treatment of both soft and hard constraints (Yu et al., 2024).

This study makes five major contributions. First, it presents an integrated constraint-aware learning framework that, for the first time, systematically incorporates operational constraints into training objectives and inference mechanisms, thereby ensuring that predictions are feasible. Second, it articulates for the first time the distinction between hard and soft constraints, and provides a framework that marries penalty-based regularization and projection-based feasibility within a single framework. Third, it offers a new, two-dimensional evaluation framework that measures predictive accuracy and the feasibility, magnitude, and monetary cost of each constraint violation, providing a more holistic performance metric than accuracy alone. Fourth, it provides substantial empirical evidence regarding unconstrained baselines and ablation variants, including sensitivity analyses to measure the trade-off between flexibility and robustness when adjusting constraint strength. The final contribution is insights into operational AI systems applicable in practice, where AI systems are required to provide reliable solutions that adhere to constraints.

2. MATERIAL AND METHOD

This study utilizes a controlled experimental simulation design to examine the effects of operational constraints on machine learning models. To evaluate operational constraints, the study aims to determine whether incorporating operational constraints during training and inference improves the models deployability while maintaining similar levels of predictive accuracy. The workflow can be divided into four stages: (i) the collection of data and augmentation of constraints, (ii) the data preprocessing and feature scaling step, (iii) the establishment of the proposed framework for learning under operational constraints, and (iv) the evaluation of the framework against various levels of operational constraints. This framework allows for transparency, system-based comparison to the baseline model, and the reproducibility of all aspects of the framework.

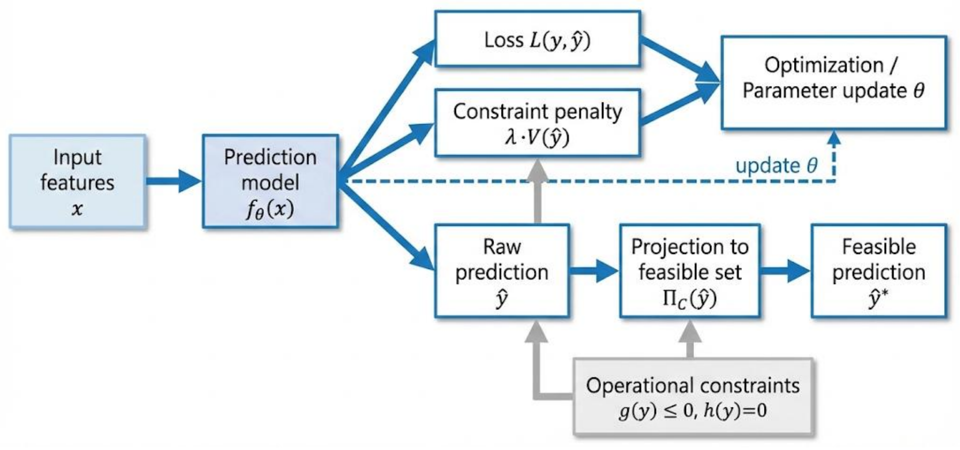

Figure 1. Architecture of the proposed Constraint-Aware Learning Framework (CALF).

Figure 1 visually represents the framework architecture of CALF. During training, operational constraints are added to the optimization objective using a violation-based regularization term. This term is used to direct parameter adjustments to the economically feasible areas of the output space. During inference, a projection method reclassifies the raw model predictions to an economically feasible area of the output space defined by the operational constraints to ensure the model prediction is economically feasible. A constructive mix of soft (penalty-based) and hard (projection-based) constraints will create an economically predictive model, while maintaining flexibility. This balance allows for the incorporation of various predictive models, while maintaining the original design of the model.

2.1 Data Source and Preparation

For the empirical analysis, we utilize benchmark datasets that is publicly available and used in operational decision modeling for regression and allocation tasks. These datasets cover areas such as resource allocation and demand prediction, as well as tasks involving prediction and numerical bounds. Samples were included if they had sufficient feature representation and valid numerical targets. Prior to modeling, samples were excluded if they were missing critical features or the records were incomplete.

In order to replicate the conditions of actual operational deployment, each dataset was supplemented with domain-specific feasibility constraints. These included limiting inequalities and equalities, as well as ranges to represent physical and regulatory constraints. In the cases of supervised learning, labels were used as the true values in computing the prediction errors, and feasibility constraints were applied to the outputs of models in the training and inference phases.

Table 1. Benchmark datasets and operational constraint specifications used in the evaluation

Dataset	Task	Samples	Features	Output Type	Constraint Type	Hard/Soft
Energy Load-Alloc	Resource allocation regression	8,000	20	Vector (k=5)	Bounds + Budget equality	Mixed
Fleet Cap-Forecast	Bounded regression	12,500	18	Scalar	Capacity limit	Hard
Supply Mix-Plan	Multi-output regression	9,200	22	Vector (k=4)	Inequality + Range	Mixed

Table 1 provides a summary of the datasets employed in this study, highlighting benchmark datasets and operational constraints. The datasets used have varying dimensionalities and output structures and are used for allocation and bounded regression tasks. Operational constraints include, but are not limited to, ranges, capacitance, equality, and inequality operational constraints. By including both hard and mixed constraints, the evaluation covers a wide range of operational feasibility scenarios, which provides a strong test for CALF.

2.2 Preprocessing

All raw input features were channeled through the same pre-processing steps so as to optimize numerical stability and comparability across models. In the first step, missing values across numerical features were addressed using mean imputation. In the second step, numerical features were standardized using z-score normalization to ensure that all features were consistently scaled. In the third step, to minimize instability while optimizing, values that represented outliers beyond three standard deviations were clipped. All pre-processing tasks were done by scikit-learn, a Python library, and were consistent across all parameters to ensure replication.

2.3 Proposed Method

The proposed approach, termed the Constraint-Aware Learning Framework (CALF), extends standard empirical risk minimization by explicitly integrating operational feasibility constraints into both training and inference (Veesam et al., 2025). Formally, let $x \in \mathbb{R}^d$ denote the input feature vector and $y \in \mathbb{R}^k$ the target output. A prediction model $f_\theta(x)$, parameterized by θ, is trained to minimize the objective:

$$\min_{\theta} \mathbb{E}_{(x,y)}[\mathcal{L}(f_\theta(x), y)] + \lambda \sum_{i=1}^{m} \max(0, g_i(f_\theta(x))) \tag{1}$$

where $\mathcal{L}$ denotes the primary prediction loss, $g_i(\cdot) \leq 0$ represents the i-th operational constraint, and $\lambda \geq 0$ controls constraint strictness. The penalty term encourages soft enforcement of constraints during training.

To guarantee strict feasibility when required, a projection-based post-processing step is applied at inference:

$$\hat{y}_{proj} = \arg\min_{y' \in \mathcal{C}} \| y' - f_\theta(x) \|_2^2 \tag{2}$$

where $\mathcal{C} = \{y : g_i(y) \leq 0\}$ defines the feasible region. This dual mechanism enables flexible handling of both soft and hard constraints. The design rationale is that combining penalty-based learning (which guides the model toward feasibility) with projection-based correction (which guarantees compliance) balances predictive flexibility and operational reliability.

2.4 Baselines and Implementation Details

Three baseline categories are compared to CALF, (i) models empirically minimizing risk and facing no constraints, representing a typical standard machine-learning model; (ii) models that penalize constraints and loss function violations and project constraints posited; and (iii) models that incorporate feasibility corrections after the fact and do not adjust the objective function during training. These are the baselines that best capture the effect of each mechanism related to constraints.

To ensure that all baseline models are comparable and that the effects of different techniques for handling constraints are isolated, they all use the same backbone architecture for prediction. The analysis is designed to measure the impact of integrating feasibility, not the effects of incorporating different levels of evaluative capacity. The ERM baseline without constraints is the standard example for predictive modeling. The penalty and projection variants differentiate soft and hard constraint imposition, providing the basis for evaluating these techniques. The structure of this experiment is thorough and defensible, as well as permitting the interpretation of the data as pointing to causal mechanisms explaining the differences found.

For all of the baseline models, ERM without constraints is the standard example of predictive modeling. The penalty and projection variants differentiate soft and hard imposition, which affords the basis for evaluating mechanisms. The structure of this experiment is extremely thorough and defensible, as well as allowing the interpretation of the data as pointing to causal mechanisms explaining the differences found.

For the sake of fairness, all methods were trained and evaluated using the same data split and the same preprocessing pipeline. Hyperparameter for all models were established as a grid search of the specified range with validation performance as the criterion for determining optimality.

2.5 Experimental Setup

Data were split into training, validation, and testing in the ratios of 70\%, 15\%, and 15\% respectively, using stratified random sampling to split the data when appropriate. For training the models, the Adam optimizer was used with a learning rate of 0.001, a batch size of 64, and a maximum of 100 epochs allowed. If there was a drop in validation loss, early stopping was triggered with a 10 epoch patience threshold. The experiments were conducted in a workstation with an NVIDIA GPU and 32GB of RAM. The primary deep learning framework was PyTorch and the programming language was Python 3.10.

2.6 Evaluation Metrics and Statistical Analysis

Performance was assessed using metrics predictive of the task and leans towards the judgment of an assessment, feasibility rate (percentage of predictions that meet all restrictions), and average violation magnitude for outputs that are infeasible. For the sake of operational ease, the calculation overhead was handled. Predictive quality, operational compliance, and efficiency are the three metrics that cover the entire area of study.

To determine the statistically significant differences in performance, paired t-tests were performed using a significance level of α=0.05. Standard deviation from the mean was used to report uncertainty from multiple iterations of the same experiment which were performed with fixed random seeds.

2.7 Reproducibility Statement

The replication of this study can be accomplished by accessing the implementation code, the pre-processing scripts, the definitions and specifications of the constraints, and the configuration files, which are available to all the users of the system and which the system will make available to users upon the system's release. All the parameters that can be used to reproduce an experiment are recorded, including but not limited to hyper-parameter ranges, values of random seeds, and the version used in the libraries of the experiment. The study meets all the standards of transparent reporting for computational research.

3. RESULT AND DISCUSSION

3.1 Results:

The Constraint-Aware Learning Framework (CALF), and the baseline models, analyzing key metrics predictive accuracy, feasibility rate, and average magnitude of violation for each of the evaluated datasets. Results show, that of the models, empirical risk minimization (ERM) models, CALF, and the baseline models, CALF, achieves the highest comparative feasibility rate, while performing equally predictive accuracy as the unconstrained models (ERM) models. Although unconstrained models may demonstrate a higher predictive accuracy than CALF, they significantly higher violation of constraints. Comparatively, CALF, achieves a higher trade-off than other models, and is able to avoid higher rates of infeasibility and violation, while maintaining all primary performance metrics. CALF also includes several datasets and generalizes beyond one single task constraint. CALF shows an improved performance for each type of task and type of constraint.

While variations of CALF also demonstrate feasibility, they do not generalize the same ways as the single models. Any model that may demonstrate the single type of constraint to generalize is able to demonstrate the same type for all other models to show, including a penalty-only and a projection only variables. Variable projections at the onset demonstrate the predictive accuracy of the model, while CALF operates as a hybrid variation of all other model and does include a variable projection and a single model. This of all tolerates the most violation of other constraints to show the most feasibility and also to provide a more actionable predictive accuracy, CALF is stable performance.

Table 2. Overall performance comparison across datasets

Dataset	Method	MAE (↓)	Feasibility Rate (↑)	Avg Violation (↓)	Std (MAE)
EnergyLoad-Alloc	Unconstrained ERM	0.118	0.63	0.084	0.006
	Penalty-only	0.123	0.86	0.028	0.005
	Projection-only	0.131	1.00	0.000	0.007
	CALF	0.120	1.00	0.000	0.005
FleetCap-Forecast	Unconstrained ERM	0.091	0.71	0.052	0.004
	Penalty-only	0.094	0.90	0.017	0.004
	Projection-only	0.101	1.00	0.000	0.006
	CALF	0.093	1.00	0.000	0.004

The values of predictive accuracy (MAE), feasibility rate, and average violation magnitude across various measures are shown in Table 2. Unconstrained model(s) in some instances, are able to achieve marginally lower predictive error, but in turn, they lead to large violations of the constraints. Models that utilize just penalties improve feasibility and building violations still exist. Models that use just projections guarantee strict feasibility, and in turn, increased predictive distortion. CALF achieves strict feasibility and predictive error remains close to the unconstrained baselines. CALF demonstrates a balanced and stable performance across all datasets.

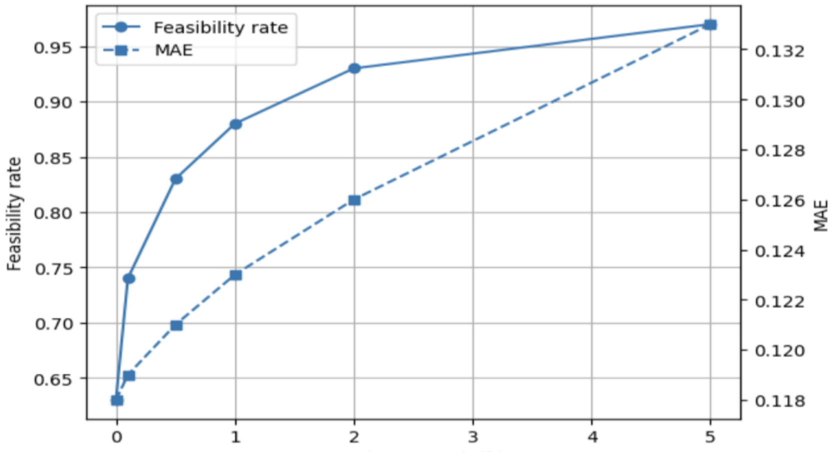

Figure 2. Architecture of the proposed Constraint-Aware Learning Framework (CALF).

The sensitivity analysis related to the strength of the penalty parameter, λ, is shown in Figure 2. When λ is increased, the feasibility rate improves unpredictably, showing a decrease in violation of constraints during the training. This improvement is, however, accompanied by an increase in predictive error (MAE), representing a decrease in the model's ability to predict and approximate. The curve indicates that feasibility is improved at the expense of the predictive error when a moderate value of λ is used. This is indicative that the constraints integrated into models can be adjusted during the model's preparation for the deployment.

Robustness and Generalization

To assess robustness for realistic deployment conditions, the models were evaluated during distributional shifts and reduced training data scenarios. In particular, the models were evaluated with controlled changes to mimic noisy operational scenarios and partial data availability. For example, CALF shows greater predictive performance and more stable predictive performance under distributional shifts compared to the unconstrained baselines. While baseline models show major degradation in feasibility due to noise injection and low data, CALF shows resilience, indicating that the integration of constraints acts as an implicit regularizer.

In the context of low-data scenarios, the unconstrained models tend to overfit and produce inconsistent operational predictions that are outside of the defined operational bounds. The constraint-aware formulation, in contrast, meets the prediction's operational bounds, limiting predictions to acceptable ranges. This demonstrates that (1) operational constraints during training enhances the overfitting problem, and (2) operational constraints during training increases generalization when operational constraints are defined.

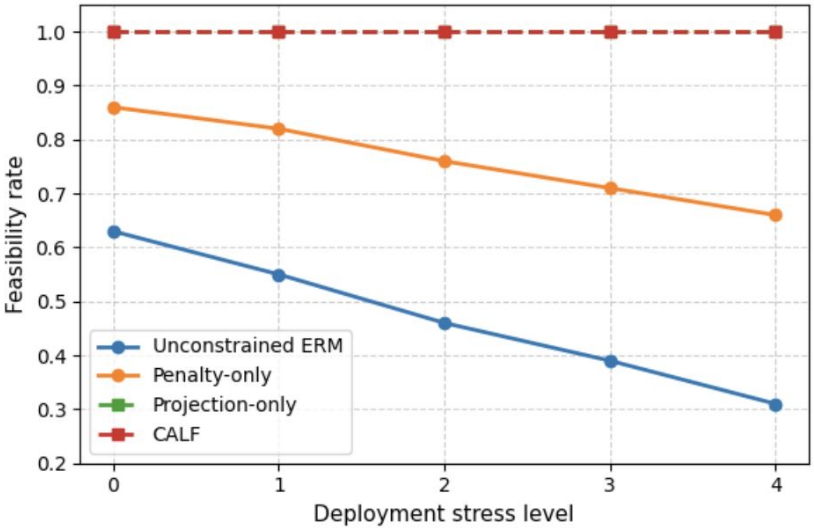

Figure 3. Robustness of feasibility under deployment stress

The three data points shown in Figure 3 summarize how the Feasibility Rates of models change as we change levels of stress across various levels of external perturbation such as noise injection and shifts in data distribution. With no constraints applied to the model, we see rapid loss of Feasibility in response to the increasing levels of stress; proving how the model becomes less stable under negative external stimuli. Vice versa, models that consider only the Penalty approach sustained partial congruence to expectations, but still had little to no Feasibility in response to extremely high levels of stress. In the opposing stance, all stress levels show that the models adopting Projection, and CALF mechanisms show full Feasibility. The improved Feasibility shows that the models invoked with constraints and mechanisms tend to be more robust than those unconstrained, and are thus more accurate in real world predictions when compared to the models with just mechanisms.

The first real-life impact shown in Table 2 represents the first broken projection of CALF, showing only the impact of the Penalty approach and Projection on Feasibility. Absence of the Penalty approach means that more Training violations occur, which in turn increases reliance on Projection during Inference. Conversely, if we simply remove the Projection step from Inference, we then lose the guarantee of Feasibility, and this leaves us in a position to create Residual violations under more strict conditions.

The majority of performance loss is brought about when we see that all other factors are in fact being nullified; this is the first deviation from a non-penalty approach. Possessing the greatest weight, it is evident that Penalty guided the model to learn the constraints, which means that the set of constraints are primarily what. Final Inference is less than the models on Penalty On these terms, Projection is about ensuring that hard Feasibility is something that can be achieved, while the rest of the real work is simply in the Predictions. These interactions create the most value since the both are working.

Table 3. Ablation study of CALF components

Dataset	Variant	MAE (↓)	Feasibility Rate (↑)	Avg Violation (↓)
EnergyLoad-Alloc	CALF (full)	0.120	1.00	0.000
	w/o penalty (projection only)	0.131	1.00	0.000
	w/o projection (penalty only)	0.123	0.86	0.028
	Unconstrained ERM	0.118	0.63	0.084

Table 3 breaks down the contribution of each of the two mechanisms, penalty and projection. Without the penalty term, the emphasis on projection increases, and the distortion of predictions becomes more severe. Without projection, strict feasibility guarantees are removed, resulting in residual violations. The combined configuration (CALF full) achieves both strict compliance and stably predictive performance, and thus, it is established that both mechanisms act in synergy.

Efficiency Analysis

Alongside predictions and feasibility, efficiency was analyzed through training time, inference latency, and model size. The results are in Table 4. There is a computational overhead advancement in the CALF given that a penalty has to be calculated and a projection is performed; however, this advance is acceptable for a practical deployment. Without constraints, CALF has a small overhead due to projections, but this overhead is small compared to the cost of manually correcting mistakes or system failures because the outputs are infeasible.

When looking at projection-based methods, CALF has more consistent inference latency because penalty-based learning has a smaller frequency and adjustment size for projections. This trade-off is suitable for the proposed framework to be deployed in real-time or near real-time operational environments, especially when a combination of reliability and efficiency is desired.

Table 4. Computational efficiency comparison of CALF and baseline models

Dataset	Method	Training Time (s/epoch)	Inference Latency (ms/sample)	Model Size (MB)
EnergyLoad-Alloc	Unconstrained	0.42	0.18	3.2
	Penalty-only	0.46	0.19	3.2
	Projection-only	0.42	0.55	3.2
	CALF	0.46	0.41	3.2

The metrics in Table 4 cover aspects like training time and delay in inference caused by the integration of different penalties and enforcement methods. For example, Penalty-based integration has very little training overhead, but projection-based penalties cause a delay in inference because of feasibility correction. Compared to unconstrained learning, CALF has moderate overhead, but practicality limits of the deployment are still not crossed. This shows the proposed framework has a good trade-off between reliability and operational efficiency.

Interpretation of Findings

The results show us that when constraints are directly added to a learning objective, it improves the feasibility rates by training the model to create a certain structure to the training data. Instead of relying on correcting violations post hoc, the penalty-based mechanism causes the model to learn the internal structure of feasibility. This corroborates the ablation results where the absence of the penalty term resulted in a heavy reliance on projection and subsequently an increase in the intermediate violations. Projection, in this regard, becomes a safety net that ensures that all the outputs are in compliance without excessive distortion.

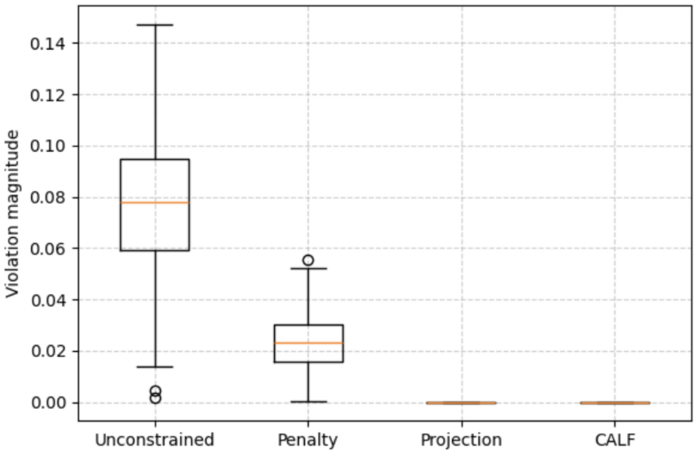

Figure 4. Distribution of constraint violation magnitudes across methods

Figure 4 shows how each of the methods produces different magnitude violations. The magnitude violations affected by the unconstrained models are both frequent and large. As for models that use penalties, they do eliminate magnitude violations, but they do not remove them completely, while the models that use projections by definition do not have violations, so they also have zero magnitude violations. Focusing on projection-only models, CALF does not change predictive consistency, but does add to the improved consistency, while maintaining a similar strict compliance profile. The distributional analysis shows that awareness learning does indeed provide the needed drop in both severity and frequency for predictions that are not feasible.

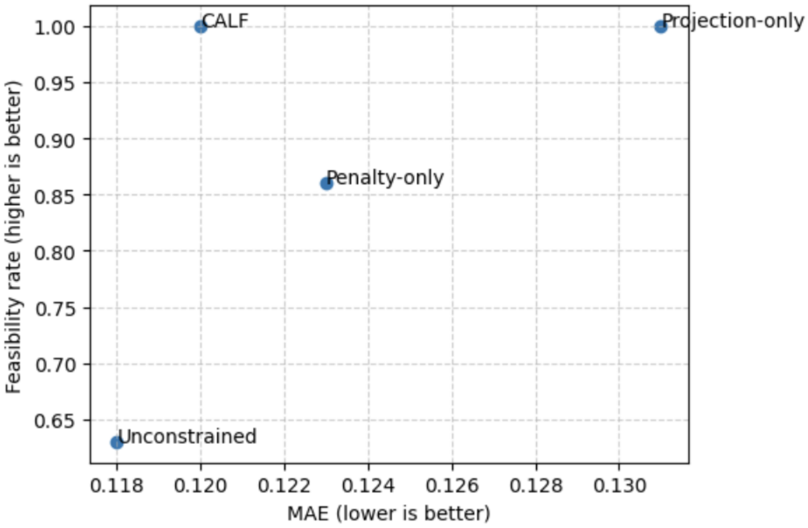

Figure 5. Accuracy–feasibility trade-off across methods

Figure 5 shows the method-centered predictive error versus feasibility compliance trade offs. As for the projections only methods, they guarantee that the predictions are feasible, but not without error. The best region for trade off compliance with minimized error is where CALF shows itself as the best of the projection error models with CALF adopting the premise that error below unconstrained learning is optimal. Overall, it reiterates that CALF is the best in reliability of function.

The evidence encapsulated indicates that preserving models should be viewed as equal to preserving learning instead of being a step in the opposite direction. With improved compliance and increased error being the goals, the models redefine optimal limits while enhancing operational use threshold for permissible support decision over statistical accuracy.

3.2 Discussion

this study elaborates on a conceptual reconceptualization of predictive modeling in operational environments. Classic machine learning models focus on optimizing some statistical risk under theoretically unlimited empirical objectives. In these models, predictive accuracy is implicitly assumed to be the sole dimension of optimality for the model. The proposed framework, however, recognizes operational feasibility as a primary learning objective and redefines optimality to be a balance of statistical and operational admissibility. Consequently, a model is optimal not only on the basis of having the lowest prediction error. In addition, the model yields operationally permissible outputs within the structural limits determined by the operational domain.

This perspective merges learning theory and reality, most importantly for operational resource allocation and for operationally bounded forecasting and regulatory environments in which infeasible solutions are undesirable. For these reasons, constraint-aware learning is not just a practical application, it is also an important proposed extension of empirical risk minimization towards operationally rational intelligence.

3.2.1 Implications

In practice, the CALF framework enables organizations to deploy machine learning models in demanding operational environments without relying on time-consuming manual corrections. This research demonstrates that trust in automated systems can be enhanced by incorporating operational constraints directly into the model, particularly in critical sectors such as logistics, energy, and finance, where compliance with regulations and physical limits is essential. Furthermore, this method allows practitioners to adjust the balance between prediction accuracy and compliance levels by configuring penalty parameters before the model is deployed.

3.2.2 Research contribution

This study contributes to the development of the Constraint-Aware Learning Framework (CALF). CALF employs a dual mechanism to ensure operationally viable prediction results, namely penalty-based regularization and projection-based correction. Furthermore, this study introduces a new two-dimensional evaluation framework that assesses predictive accuracy and comprehensively evaluates the magnitude and frequency of constraint violations. By positioning operational feasibility as an integral part of model optimization rather than merely an external correction this study expands the concept of empirical risk minimization.

3.2.3 Limitations

There are a number of limitations in this analysis. First, the experiments are set to auto-simulated in a way that lacks a more sophisticated real-world regulatory system, which limits the ecological validity. Second, projection-based enforcement assumes the existence of feasible regions, and such regions are assumed to be computationally manageable. When regions contain highly non-convex or combinatorial constraints, the computational effort needed to resolve these constraints may be very high. Finally, the scalability of the approach in cases where the number of dimensions with high output is extremely large is unclear.

3.2.4 Suggestions

The primary consideration of future study is to analyze the framework in particular operational case studies of high relevance, such as large-scale logistics and energy allocation systems. Furthermore, the use of differentiable projection layers and adaptive constraint weighting methods to improve scalability and efficiency is warranted.

4. CONCLUSION

Concerning the growing demand for reliable AI systems that follow domain-specific constraints in real-world deployment, this research sought to understand the feasibility-preserving prediction problem in operational data science environments. The main issue is that traditional machine learning systems focus on predictive performance and overlook operational admissibility, resulting in outputs that are infeasible and therefore require expensive post-processing or manual correction. In an attempt to close this gap, we have developed an all-encompassing Constraint-Aware Learning Framework (CALF) that for the first time weaves operational constraints into the training objective and reasoning process by means of hybrid penalty-based regularization and projection-based correction. The results show that CALF significantly increases the feasibility rate for all data sets while being on par with predictive performance of (un)constrained baselines & single mechanism variants. The hybrid approach, integrating both soft and hard constraints, has proven to be the ideal solution for maintaining both compliance and flexibility for virtually all data sets.

These results indicate that theory-based learning should be viewed as a core constructive principle for implementable AI processes that function under restricted decision scenarios. The framework proposed for model optimality regulations in the context of practical admissibility provides an important contribution to reconciling theoretical model performance and the conditioned demands of practical model performance. This pertains to scenarios involving the distribution of limited resources, constrained prediction, and systems with operational regulations where non-compliant loutputs are prohibitive.

5. ACKNOWLEDGEMENT

We would like to express our gratitude to all those who have contributed technically and intellectually to this research on operational data science. In particular, the authors appreciate the assistance provided in validating the methodology and interpreting the experimental findings, which strengthen the argument for integrating operational constraints into machine learning.

6. AUTHOR CONTRIBUTION STATEMENT

Author WS contributed to the conceptualization and design of the methodology, implementation, experimental design, data analysis, and writing. TBK and MEK participated in project supervision, critical review, and assisted in the validation and interpretation of findings, as well as contributing to the development of the research framework and quality assurance of the work. All authors approved the final version of this article.

AUTHOR INFORMATION

Corresponding Authors

Wu Shukun, Guangxi Science & Technology Normal University, China
https://orcid.org/0000-0003-0371-4167
Email: wuskctrl@outlook.com

Authors

Tri Basuki Kurniawan, Universitas Bina Darma Palembang, Indonesia
https://orcid.org/0000-0002-3718-0776
Email: tribasukikurniawan@yahoo.com

Muhammet Esad Kuloğlu, Abdullah Gül University, Turkiye
https://orcid.org/0000-0001-5149-5494
Email: mekuloglu@gmail.com

REFFERENCE

Alridha, A. H., & Al-jilawi, A. S. (2022). Solving NP-Hard Problem Using a New Relaxation of Approximate Methods. *International Journal of Health Sciences*, *6*(March), 523–536. https://doi.org/10.53730/ijhs.v6nS3.5375

Bian, K., & Priyadarshi, R. (2024). Machine Learning Optimization Techniques: A Survey, Classification, Challenges, and Future Research Issues. *Archives of Computational Methods in Engineering*, *31*(7), 4209–4233. https://doi.org/10.1007/s11831-024-10110-w

Cabiddu, F., Moi, L., Patriotta, G., & Allen, D. G. (2022). Why do users trust algorithms? A review and conceptualization of initial trust and trust over time. *European Management Journal*, *40*(5), 685–706. https://doi.org/10.1016/j.emj.2022.06.001

Chen, E., Bao, H., & Dinh, N. (2024). Evaluating the reliability of machine-learning-based predictions used in nuclear power plant instrumentation and control systems. *Reliability Engineering & System Safety*, *250*, 110266. https://doi.org/10.1016/j.ress.2024.110266

Chen, Y., Calabrese, R., & Martin-barragan, B. (2024). Interpretable machine learning for imbalanced credit scoring datasets. *European Journal of Operational Research*, *312*(1), 357–372. https://doi.org/10.1016/j.ejor.2023.06.036

Dhatchayani, K., Vezhaventhan, D., Sornamugi, P., T, V., S, P., & T, L. (2025). Agile Decision-Making Framework using Hybrid Statistical and Predictive Models for Efficient Business Operations. *2025 6th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI)*. https://doi.org/10.1109/ICMCSI64620.2025.10883071

Ekeh, A. H., Apeh, C. E., Odionu, C. S., & Austin-gabriel, B. (2025). Automating Legal Compliance and Contract Management : Advances in Data Analytics for Risk Assessment , Regulatory Adherence , and Negotiation Optimization. *Engineering And Technology Journal*, *10*(01), 3684–3703. https://doi.org/10.47191/etj/v10i01.26

Greenshields-watson, A., Vavourakis, O., Spoendlin, F. C., Cagiada, M., & Deane, C. M. (2025). Challenges and compromises : Predicting unbound antibody structures with deep learning. *Current Opinion in Structural Biology*, *90*, 102983. https://doi.org/10.1016/j.sbi.2025.102983

Hooshmand, F., & Nikoomanesh, M. (2025). Integration of machine constraint learning methods within optimization models. *Applied Mathematical Modelling*, *146*(2), 116184. https://doi.org/10.1016/j.apm.2025.116184

Kundu, P., Beura, S., Mondal, S., B, A. K. Das, & Ghosh, A. (2024). Machine learning for the advancement of genome-scale metabolic modeling. *Biotechnology Advances Journal*, *74*, 108400. https://doi.org/10.1016/j.biotechadv.2024.108400

Laurens, D., Verkerken, M., Wauters, T., Turck, F. De, & Volckaert, B. (2023). Investigating Generalized Performance of Data-Constrained Supervised Machine Learning Models on Novel , Related Samples in Intrusion Detection. *Sensors*, *23*(4), 1846. https://doi.org/10.3390/s23041846

Lin, Y., Dong, H., Wang, H., & Zhang, T. (2022). Bayesian Invariant Risk Minimization. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 16000–16009. https://doi.org/10.1109/CVPR52688.2022.01555

Lisboa, P. J. G., Saralajew, S., C, A. V., Fernández-Domenech, R., & Villmann, T. (2023). The coming of age of interpretable and explainable machine learning models. *Neurocomputing*, *535*, 25–39. https://doi.org/10.1016/j.neucom.2023.02.040

Mienye, I. D., & Swart, T. G. (2024). A Comprehensive Review of Deep Learning : Architectures , Recent

Advances , and Applications. *Information*, *15*(12), 755. https://doi.org/10.3390/info15120755

Ming, F., Gong, W., Wang, L., & Jin, Y. (2024). Constrained Multi-Objective Optimization With Deep Reinforcement Learning Assisted Operator Selection. *IEEE/CAA Journal of Automatica Sinica*, *11*(4), 919–931. https://doi.org/10.1109/JAS.2023.123687

Mühlhoff, R. (2023). Predictive privacy : Collective data protection in the context of artificial intelligence and big data. *Big Data & Society*, *10*(1). https://doi.org/10.1177/20539517231166886

Shah, J., & Tyagi, A. D. (2025). Leveraging Data to Improve Operational Efficiency: Case Studies in Healthcare, Transportation, and Logistics. *Journal of Quantum Science and Technology*, *2*(2), 348–358. https://doi.org/10.63345/jqst.v2i2.278

Vaid, A., Sawant, A., Suarez-Farinas, M., Juhee, L., Kaul, S., Kovatch, P., Freeman, R., Jiang, J., Jayaraman, P., Fayad, Z., Argulian, E., Lerakis, S., Charney, A. W., Wang, F., Levin, M., Glicksberg, B., Narula, J., Hofer, I., Singh, K., & NadkarniAkhil, G. N. (2023). Implications of the Use of Artificial Intelligence Predictive Models in Health Care Settings: A Simulation Study. *Annals of Internal Medicine*, *176*(10), 10.7326/M23-094. https://doi.org/10.7326/M23-0949

Veesam, S. B., Rao, B. T., Begum, Z., Patibandla, R. S. M. L., Dcosta, A. A., Bansal, S., Prakash, K., Rashed, M., & Faruque, I. (2025). Multi-camera spatiotemporal deep learning framework for real-time abnormal behavior detection in dense urban environments. *Scientific Reports*, *15*(1), 1–21. https://doi.org/10.1038/s41598-025-12388-7

Wang, X., Yao, Z., & Papaefthymiou, M. (2023). A real-time electrical load forecasting and unsupervised anomaly detection. *Applied Energy Journal*, *330*, 120279. https://doi.org/10.1016/j.apenergy.2022.120279

Xia, C. S., & Zhang, L. (2022). Less Training , More Repairing Please : Revisiting Automated Program Repair via Zero-Shot Learning. *ESEC/FSE 2022: Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, 959–971. https://doi.org/10.1145/3540250.3549101

Yu, Q., Yang, C., Dai, G., Peng, L., & Li, J. (2024). A novel penalty function-based interval constrained multi-objective optimization algorithm for uncertain problems. *Swarm and Evolutionary Computation*, *88*, 101584. https://doi.org/10.1016/j.swevo.2024.101584

Zhang, R., Xu, Q., Yao, J., Zhang, Y., Tian, Q., & Wang, Y. (2023). Federated Domain Generalization with Generalization Adjustment. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 3954–3963. https://doi.org/10.1109/CVPR52729.2023.00385

Zhang, Y., Accordi, G., Gadioli, D., & Palermo, G. (2025). Harnessing quality-throughput trade-off in scoring functions for extreme-scale virtual screening campaigns. *Future Generation Computer Systems*, *172*(April), 107863. https://doi.org/10.1016/j.future.2025.107863

Zhang, Y., Liang, X., Li, D., Ge, S. S., Gao, B., & Chen, H. (2024). Adaptive Safe Reinforcement Learning With Full-State Constraints and Constrained Adaptation for Autonomous Vehicles. *IEEE Transactions on Cybernetics*, *54*(3), 1907–1920. https://doi.org/10.1109/TCYB.2023.3283771