FETNAS: JOINT OPTIMIZATION OF FEATURE SELECTION AND NEURAL ARCHITECTURE SEARCH

Anonymous authors
Paper under double-blind review

ABSTRACT

Deep learning is increasingly deployed in resource-constrained settings such as IoT and industrial sensing, where memory and compute are limited. In these settings, input features often correspond to physical sensors that are costly to install and read. Because feature relevance is rarely known a priori and datasets often contain irrelevant or redundant variables, an effective model must select as few input features as necessary while also fitting a constrained memory footprint. Existing methods address only one side of this problem: Neural Architecture Search (NAS) treats the input features as fixed, whereas embedded feature selection assumes a fixed architecture, leaving the joint search space largely unexplored. We address this gap with FETNAS, a weight-sharing NAS method that treats the number of input features as an elastic dimension alongside network depth and width. FETNAS trains a single over-parameterized supernet using progressive shrinking, where each subnetwork uses the top-ranked features according to a gradient-based importance ranking rather than a random ordering. Once the supernet is trained, subnetworks are evaluated without retraining and a multi-objective evolutionary search identifies the Pareto front over predictive performance, parameter count and input dimensionality. Across the TabArena benchmark, FETNAS achieves a better performance-complexity trade-off than the baselines for classification. Up to the third quartile of model size, it typically uses 2–4$\times$ fewer parameters and 1.8–2.7$\times$ fewer input features. For regression, gains are more modest and budget-dependent. FETNAS improves performance at moderate feature budgets but retains a larger share of the features than for classification. Beyond reducing model complexity, the explicit feature subset lowers data-acquisition cost and eases edge deployment.

1 INTRODUCTION

Deep learning models are increasingly deployed in resource-constrained environments, where inference must operate under tight memory and computational budgets. These challenges are particularly pronounced at the edge, where devices often have limited resources. In many edge applications, especially in IoT and industrial sensing, an additional constraint arises: input features correspond to physical sensors whose acquisition, transmission and processing incur a direct cost. Consequently, an effective model should not only achieve high predictive accuracy but also use fewer input features and operate within limited computational resources. Despite the practical importance of both objectives, they are rarely considered jointly in standard model design.

In many sensing applications, measurements are collected as tabular records, where each sensor contributes one or more features. A deployment may therefore involve many candidate variables whose relevance to the target is not known a priori. Irrelevant or redundant features can substantially degrade predictive performance (Liu and Motoda, 2012), especially for compact models, making the identification of an informative feature subset valuable. When a retained feature corresponds to a sensor that must be installed, maintained and read, reducing the number of features becomes even more important. In practice, feature selection is typically performed as a separate preprocessing step, after which a model is designed for the selected features.

In contrast, embedded approaches integrate feature selection directly into model training. TabNet learns instance-wise feature masks through sequential attention (Arik and Pfister, 2021), while other methods rely on stochastic gates (Yamada et al., 2020) or sparsity-inducing regularization (Lemhadri

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

et al., 2021). Although these methods learn feature relevance end-to-end, they rely on a fixed architecture that cannot adapt to constrained hardware. Instance-wise approaches such as TabNet also still require every feature at inference time, even if some contribute little to the prediction.

In parallel, Neural Architecture Search (NAS) has emerged as an effective approach for automating network design, enabling the discovery of models that match or outperform manually designed architectures while satisfying strict latency and resource constraints (Cai et al., 2019; Wu et al., 2019; Wan et al., 2020; De Ruyter et al., 2026; Robinson et al., 2026). However, most NAS methods assume a fixed input representation, including those developed for tabular data (Yang et al., 2022; Egele et al., 2021; Xing et al., 2026). As a result, relatively little attention has been paid to jointly optimizing feature selection and network architecture (Hafiz et al., 2023). This is an important limitation, as feature selection and architecture design are typically performed as separate stages despite being closely coupled. Optimizing them independently can lead to feature subsets and architectures that perform well in isolation but are suboptimal when considered together.

Bridging this gap, we propose FETNAS, a method that jointly optimizes feature selection and architecture design within a single weight-sharing NAS framework. We construct an over-parameterized supernet whose subnetworks vary in the number of input features, network depth and layer width, where each subnetwork uses the k most important features according to a gradient-based ranking. To train this supernet, we extend progressive shrinking with an elastic-features stage, in which subnetworks with different input dimensionalities share weights. The importance-based ranking keeps these subnetworks consistent: smaller subnetworks always retain the most informative inputs, allowing the shared weights to receive coherent training signals across feature budgets. Once trained, the supernet is searched using a multi-objective evolutionary algorithm over predictive performance, parameter count and number of input features, enabling the selection of a model for any feature or memory budget without retraining and without training a separate supernet for each budget.

Our contributions are as follows:

- **Elastic input dimensionality.** We introduce FETNAS, a weight-sharing NAS method that treats the number of input features as a searchable dimension alongside network depth and width. A single supernet therefore yields models for a wide range of feature and parameter budgets without retraining.
- **Ranked feature sampling for weight sharing.** We show that effective weight sharing across input dimensionalities requires nested feature subsets ordered by feature importance. We demonstrate that sampling random feature subsets substantially degrades performance during supernet training.
- **An end-to-end, sweep-free pipeline.** FETNAS removes the need for a separate feature-selection stage and its threshold or retention sweeps. A single supernet spans all feature counts and the desired feature budget can be selected after training from the Pareto front.

2 Related Work

2.1 Feature Selection

Feature selection identifies an informative subset of the original features, improving generalization and interpretability and mitigating the curse of dimensionality in high-dimensional settings. Early approaches relied on domain experts to manually prune irrelevant features (Guyon and Elisseeff, 2003), whereas modern approaches are predominantly data-driven and fall into *filter* and *wrapper* methods.

Filter methods rank features using predefined statistical criteria, independently of any downstream learning algorithm (Jović et al., 2015). Univariate filters, such as Pearson correlation, are computationally efficient but ignore redundancy and interactions between features. In contrast, multivariate filters model feature dependencies jointly, typically at a higher computational cost. Wrapper methods evaluate feature subsets based on the performance of a predictive model, relying on heuristic search because exhaustive search is generally infeasible (Jović et al., 2015; Chandrashekar and Sahin, 2014). Although they can capture complex feature interactions, they are computationally expensive and are therefore typically paired with fast-to-evaluate models. AutoML frameworks such as auto-sklearn and TPOT reflect this trade-off: they automate feature selection but favor computationally inexpensive

models during search (Feurer et al., 2015; Olson et al., 2016), so multilayer perceptrons (MLPs) are rarely included by default. AutoGluon includes MLPs, but only as components of ensemble models (Erickson et al., 2020).

More recently, some deep learning architectures integrate feature selection directly into training, using sequential attention (Arik and Pfister, 2021), stochastic gates (Yamada et al., 2020) or differentiable ℓ_1-based relaxations (Lemhadri et al., 2021). However, these embedded methods operate within a fixed, hand-designed architecture.

2.2 Neural Architecture Search

Before automated model design, neural architectures were built from hand-crafted heuristics and rules of thumb (Huang, 2003; Ripley, 1996). These are easy to apply but often yield suboptimal architectures that underuse model capacity for the task at hand.

NAS automates this process and has matched or outperformed manually designed architectures across benchmarks (Chen et al., 2018; Ghiasi et al., 2019), in both single-task and multi-task settings (Gao et al., 2020; Danneels and Verbeke, 2026). Research has since shifted toward efficient models with low inference latency and resource usage, suited to deployment on constrained devices (Cai et al., 2019; De Ruyter et al., 2026; Robinson et al., 2026; Wan et al., 2020; Wu et al., 2019). However, these methods treat the input as fixed and optimize only the architecture.

Early NAS methods relied on computationally expensive reinforcement learning or evolutionary algorithms, requiring many candidate architectures to be trained from scratch. Weight-sharing approaches substantially reduce this cost. DARTS (Liu et al., 2019) relaxes the discrete search space into a continuous one that can be optimized using gradient-based methods, while later work trains a supernet whose sampled subnetworks inherit shared weights (Bender et al., 2018; Yu et al., 2019; Tan and Le, 2019; Wang et al., 2021). This enables many candidate architectures to be evaluated without retraining, making larger search spaces tractable.

2.3 Joint Optimization of Features and Architecture

Feature selection and neural architecture design are traditionally treated as separate problems: informative features are selected first and the architecture is optimized afterward. This sequential approach can be suboptimal, since feature relevance and architectural design are fundamentally connected.

Few studies address the joint optimization of feature selection and architecture design, with most work focusing on specific application domains. Hafiz et al. (2023) co-evolve feature subsets and architectures in a multi-objective framework for stock market forecasting, but must train every candidate from scratch, which is costly. NAS has also been applied directly to tabular data over fully connected or tree-inspired architectures (Yang et al., 2022; Xing et al., 2026; Egele et al., 2021), but these methods do not include feature selection in the search space. To the best of our knowledge, FETNAS is the first method to address this gap by making the number of input features an elastic dimension of a weight-sharing supernet.

3 Methodology

FETNAS proceeds in two phases. In the supernet training phase, an MLP supernet is trained in four stages that progressively enable variable depth, variable width and variable input dimensionality, allowing any subnetwork to be evaluated using inherited weights. In the search phase, an evolutionary algorithm samples subnetworks, each defined by a feature count k, a depth and per-layer widths, and returns a Pareto front over validation performance, parameter count and number of input features.

3.1 Problem Definition

Let $\mathcal{F} = \{f_1, \ldots, f_d\}$ be the set of d input features and $\mathcal{S}$ the space of architectures. For an architecture $A \in \mathcal{S}$ and a feature subset $U \subseteq \mathcal{F}$, let $w^*(A, U)$ denote the weights obtained by training the model on the training set, let P_{val} denote its validation performance (F1-macro for classification, R^2 for regression) and let $C(A, U)$ denote its number of parameters, which depends

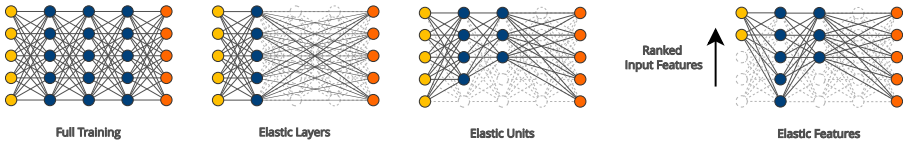


Figure 1: **FETNAS supernet training stages.** Supernet training is divided into successive stages in which additional degrees of freedom in the subnetwork space are gradually unlocked. Starting from the full supernet trained with all input features, the training procedure subsequently enables a variable number of layers, number of hidden units per layer and feature subsets. This allows the shared weights to adapt smoothly to an increasingly diverse set of subnetworks.

on both the architecture and the number of input features. We seek the Pareto-optimal pairs of the multi-objective problem

$$\min_{A\in\mathcal{S},\, U\subseteq\mathcal{F}} \Big(-P_{\text{val}}\big(w^*(A,U)\big),\ C(A,U),\ |U|\Big), \qquad w^*(A,U) = \arg\min_{w} \mathcal{L}_{\text{train}}(w; A, U), \quad (1)$$

from which a model can be selected for any given feature or parameter budget.

Solving Equation 1 exactly requires training every candidate pair. Weight sharing has emerged as the de facto standard in NAS (Liu et al., 2019; Cai et al., 2020), as it removes the computationally prohibitive requirement of retraining every sampled architecture from scratch.

We approximate $w^*(A,U)$ by the weights that (A,U) inherits from a jointly trained supernet and restrict U to the family $\{U_k\}_{k=k_{\min}}^{d}$ of nested subsets, where U_k consists of the k highest-ranked features (Section 3.4) and $k_{\min}$ is the minimum number of input features a subnetwork may use. The search over U thus reduces to a search over the feature count $k = |U|$.

3.2 Search Space

In this work, the supernet is a fully connected MLP with at most $L_{\max}$ hidden layers of at most $N_{\max}$ units, taking all d features as input, followed by a task-specific output head. A subnetwork is specified by $(k, \ell, n_1, \ldots, n_\ell)$ with $k \le d$ input features, $\ell \le L_{\max}$ layers and $n_j \le N_{\max}$ units in layer j. It uses the first k input columns of the first weight matrix (in ranked order), the first ℓ layers and the first n_j units of each layer, thereby preserving weight sharing while preventing interference between subnetworks (Cai et al., 2020). The output head consumes the first n_ℓ units of the last active layer and is not subject to architectural modifications.

Following several recent approaches (Cai et al., 2020; Tan and Le, 2019; Wang et al., 2021; Yu et al., 2019), we decouple model training and architecture search into two distinct phases. In the first phase, the supernet weights are trained by sampling subnetworks that are individually optimized for the prediction task. In the second phase, architectures are sampled using an evolutionary algorithm that balances predictive performance, model complexity and the number of input features.

3.3 Supernet Training

Training is divided into four stages that progressively enlarge the set of subnetworks the shared weights must support (Figure 1). The ordering is not arbitrary: depth and width are optimized first because they do not affect the input. In contrast, varying the input dimensionality depends on the feature ranking introduced in Section 3.4, which is computed from the supernet after the depth and width stages. Across all stages, subnetworks are constructed in a nested manner to maximize weight sharing and avoid interference between subnetworks.

- **Full Training**. First, the full MLP supernet $\mathcal{S}$ is trained using all available input features $\mathcal{F}$. This initial stage ensures that all supernet weights are, to some extent, optimized for the task and prevents subnetworks sampled in later phases from suffering from a cold start due to untrained parameters.
- **Elastic Layers**. In this phase, the weights are optimized for a variable number of layers by dynamically sampling shallower subnetworks. These subnetworks are always constructed starting from the initial layers of the model, ensuring that the earliest hidden representations are consistently shared.

- **Elastic Units**. Subsequently, the weights are trained to support narrower architectures by allowing each layer to activate only a subset of its hidden units, with the number of active units per layer independently sampled for each subnetwork. As with layer selection, the active units correspond to the first n_j units of the layer, forming a contiguous block rather than an arbitrary subset of units.
- **Elastic Features**. Finally, the supernet is trained with subnetworks that use only the top-k ranked features, with k sampled per subnetwork from $\{k_{\min}, \ldots, d\}$.

Each elastic stage uses sandwich sampling (Yu et al., 2019): per batch, the largest subnetwork, the smallest subnetwork ($k_{\min}$ features, $\ell_{\min}$ layers, $n_{\min}$ units) and m random intermediate subnetworks are sampled from the supernet and their shared weights are updated sequentially on the same batch. The largest subnetwork is trained on $\mathcal{L}_{\text{task}}$, whereas the remaining subnetworks are trained on $\mathcal{L}_{\text{task}}$ combined with in-place knowledge distillation from the largest subnetwork.

3.4 Feature Sampling

Although the sampling of input features during supernet training may seem straightforward, it strongly influences how effectively weights can be shared across subnetworks operating on different feature subsets. If feature subsets are sampled uniformly at random, subnetworks may be trained on substantially different inputs, causing the shared weights to receive inconsistent optimization signals. Consequently, the supernet must accommodate a wide range of feature configurations, which can degrade the quality of the shared weights and ultimately reduce the performance of sampled architectures. To mitigate this issue, we introduce an importance-based feature ordering strategy that serves as a heuristic for feature informativeness and promotes more consistent training signals across subnetworks.

Feature importance is estimated once, after the elastic-units stage, based on the sensitivity of the trained supernet to each input feature. The goal is to ensure that smaller feature subsets retain the most important inputs, while less important features are removed first. Since lower-ranked features contribute relatively little to the learned representations, removing them is less likely to alter the supernet's behavior. This, in turn, facilitates weight sharing among subnetworks with different numbers of input features and helps preserve predictive performance across the search space.

Specifically, we rank features by an importance score s_i and define U_k as the set of the k highest-ranked features, where $k_{\min} \leq k \leq d$. All subnetworks therefore share the most informative inputs and differ only in the number of lower-ranked features they discard. During the elastic-features stage, k is sampled uniformly from $\{k_{\min}, \ldots, d\}$ for each subnetwork. We compute the scores once from the trained supernet as the expected absolute gradient over the training set:

$$s_i = \mathbb{E}_{(x,y)\sim\mathcal{D}}\left[\, \left| \frac{\partial \mathcal{L}_{\text{task}}(x,y)}{\partial x_i} \right| \right], \tag{2}$$

We also evaluate other scoring methods in Section 4.5, such as the input-gradient product $\mathbb{E}[\,|x_i\, \partial \mathcal{L}_{\text{task}}/\partial x_i|\,]$, the ℓ_2 norm of the first-layer weights of feature i and random ranking.

4 Experimental Setup

4.1 Datasets

To evaluate FETNAS across a wide range of input dimensionalities, we use all 51 datasets from TabArena (Erickson et al., 2025). The datasets range from 4 to 1,776 input features and from 748 to 150,000 instances, covering 38 classification and 13 regression tasks from domains such as finance, healthcare and industrial process monitoring. This diversity allows us to study how the jointly selected feature count and architecture adapt to both low- and high-dimensional problems. Per-dataset characteristics are listed in Appendix A. All experiments follow the training procedure described in Appendix C.

4.2 BASELINES

We compare against six baselines (Table 1). Four of them are sequential pipelines that combine one of two feature-selection methods with one of two architecture-design methods. The remaining two apply the two architecture-design steps to the full feature set.

Feature-selection steps. *Univariate filter (UF).* Each feature is scored independently using the absolute Pearson correlation with the target for regression and mutual information (MI) for classification. A feature is retained if its score is at least $\tau = 0.4$, applied directly to the correlation for regression ($|\rho(x_i, y)| \geq \tau$) and relative to the maximum MI score for classification ($I(x_i, y) \geq \tau \cdot \max_j I(x_j, y)$). For regression, this threshold corresponds to at least a moderate Pearson correlation strength (Schober et al., 2018), selecting features that are moderately to strongly correlated with the target. The filter is computationally inexpensive but does not account for feature redundancy or interactions. *Minimum Redundancy Maximum Relevance (mRMR).* Features are selected greedily to maximize relevance to the target while minimizing redundancy among the selected features (Peng et al., 2005). Because MRMR requires the number of retained features to be specified in advance, we evaluate retention fractions of $f \in \{20, 40, 60, 80\}\%$.

Architecture-design steps. *Huang's rule of thumb (RoT).* This handcrafted baseline builds a two-hidden-layer network whose widths depend on the number of output units and the number of training samples (Huang, 2003; see Appendix B). This represents a fully heuristic pipeline. *Evolutionary NAS (Evo).* EVO uses the same supernet, elastic training procedure and evolutionary search as FETNAS, but with a fixed feature subset, i.e., without the elastic-features stage.

Resulting pipelines. ROT-UF and ROT-MRMR combine each feature-selection method with the heuristic architecture, while ROT-ALL uses the full feature set. EVO-UF and EVO-MRMR combine them with architecture search and test whether a separate feature-selection stage adds value on top of NAS. EVO-ALL applies architecture search to the full feature set and differs from FETNAS only in the absence of the searchable feature dimension.

4.3 SEARCH SPACE

All methods, except Huang's rule of thumb, share the same architectural search space to ensure a fair comparison: fully connected networks with 1 to 10 hidden layers, where the width of each layer is sampled between 8 and 256 neurons. This search space covers both shallow, narrow models and deep, high-capacity architectures. FETNAS additionally treats the number of input features, $k \in \{k_{\min}, \ldots, d\}$, with $k_{\min} = 1$, as a searchable dimension (Section 3.4), whereas all baselines fix the feature subset before architecture selection (using mRMR, UF or the full feature set). After training the supernet, we run a multi-objective evolutionary search over all candidate architectures using three objectives: the validation performance of each subnetwork with inherited weights, the number of parameters and the number of input features. This produces a Pareto front from which a model can be selected for any desired complexity budget without requiring additional supernet training.

Table 1: **Baselines**, decomposed into their feature-selection and architecture-design steps. For mRMR, one model or supernet is trained for each retention fraction f. The univariate filter retains features with absolute Pearson correlation $|\rho| \geq 0.4$ for regression (Schober et al., 2018) or mutual information at least 0.4 times the maximum across all features (classification).

Method	Feature-selection step	Architecture-design step
RoT-All	none (all features)	Huang's rule of thumb
RoT-UF	univariate filter ($\lvert\rho\rvert \geq 0.4$ / $I \geq 0.4 \cdot \max I$)	Huang's rule of thumb
RoT-mRMR	mRMR, top-f%, $f \in \{20, 40, 60, 80\}$	Huang's rule of thumb
Evo-All	none (all features)	supernet + evolutionary search
Evo-UF	univariate filter ($\lvert\rho\rvert \geq 0.4$ / $I \geq 0.4 \cdot \max I$)	supernet + evolutionary search
Evo-mRMR	mRMR, top-f%, $f \in \{20, 40, 60, 80\}$	supernet + evolutionary search
FetNAS (ours)	*jointly searched feature subset U*	*supernet + evolutionary search*

4.4 Performance at Matched Feature Budgets

We compare against baselines using *zero-shot* results, i.e., subnetworks sampled directly from the trained supernet without any retraining. To compare FETNAS against the baselines, we match methods by feature cost. Each baseline, at its fraction of selected features (reported as k/d)[1], defines a Pareto front over parameter count and *zero-shot* validation performance. We take the parameter counts at five positions along each baseline front: the minimum, first quartile (Q1), median (Q2), third quartile (Q3) and maximum. For each position, we report the FETNAS subnetwork with the best validation performance within that parameter budget and using at most the baseline's number of features. The reported k/d for FETNAS is the mean over these five subnetworks. All selected models are evaluated on the test set (Tables 2 and 3).

For classification, FETNAS achieves a better performance–complexity trade-off than EVO with all features and with MRMR. Up to Q3, it improves F1-macro by 0.01—0.03 while using 1.8—$2.7\times$ fewer features and typically 2–4$\times$ fewer parameters. Against EVO-UF, FETNAS matches performance with fewer features and parameters. At the largest model sizes, EVO closes the gap but requires roughly $3\times$ more parameters. ROT reaches the same peak F1 as the searched models. However, FETNAS matches it with about $3\times$ fewer parameters and stays within 0.02–0.03 with 14–$41\times$ fewer parameters under all-feature and MRMR budgets.

For regression, the trade-off is less pronounced. Under moderate feature budgets (MRMR with $f = 40$ and 60), FETNAS improves upon EVO by up to 0.04 across all model sizes while using fewer parameters. Under stricter feature budgets (MRMR with $f = 20$ and UF), and with all features at larger model sizes, EVO often performs slightly better. ROT is also a stronger method in this setting, with most searched models matching its performance only at comparable or greater complexity.

4.5 Alternative Importance Ranking Metrics

We further evaluate two alternative feature importance metrics alongside a random sampling baseline to assess the benefit of ranking-based feature selection. Specifically, we consider the following variants:

$$\begin{aligned} &\text{(G)} \quad s_i = \mathbb{E}_{(x,y)\sim\mathcal{D}}\Big[\Big|\tfrac{\partial \mathcal{L}_{\text{task}}(x,y)}{\partial x_i}\Big|\Big], \quad \text{(GIN)} \quad s_i = \mathbb{E}_{(x,y)\sim\mathcal{D}}\Big[\Big|x_i\,\tfrac{\partial \mathcal{L}_{\text{task}}(x,y)}{\partial x_i}\Big|\Big], \\ &\text{(WN)} \quad s_i = \|W^{(1)}_{:,i}\|_2, \qquad\qquad \text{(R)} \quad s_i \sim \text{Uniform}(0,1). \end{aligned} \tag{3}$$

To compare these feature-sampling strategies, we construct the three-dimensional Pareto fronts defined by test performance, model complexity (measured by the number of parameters) and the number of selected features using the *zero-shot* evaluations. As this analysis compares ranking strategies rather than selecting a final model, we construct the fronts directly on test performance. We quantify the quality of each front by computing its dominated hypervolume. To ensure a fair comparison, the hypervolume is evaluated over a common range spanning from the first Pareto-optimal point to the furthest point reached by any of the four variants. This avoids penalizing shorter but higher-performing fronts relative to longer but lower-performing ones. The resulting hypervolume values are used to rank the variants for each dataset and to construct critical difference (CD) diagrams. In the CD diagrams, bold horizontal lines connect methods whose mean ranks are not statistically different according to the Friedman-Nemenyi test.

Based on the *zero-shot* results, FETNAS_R achieves the worst average rank, with its critical-difference interval not overlapping with that of any other method. This suggests that randomly selecting features during supernet training provides a poor training signal, limiting the supernet's ability to learn reliable representations across candidate architectures. Among the importance-based approaches, gradient-based methods outperform the weight-norm alternative on average, although weight-norm ranks higher on several individual datasets (Appendix E). A possible explanation is that gradient-based importance scores directly capture the sensitivity of the prediction with respect to each feature, thereby providing a more faithful estimate of feature influence. In contrast, the weight norm reflects only the magnitude of the parameters associated with a feature, which does not necessarily translate into a strong effect on the model output due to interactions across layers, activation functions and information propagation throughout the network.

[1]For MRMR, k/d may exceed the requested fraction f because $f \times d$ is rounded up when it is not an integer.

Table 2: **Classification results at matched feature budgets.** For each feature budget, the columns correspond to the parameter counts at the minimum, Q1, median (Q2), Q3 and maximum of the baseline's (Evo-*) Pareto front. For FETNAS, we report the subnetwork with the best validation performance within that parameter budget, using at most the baseline's number of features. Each cell gives mean ± std F1-macro across the 38 classification datasets (top) and mean ± std parameter count (bottom, in parentheses). Evo-* and FETNAS are evaluated *zero-shot* with inherited supernet weights. RoT is trained from scratch and yields a single architecture. Underlined values mark the better of Evo-* and FETNAS in each column.

			F1-macro (parameter budget in parentheses)				
Feature budget	Method	k/d	min	Q1	Q2	Q3	max
all features, $k/d = 1$	Evo-All	1.00	0.59 ± 0.01 (1.1K ± 29)	0.61 ± 0.01 (4.8K ± 3.0K)	0.61 ± 0.01 (5.7K ± 3.1K)	0.62 ± 0.01 (8.0K ± 4.4K)	0.66 ± 0.01 (54.7K ± 5.7K)
	FetNAS	0.37	0.61 ± 0.01 (413 ± 105)	0.62 ± 0.01 (1.5K ± 1.5K)	0.63 ± 0.01 (2.1K ± 2.2K)	0.63 ± 0.01 (3.0K ± 2.3K)	0.66 ± 0.01 (15.0K ± 3.1K)
	RoT-All	1.00					0.66 ± 0.01 (46.1K ± 0)
mRMR, $f = 80$	Evo-mRMR$_{80}$	0.85	0.58 ± 0.01 (896 ± 24)	0.61 ± 0.01 (3.5K ± 1.4K)	0.61 ± 0.01 (4.2K ± 1.3K)	0.62 ± 0.01 (7.8K ± 2.2K)	0.66 ± 0.01 (48.8K ± 1.2K)
	FetNAS	0.35	0.60 ± 0.01 (368 ± 95)	0.62 ± 0.01 (1.4K ± 1.4K)	0.63 ± 0.00 (1.8K ± 1.3K)	0.63 ± 0.01 (3.0K ± 1.7K)	0.66 ± 0.01 (15.3K ± 2.5K)
	RoT-mRMR$_{80}$	0.85					0.66 ± 0.00 (42.3K ± 24)
mRMR, $f = 60$	Evo-mRMR$_{60}$	0.67	0.58 ± 0.01 (716 ± 8)	0.59 ± 0.01 (3.3K ± 3.0K)	0.60 ± 0.01 (3.8K ± 3.0K)	0.61 ± 0.01 (5.3K ± 2.8K)	0.65 ± 0.01 (42.9K ± 4.3K)
	FetNAS	0.30	0.60 ± 0.01 (320 ± 74)	0.62 ± 0.01 (1.4K ± 1.6K)	0.63 ± 0.01 (1.6K ± 1.6K)	0.62 ± 0.01 (2.5K ± 1.4K)	0.65 ± 0.01 (15.1K ± 1.7K)
	RoT-mRMR$_{60}$	0.67					0.65 ± 0.01 (38.4K ± 42)
mRMR, $f = 40$	Evo-mRMR$_{40}$	0.48	0.58 ± 0.01 (566 ± 28)	0.60 ± 0.01 (1.6K ± 287)	0.60 ± 0.01 (2.1K ± 319)	0.61 ± 0.01 (4.5K ± 922)	0.65 ± 0.00 (41.4K ± 4.0K)
	FetNAS	0.25	0.59 ± 0.01 (292 ± 70)	0.62 ± 0.00 (676 ± 313)	0.62 ± 0.01 (834 ± 349)	0.63 ± 0.01 (1.4K ± 408)	0.65 ± 0.01 (13.7K ± 1.8K)
	RoT-mRMR$_{40}$	0.48					0.65 ± 0.01 (34.4K ± 57)
mRMR, $f = 20$	Evo-mRMR$_{20}$	0.28	0.56 ± 0.01 (325 ± 4)	0.58 ± 0.01 (1.7K ± 1.1K)	0.59 ± 0.01 (2.0K ± 1.3K)	0.59 ± 0.01 (2.7K ± 1.5K)	0.64 ± 0.01 (38.3K ± 5.1K)
	FetNAS	0.16	0.58 ± 0.02 (126 ± 70)	0.60 ± 0.02 (450 ± 182)	0.60 ± 0.02 (528 ± 172)	0.61 ± 0.01 (716 ± 89)	0.63 ± 0.01 (11.1K ± 3.4K)
	RoT-mRMR$_{20}$	0.28					0.63 ± 0.00 (29.5K ± 45)
UF, $I \geq 0.4 \cdot \max I$	Evo-UF	0.28	0.57 ± 0.01 (152 ± 13)	0.58 ± 0.01 (1.3K ± 966)	0.59 ± 0.02 (1.4K ± 964)	0.59 ± 0.02 (2.9K ± 968)	0.65 ± 0.01 (30.2K ± 5.8K)
	FetNAS	0.16	0.57 ± 0.01 (75 ± 11)	0.59 ± 0.01 (488 ± 454)	0.59 ± 0.01 (549 ± 443)	0.59 ± 0.01 (969 ± 442)	0.63 ± 0.01 (8.7K ± 376)
	RoT-UF	0.28					0.64 ± 0.00 (26.6K ± 139)

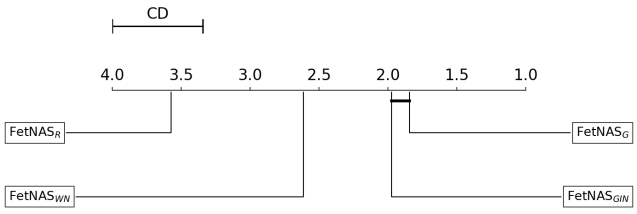


Figure 2: **Critical difference (CD) diagram comparing the mean ranks of the FETNAS variants on the *zero-shot* test-set results across all 51 datasets.** Methods connected by a bold horizontal line are not statistically significantly different (Friedman-Nemenyi test, $\alpha = 0.05$). Gradient-based rankings serve as a better heuristic for feature ranking during supernet training than random and weight-norm rankings.

5 CONCLUSION

We introduced FETNAS, a weight-sharing NAS method that treats the number of input features as an elastic search dimension alongside network depth and width. A single supernet is trained once

Table 3: **Regression results at matched feature budgets.** Same protocol as Table 2, reporting mean $\pm$ std R^2. UF rows average over the 9 datasets on which UF retains at least one feature. (UF did not retain features for *airfoil_self_noise*, *food_delivery_time*, *physiochemical_protein* and *qsar_tid_11*.)

			R^2 (parameter budget in parentheses)				
Feature budget	Method	k/d	min	Q1	Q2	Q3	max
all features, $k/d = 1$	Evo-All	1.00	0.61 ± 0.02 (841 ± 86)	0.65 ± 0.01 (4.3K ± 1.3K)	0.67 ± 0.01 (9.3K ± 2.3K)	0.69 ± 0.00 (25.6K ± 6.4K)	0.70 ± 0.00 (93.3K ± 16.3K)
	FetNAS	0.70	0.61 ± 0.01 (583 ± 123)	0.64 ± 0.01 (3.0K ± 494)	0.65 ± 0.02 (7.9K ± 2.1K)	0.68 ± 0.01 (18.1K ± 2.4K)	0.68 ± 0.01 (51.7K ± 13.2K)
	RoT-All	1.00					0.72 ± 0.00 (31.7K ± 0)
mRMR, $f = 80$	Evo-mRMR$_{80}$	0.84	0.60 ± 0.01 (642 ± 40)	0.63 ± 0.01 (4.6K ± 724)	0.66 ± 0.02 (9.1K ± 1.6K)	0.68 ± 0.02 (25.1K ± 3.0K)	0.69 ± 0.01 (96.1K ± 16.5K)
	FetNAS	0.66	0.61 ± 0.01 (533 ± 75)	0.64 ± 0.00 (3.5K ± 642)	0.65 ± 0.02 (7.4K ± 1.6K)	0.67 ± 0.01 (19.2K ± 4.1K)	0.68 ± 0.01 (47.6K ± 10.0K)
	RoT-mRMR$_{80}$	0.84					0.71 ± 0.01 (28.6K ± 0)
mRMR, $f = 60$	Evo-mRMR$_{60}$	0.64	0.57 ± 0.01 (479 ± 22)	0.59 ± 0.01 (4.1K ± 1.5K)	0.62 ± 0.01 (8.8K ± 2.4K)	0.64 ± 0.01 (21.9K ± 4.8K)	0.65 ± 0.01 (78.6K ± 9.3K)
	FetNAS	0.56	0.58 ± 0.02 (412 ± 67)	0.63 ± 0.01 (3.0K ± 1.5K)	0.64 ± 0.02 (6.8K ± 1.9K)	0.66 ± 0.01 (16.5K ± 2.3K)	0.67 ± 0.01 (42.4K ± 6.8K)
	RoT-mRMR$_{60}$	0.64					0.66 ± 0.01 (25.3K ± 0)
mRMR, $f = 40$	Evo-mRMR$_{40}$	0.45	0.54 ± 0.02 (339 ± 22)	0.57 ± 0.02 (3.3K ± 623)	0.59 ± 0.01 (8.6K ± 1.9K)	0.60 ± 0.01 (25.0K ± 5.2K)	0.61 ± 0.01 (75.7K ± 12.1K)
	FetNAS	0.41	0.56 ± 0.02 (288 ± 17)	0.59 ± 0.01 (2.8K ± 234)	0.61 ± 0.01 (6.1K ± 769)	0.62 ± 0.01 (16.1K ± 3.8K)	0.63 ± 0.01 (35.0K ± 5.2K)
	RoT-mRMR$_{40}$	0.45					0.63 ± 0.02 (22.0K ± 0)
mRMR, $f = 20$	Evo-mRMR$_{20}$	0.24	0.48 ± 0.01 (177 ± 10)	0.50 ± 0.00 (2.0K ± 819)	0.51 ± 0.01 (5.8K ± 2.7K)	0.52 ± 0.01 (15.6K ± 5.4K)	0.52 ± 0.01 (58.4K ± 10.3K)
	FetNAS	0.23	0.46 ± 0.01 (143 ± 10)	0.49 ± 0.01 (1.1K ± 684)	0.51 ± 0.01 (3.4K ± 1.8K)	0.51 ± 0.01 (9.9K ± 2.6K)	0.52 ± 0.01 (29.9K ± 4.4K)
	RoT-mRMR$_{20}$	0.24					0.53 ± 0.01 (18.8K ± 0)
UF, $\lvert\rho\rvert \geq 0.4$	Evo-UF	0.24	0.53 ± 0.04 (66 ± 5)	0.54 ± 0.03 (830 ± 153)	0.53 ± 0.05 (3.4K ± 1.1K)	0.55 ± 0.02 (13.7K ± 1.9K)	0.56 ± 0.02 (55.3K ± 20.7K)
	FetNAS	0.23	0.51 ± 0.04 (58 ± 4)	0.52 ± 0.02 (526 ± 69)	0.54 ± 0.01 (2.4K ± 853)	0.55 ± 0.02 (7.6K ± 2.9K)	0.55 ± 0.01 (22.9K ± 6.5K)
	RoT-UF	0.24					0.54 ± 0.04 (13.9K ± 32)

and searched via a multi-objective evolutionary algorithm, yielding a Pareto front of subnetworks that can be deployed for any feature or parameter budget. Across all TabArena datasets, FETNAS improves the performance-complexity trade-off for classification up to the third quartile of model size, typically using 2–4$\times$ fewer parameters and 1.8–2.7$\times$ fewer input features than other NAS pipelines. For regression, gains are more modest and budget-dependent. We further showed that gradient-based feature ranking is essential for effective weight sharing across input dimensionalities, as random feature sampling substantially degrades the supernet. Overall, our results suggest that feature dimensionality should not be selected separately from the neural architecture. Instead, both can be optimized within the same search procedure, particularly in settings where feature acquisition and model size are constrained.

Limitations and Future Work. FETNAS is intended as a building block for resource-aware model design in sensor-driven applications rather than a general-purpose AutoML replacement. The current search space is limited to MLPs, and the feature ranking is computed once and kept fixed during the elastic-features stage, meaning that interactions between features are only captured implicitly. In addition, all subnetworks are evaluated *zero-shot* and the gap with fine-tuned or retrained models remains to be quantified. Future work includes extending the elastic-features dimension to other architectures, investigating other feature ranking methods and accounting for per-sensor acquisition costs rather than feature counts.

REFERENCES

Arik, S. and Pfister, T. (2021). TabNet: Attentive interpretable tabular learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35:6679–6687.

Bender, G., Kindermans, P.-J., Zoph, B., Vasudevan, V., and Le, Q. (2018). Understanding and simplifying one-shot architecture search. In *Proceedings of the 35th International Conference on Machine Learning (ICML)*, pages 550–559. PMLR.

Cai, H., Gan, C., Wang, T., Zhang, Z., and Han, S. (2020). Once-for-all: Train one network and specialize it for efficient deployment. In *International Conference on Learning Representations*.

Cai, H., Zhu, L., and Han, S. (2019). ProxylessNAS: Direct neural architecture search on target task and hardware. In *International Conference on Learning Representations*.

Chandrashekar, G. and Sahin, F. (2014). A survey on feature selection methods. *Computers & Electrical Engineering*, 40(1):16–28.

Chen, L., Collins, M. D., Zhu, Y., Papandreou, G., Zoph, B., Schroff, F., Adam, H., and Shlens, J. (2018). Searching for efficient multi-scale architectures for dense image prediction. *CoRR*, abs/1809.04184.

Danneels, Q. and Verbeke, M. (2026). MTL-SIMNAS: Task similarity-driven neural architecture search for enhanced multi-task learning. In *Artificial Neural Networks and Machine Learning – ICANN 2025*, pages 272–284, Cham. Springer Nature Switzerland.

De Ruyter, G., Verbeke, M., and Hallez, H. (2026). Constraint-guided hardware-aware NAS through gradient modification. In *The Fourteenth International Conference on Learning Representations*.

Egele, R., Balaprakash, P., Guyon, I., Vishwanath, V., Xia, F., Stevens, R., and Liu, Z. (2021). AgEBO-Tabular: Joint neural architecture and hyperparameter search with autotuned data-parallel training for tabular data. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC '21)*.

Erickson, N., Mueller, J., Shirkov, A., Zhang, H., Larroy, P., Li, M., and Smola, A. (2020). AutoGluon-Tabular: Robust and accurate automl for structured data. *arXiv preprint arXiv:2003.06505*.

Erickson, N., Purucker, L., Tschalzev, A., Holzmüller, D., Desai, P., Salinas, D., and Hutter, F. (2025). TabArena: A living benchmark for machine learning on tabular data. In Belgrave, D., Zhang, C., Lin, H., Pascanu, R., Koniusz, P., Ghassemi, M., and Chen, N., editors, *Advances in Neural Information Processing Systems*, volume 38, Main Conference. Curran Associates, Inc.

Feurer, M., Klein, A., Eggensperger, K., Springenberg, J., Blum, M., and Hutter, F. (2015). Efficient and robust automated machine learning. In *Advances in Neural Information Processing Systems 28 (2015)*, pages 2962–2970.

Gao, Y., Bai, H., Jie, Z., and et al. (2020). MTL-NAS: Task-agnostic neural architecture search towards general-purpose multi-task learning. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11540–11549.

Ghiasi, G., Lin, T., Pang, R., and Le, Q. V. (2019). NAS-FPN: Learning scalable feature pyramid architecture for object detection. *CoRR*, abs/1904.07392.

Guyon, I. and Elisseeff, A. (2003). An introduction to variable and feature selection. *J. Mach. Learn. Res.*, 3:1157–1182.

Hafiz, F., Broekaert, J., La Torre, D., and Swain, A. (2023). Co-evolution of neural architectures and features for stock market forecasting: A multi-objective decision perspective. *Decision Support Systems*, 174:114015.

Huang, G.-B. (2003). Learning capability and storage capacity of two-hidden-layer feedforward networks. *IEEE Transactions on Neural Networks*, 14(2):274–281.

Jović, A., Brkić, K., and Bogunović, N. (2015). A review of feature selection methods with applications. In *2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pages 1200–1205.

Lemhadri, I., Ruan, F., Abraham, L., and Tibshirani, R. (2021). LassoNet: A neural network with feature sparsity. *Journal of Machine Learning Research*, 22(127):1–29.

Liu, H. and Motoda, H. (2012). *Feature Selection for Knowledge Discovery and Data Mining*. Springer, Boston, MA.

Liu, H., Simonyan, K., and Yang, Y. (2019). DARTS: Differentiable architecture search. In *International Conference on Learning Representations*.

Olson, R. S., Bartley, N., Urbanowicz, R. J., and Moore, J. H. (2016). Evaluation of a tree-based pipeline optimization tool for automating data science. In *Proceedings of the Genetic and Evolutionary Computation Conference 2016*, GECCO '16, pages 485–492. ACM.

Peng, H., Long, F., and Ding, C. (2005). Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(8):1226–1238.

Ripley, B. D. (1996). *Pattern Recognition and Neural Networks*. Cambridge University Press.

Robinson, I., Robicheaux, P., Popov, M., Ramanan, D., and Peri, N. (2026). RF-DETR: Neural architecture search for real-time detection transformers. In Vondrick, C., Hariharan, B., Raffel, C., Pinto, L., Yang, D., and Faust, A., editors, *International Conference on Learning Representations*, volume 2026, pages 56087–56100.

Schober, P., Boer, C., and Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia & Analgesia*, 126(5):1763–1768.

Tan, M. and Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In *International Conference on Machine Learning (ICML)*.

Wan, A., Dai, X., Zhang, P., He, Z., Tian, Y., Xie, S., Wu, B., Yu, M., Xu, T., Chen, K., Vajda, P., and Gonzalez, J. E. (2020). FBNetV2: Differentiable neural architecture search for spatial and channel dimensions. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12962–12971.

Wang, D., Gong, C., Li, M., Liu, Q., and Chandra, V. (2021). AlphaNet: Improved training of supernets with alpha-divergence. In Meila, M. and Zhang, T., editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 10760–10771. PMLR.

Wu, B., Dai, X., Zhang, P., Wang, Y., Sun, F., Wu, Y., Tian, Y., Vajda, P., Jia, Y., and Keutzer, K. (2019). FBNet: Hardware-aware efficient ConvNet design via differentiable neural architecture search. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10726–10734.

Xing, N., Cai, S., Zeng, L., Zhu, J., LU, P., Pei, J., and Ooi, B. C. (2026). pTNAS: Progressive neural architecture search for tabular data. In *Forty-third International Conference on Machine Learning*.

Yamada, Y., Lindenbaum, O., Negahban, S., and Kluger, Y. (2020). Feature selection using stochastic gates. In *Proceedings of the 37th International Conference on Machine Learning*, ICML'20. JMLR.org.

Yang, C., Bender, G., Liu, H., Kindermans, P.-J., Udell, M., Lu, Y., Le, Q. V., and Huang, D. (2022). TabNAS: Rejection sampling for neural architecture search on tabular datasets. In *Advances in Neural Information Processing Systems*.

Yu, J., Yang, L., Xu, N., Yang, J., and Huang, T. (2019). Slimmable neural networks. In *International Conference on Learning Representations*.

A DATASETS

Table 4: Overview of the 51 TabArena datasets (Erickson et al., 2025) used in our experiments. Tasks are categorized as classification (Clf.) or regression (Reg.).

Dataset	#Instances	#Features	#Classes	Task
airfoil_self_noise	1,503	5	1	Reg.
amazon_employee_access	32,769	9	1	Clf.
anneal	898	38	5	Clf.
another_dataset_on_used_fiat_500	1,538	7	1	Reg.
apsfailure	76,000	170	1	Clf.
bank_customer_churn	10,000	10	1	Clf.
bioresponse	3,751	1776	1	Clf.
blood_transfusion_service_center	748	4	1	Clf.
churn	5,000	19	1	Clf.
coil2000_insurance_policies	9,822	85	1	Clf.
concrete_compressive_strength	1,030	8	1	Reg.
credit_card_clients_default	30,000	23	1	Clf.
credit_g	1,000	20	1	Clf.
customer_satisfaction_in_airline	129,880	21	1	Clf.
diabetes130us	71,518	47	1	Clf.
diamonds	53,940	9	1	Reg.
e_commereshippingdata	10,999	10	1	Clf.
fitness_club	1,500	6	1	Clf.
food_delivery_time	45,451	9	1	Reg.
givemesomecredit	150,000	10	1	Clf.
hazelnut_spread_contaminant_detection	2,400	30	1	Clf.
healthcare_insurance_expenses	1,338	6	1	Reg.
heloc	10,459	23	1	Clf.
hiva_agnostic	3,845	1617	3	Clf.
houses	20,640	8	1	Reg.
hr_analytics_job_change_of_data_scientists	19,158	12	1	Clf.
in_vehicle_coupon_recommendation	12,684	24	1	Clf.
is_this_a_good_customer	1,723	13	1	Clf.
jm1	10,885	21	1	Clf.
kddcup09_appetency	50,000	212	1	Clf.
marketing_campaign	2,240	25	1	Clf.
maternal_health_risk	1,014	6	3	Clf.
miami_housing	13,776	15	1	Reg.
mic	1,699	111	8	Clf.
naticusdroid	7,491	86	1	Clf.
online_shoppers_intention	12,330	17	1	Clf.
physiochemical_protein	45,730	9	1	Reg.
polish_companies_bankruptcy	5,910	64	1	Clf.
qsar_biodeg	1,054	41	1	Clf.
qsar_fish_toxicity	907	6	1	Reg.
qsar_tid_11	5,742	1024	1	Reg.
sdss17	78,053	11	3	Clf.
seismic_bumps	2,584	15	1	Clf.
splice	3,190	60	3	Clf.
students_dropout_and_academic_success	4,424	36	3	Clf.
superconductivity	21,263	81	1	Reg.
tabarena_bank_marketing	45,211	13	1	Clf.
tabarena_diabetes	768	8	1	Clf.
taiwanese_bankruptcy_prediction	6,819	94	1	Clf.
website_phishing	1,353	9	3	Clf.
wine_quality	6,497	12	1	Reg.

B Architecture Details

B.1 Handcrafted Baselines

The architecture of the handcrafted baselines was determined according to the heuristic proposed by Huang (2003), which estimates the number of neurons in a two-hidden-layer feedforward network as a function of the number of output units m and the number of training samples N. The numbers of neurons in the first and second hidden layers are computed as:

$$h_1 = \left\lceil \sqrt{(m+2)N} + 2\sqrt{\frac{N}{m+2}} \right\rceil, \tag{4}$$

$$h_2 = \left\lceil m\sqrt{\frac{N}{m+2}} \right\rceil, \tag{5}$$

This heuristic aims to provide sufficient network capacity while maintaining good generalization performance by scaling model complexity with the amount of available training data.

C Additional Training Details

This appendix provides additional training details to reproduce our experiments.[2]

Feature Selection Thresholds

For the univariate filter (UF) baselines, the threshold τ is set to 0.4 for all experiments. For regression tasks, features are retained when their absolute Pearson correlation with the target satisfies $|\rho| \geq 0.4$. For classification tasks, mutual information scores are normalized by the maximum mutual information value observed for the dataset, and features are retained when $I \geq 0.4 \cdot \max(I)$. Four regression datasets (*airfoil_self_noise*, *food_delivery_time*, *physiochemical_protein* and *qsar_tid_11*) yield no selected features under this criterion and are therefore excluded from the UF evaluation.

For the mRMR-based baselines, we evaluate feature subsets obtained by retaining the top 20%, 40%, 60% and 80% of the ranked features. Each subset contains $k = \lceil f \cdot d \rceil$ features, where d denotes the number of input features and $f \in \{20, 40, 60, 80\}\%$. For multi-class classification tasks, feature rankings are computed for each class separately and the resulting top-k sets are merged, so the final number of retained features may exceed k.

Supernet Training

All experiments are repeated on five cross-validation splits. Within each split, 20% of each training portion is held out for validation and used for all model-selection decisions.

Training proceeds in four stages: full training, elastic layers, elastic units and elastic features. The full-training stage runs for 150 epochs and each elastic stage for 120 epochs. Evo baselines skip the elastic-features stage, so FETNAS trains for 120 additional epochs per supernet, whereas EVO-MRMR trains one supernet per retention fraction.

All stages use a batch size of 128 and plain SGD without momentum. A weight decay of 10^{-4} is applied to model parameters. The learning rate is initialized at 10^{-2} and annealed to 10^{-6} using cosine schedule, with the schedule restarted at the beginning of each training stage.

We adopt the sandwich sampling during training (except for the full-training stage) by sampling the largest subnetwork, the smallest subnetwork and two random subnetworks per mini-batch. The largest subnetwork is optimized using only the task loss. The smallest subnetwork and the two random subnetworks are trained using the task loss combined with distillation from the largest subnetwork.

Evolutionary Search

Evolutionary search is performed for 20 generations with a population size of 512. We employ a crossover probability of 0.5 and a mutation probability of 0.2, with both crossover and mutation sizes set to 128.

The search simultaneously maximizes validation performance and minimizes model complexity. Validation F1-macro is used for classification tasks, while validation R^2 is used for regression tasks. Parameter count and the number of selected features are treated as minimization objectives.

Before each candidate subnet is evaluated, its Batch Normalization statistics are recalibrated. All running statistics are reset and recomputed as a cumulative average over the first 64 shuffled training batches. During this procedure, only the Batch Normalization layers are updated, while all remaining network parameters remain in evaluation mode.

For *zero-shot* evaluation, candidate architectures inherit the supernet weights, undergo Batch Normalization recalibration as described above and are evaluated on the test set.

[2]The source code will be made available upon acceptance.

HANDCRAFTED BASELINES

Handcrafted models are trained for 150 epochs with a batch size of 16. Adam is used as the optimizer with a learning rate of 10^{-2} and a weight decay of 10^{-4} applied to all parameters. The checkpoint achieving the best validation performance is retained.

D PER-DATASET RESULTS

Table 5: **Classification results at matched feature budgets for each dataset within the TabArena benchmark.** For each feature budget, the columns correspond to the parameter counts at the minimum, Q1, median (Q2), Q3 and maximum of the baseline's (Evo-*) Pareto front. For FETNAS, we report the subnetwork with the best validation performance within that parameter budget, using at most the baseline's number of features. Each cell gives mean ± std F1-macro over the five cross-validation splits (top) and mean ± std parameter count (bottom, in parentheses). Evo-* and FETNAS are evaluated *zero-shot* with inherited supernet weights. RoT is trained from scratch and yields a single architecture. Underlined values mark the better of Evo-* and FETNAS in each column.

		F1-macro (parameter budget in parentheses)				
Method	k/d	min	Q1	Q2	Q3	max
amazon_employee_access						
Evo-All	1.00	0.49 ± 0.00 (89 ± 0)	0.49 ± 0.00 (1.1K ± 2.2K)	0.49 ± 0.00 (1.1K ± 2.2K)	0.49 ± 0.00 (1.1K ± 2.2K)	0.49 ± 0.01 (89.6K ± 91.6K)
FetNAS	0.12	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (13.9K ± 31.0K)
RoT-All	1.00					0.49 ± 0.00 (33.4K ± 0)
Evo-mRMR$_{80}$	0.89	0.49 ± 0.01 (81 ± 0)	0.49 ± 0.01 (105 ± 54)	0.49 ± 0.01 (131 ± 112)	0.49 ± 0.02 (2.6K ± 5.7K)	0.49 ± 0.02 (3.1K ± 5.8K)
FetNAS	0.11	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)
RoT-mRMR$_{80}$	0.89					0.49 ± 0.00 (33.0K ± 0)
Evo-mRMR$_{60}$	0.67	0.49 ± 0.00 (65 ± 0)	0.49 ± 0.00 (65 ± 0)	0.49 ± 0.00 (65 ± 0)	0.49 ± 0.00 (65 ± 0)	0.49 ± 0.00 (35.4K ± 79.1K)
FetNAS	0.12	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (13.9K ± 31.0K)
RoT-mRMR$_{60}$	0.67					0.49 ± 0.00 (32.2K ± 0)
Evo-mRMR$_{40}$	0.44	0.49 ± 0.00 (49 ± 0)	0.49 ± 0.00 (49 ± 0)	0.49 ± 0.00 (49 ± 0)	0.49 ± 0.00 (49 ± 0)	0.49 ± 0.00 (8.8K ± 19.7K)
FetNAS	0.12	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (3.4K ± 7.5K)
RoT-mRMR$_{40}$	0.44					0.49 ± 0.00 (31.5K ± 0)
Evo-mRMR$_{20}$	0.22	0.49 ± 0.00 (33 ± 0)	0.49 ± 0.00 (33 ± 0)	0.49 ± 0.00 (33 ± 0)	0.49 ± 0.00 (33 ± 0)	0.49 ± 0.00 (33 ± 0)
FetNAS	0.11	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)
RoT-mRMR$_{20}$	0.22					0.49 ± 0.00 (30.7K ± 0)
Evo-UF	0.33	0.49 ± 0.00 (41 ± 0)	0.49 ± 0.00 (41 ± 0)	0.49 ± 0.00 (41 ± 0)	0.49 ± 0.00 (41 ± 0)	0.49 ± 0.00 (41 ± 0)
FetNAS	0.11	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)	0.49 ± 0.00 (26 ± 1)
RoT-UF	0.33					0.49 ± 0.00 (31.1K ± 0)
anneal						
Evo-All	1.00	0.31 ± 0.08 (366 ± 20)	0.35 ± 0.07 (1.4K ± 1.5K)	0.35 ± 0.07 (1.4K ± 1.5K)	0.38 ± 0.11 (3.1K ± 3.8K)	0.72 ± 0.03 (157.2K ± 38.8K)
FetNAS	0.40	0.37 ± 0.05 (230 ± 81)	0.39 ± 0.07 (874 ± 631)	0.39 ± 0.07 (874 ± 631)	0.43 ± 0.10 (2.6K ± 3.8K)	0.70 ± 0.03 (53.4K ± 35.8K)
RoT-All	1.00					0.91 ± 0.04 (6.3K ± 0)
Evo-mRMR$_{80}$	0.98	0.30 ± 0.09 (352 ± 4)	0.35 ± 0.11 (1.9K ± 1.6K)	0.35 ± 0.11 (1.9K ± 1.6K)	0.36 ± 0.13 (2.9K ± 3.2K)	0.71 ± 0.03 (150.1K ± 37.5K)
FetNAS	0.40	0.36 ± 0.04 (220 ± 62)	0.38 ± 0.03 (1.2K ± 993)	0.38 ± 0.03 (1.2K ± 993)	0.42 ± 0.07 (2.2K ± 3.0K)	0.70 ± 0.03 (53.4K ± 35.8K)
RoT-mRMR$_{80}$	0.98					0.87 ± 0.05 (6.3K ± 41)
Evo-mRMR$_{60}$	0.96	0.23 ± 0.08 (353 ± 17)	0.26 ± 0.13 (442 ± 199)	0.29 ± 0.18 (2.4K ± 3.9K)	0.36 ± 0.22 (14.5K ± 30.6K)	0.69 ± 0.05 (165.4K ± 36.9K)
FetNAS	0.44	0.36 ± 0.04 (220 ± 62)	0.38 ± 0.01 (261 ± 99)	0.44 ± 0.08 (1.8K ± 3.0K)	0.46 ± 0.09 (11.1K ± 23.5K)	0.70 ± 0.03 (53.4K ± 35.8K)
RoT-mRMR$_{60}$	0.96					0.92 ± 0.03 (6.2K ± 41)
Evo-mRMR$_{40}$	0.88	0.29 ± 0.07 (322 ± 4)	0.32 ± 0.07 (712 ± 546)	0.33 ± 0.08 (893 ± 535)	0.36 ± 0.08 (1.1K ± 643)	0.69 ± 0.05 (112.4K ± 66.4K)
FetNAS	0.46	0.36 ± 0.04 (220 ± 62)	0.41 ± 0.07 (443 ± 386)	0.44 ± 0.07 (615 ± 408)	0.43 ± 0.07 (889 ± 624)	0.70 ± 0.03 (53.4K ± 35.8K)
RoT-mRMR$_{40}$	0.88					0.92 ± 0.04 (6.0K ± 41)
Evo-mRMR$_{20}$	0.64	0.30 ± 0.08 (259 ± 28)	0.31 ± 0.08 (584 ± 731)	0.31 ± 0.08 (584 ± 731)	0.36 ± 0.11 (1.2K ± 1.4K)	0.70 ± 0.06 (116.3K ± 63.4K)
FetNAS	0.42	0.34 ± 0.04 (208 ± 45)	0.36 ± 0.07 (549 ± 725)	0.36 ± 0.07 (549 ± 725)	0.38 ± 0.06 (884 ± 889)	0.67 ± 0.07 (37.6K ± 24.0K)

Method	k/d	F1-macro (parameter budget in parentheses) min	Q1	Q2	Q3	max
RoT-mRMR$_{20}$	0.64					0.91 ± 0.05 (5.3K ± 144)
Evo-UF	0.17	0.20 ± 0.04 (104 ± 7)	0.25 ± 0.13 (902 ± 1.8K)	0.26 ± 0.12 (1.0K ± 1.7K)	0.40 ± 0.15 (2.5K ± 1.6K)	0.72 ± 0.03 (99.4K ± 54.5K)
FetNAS	0.11	0.26 ± 0.08 (86 ± 8)	0.33 ± 0.11 (581 ± 1.1K)	0.34 ± 0.11 (658 ± 1.1K)	0.38 ± 0.15 (1.6K ± 983)	0.51 ± 0.08 (22.0K ± 14.8K)
RoT-UF	0.16					0.90 ± 0.09 (3.9K ± 63)
apsfailure						
Evo-All	1.00	0.65 ± 0.01 (1.4K ± 154)	0.74 ± 0.02 (7.4K ± 2.4K)	0.76 ± 0.01 (22.1K ± 11.6K)	0.77 ± 0.01 (40.6K ± 16.1K)	0.79 ± 0.02 (125.1K ± 64.4K)
FetNAS	0.21	0.74 ± 0.02 (926 ± 259)	0.75 ± 0.02 (3.0K ± 2.2K)	0.75 ± 0.02 (10.9K ± 10.0K)	0.75 ± 0.02 (10.9K ± 10.0K)	0.75 ± 0.02 (10.9K ± 10.0K)
RoT-All	1.00					0.53 ± 0.06 (167.9K ± 0)
Evo-mRMR$_{80}$	0.80	0.65 ± 0.07 (1.2K ± 123)	0.75 ± 0.04 (3.1K ± 1.7K)	0.78 ± 0.03 (6.3K ± 4.0K)	0.80 ± 0.03 (23.1K ± 15.0K)	0.81 ± 0.02 (52.0K ± 27.6K)
FetNAS	0.10	0.74 ± 0.02 (796 ± 264)	0.74 ± 0.02 (1.2K ± 750)	0.75 ± 0.02 (1.9K ± 1.4K)	0.75 ± 0.02 (11.4K ± 10.1K)	0.75 ± 0.02 (12.2K ± 10.5K)
RoT-mRMR$_{80}$	0.80					0.59 ± 0.14 (148.1K ± 0)
Evo-mRMR$_{60}$	0.60	0.58 ± 0.08 (895 ± 93)	0.66 ± 0.09 (3.1K ± 1.6K)	0.70 ± 0.12 (4.9K ± 2.7K)	0.71 ± 0.12 (12.9K ± 12.0K)	0.77 ± 0.02 (67.1K ± 75.0K)
FetNAS	0.11	0.75 ± 0.02 (731 ± 152)	0.75 ± 0.02 (2.3K ± 1.4K)	0.75 ± 0.02 (3.0K ± 2.2K)	0.75 ± 0.02 (7.1K ± 7.6K)	0.75 ± 0.02 (12.2K ± 10.5K)
RoT-mRMR$_{60}$	0.60					0.55 ± 0.08 (128.3K ± 0)
Evo-mRMR$_{40}$	0.40	0.65 ± 0.08 (561 ± 0)	0.71 ± 0.10 (2.0K ± 1.4K)	0.73 ± 0.11 (4.7K ± 2.8K)	0.74 ± 0.12 (10.9K ± 7.9K)	0.77 ± 0.09 (84.8K ± 78.6K)
FetNAS	0.10	0.71 ± 0.07 (473 ± 65)	0.74 ± 0.02 (912 ± 435)	0.75 ± 0.02 (1.7K ± 1.3K)	0.75 ± 0.02 (4.7K ± 7.4K)	0.75 ± 0.02 (7.1K ± 7.6K)
RoT-mRMR$_{40}$	0.40					0.67 ± 0.10 (108.5K ± 0)
Evo-mRMR$_{20}$	0.20	0.57 ± 0.10 (296 ± 16)	0.59 ± 0.13 (673 ± 537)	0.60 ± 0.14 (1.3K ± 1.6K)	0.60 ± 0.15 (2.5K ± 3.2K)	0.79 ± 0.01 (37.7K ± 19.2K)
FetNAS	0.08	0.68 ± 0.06 (213 ± 93)	0.71 ± 0.03 (504 ± 488)	0.71 ± 0.03 (504 ± 488)	0.71 ± 0.03 (504 ± 488)	0.75 ± 0.02 (7.1K ± 7.6K)
RoT-mRMR$_{20}$	0.20					0.63 ± 0.12 (88.7K ± 0)
Evo-UF	0.55	0.59 ± 0.10 (780 ± 49)	0.61 ± 0.13 (2.1K ± 1.8K)	0.61 ± 0.13 (2.5K ± 2.1K)	0.64 ± 0.13 (5.8K ± 7.1K)	0.69 ± 0.10 (20.0K ± 13.7K)
FetNAS	0.11	0.72 ± 0.07 (553 ± 166)	0.75 ± 0.02 (1.2K ± 995)	0.75 ± 0.02 (1.2K ± 995)	0.75 ± 0.02 (3.4K ± 5.9K)	0.75 ± 0.02 (7.1K ± 7.6K)
RoT-UF	0.55					0.52 ± 0.06 (123.0K ± 1.0K)
bank_customer_churn						
Evo-All	1.00	0.69 ± 0.05 (97 ± 0)	0.71 ± 0.05 (2.4K ± 4.6K)	0.72 ± 0.03 (2.7K ± 4.5K)	0.72 ± 0.03 (3.6K ± 4.7K)	0.75 ± 0.01 (29.6K ± 36.3K)
FetNAS	0.68	0.71 ± 0.04 (87 ± 12)	0.73 ± 0.03 (207 ± 115)	0.74 ± 0.02 (382 ± 256)	0.74 ± 0.01 (1.4K ± 2.4K)	0.75 ± 0.01 (5.9K ± 5.9K)
RoT-All	1.00					0.74 ± 0.02 (11.5K ± 0)
Evo-mRMR$_{80}$	0.80	0.70 ± 0.05 (83 ± 4)	0.70 ± 0.05 (294 ± 139)	0.70 ± 0.05 (324 ± 159)	0.71 ± 0.06 (1.5K ± 2.7K)	0.74 ± 0.01 (22.1K ± 17.2K)
FetNAS	0.62	0.71 ± 0.04 (74 ± 6)	0.73 ± 0.04 (181 ± 131)	0.73 ± 0.04 (181 ± 131)	0.73 ± 0.04 (285 ± 355)	0.75 ± 0.01 (6.5K ± 6.4K)
RoT-mRMR$_{80}$	0.80					0.74 ± 0.01 (11.1K ± 0)
Evo-mRMR$_{60}$	0.60	0.71 ± 0.04 (68 ± 4)	0.73 ± 0.01 (318 ± 251)	0.73 ± 0.01 (824 ± 992)	0.73 ± 0.01 (937 ± 1.0K)	0.74 ± 0.00 (15.5K ± 9.5K)
FetNAS	0.52	0.69 ± 0.05 (65 ± 6)	0.72 ± 0.04 (198 ± 182)	0.73 ± 0.04 (369 ± 347)	0.73 ± 0.03 (414 ± 384)	0.74 ± 0.02 (4.9K ± 6.4K)
RoT-mRMR$_{60}$	0.60					0.74 ± 0.01 (10.6K ± 0)
Evo-mRMR$_{40}$	0.40	0.69 ± 0.05 (50 ± 3)	0.71 ± 0.06 (1.7K ± 2.6K)	0.71 ± 0.06 (1.8K ± 2.6K)	0.71 ± 0.06 (1.8K ± 2.6K)	0.74 ± 0.02 (32.3K ± 26.1K)
FetNAS	0.33	0.66 ± 0.07 (44 ± 10)	0.69 ± 0.08 (960 ± 1.9K)	0.70 ± 0.07 (1.1K ± 1.8K)	0.70 ± 0.07 (1.1K ± 1.8K)	0.73 ± 0.01 (8.3K ± 7.5K)
RoT-mRMR$_{40}$	0.40					0.74 ± 0.02 (10.2K ± 0)
Evo-mRMR$_{20}$	0.20	0.64 ± 0.05 (33 ± 0)	0.67 ± 0.06 (109 ± 84)	0.67 ± 0.06 (119 ± 77)	0.68 ± 0.06 (140 ± 74)	0.73 ± 0.00 (30.2K ± 46.6K)
FetNAS	0.19	0.59 ± 0.06 (30 ± 4)	0.66 ± 0.06 (56 ± 29)	0.66 ± 0.06 (61 ± 25)	0.66 ± 0.06 (61 ± 25)	0.68 ± 0.05 (6.8K ± 13.3K)
RoT-mRMR$_{20}$	0.20					0.71 ± 0.01 (9.8K ± 0)
Evo-UF	0.20	0.64 ± 0.05 (33 ± 0)	0.67 ± 0.06 (109 ± 84)	0.67 ± 0.06 (119 ± 77)	0.67 ± 0.06 (140 ± 74)	0.73 ± 0.00 (30.2K ± 46.6K)
FetNAS	0.19	0.59 ± 0.06 (30 ± 4)	0.66 ± 0.06 (56 ± 29)	0.66 ± 0.06 (61 ± 25)	0.66 ± 0.06 (61 ± 25)	0.68 ± 0.05 (6.8K ± 13.3K)
RoT-UF	0.20					0.71 ± 0.01 (9.8K ± 0)
bioresponse						
Evo-All	1.00	0.68 ± 0.02 (14.6K ± 795)	0.71 ± 0.03 (119.2K ± 111.4K)	0.71 ± 0.03 (119.2K ± 111.4K)	0.71 ± 0.03 (136.2K ± 143.2K)	0.72 ± 0.02 (327.2K ± 114.6K)
FetNAS	0.05	0.74 ± 0.02 (702 ± 919)	0.73 ± 0.01 (24.8K ± 53.6K)	0.73 ± 0.01 (24.8K ± 53.6K)	0.73 ± 0.01 (24.8K ± 53.6K)	0.73 ± 0.01 (26.1K ± 56.4K)

Method	k/d	F1-macro (parameter budget in parentheses) min	Q1	Q2	Q3	max
RoT-All	1.00					0.71 ± 0.03 (234.4K ± 0)
Evo-mRMR$_{80}$	0.80	0.69 ± 0.01 (11.7K ± 636)	0.70 ± 0.02 (81.9K ± 58.0K)	0.70 ± 0.02 (81.9K ± 58.0K)	0.71 ± 0.02 (119.1K ± 80.4K)	0.71 ± 0.01 (245.2K ± 92.9K)
FetNAS	0.05	0.74 ± 0.02 (702 ± 919)	0.73 ± 0.01 (26.1K ± 56.4K)	0.73 ± 0.01 (26.1K ± 56.4K)	0.73 ± 0.01 (26.1K ± 56.4K)	0.73 ± 0.01 (26.1K ± 56.4K)
RoT-mRMR$_{80}$	0.80					0.73 ± 0.01 (188.3K ± 0)
Evo-mRMR$_{60}$	0.60	0.68 ± 0.02 (8.5K ± 0)	0.70 ± 0.03 (70.2K ± 76.1K)	0.70 ± 0.03 (70.2K ± 76.1K)	0.70 ± 0.03 (71.0K ± 77.2K)	0.72 ± 0.01 (226.3K ± 46.4K)
FetNAS	0.05	0.74 ± 0.02 (702 ± 919)	0.73 ± 0.01 (26.1K ± 56.4K)	0.73 ± 0.01 (26.1K ± 56.4K)	0.73 ± 0.01 (26.1K ± 56.4K)	0.73 ± 0.01 (26.1K ± 56.4K)
RoT-mRMR$_{60}$	0.60					0.73 ± 0.03 (142.1K ± 0)
Evo-mRMR$_{40}$	0.40	0.68 ± 0.04 (5.7K ± 46)	0.69 ± 0.04 (19.5K ± 15.8K)	0.70 ± 0.03 (19.6K ± 15.7K)	0.69 ± 0.03 (44.8K ± 56.0K)	0.71 ± 0.01 (134.5K ± 40.7K)
FetNAS	0.01	0.74 ± 0.02 (702 ± 919)	0.74 ± 0.02 (702 ± 919)	0.74 ± 0.02 (702 ± 919)	0.74 ± 0.02 (702 ± 919)	0.73 ± 0.01 (26.1K ± 56.4K)
RoT-mRMR$_{40}$	0.40					0.72 ± 0.01 (96.0K ± 0)
Evo-mRMR$_{20}$	0.20	0.70 ± 0.03 (2.9K ± 0)	0.70 ± 0.04 (15.4K ± 16.1K)	0.70 ± 0.04 (15.4K ± 16.1K)	0.70 ± 0.04 (15.4K ± 16.1K)	0.71 ± 0.04 (54.3K ± 47.0K)
FetNAS	0.00	0.74 ± 0.02 (702 ± 919)	0.74 ± 0.02 (702 ± 919)	0.74 ± 0.02 (702 ± 919)	0.74 ± 0.02 (702 ± 919)	0.74 ± 0.02 (702 ± 919)
RoT-mRMR$_{20}$	0.20					0.73 ± 0.01 (49.8K ± 0)
Evo-UF	0.02	0.72 ± 0.02 (259 ± 90)	0.73 ± 0.02 (642 ± 327)	0.72 ± 0.03 (811 ± 447)	0.73 ± 0.03 (4.5K ± 6.2K)	0.74 ± 0.01 (53.8K ± 50.4K)
FetNAS	0.00	0.73 ± 0.03 (54 ± 32)	0.73 ± 0.02 (160 ± 250)	0.73 ± 0.02 (160 ± 250)	0.73 ± 0.02 (473 ± 712)	0.74 ± 0.02 (702 ± 919)
RoT-UF	0.02					0.75 ± 0.02 (7.5K ± 1.5K)
blood_transfusion_service_center						
Evo-All	1.00	0.49 ± 0.04 (50 ± 3)	0.51 ± 0.05 (86 ± 48)	0.53 ± 0.08 (151 ± 131)	0.53 ± 0.09 (226 ± 295)	0.67 ± 0.04 (19.7K ± 18.6K)
FetNAS	0.43	0.45 ± 0.03 (30 ± 5)	0.46 ± 0.03 (47 ± 38)	0.46 ± 0.03 (47 ± 38)	0.49 ± 0.07 (164 ± 256)	0.59 ± 0.06 (6.5K ± 5.2K)
RoT-All	1.00					0.57 ± 0.11 (1.0K ± 0)
Evo-mRMR$_{80}$	1.00	0.48 ± 0.04 (50 ± 3)	0.52 ± 0.06 (86 ± 48)	0.54 ± 0.08 (151 ± 131)	0.54 ± 0.09 (226 ± 295)	0.66 ± 0.03 (19.7K ± 18.6K)
FetNAS	0.43	0.45 ± 0.03 (30 ± 5)	0.46 ± 0.03 (47 ± 38)	0.46 ± 0.03 (47 ± 38)	0.49 ± 0.07 (164 ± 256)	0.59 ± 0.06 (6.5K ± 5.2K)
RoT-mRMR$_{80}$	1.00					0.62 ± 0.08 (1.0K ± 0)
Evo-mRMR$_{60}$	0.75	0.47 ± 0.03 (42 ± 2)	0.48 ± 0.05 (73 ± 69)	0.48 ± 0.04 (75 ± 68)	0.48 ± 0.04 (75 ± 68)	0.66 ± 0.02 (27.3K ± 24.7K)
FetNAS	0.40	0.45 ± 0.03 (30 ± 5)	0.46 ± 0.03 (47 ± 38)	0.46 ± 0.03 (47 ± 38)	0.46 ± 0.03 (47 ± 38)	0.60 ± 0.04 (8.1K ± 6.0K)
RoT-mRMR$_{60}$	0.75					0.62 ± 0.11 (953 ± 0)
Evo-mRMR$_{40}$	0.50	0.44 ± 0.01 (33 ± 0)	0.44 ± 0.01 (33 ± 0)	0.45 ± 0.02 (35 ± 4)	0.45 ± 0.02 (35 ± 4)	0.61 ± 0.03 (40.2K ± 22.5K)
FetNAS	0.32	0.44 ± 0.01 (28 ± 3)	0.44 ± 0.01 (28 ± 3)	0.45 ± 0.03 (30 ± 5)	0.45 ± 0.03 (30 ± 5)	0.56 ± 0.05 (20.7K ± 27.0K)
RoT-mRMR$_{40}$	0.50					0.53 ± 0.05 (895 ± 0)
Evo-mRMR$_{20}$	0.25	0.44 ± 0.02 (25 ± 0)	0.44 ± 0.02 (25 ± 0)	0.44 ± 0.02 (25 ± 0)	0.44 ± 0.02 (25 ± 0)	0.53 ± 0.06 (31.3K ± 45.0K)
FetNAS	0.25	0.43 ± 0.00 (25 ± 0)	0.43 ± 0.00 (25 ± 0)	0.43 ± 0.00 (25 ± 0)	0.43 ± 0.00 (25 ± 0)	0.47 ± 0.06 (306 ± 485)
RoT-mRMR$_{20}$	0.25					0.43 ± 0.00 (837 ± 0)
Evo-UF	0.60	0.43 ± 0.00 (36 ± 4)	0.43 ± 0.00 (36 ± 4)	0.44 ± 0.01 (38 ± 8)	0.44 ± 0.01 (38 ± 8)	0.61 ± 0.05 (41.2K ± 21.8K)
FetNAS	0.36	0.45 ± 0.03 (30 ± 5)	0.45 ± 0.03 (30 ± 5)	0.45 ± 0.03 (30 ± 5)	0.45 ± 0.03 (30 ± 5)	0.59 ± 0.05 (17.8K ± 24.8K)
RoT-UF	0.60					0.56 ± 0.06 (918 ± 32)
churn						
Evo-All	1.00	0.62 ± 0.11 (169 ± 0)	0.62 ± 0.11 (190 ± 21)	0.63 ± 0.11 (228 ± 98)	0.65 ± 0.12 (249 ± 85)	0.80 ± 0.03 (78.8K ± 69.7K)
FetNAS	0.41	0.73 ± 0.05 (152 ± 18)	0.73 ± 0.05 (169 ± 29)	0.74 ± 0.05 (204 ± 82)	0.74 ± 0.05 (204 ± 82)	0.82 ± 0.03 (21.7K ± 14.8K)
RoT-All	1.00					0.85 ± 0.02 (7.6K ± 0)
Evo-mRMR$_{80}$	0.84	0.71 ± 0.03 (149 ± 8)	0.76 ± 0.04 (374 ± 370)	0.77 ± 0.04 (718 ± 729)	0.79 ± 0.02 (3.8K ± 3.6K)	0.80 ± 0.02 (38.8K ± 46.2K)
FetNAS	0.43	0.72 ± 0.06 (122 ± 47)	0.73 ± 0.07 (272 ± 307)	0.77 ± 0.06 (520 ± 509)	0.78 ± 0.05 (3.3K ± 3.1K)	0.82 ± 0.03 (17.2K ± 23.4K)
RoT-mRMR$_{80}$	0.84					0.84 ± 0.02 (7.1K ± 0)
Evo-mRMR$_{60}$	0.63	0.68 ± 0.05 (113 ± 0)	0.69 ± 0.04 (127 ± 14)	0.69 ± 0.05 (189 ± 138)	0.69 ± 0.05 (225 ± 138)	0.78 ± 0.03 (16.1K ± 12.8K)
FetNAS	0.32	0.71 ± 0.07 (95 ± 31)	0.71 ± 0.06 (103 ± 35)	0.71 ± 0.07 (166 ± 154)	0.72 ± 0.07 (203 ± 160)	0.81 ± 0.03 (10.6K ± 10.6K)
RoT-mRMR$_{60}$	0.63					0.79 ± 0.03 (6.5K ± 0)

Method	k/d	F1-macro (parameter budget in parentheses) min	Q1	Q2	Q3	max
Evo-mRMR$_{40}$	0.42	0.65 ± 0.02 (81 ± 0)	0.66 ± 0.01 (187 ± 92)	0.66 ± 0.01 (187 ± 92)	0.67 ± 0.03 (1.3K ± 2.5K)	0.77 ± 0.02 (40.0K ± 31.5K)
FetNAS	0.30	0.67 ± 0.05 (61 ± 17)	0.71 ± 0.07 (123 ± 54)	0.71 ± 0.07 (123 ± 54)	0.74 ± 0.06 (552 ± 918)	0.80 ± 0.04 (18.5K ± 14.7K)
RoT-mRMR$_{40}$	0.42					0.78 ± 0.05 (5.9K ± 0)
Evo-mRMR$_{20}$	0.21	0.57 ± 0.07 (49 ± 0)	0.63 ± 0.08 (386 ± 480)	0.63 ± 0.08 (421 ± 539)	0.64 ± 0.06 (427 ± 534)	0.69 ± 0.03 (40.7K ± 58.3K)
FetNAS	0.15	0.63 ± 0.01 (43 ± 6)	0.68 ± 0.05 (263 ± 286)	0.68 ± 0.05 (263 ± 286)	0.69 ± 0.05 (267 ± 282)	0.70 ± 0.04 (11.7K ± 23.6K)
RoT-mRMR$_{20}$	0.21					0.67 ± 0.06 (5.3K ± 0)
Evo-UF	0.21	0.58 ± 0.06 (49 ± 6)	0.60 ± 0.09 (330 ± 584)	0.62 ± 0.08 (390 ± 563)	0.62 ± 0.09 (439 ± 545)	0.74 ± 0.02 (8.7K ± 5.4K)
FetNAS	0.15	0.63 ± 0.02 (42 ± 7)	0.66 ± 0.05 (169 ± 271)	0.66 ± 0.05 (169 ± 271)	0.67 ± 0.05 (173 ± 269)	0.71 ± 0.03 (4.0K ± 6.1K)
RoT-UF	0.21					0.73 ± 0.03 (5.3K ± 106)
coil2000_insurance_policies						
Evo-All	1.00	0.49 ± 0.00 (697 ± 0)	0.50 ± 0.01 (3.9K ± 4.9K)	0.50 ± 0.01 (3.9K ± 4.9K)	0.50 ± 0.01 (4.5K ± 6.2K)	0.51 ± 0.02 (12.6K ± 8.2K)
FetNAS	0.28	0.49 ± 0.01 (280 ± 252)	0.50 ± 0.01 (617 ± 651)	0.50 ± 0.01 (617 ± 651)	0.50 ± 0.01 (617 ± 651)	0.51 ± 0.03 (1.2K ± 791)
RoT-All	1.00					0.49 ± 0.00 (26.8K ± 0)
Evo-mRMR$_{80}$	0.80	0.50 ± 0.02 (561 ± 0)	0.50 ± 0.02 (1.2K ± 744)	0.51 ± 0.03 (1.8K ± 1.1K)	0.51 ± 0.03 (2.2K ± 1.3K)	0.53 ± 0.03 (42.3K ± 41.7K)
FetNAS	0.35	0.49 ± 0.01 (248 ± 205)	0.50 ± 0.01 (761 ± 583)	0.51 ± 0.03 (1.2K ± 791)	0.51 ± 0.03 (1.2K ± 791)	0.51 ± 0.03 (1.2K ± 791)
RoT-mRMR$_{80}$	0.80					0.48 ± 0.00 (23.3K ± 0)
Evo-mRMR$_{60}$	0.60	0.50 ± 0.02 (425 ± 0)	0.50 ± 0.02 (1.0K ± 1.2K)	0.50 ± 0.02 (1.9K ± 2.0K)	0.50 ± 0.02 (2.0K ± 1.9K)	0.52 ± 0.03 (10.5K ± 8.1K)
FetNAS	0.20	0.49 ± 0.01 (163 ± 183)	0.50 ± 0.01 (500 ± 701)	0.50 ± 0.01 (585 ± 654)	0.50 ± 0.01 (585 ± 654)	0.50 ± 0.03 (1.8K ± 1.7K)
RoT-mRMR$_{60}$	0.60					0.48 ± 0.00 (19.7K ± 0)
Evo-mRMR$_{40}$	0.40	0.50 ± 0.01 (289 ± 0)	0.50 ± 0.01 (1.2K ± 1.8K)	0.50 ± 0.01 (1.2K ± 1.8K)	0.50 ± 0.01 (1.5K ± 2.2K)	0.51 ± 0.01 (17.0K ± 25.2K)
FetNAS	0.08	0.49 ± 0.01 (109 ± 109)	0.49 ± 0.01 (136 ± 161)	0.49 ± 0.01 (161 ± 179)	0.49 ± 0.01 (161 ± 179)	0.49 ± 0.02 (1.2K ± 2.5K)
RoT-mRMR$_{40}$	0.40					0.48 ± 0.00 (16.2K ± 0)
Evo-mRMR$_{20}$	0.20	0.49 ± 0.01 (153 ± 0)	0.49 ± 0.01 (400 ± 243)	0.49 ± 0.01 (400 ± 243)	0.49 ± 0.01 (810 ± 1.1K)	0.50 ± 0.01 (51.9K ± 24.0K)
FetNAS	0.03	0.49 ± 0.01 (44 ± 31)	0.49 ± 0.01 (44 ± 31)	0.49 ± 0.01 (44 ± 31)	0.49 ± 0.01 (44 ± 31)	0.50 ± 0.02 (643 ± 748)
RoT-mRMR$_{20}$	0.20					0.48 ± 0.00 (12.6K ± 0)
Evo-UF	0.16	0.48 ± 0.00 (134 ± 54)	0.48 ± 0.00 (134 ± 54)	0.49 ± 0.00 (283 ± 319)	0.49 ± 0.00 (283 ± 319)	0.53 ± 0.02 (36.0K ± 33.8K)
FetNAS	0.05	0.49 ± 0.01 (68 ± 85)	0.49 ± 0.01 (68 ± 85)	0.49 ± 0.01 (126 ± 132)	0.49 ± 0.01 (126 ± 132)	0.49 ± 0.02 (306 ± 472)
RoT-UF	0.16					0.48 ± 0.00 (11.8K ± 916)
credit_card_clients_default						
Evo-All	1.00	0.66 ± 0.02 (216 ± 22)	0.69 ± 0.00 (498 ± 152)	0.69 ± 0.01 (810 ± 426)	0.70 ± 0.00 (3.5K ± 4.5K)	0.70 ± 0.01 (16.3K ± 16.6K)
FetNAS	0.38	0.67 ± 0.01 (138 ± 106)	0.68 ± 0.01 (252 ± 150)	0.68 ± 0.01 (454 ± 462)	0.69 ± 0.01 (2.2K ± 3.5K)	0.69 ± 0.00 (4.0K ± 3.6K)
RoT-All	1.00					0.67 ± 0.01 (36.0K ± 0)
Evo-mRMR$_{80}$	0.83	0.66 ± 0.01 (169 ± 0)	0.68 ± 0.00 (317 ± 76)	0.69 ± 0.00 (752 ± 533)	0.69 ± 0.00 (2.9K ± 4.4K)	0.69 ± 0.00 (7.3K ± 6.1K)
FetNAS	0.35	0.67 ± 0.01 (93 ± 65)	0.68 ± 0.01 (200 ± 98)	0.68 ± 0.01 (378 ± 342)	0.68 ± 0.01 (2.1K ± 3.1K)	0.69 ± 0.01 (3.2K ± 3.4K)
RoT-mRMR$_{80}$	0.83					0.68 ± 0.01 (34.6K ± 0)
Evo-mRMR$_{60}$	0.61	0.65 ± 0.02 (135 ± 9)	0.68 ± 0.01 (283 ± 134)	0.69 ± 0.00 (1.2K ± 1.3K)	0.69 ± 0.00 (3.0K ± 2.1K)	0.69 ± 0.00 (6.0K ± 4.6K)
FetNAS	0.31	0.67 ± 0.00 (65 ± 55)	0.68 ± 0.01 (189 ± 96)	0.68 ± 0.01 (474 ± 460)	0.69 ± 0.01 (1.2K ± 1.1K)	0.69 ± 0.01 (2.5K ± 2.5K)
RoT-mRMR$_{60}$	0.61					0.68 ± 0.01 (32.7K ± 0)
Evo-mRMR$_{40}$	0.43	0.66 ± 0.01 (102 ± 7)	0.67 ± 0.01 (196 ± 77)	0.68 ± 0.01 (348 ± 162)	0.69 ± 0.00 (1.1K ± 609)	0.69 ± 0.00 (2.5K ± 1.3K)
FetNAS	0.24	0.67 ± 0.00 (37 ± 27)	0.67 ± 0.01 (104 ± 95)	0.68 ± 0.01 (167 ± 186)	0.69 ± 0.00 (906 ± 565)	0.69 ± 0.00 (1.8K ± 841)
RoT-mRMR$_{40}$	0.43					0.68 ± 0.01 (31.3K ± 0)
Evo-mRMR$_{20}$	0.22	0.66 ± 0.01 (57 ± 0)	0.66 ± 0.01 (185 ± 92)	0.67 ± 0.00 (261 ± 100)	0.67 ± 0.00 (679 ± 686)	0.68 ± 0.01 (4.8K ± 2.8K)
FetNAS	0.09	0.67 ± 0.00 (25 ± 0)	0.67 ± 0.00 (69 ± 70)	0.67 ± 0.00 (69 ± 70)	0.67 ± 0.00 (69 ± 70)	0.68 ± 0.01 (2.8K ± 2.0K)
RoT-mRMR$_{20}$	0.22					0.67 ± 0.00 (29.4K ± 0)
Evo-UF	0.20	0.65 ± 0.02 (56 ± 4)	0.68 ± 0.03 (1.8K ± 2.5K)	0.68 ± 0.03 (1.8K ± 2.4K)	0.68 ± 0.03 (2.7K ± 2.6K)	0.69 ± 0.00 (20.0K ± 21.7K)

Method	k/d	F1-macro (parameter budget in parentheses) min	Q1	Q2	Q3	max
FetNAS	0.09	0.67 ± 0.00 (25 ± 0)	0.67 ± 0.01 (1.2K ± 2.4K)	0.67 ± 0.01 (1.2K ± 2.4K)	0.68 ± 0.01 (1.6K ± 2.4K)	0.68 ± 0.01 (9.8K ± 13.1K)
RoT-UF	0.20					0.69 ± 0.01 (29.3K ± 200)
credit_g						
Evo-All	1.00	0.59 ± 0.07 (186 ± 12)	0.61 ± 0.06 (765 ± 372)	0.61 ± 0.06 (765 ± 372)	0.61 ± 0.06 (799 ± 392)	0.63 ± 0.02 (15.1K ± 10.5K)
FetNAS	0.41	0.61 ± 0.03 (90 ± 30)	0.61 ± 0.03 (234 ± 176)	0.61 ± 0.03 (234 ± 176)	0.61 ± 0.03 (234 ± 176)	0.64 ± 0.03 (3.6K ± 2.5K)
RoT-All	1.00					0.53 ± 0.08 (2.4K ± 0)
Evo-mRMR$_{80}$	0.80	0.51 ± 0.09 (145 ± 0)	0.53 ± 0.08 (203 ± 111)	0.53 ± 0.07 (673 ± 759)	0.54 ± 0.08 (1.9K ± 2.9K)	0.59 ± 0.03 (45.8K ± 45.6K)
FetNAS	0.25	0.61 ± 0.03 (90 ± 30)	0.60 ± 0.04 (123 ± 78)	0.62 ± 0.03 (465 ± 836)	0.63 ± 0.04 (677 ± 1.3K)	0.62 ± 0.02 (8.7K ± 8.8K)
RoT-mRMR$_{80}$	0.80					0.53 ± 0.09 (2.1K ± 0)
Evo-mRMR$_{60}$	0.60	0.52 ± 0.10 (119 ± 13)	0.52 ± 0.10 (152 ± 55)	0.53 ± 0.11 (427 ± 564)	0.54 ± 0.12 (2.1K ± 2.6K)	0.61 ± 0.02 (14.2K ± 9.8K)
FetNAS	0.19	0.61 ± 0.03 (75 ± 17)	0.61 ± 0.03 (90 ± 30)	0.62 ± 0.04 (742 ± 1.4K)	0.62 ± 0.04 (865 ± 1.4K)	0.64 ± 0.03 (3.3K ± 2.7K)
RoT-mRMR$_{60}$	0.60					0.55 ± 0.07 (1.8K ± 0)
Evo-mRMR$_{40}$	0.40	0.51 ± 0.07 (81 ± 0)	0.54 ± 0.08 (158 ± 75)	0.54 ± 0.09 (1.3K ± 1.7K)	0.55 ± 0.09 (10.8K ± 22.5K)	0.61 ± 0.04 (53.0K ± 40.9K)
FetNAS	0.21	0.60 ± 0.02 (64 ± 9)	0.61 ± 0.03 (72 ± 16)	0.62 ± 0.02 (351 ± 615)	0.61 ± 0.01 (3.8K ± 7.6K)	0.63 ± 0.02 (22.8K ± 15.6K)
RoT-mRMR$_{40}$	0.40					0.56 ± 0.10 (1.6K ± 0)
Evo-mRMR$_{20}$	0.20	0.45 ± 0.05 (49 ± 0)	0.50 ± 0.08 (357 ± 584)	0.53 ± 0.09 (476 ± 580)	0.54 ± 0.09 (537 ± 540)	0.60 ± 0.03 (11.6K ± 3.2K)
FetNAS	0.14	0.55 ± 0.09 (40 ± 10)	0.60 ± 0.05 (166 ± 225)	0.60 ± 0.05 (288 ± 303)	0.60 ± 0.05 (303 ± 290)	0.61 ± 0.06 (3.7K ± 5.3K)
RoT-mRMR$_{20}$	0.20					0.57 ± 0.08 (1.3K ± 0)
Evo-UF	0.22	0.46 ± 0.09 (52 ± 11)	0.46 ± 0.09 (52 ± 11)	0.45 ± 0.09 (67 ± 40)	0.49 ± 0.11 (147 ± 138)	0.62 ± 0.04 (69.0K ± 70.9K)
FetNAS	0.14	0.57 ± 0.06 (47 ± 13)	0.57 ± 0.06 (47 ± 13)	0.57 ± 0.06 (47 ± 13)	0.57 ± 0.06 (66 ± 42)	0.62 ± 0.05 (13.6K ± 14.3K)
RoT-UF	0.22					0.52 ± 0.07 (1.3K ± 90)
customer_satisfaction_in_airline						
Evo-All	1.00	0.56 ± 0.05 (190 ± 10)	0.62 ± 0.07 (1.1K ± 955)	0.69 ± 0.06 (12.8K ± 22.8K)	0.73 ± 0.05 (23.7K ± 24.6K)	0.74 ± 0.05 (54.5K ± 39.4K)
FetNAS	0.47	0.46 ± 0.15 (114 ± 85)	0.67 ± 0.19 (1.1K ± 938)	0.76 ± 0.11 (10.4K ± 17.1K)	0.77 ± 0.11 (16.3K ± 18.6K)	0.77 ± 0.12 (27.5K ± 25.6K)
RoT-All	1.00					0.91 ± 0.01 (132.5K ± 0)
Evo-mRMR$_{80}$	0.81	0.56 ± 0.07 (161 ± 10)	0.66 ± 0.08 (3.4K ± 3.9K)	0.74 ± 0.02 (10.0K ± 6.2K)	0.76 ± 0.02 (33.8K ± 16.7K)	0.77 ± 0.02 (70.4K ± 25.8K)
FetNAS	0.49	0.42 ± 0.16 (48 ± 50)	0.66 ± 0.23 (2.1K ± 1.9K)	0.76 ± 0.11 (7.1K ± 7.0K)	0.77 ± 0.12 (23.8K ± 20.5K)	0.77 ± 0.12 (36.6K ± 34.5K)
RoT-mRMR$_{80}$	0.81					0.92 ± 0.00 (129.5K ± 0)
Evo-mRMR$_{60}$	0.62	0.56 ± 0.09 (127 ± 8)	0.63 ± 0.06 (1.6K ± 1.2K)	0.68 ± 0.02 (5.2K ± 3.7K)	0.71 ± 0.02 (12.9K ± 8.0K)	0.72 ± 0.02 (47.1K ± 27.8K)
FetNAS	0.32	0.42 ± 0.15 (43 ± 39)	0.65 ± 0.19 (1.2K ± 698)	0.69 ± 0.15 (2.0K ± 1.1K)	0.69 ± 0.15 (4.4K ± 5.8K)	0.72 ± 0.13 (20.1K ± 25.9K)
RoT-mRMR$_{60}$	0.62					0.91 ± 0.00 (126.5K ± 0)
Evo-mRMR$_{40}$	0.43	0.65 ± 0.10 (96 ± 10)	0.75 ± 0.05 (984 ± 324)	0.80 ± 0.04 (3.1K ± 743)	0.83 ± 0.03 (10.8K ± 3.4K)	0.84 ± 0.03 (62.9K ± 24.4K)
FetNAS	0.29	0.42 ± 0.15 (38 ± 28)	0.68 ± 0.11 (832 ± 356)	0.74 ± 0.09 (2.1K ± 291)	0.74 ± 0.09 (2.1K ± 291)	0.76 ± 0.10 (31.1K ± 29.4K)
RoT-mRMR$_{40}$	0.43					0.89 ± 0.01 (123.4K ± 0)
Evo-mRMR$_{20}$	0.24	0.66 ± 0.05 (61 ± 6)	0.77 ± 0.07 (1.3K ± 1.4K)	0.81 ± 0.05 (3.4K ± 2.3K)	0.84 ± 0.04 (10.4K ± 3.8K)	0.86 ± 0.03 (56.5K ± 20.5K)
FetNAS	0.18	0.41 ± 0.12 (29 ± 7)	0.64 ± 0.14 (734 ± 486)	0.70 ± 0.06 (1.4K ± 224)	0.70 ± 0.06 (4.0K ± 5.6K)	0.71 ± 0.07 (22.9K ± 21.1K)
RoT-mRMR$_{20}$	0.24					0.88 ± 0.00 (120.4K ± 0)
Evo-UF	0.19	0.45 ± 0.09 (49 ± 0)	0.47 ± 0.11 (86 ± 46)	0.49 ± 0.12 (153 ± 152)	0.51 ± 0.15 (737 ± 1.1K)	0.72 ± 0.05 (16.7K ± 25.6K)
FetNAS	0.10	0.41 ± 0.12 (29 ± 7)	0.41 ± 0.12 (29 ± 7)	0.43 ± 0.14 (123 ± 150)	0.48 ± 0.18 (423 ± 617)	0.69 ± 0.06 (1.1K ± 201)
RoT-UF	0.19					0.88 ± 0.00 (119.6K ± 0)
diabetes130us						
Evo-All	1.00	0.48 ± 0.00 (403 ± 22)	0.48 ± 0.00 (403 ± 22)	0.48 ± 0.00 (529 ± 254)	0.48 ± 0.01 (740 ± 515)	0.49 ± 0.01 (24.4K ± 29.0K)
FetNAS	0.06	0.49 ± 0.01 (157 ± 158)	0.49 ± 0.01 (157 ± 158)	0.49 ± 0.01 (157 ± 158)	0.49 ± 0.01 (519 ± 780)	0.49 ± 0.01 (3.6K ± 5.6K)
RoT-All	1.00					0.48 ± 0.00 (91.0K ± 0)
Evo-mRMR$_{80}$	0.81	0.48 ± 0.00 (337 ± 36)	0.48 ± 0.00 (1.2K ± 1.1K)	0.48 ± 0.00 (1.4K ± 1.7K)	0.48 ± 0.01 (2.4K ± 2.8K)	0.50 ± 0.01 (48.2K ± 71.6K)

Method	k/d	F1-macro (parameter budget in parentheses)				
		min	Q1	Q2	Q3	max
FetNAS	0.10	0.48 ± 0.01 (97 ± 118)	0.49 ± 0.01 (790 ± 1.1K)	0.49 ± 0.01 (790 ± 1.1K)	0.49 ± 0.01 (790 ± 1.1K)	0.49 ± 0.01 (28.9K ± 49.0K)
RoT-mRMR$_{80}$	0.81					0.48 ± 0.00 (86.0K ± 0)
Evo-mRMR$_{60}$	0.62	0.48 ± 0.00 (249 ± 0)	0.48 ± 0.00 (954 ± 1.2K)	0.48 ± 0.00 (1.1K ± 1.2K)	0.48 ± 0.00 (1.1K ± 1.2K)	0.49 ± 0.01 (11.1K ± 7.9K)
FetNAS	0.07	0.48 ± 0.01 (61 ± 42)	0.49 ± 0.01 (605 ± 1.1K)	0.49 ± 0.01 (737 ± 1.0K)	0.49 ± 0.01 (737 ± 1.0K)	0.49 ± 0.01 (2.4K ± 1.6K)
RoT-mRMR$_{60}$	0.62					0.48 ± 0.00 (80.9K ± 0)
Evo-mRMR$_{40}$	0.40	0.48 ± 0.00 (169 ± 0)	0.48 ± 0.00 (169 ± 0)	0.48 ± 0.00 (169 ± 0)	0.48 ± 0.00 (169 ± 0)	0.48 ± 0.01 (39.1K ± 37.5K)
FetNAS	0.07	0.48 ± 0.01 (61 ± 42)	0.48 ± 0.01 (61 ± 42)	0.48 ± 0.01 (61 ± 42)	0.48 ± 0.01 (61 ± 42)	0.49 ± 0.01 (4.5K ± 5.0K)
RoT-mRMR$_{40}$	0.40					0.48 ± 0.00 (75.2K ± 0)
Evo-mRMR$_{20}$	0.21	0.48 ± 0.00 (99 ± 5)	0.48 ± 0.00 (1.2K ± 2.1K)	0.48 ± 0.00 (1.2K ± 2.1K)	0.48 ± 0.01 (2.7K ± 3.7K)	0.48 ± 0.01 (25.9K ± 16.9K)
FetNAS	0.06	0.48 ± 0.01 (47 ± 21)	0.48 ± 0.01 (667 ± 1.1K)	0.48 ± 0.01 (667 ± 1.1K)	0.49 ± 0.01 (1.4K ± 2.0K)	0.49 ± 0.01 (7.7K ± 7.6K)
RoT-mRMR$_{20}$	0.21					0.48 ± 0.00 (70.2K ± 0)
Evo-UF	0.10	0.48 ± 0.00 (59 ± 17)	0.48 ± 0.00 (308 ± 345)	0.48 ± 0.00 (308 ± 345)	0.48 ± 0.00 (476 ± 618)	0.49 ± 0.01 (16.3K ± 17.3K)
FetNAS	0.06	0.48 ± 0.00 (43 ± 22)	0.48 ± 0.00 (43 ± 22)	0.48 ± 0.00 (43 ± 22)	0.48 ± 0.01 (193 ± 356)	0.49 ± 0.01 (7.2K ± 8.2K)
RoT-UF	0.10					0.48 ± 0.00 (67.2K ± 1.0K)
e_commereshippingdata						
Evo-All	1.00	0.61 ± 0.03 (99 ± 5)	0.63 ± 0.04 (462 ± 537)	0.64 ± 0.02 (1.3K ± 1.6K)	0.65 ± 0.02 (2.5K ± 2.5K)	0.66 ± 0.01 (28.9K ± 43.3K)
FetNAS	0.48	0.65 ± 0.01 (72 ± 9)	0.66 ± 0.01 (217 ± 279)	0.66 ± 0.01 (1.3K ± 1.8K)	0.66 ± 0.01 (1.4K ± 1.7K)	0.66 ± 0.01 (6.4K ± 3.9K)
RoT-All	1.00					0.68 ± 0.01 (12.5K ± 0)
Evo-mRMR$_{80}$	0.80	0.60 ± 0.03 (85 ± 9)	0.63 ± 0.02 (454 ± 432)	0.64 ± 0.03 (1.1K ± 1.2K)	0.65 ± 0.03 (4.4K ± 7.4K)	0.66 ± 0.01 (34.9K ± 28.8K)
FetNAS	0.48	0.65 ± 0.01 (67 ± 6)	0.66 ± 0.01 (168 ± 133)	0.66 ± 0.01 (1.1K ± 1.9K)	0.66 ± 0.01 (2.6K ± 4.7K)	0.66 ± 0.01 (7.1K ± 4.0K)
RoT-mRMR$_{80}$	0.80					0.67 ± 0.00 (12.1K ± 0)
Evo-mRMR$_{60}$	0.60	0.61 ± 0.02 (68 ± 4)	0.62 ± 0.03 (14.6K ± 31.2K)	0.62 ± 0.03 (16.1K ± 33.6K)	0.63 ± 0.03 (16.3K ± 33.5K)	0.66 ± 0.01 (45.4K ± 26.8K)
FetNAS	0.43	0.65 ± 0.01 (64 ± 3)	0.66 ± 0.00 (552 ± 596)	0.66 ± 0.00 (552 ± 596)	0.66 ± 0.00 (1.1K ± 1.7K)	0.66 ± 0.01 (18.7K ± 20.8K)
RoT-mRMR$_{60}$	0.60					0.67 ± 0.01 (11.6K ± 0)
Evo-mRMR$_{40}$	0.40	0.58 ± 0.07 (50 ± 3)	0.63 ± 0.03 (136 ± 161)	0.63 ± 0.03 (136 ± 161)	0.63 ± 0.03 (136 ± 161)	0.65 ± 0.02 (23.4K ± 30.0K)
FetNAS	0.32	0.65 ± 0.01 (45 ± 5)	0.65 ± 0.01 (110 ± 136)	0.65 ± 0.01 (110 ± 136)	0.65 ± 0.01 (110 ± 136)	0.67 ± 0.00 (13.2K ± 21.4K)
RoT-mRMR$_{40}$	0.40					0.65 ± 0.02 (11.2K ± 0)
Evo-mRMR$_{20}$	0.20	0.59 ± 0.06 (35 ± 2)	0.61 ± 0.02 (55 ± 44)	0.61 ± 0.03 (55 ± 44)	0.63 ± 0.02 (393 ± 514)	0.63 ± 0.02 (29.5K ± 16.1K)
FetNAS	0.16	0.58 ± 0.12 (31 ± 3)	0.59 ± 0.12 (41 ± 25)	0.59 ± 0.12 (41 ± 25)	0.59 ± 0.12 (278 ± 429)	0.65 ± 0.01 (8.0K ± 4.8K)
RoT-mRMR$_{20}$	0.20					0.64 ± 0.00 (10.7K ± 0)
Evo-UF	0.20	0.54 ± 0.12 (34 ± 2)	0.60 ± 0.13 (256 ± 244)	0.60 ± 0.13 (256 ± 244)	0.60 ± 0.13 (256 ± 244)	0.67 ± 0.00 (30.3K ± 23.0K)
FetNAS	0.16	0.58 ± 0.12 (31 ± 3)	0.59 ± 0.12 (141 ± 205)	0.59 ± 0.12 (141 ± 205)	0.59 ± 0.12 (141 ± 205)	0.65 ± 0.01 (5.2K ± 4.3K)
RoT-UF	0.20					0.67 ± 0.01 (10.7K ± 0)
fitness_club						
Evo-All	1.00	0.69 ± 0.07 (65 ± 0)	0.70 ± 0.03 (640 ± 1.0K)	0.70 ± 0.03 (714 ± 1.0K)	0.70 ± 0.02 (833 ± 1.0K)	0.69 ± 0.02 (45.9K ± 41.2K)
FetNAS	0.29	0.72 ± 0.01 (35 ± 10)	0.72 ± 0.01 (43 ± 18)	0.72 ± 0.01 (43 ± 18)	0.72 ± 0.01 (43 ± 18)	0.72 ± 0.01 (17.7K ± 18.1K)
RoT-All	1.00					0.72 ± 0.02 (2.0K ± 0)
Evo-mRMR$_{80}$	0.83	0.69 ± 0.03 (58 ± 3)	0.71 ± 0.02 (144 ± 109)	0.71 ± 0.02 (147 ± 107)	0.71 ± 0.02 (231 ± 176)	0.70 ± 0.03 (28.4K ± 60.3K)
FetNAS	0.34	0.72 ± 0.01 (35 ± 10)	0.72 ± 0.01 (70 ± 61)	0.72 ± 0.01 (70 ± 61)	0.72 ± 0.01 (70 ± 61)	0.72 ± 0.02 (5.0K ± 10.3K)
RoT-mRMR$_{80}$	0.83					0.71 ± 0.02 (1.9K ± 0)
Evo-mRMR$_{60}$	0.67	0.69 ± 0.04 (50 ± 3)	0.70 ± 0.04 (563 ± 948)	0.70 ± 0.04 (563 ± 948)	0.71 ± 0.03 (848 ± 968)	0.70 ± 0.03 (8.1K ± 15.8K)
FetNAS	0.25	0.72 ± 0.01 (35 ± 10)	0.72 ± 0.01 (334 ± 648)	0.72 ± 0.01 (334 ± 648)	0.72 ± 0.01 (334 ± 648)	0.72 ± 0.01 (3.5K ± 6.5K)
RoT-mRMR$_{60}$	0.67					0.72 ± 0.02 (1.8K ± 0)
Evo-mRMR$_{40}$	0.50	0.71 ± 0.02 (43 ± 4)	0.70 ± 0.03 (245 ± 213)	0.71 ± 0.02 (295 ± 240)	0.71 ± 0.02 (348 ± 203)	0.71 ± 0.02 (6.9K ± 12.2K)
FetNAS	0.27	0.73 ± 0.01 (32 ± 6)	0.72 ± 0.01 (43 ± 18)	0.72 ± 0.01 (43 ± 18)	0.72 ± 0.01 (334 ± 648)	0.72 ± 0.02 (766 ± 1.0K)

Method	k/d	F1-macro (parameter budget in parentheses) min	Q1	Q2	Q3	max
RoT-mRMR$_{40}$	0.50					0.72 ± 0.02 (1.8K ± 0)
Evo-mRMR$_{20}$	0.33	0.71 ± 0.03 (34 ± 2)	0.71 ± 0.02 (219 ± 352)	0.71 ± 0.02 (219 ± 352)	0.71 ± 0.02 (219 ± 352)	0.72 ± 0.02 (10.2K ± 8.6K)
FetNAS	0.21	0.73 ± 0.01 (30 ± 3)	0.73 ± 0.01 (34 ± 9)	0.73 ± 0.01 (34 ± 9)	0.73 ± 0.01 (34 ± 9)	0.72 ± 0.02 (3.8K ± 7.3K)
RoT-mRMR$_{20}$	0.33					0.72 ± 0.01 (1.7K ± 0)
Evo-UF	0.23	0.65 ± 0.14 (28 ± 4)	0.65 ± 0.14 (28 ± 4)	0.65 ± 0.14 (28 ± 4)	0.65 ± 0.14 (30 ± 7)	0.71 ± 0.03 (20.1K ± 29.7K)
FetNAS	0.18	0.73 ± 0.02 (27 ± 3)	0.73 ± 0.02 (27 ± 3)	0.73 ± 0.02 (27 ± 3)	0.73 ± 0.02 (27 ± 3)	0.72 ± 0.01 (882 ± 995)
RoT-UF	0.23					0.71 ± 0.02 (1.6K ± 45)
givemesomecredit						
Evo-All	1.00	0.52 ± 0.03 (99 ± 5)	0.52 ± 0.03 (138 ± 72)	0.53 ± 0.02 (415 ± 291)	0.54 ± 0.03 (502 ± 326)	0.54 ± 0.02 (965 ± 945)
FetNAS	0.40	0.51 ± 0.01 (69 ± 32)	0.52 ± 0.03 (141 ± 122)	0.55 ± 0.07 (344 ± 326)	0.56 ± 0.07 (367 ± 302)	0.56 ± 0.07 (482 ± 386)
RoT-All	1.00					0.53 ± 0.07 (143.3K ± 0)
Evo-mRMR$_{80}$	0.80	0.54 ± 0.02 (87 ± 9)	0.54 ± 0.02 (123 ± 35)	0.55 ± 0.02 (243 ± 139)	0.56 ± 0.01 (335 ± 73)	0.56 ± 0.02 (949 ± 821)
FetNAS	0.36	0.51 ± 0.01 (63 ± 27)	0.51 ± 0.01 (67 ± 32)	0.55 ± 0.07 (144 ± 101)	0.55 ± 0.07 (230 ± 81)	0.56 ± 0.06 (493 ± 289)
RoT-mRMR$_{80}$	0.80					0.58 ± 0.09 (141.7K ± 0)
Evo-mRMR$_{60}$	0.60	0.53 ± 0.01 (65 ± 0)	0.54 ± 0.01 (127 ± 43)	0.56 ± 0.03 (673 ± 546)	0.56 ± 0.03 (2.4K ± 4.0K)	0.58 ± 0.03 (3.0K ± 4.2K)
FetNAS	0.29	0.51 ± 0.01 (45 ± 17)	0.52 ± 0.03 (131 ± 125)	0.56 ± 0.06 (559 ± 490)	0.58 ± 0.06 (816 ± 885)	0.61 ± 0.05 (1.0K ± 758)
RoT-mRMR$_{60}$	0.60					0.58 ± 0.07 (140.0K ± 0)
Evo-mRMR$_{40}$	0.40	0.50 ± 0.01 (49 ± 0)	0.50 ± 0.01 (146 ± 167)	0.51 ± 0.01 (176 ± 168)	0.52 ± 0.03 (232 ± 215)	0.54 ± 0.03 (2.0K ± 3.3K)
FetNAS	0.19	0.50 ± 0.01 (32 ± 9)	0.51 ± 0.01 (51 ± 40)	0.55 ± 0.06 (155 ± 148)	0.55 ± 0.07 (179 ± 160)	0.58 ± 0.07 (369 ± 291)
RoT-mRMR$_{40}$	0.40					0.60 ± 0.06 (138.4K ± 0)
Evo-mRMR$_{20}$	0.20	0.48 ± 0.00 (33 ± 0)	0.48 ± 0.00 (97 ± 133)	0.48 ± 0.00 (97 ± 133)	0.48 ± 0.00 (99 ± 132)	0.48 ± 0.00 (619 ± 911)
FetNAS	0.11	0.50 ± 0.00 (26 ± 3)	0.50 ± 0.00 (74 ± 95)	0.50 ± 0.00 (74 ± 95)	0.50 ± 0.00 (74 ± 95)	0.51 ± 0.01 (408 ± 508)
RoT-mRMR$_{20}$	0.20					0.48 ± 0.00 (136.8K ± 0)
Evo-UF	0.40	0.50 ± 0.01 (51 ± 5)	0.51 ± 0.01 (153 ± 138)	0.52 ± 0.02 (1.1K ± 1.4K)	0.52 ± 0.02 (2.4K ± 3.9K)	0.53 ± 0.02 (11.6K ± 15.0K)
FetNAS	0.20	0.51 ± 0.01 (37 ± 14)	0.51 ± 0.01 (186 ± 197)	0.55 ± 0.06 (313 ± 324)	0.56 ± 0.06 (567 ± 729)	0.60 ± 0.06 (5.0K ± 7.5K)
RoT-UF	0.40					0.61 ± 0.05 (138.4K ± 0)
hazelnut_spread_contaminant_detection						
Evo-All	1.00	0.75 ± 0.07 (263 ± 14)	0.77 ± 0.05 (437 ± 284)	0.77 ± 0.05 (437 ± 284)	0.77 ± 0.05 (485 ± 355)	0.89 ± 0.01 (41.4K ± 37.6K)
FetNAS	0.38	0.81 ± 0.03 (156 ± 34)	0.81 ± 0.03 (276 ± 278)	0.81 ± 0.03 (276 ± 278)	0.82 ± 0.03 (338 ± 416)	0.83 ± 0.02 (12.1K ± 16.4K)
RoT-All	1.00					0.92 ± 0.01 (5.5K ± 0)
Evo-mRMR$_{80}$	0.80	0.79 ± 0.04 (219 ± 23)	0.79 ± 0.04 (268 ± 117)	0.80 ± 0.05 (1.2K ± 2.2K)	0.81 ± 0.05 (2.1K ± 2.6K)	0.87 ± 0.02 (82.0K ± 42.2K)
FetNAS	0.38	0.81 ± 0.03 (156 ± 34)	0.81 ± 0.03 (156 ± 34)	0.82 ± 0.03 (301 ± 339)	0.82 ± 0.02 (676 ± 835)	0.82 ± 0.03 (21.1K ± 17.1K)
RoT-mRMR$_{80}$	0.80					0.92 ± 0.01 (4.8K ± 0)
Evo-mRMR$_{60}$	0.60	0.74 ± 0.07 (161 ± 0)	0.78 ± 0.05 (4.7K ± 8.4K)	0.78 ± 0.05 (4.7K ± 8.4K)	0.78 ± 0.05 (4.7K ± 8.4K)	0.85 ± 0.04 (65.7K ± 62.9K)
FetNAS	0.34	0.81 ± 0.04 (129 ± 15)	0.81 ± 0.03 (3.2K ± 6.3K)	0.81 ± 0.03 (3.2K ± 6.3K)	0.81 ± 0.03 (3.2K ± 6.3K)	0.82 ± 0.03 (34.8K ± 42.4K)
RoT-mRMR$_{60}$	0.60					0.90 ± 0.01 (4.2K ± 0)
Evo-mRMR$_{40}$	0.40	0.76 ± 0.05 (119 ± 13)	0.78 ± 0.04 (329 ± 379)	0.77 ± 0.04 (958 ± 891)	0.80 ± 0.04 (3.0K ± 5.2K)	0.83 ± 0.03 (67.4K ± 70.4K)
FetNAS	0.25	0.78 ± 0.04 (93 ± 20)	0.80 ± 0.03 (117 ± 32)	0.81 ± 0.01 (376 ± 420)	0.81 ± 0.01 (376 ± 420)	0.81 ± 0.02 (31.2K ± 46.4K)
RoT-mRMR$_{40}$	0.40					0.88 ± 0.01 (3.6K ± 0)
Evo-mRMR$_{20}$	0.20	0.71 ± 0.05 (71 ± 7)	0.72 ± 0.05 (441 ± 770)	0.72 ± 0.05 (594 ± 762)	0.72 ± 0.05 (594 ± 762)	0.78 ± 0.03 (29.2K ± 30.3K)
FetNAS	0.14	0.73 ± 0.06 (57 ± 17)	0.74 ± 0.07 (61 ± 20)	0.74 ± 0.07 (61 ± 20)	0.74 ± 0.07 (61 ± 20)	0.79 ± 0.04 (8.3K ± 9.8K)
RoT-mRMR$_{20}$	0.20					0.83 ± 0.01 (3.0K ± 0)
Evo-UF	0.60	0.74 ± 0.10 (165 ± 17)	0.80 ± 0.02 (963 ± 1.6K)	0.80 ± 0.02 (1.1K ± 1.6K)	0.82 ± 0.02 (1.2K ± 1.5K)	0.86 ± 0.01 (59.4K ± 53.8K)
FetNAS	0.39	0.81 ± 0.04 (125 ± 7)	0.81 ± 0.03 (622 ± 1.0K)	0.82 ± 0.03 (742 ± 989)	0.82 ± 0.02 (783 ± 994)	0.82 ± 0.03 (11.5K ± 6.8K)

Method	k/d	F1-macro (parameter budget in parentheses) min	Q1	Q2	Q3	max
RoT-UF	0.60					0.90 ± 0.02 (4.2K ± 195)
heloc						
Evo-All	1.00	0.71 ± 0.01 (206 ± 11)	0.71 ± 0.01 (406 ± 182)	0.71 ± 0.01 (406 ± 182)	0.71 ± 0.01 (406 ± 182)	0.70 ± 0.01 (21.7K ± 23.2K)
FetNAS	0.68	0.71 ± 0.01 (132 ± 47)	0.71 ± 0.01 (249 ± 183)	0.71 ± 0.01 (249 ± 183)	0.71 ± 0.01 (249 ± 183)	0.72 ± 0.01 (15.3K ± 16.7K)
RoT-All	1.00					0.72 ± 0.01 (14.8K ± 0)
Evo-mRMR$_{80}$	0.83	0.71 ± 0.01 (177 ± 19)	0.71 ± 0.01 (2.6K ± 3.9K)	0.71 ± 0.01 (2.6K ± 3.9K)	0.71 ± 0.01 (2.6K ± 3.9K)	0.71 ± 0.01 (5.1K ± 3.8K)
FetNAS	0.65	0.71 ± 0.01 (125 ± 35)	0.72 ± 0.01 (1.8K ± 3.3K)	0.72 ± 0.01 (1.8K ± 3.3K)	0.72 ± 0.01 (1.8K ± 3.3K)	0.72 ± 0.01 (3.9K ± 3.9K)
RoT-mRMR$_{80}$	0.83					0.71 ± 0.01 (13.9K ± 0)
Evo-mRMR$_{60}$	0.61	0.71 ± 0.01 (129 ± 0)	0.71 ± 0.01 (651 ± 292)	0.71 ± 0.01 (654 ± 286)	0.71 ± 0.01 (654 ± 286)	0.70 ± 0.00 (34.4K ± 37.9K)
FetNAS	0.42	0.71 ± 0.01 (103 ± 21)	0.71 ± 0.01 (181 ± 189)	0.71 ± 0.01 (181 ± 189)	0.71 ± 0.01 (181 ± 189)	0.71 ± 0.01 (17.3K ± 21.1K)
RoT-mRMR$_{60}$	0.61					0.71 ± 0.01 (12.8K ± 0)
Evo-mRMR$_{40}$	0.43	0.71 ± 0.01 (99 ± 5)	0.71 ± 0.01 (385 ± 343)	0.71 ± 0.01 (385 ± 343)	0.71 ± 0.01 (385 ± 343)	0.71 ± 0.00 (9.1K ± 13.2K)
FetNAS	0.32	0.71 ± 0.01 (78 ± 17)	0.71 ± 0.01 (129 ± 104)	0.71 ± 0.01 (129 ± 104)	0.71 ± 0.01 (129 ± 104)	0.71 ± 0.01 (3.8K ± 8.2K)
RoT-mRMR$_{40}$	0.43					0.71 ± 0.00 (12.0K ± 0)
Evo-mRMR$_{20}$	0.22	0.70 ± 0.00 (61 ± 6)	0.70 ± 0.00 (771 ± 1.1K)	0.70 ± 0.00 (864 ± 1.1K)	0.70 ± 0.00 (1.3K ± 1.3K)	0.70 ± 0.00 (18.0K ± 21.4K)
FetNAS	0.18	0.70 ± 0.01 (56 ± 1)	0.69 ± 0.01 (280 ± 346)	0.70 ± 0.01 (385 ± 360)	0.70 ± 0.01 (410 ± 385)	0.70 ± 0.01 (3.7K ± 6.2K)
RoT-mRMR$_{20}$	0.22					0.71 ± 0.00 (10.9K ± 0)
Evo-UF	0.17	0.70 ± 0.01 (56 ± 14)	0.70 ± 0.01 (1.1K ± 1.8K)	0.70 ± 0.01 (1.1K ± 1.8K)	0.70 ± 0.01 (1.1K ± 1.8K)	0.70 ± 0.01 (27.0K ± 41.2K)
FetNAS	0.15	0.68 ± 0.04 (50 ± 8)	0.68 ± 0.04 (213 ± 368)	0.68 ± 0.04 (213 ± 368)	0.68 ± 0.04 (213 ± 368)	0.68 ± 0.04 (6.0K ± 11.2K)
RoT-UF	0.17					0.70 ± 0.01 (10.7K ± 216)
hiva_agnostic						
Evo-All	1.00	0.33 ± 0.00 (13.0K ± 0)	0.33 ± 0.00 (13.0K ± 0)	0.33 ± 0.00 (13.0K ± 0)	0.33 ± 0.00 (13.0K ± 0)	0.33 ± 0.01 (160.7K ± 151.2K)
FetNAS	0.09	0.33 ± 0.00 (8.4K ± 4.1K)	0.33 ± 0.00 (8.4K ± 4.1K)	0.33 ± 0.00 (8.4K ± 4.1K)	0.33 ± 0.00 (8.4K ± 4.1K)	0.33 ± 0.01 (36.5K ± 16.7K)
RoT-All	1.00					0.33 ± 0.00 (238.7K ± 0)
Evo-mRMR$_{80}$	0.89	0.33 ± 0.00 (11.9K ± 707)	0.33 ± 0.00 (11.9K ± 707)	0.33 ± 0.00 (11.9K ± 707)	0.33 ± 0.00 (11.9K ± 707)	0.33 ± 0.01 (84.5K ± 71.4K)
FetNAS	0.09	0.33 ± 0.00 (7.4K ± 3.3K)	0.33 ± 0.00 (7.4K ± 3.3K)	0.33 ± 0.00 (7.4K ± 3.3K)	0.33 ± 0.00 (7.4K ± 3.3K)	0.33 ± 0.01 (32.0K ± 31.5K)
RoT-mRMR$_{80}$	0.89					0.33 ± 0.00 (214.4K ± 1.1K)
Evo-mRMR$_{60}$	0.78	0.33 ± 0.00 (10.2K ± 72)	0.33 ± 0.00 (10.2K ± 72)	0.33 ± 0.00 (10.2K ± 72)	0.33 ± 0.00 (10.2K ± 72)	0.33 ± 0.01 (105.8K ± 97.5K)
FetNAS	0.09	0.33 ± 0.00 (6.5K ± 2.4K)	0.33 ± 0.00 (6.5K ± 2.4K)	0.33 ± 0.00 (6.5K ± 2.4K)	0.33 ± 0.00 (6.5K ± 2.4K)	0.33 ± 0.01 (35.0K ± 19.3K)
RoT-mRMR$_{60}$	0.78					0.33 ± 0.00 (188.9K ± 1.3K)
Evo-mRMR$_{40}$	0.66	0.33 ± 0.00 (9.4K ± 880)	0.33 ± 0.00 (9.4K ± 880)	0.33 ± 0.00 (9.4K ± 880)	0.33 ± 0.00 (9.4K ± 880)	0.33 ± 0.01 (61.4K ± 71.4K)
FetNAS	0.09	0.33 ± 0.00 (6.5K ± 2.4K)	0.33 ± 0.00 (6.5K ± 2.4K)	0.33 ± 0.00 (6.5K ± 2.4K)	0.33 ± 0.00 (6.5K ± 2.4K)	0.33 ± 0.01 (24.7K ± 24.7K)
RoT-mRMR$_{40}$	0.66					0.33 ± 0.00 (160.5K ± 2.2K)
Evo-mRMR$_{20}$	0.42	0.33 ± 0.00 (5.4K ± 95)	0.33 ± 0.00 (5.4K ± 95)	0.33 ± 0.00 (5.4K ± 95)	0.33 ± 0.00 (5.4K ± 95)	0.33 ± 0.00 (87.5K ± 39.7K)
FetNAS	0.04	0.33 ± 0.00 (1.4K ± 2.0K)	0.33 ± 0.00 (1.4K ± 2.0K)	0.33 ± 0.00 (1.4K ± 2.0K)	0.33 ± 0.00 (1.4K ± 2.0K)	0.33 ± 0.01 (35.9K ± 23.3K)
RoT-mRMR$_{20}$	0.42					0.33 ± 0.00 (105.1K ± 1.7K)
Evo-UF	0.13	0.33 ± 0.00 (1.8K ± 283)	0.33 ± 0.00 (1.8K ± 283)	0.33 ± 0.00 (1.8K ± 283)	0.33 ± 0.00 (1.8K ± 283)	0.33 ± 0.00 (33.7K ± 32.4K)
FetNAS	0.01	0.33 ± 0.00 (61 ± 11)	0.33 ± 0.00 (61 ± 11)	0.33 ± 0.00 (61 ± 11)	0.33 ± 0.00 (61 ± 11)	0.33 ± 0.00 (19.7K ± 26.1K)
RoT-UF	0.09					0.33 ± 0.00 (28.7K ± 6.6K)
hr_analytics_job_change_of_data_scientists						
Evo-All	1.00	0.64 ± 0.01 (113 ± 0)	0.69 ± 0.02 (290 ± 97)	0.70 ± 0.01 (610 ± 447)	0.70 ± 0.01 (1.2K ± 897)	0.71 ± 0.01 (5.7K ± 3.2K)
FetNAS	0.44	0.68 ± 0.01 (66 ± 32)	0.68 ± 0.01 (184 ± 86)	0.69 ± 0.01 (283 ± 180)	0.70 ± 0.01 (689 ± 728)	0.70 ± 0.01 (1.8K ± 1.1K)
RoT-All	1.00					0.71 ± 0.01 (21.1K ± 0)
Evo-mRMR$_{80}$	0.83	0.56 ± 0.05 (97 ± 0)	0.62 ± 0.06 (316 ± 124)	0.64 ± 0.04 (519 ± 343)	0.65 ± 0.04 (937 ± 495)	0.65 ± 0.04 (3.2K ± 2.0K)
FetNAS	0.45	0.68 ± 0.01 (65 ± 30)	0.68 ± 0.01 (234 ± 114)	0.69 ± 0.01 (286 ± 184)	0.69 ± 0.01 (436 ± 367)	0.70 ± 0.01 (1.8K ± 2.2K)

Method	k/d	F1-macro (parameter budget in parentheses) min	Q1	Q2	Q3	max
RoT-mRMR$_{80}$	0.83					0.60 ± 0.11 (20.6K ± 0)
Evo-mRMR$_{60}$	0.67	0.55 ± 0.06 (81 ± 0)	0.59 ± 0.08 (229 ± 138)	0.61 ± 0.08 (717 ± 619)	0.64 ± 0.08 (1.6K ± 1.1K)	0.67 ± 0.04 (15.8K ± 18.0K)
FetNAS	0.39	0.67 ± 0.01 (51 ± 21)	0.68 ± 0.01 (172 ± 129)	0.70 ± 0.01 (337 ± 203)	0.70 ± 0.01 (337 ± 203)	0.71 ± 0.01 (8.2K ± 10.6K)
RoT-mRMR$_{60}$	0.67					0.61 ± 0.07 (20.0K ± 0)
Evo-mRMR$_{40}$	0.42	0.51 ± 0.09 (61 ± 6)	0.54 ± 0.11 (198 ± 164)	0.56 ± 0.12 (702 ± 842)	0.57 ± 0.13 (1.7K ± 2.2K)	0.67 ± 0.02 (14.2K ± 9.2K)
FetNAS	0.25	0.66 ± 0.02 (36 ± 7)	0.67 ± 0.01 (107 ± 74)	0.68 ± 0.01 (206 ± 210)	0.68 ± 0.01 (577 ± 944)	0.70 ± 0.01 (5.2K ± 2.7K)
RoT-mRMR$_{40}$	0.42					0.55 ± 0.13 (19.1K ± 0)
Evo-mRMR$_{20}$	0.25	0.50 ± 0.10 (41 ± 0)	0.52 ± 0.12 (126 ± 175)	0.52 ± 0.12 (126 ± 175)	0.52 ± 0.12 (126 ± 175)	0.66 ± 0.02 (48.3K ± 47.4K)
FetNAS	0.13	0.67 ± 0.01 (34 ± 5)	0.68 ± 0.01 (67 ± 60)	0.68 ± 0.01 (67 ± 60)	0.68 ± 0.01 (67 ± 60)	0.68 ± 0.01 (9.2K ± 15.4K)
RoT-mRMR$_{20}$	0.25					0.64 ± 0.04 (18.5K ± 0)
Evo-UF	0.20	0.66 ± 0.01 (36 ± 4)	0.67 ± 0.02 (9.2K ± 20.5K)	0.67 ± 0.02 (9.2K ± 20.5K)	0.67 ± 0.02 (9.2K ± 20.5K)	0.69 ± 0.02 (13.6K ± 20.4K)
FetNAS	0.13	0.67 ± 0.01 (29 ± 4)	0.68 ± 0.01 (203 ± 349)	0.68 ± 0.01 (203 ± 349)	0.68 ± 0.01 (203 ± 349)	0.68 ± 0.01 (203 ± 349)
RoT-UF	0.20					0.68 ± 0.01 (18.3K ± 160)
in_vehicle_coupon_recommendation						
Evo-All	1.00	0.62 ± 0.01 (219 ± 23)	0.63 ± 0.01 (1.2K ± 848)	0.63 ± 0.01 (1.2K ± 840)	0.63 ± 0.01 (1.2K ± 841)	0.63 ± 0.01 (18.0K ± 12.3K)
FetNAS	0.51	0.63 ± 0.01 (171 ± 49)	0.63 ± 0.01 (571 ± 457)	0.63 ± 0.01 (571 ± 457)	0.63 ± 0.01 (571 ± 457)	0.63 ± 0.01 (9.0K ± 6.4K)
RoT-All	1.00					0.60 ± 0.03 (17.5K ± 0)
Evo-mRMR$_{80}$	0.83	0.61 ± 0.01 (186 ± 20)	0.61 ± 0.01 (1.8K ± 3.6K)	0.61 ± 0.01 (1.8K ± 3.6K)	0.61 ± 0.01 (1.8K ± 3.6K)	0.64 ± 0.01 (55.5K ± 48.9K)
FetNAS	0.47	0.62 ± 0.01 (144 ± 40)	0.63 ± 0.00 (1.7K ± 3.4K)	0.63 ± 0.00 (1.7K ± 3.4K)	0.63 ± 0.00 (1.7K ± 3.4K)	0.64 ± 0.01 (17.2K ± 6.6K)
RoT-mRMR$_{80}$	0.83					0.61 ± 0.03 (16.5K ± 0)
Evo-mRMR$_{60}$	0.62	0.62 ± 0.01 (137 ± 0)	0.62 ± 0.01 (1.3K ± 1.9K)	0.62 ± 0.01 (1.3K ± 1.9K)	0.62 ± 0.01 (1.3K ± 1.9K)	0.64 ± 0.00 (20.1K ± 14.9K)
FetNAS	0.42	0.62 ± 0.00 (104 ± 19)	0.62 ± 0.01 (1.1K ± 1.8K)	0.62 ± 0.01 (1.1K ± 1.8K)	0.62 ± 0.01 (1.1K ± 1.8K)	0.64 ± 0.01 (13.6K ± 9.6K)
RoT-mRMR$_{60}$	0.62					0.64 ± 0.01 (15.3K ± 0)
Evo-mRMR$_{40}$	0.42	0.61 ± 0.01 (97 ± 0)	0.61 ± 0.01 (234 ± 147)	0.61 ± 0.01 (234 ± 147)	0.61 ± 0.01 (234 ± 147)	0.64 ± 0.02 (17.3K ± 3.9K)
FetNAS	0.25	0.61 ± 0.01 (86 ± 14)	0.62 ± 0.01 (184 ± 134)	0.62 ± 0.01 (184 ± 134)	0.62 ± 0.01 (184 ± 134)	0.63 ± 0.01 (9.3K ± 8.0K)
RoT-mRMR$_{40}$	0.42					0.64 ± 0.01 (14.1K ± 0)
Evo-mRMR$_{20}$	0.21	0.57 ± 0.04 (58 ± 3)	0.58 ± 0.01 (85 ± 55)	0.58 ± 0.02 (166 ± 174)	0.59 ± 0.01 (235 ± 317)	0.63 ± 0.02 (48.2K ± 35.3K)
FetNAS	0.17	0.59 ± 0.01 (51 ± 15)	0.60 ± 0.01 (64 ± 37)	0.61 ± 0.02 (93 ± 52)	0.61 ± 0.02 (215 ± 318)	0.63 ± 0.02 (20.5K ± 23.8K)
RoT-mRMR$_{20}$	0.21					0.61 ± 0.04 (12.9K ± 0)
Evo-UF	0.09	0.56 ± 0.02 (36 ± 12)	0.56 ± 0.02 (37 ± 15)	0.56 ± 0.02 (44 ± 31)	0.56 ± 0.02 (44 ± 31)	0.61 ± 0.01 (23.0K ± 46.9K)
FetNAS	0.08	0.56 ± 0.03 (32 ± 9)	0.56 ± 0.03 (35 ± 15)	0.56 ± 0.03 (40 ± 25)	0.56 ± 0.03 (40 ± 25)	0.59 ± 0.04 (17.5K ± 39.0K)
RoT-UF	0.09					0.61 ± 0.01 (12.3K ± 391)
is_this_a_good_customer						
Evo-All	1.00	0.47 ± 0.00 (121 ± 0)	0.48 ± 0.01 (445 ± 724)	0.48 ± 0.01 (445 ± 724)	0.48 ± 0.01 (445 ± 724)	0.56 ± 0.05 (27.3K ± 16.9K)
FetNAS	0.33	0.48 ± 0.02 (42 ± 35)	0.49 ± 0.03 (253 ± 465)	0.49 ± 0.03 (253 ± 465)	0.49 ± 0.03 (253 ± 465)	0.52 ± 0.03 (6.0K ± 9.7K)
RoT-All	1.00					0.47 ± 0.00 (2.9K ± 0)
Evo-mRMR$_{80}$	0.85	0.47 ± 0.00 (110 ± 7)	0.47 ± 0.00 (360 ± 563)	0.47 ± 0.00 (360 ± 563)	0.47 ± 0.00 (360 ± 563)	0.52 ± 0.04 (41.8K ± 38.5K)
FetNAS	0.27	0.48 ± 0.02 (42 ± 35)	0.49 ± 0.03 (253 ± 465)	0.49 ± 0.03 (253 ± 465)	0.49 ± 0.03 (253 ± 465)	0.52 ± 0.05 (8.6K ± 12.1K)
RoT-mRMR$_{80}$	0.85					0.47 ± 0.00 (2.7K ± 0)
Evo-mRMR$_{60}$	0.62	0.47 ± 0.01 (81 ± 0)	0.49 ± 0.04 (141 ± 134)	0.49 ± 0.04 (141 ± 134)	0.49 ± 0.04 (141 ± 134)	0.49 ± 0.03 (13.4K ± 23.3K)
FetNAS	0.18	0.48 ± 0.03 (35 ± 21)	0.48 ± 0.03 (98 ± 135)	0.48 ± 0.03 (98 ± 135)	0.48 ± 0.03 (98 ± 135)	0.50 ± 0.05 (1.9K ± 4.2K)
RoT-mRMR$_{60}$	0.62					0.47 ± 0.00 (2.4K ± 0)
Evo-mRMR$_{40}$	0.46	0.47 ± 0.00 (65 ± 0)	0.47 ± 0.00 (65 ± 0)	0.47 ± 0.00 (65 ± 0)	0.47 ± 0.00 (65 ± 0)	0.49 ± 0.03 (13.9K ± 25.6K)
FetNAS	0.10	0.47 ± 0.00 (26 ± 1)	0.47 ± 0.00 (26 ± 1)	0.47 ± 0.00 (26 ± 1)	0.47 ± 0.00 (26 ± 1)	0.50 ± 0.05 (3.8K ± 6.7K)
RoT-mRMR$_{40}$	0.46					0.47 ± 0.00 (2.2K ± 0)

Method	k/d	F1-macro (parameter budget in parentheses) min	Q1	Q2	Q3	max
Evo-mRMR$_{20}$	0.23	0.47 ± 0.00 (41 ± 0)	0.47 ± 0.00 (41 ± 0)	0.47 ± 0.00 (41 ± 0)	0.47 ± 0.00 (41 ± 0)	0.49 ± 0.04 (6.3K ± 10.3K)
FetNAS	0.08	0.47 ± 0.00 (26 ± 1)	0.47 ± 0.00 (26 ± 1)	0.47 ± 0.00 (26 ± 1)	0.47 ± 0.00 (26 ± 1)	0.47 ± 0.00 (524 ± 950)
RoT-mRMR$_{20}$	0.23					0.47 ± 0.00 (2.0K ± 0)
Evo-UF	0.22	0.47 ± 0.00 (39 ± 15)	0.47 ± 0.00 (39 ± 15)	0.47 ± 0.00 (39 ± 15)	0.47 ± 0.00 (39 ± 15)	0.47 ± 0.00 (39 ± 15)
FetNAS	0.08	0.47 ± 0.00 (26 ± 1)	0.47 ± 0.00 (26 ± 1)	0.47 ± 0.00 (26 ± 1)	0.47 ± 0.00 (26 ± 1)	0.47 ± 0.00 (26 ± 1)
RoT-UF	0.22					0.47 ± 0.00 (2.0K ± 169)
jm1						
Evo-All	1.00	0.50 ± 0.03 (185 ± 0)	0.53 ± 0.05 (432 ± 326)	0.53 ± 0.05 (460 ± 386)	0.53 ± 0.05 (460 ± 386)	0.62 ± 0.02 (25.2K ± 13.8K)
FetNAS	0.26	0.52 ± 0.03 (145 ± 49)	0.53 ± 0.03 (272 ± 265)	0.53 ± 0.03 (272 ± 265)	0.53 ± 0.03 (272 ± 265)	0.56 ± 0.03 (5.2K ± 6.5K)
RoT-All	1.00					0.53 ± 0.03 (14.6K ± 0)
Evo-mRMR$_{80}$	0.81	0.50 ± 0.03 (157 ± 8)	0.51 ± 0.03 (202 ± 69)	0.52 ± 0.05 (300 ± 174)	0.52 ± 0.05 (735 ± 1.1K)	0.61 ± 0.02 (12.3K ± 14.8K)
FetNAS	0.23	0.52 ± 0.02 (92 ± 22)	0.53 ± 0.03 (145 ± 86)	0.53 ± 0.02 (195 ± 130)	0.53 ± 0.02 (382 ± 522)	0.55 ± 0.02 (1.3K ± 1.4K)
RoT-mRMR$_{80}$	0.81					0.54 ± 0.05 (13.7K ± 0)
Evo-mRMR$_{60}$	0.62	0.50 ± 0.01 (121 ± 0)	0.52 ± 0.03 (222 ± 145)	0.52 ± 0.04 (359 ± 444)	0.53 ± 0.05 (721 ± 1.3K)	0.60 ± 0.01 (32.8K ± 27.0K)
FetNAS	0.22	0.52 ± 0.02 (92 ± 22)	0.52 ± 0.02 (168 ± 105)	0.53 ± 0.03 (228 ± 232)	0.53 ± 0.03 (228 ± 232)	0.57 ± 0.03 (8.3K ± 7.7K)
RoT-mRMR$_{60}$	0.62					0.55 ± 0.03 (12.8K ± 0)
Evo-mRMR$_{40}$	0.43	0.51 ± 0.00 (89 ± 0)	0.52 ± 0.02 (286 ± 273)	0.52 ± 0.02 (286 ± 273)	0.52 ± 0.02 (288 ± 271)	0.58 ± 0.03 (29.9K ± 26.5K)
FetNAS	0.25	0.52 ± 0.02 (76 ± 5)	0.52 ± 0.02 (179 ± 146)	0.52 ± 0.02 (179 ± 146)	0.52 ± 0.02 (179 ± 146)	0.56 ± 0.03 (1.8K ± 1.8K)
RoT-mRMR$_{40}$	0.43					0.52 ± 0.02 (12.0K ± 0)
Evo-mRMR$_{20}$	0.24	0.49 ± 0.01 (57 ± 0)	0.49 ± 0.01 (92 ± 49)	0.50 ± 0.02 (168 ± 202)	0.50 ± 0.02 (182 ± 218)	0.55 ± 0.01 (17.8K ± 9.5K)
FetNAS	0.14	0.50 ± 0.03 (47 ± 10)	0.50 ± 0.03 (78 ± 48)	0.50 ± 0.03 (111 ± 98)	0.51 ± 0.04 (118 ± 93)	0.56 ± 0.03 (5.2K ± 6.5K)
RoT-mRMR$_{20}$	0.24					0.51 ± 0.04 (11.1K ± 0)
Evo-UF	0.90	0.50 ± 0.01 (172 ± 12)	0.51 ± 0.01 (492 ± 339)	0.51 ± 0.01 (492 ± 339)	0.54 ± 0.04 (2.7K ± 4.7K)	0.61 ± 0.03 (17.6K ± 22.4K)
FetNAS	0.27	0.52 ± 0.03 (121 ± 46)	0.53 ± 0.02 (231 ± 112)	0.53 ± 0.02 (231 ± 112)	0.55 ± 0.03 (499 ± 360)	0.55 ± 0.02 (4.4K ± 7.0K)
RoT-UF	0.90					0.54 ± 0.05 (14.1K ± 184)
kddcup09_appetency						
Evo-All	1.00	0.51 ± 0.02 (1.8K ± 96)	0.51 ± 0.02 (2.7K ± 1.2K)	0.51 ± 0.01 (4.3K ± 2.6K)	0.51 ± 0.01 (7.5K ± 6.4K)	0.51 ± 0.01 (31.0K ± 26.1K)
FetNAS	0.12	0.52 ± 0.03 (251 ± 300)	0.52 ± 0.03 (251 ± 300)	0.52 ± 0.03 (627 ± 968)	0.52 ± 0.03 (2.7K ± 5.6K)	0.52 ± 0.03 (7.8K ± 14.7K)
RoT-All	1.00					0.50 ± 0.00 (145.6K ± 0)
Evo-mRMR$_{80}$	0.80	0.52 ± 0.01 (1.4K ± 77)	0.52 ± 0.02 (1.6K ± 186)	0.52 ± 0.02 (7.1K ± 10.8K)	0.52 ± 0.02 (13.4K ± 15.1K)	0.52 ± 0.02 (19.7K ± 15.1K)
FetNAS	0.02	0.52 ± 0.03 (251 ± 300)	0.52 ± 0.03 (251 ± 300)	0.52 ± 0.03 (251 ± 300)	0.52 ± 0.03 (627 ± 968)	0.52 ± 0.03 (627 ± 968)
RoT-mRMR$_{80}$	0.80					0.50 ± 0.00 (125.7K ± 0)
Evo-mRMR$_{60}$	0.60	0.52 ± 0.02 (1.1K ± 71)	0.52 ± 0.02 (1.3K ± 250)	0.52 ± 0.02 (2.2K ± 1.6K)	0.52 ± 0.02 (3.5K ± 3.3K)	0.53 ± 0.02 (16.2K ± 25.9K)
FetNAS	0.03	0.52 ± 0.03 (251 ± 300)	0.52 ± 0.03 (251 ± 300)	0.52 ± 0.03 (627 ± 968)	0.52 ± 0.03 (627 ± 968)	0.52 ± 0.03 (627 ± 968)
RoT-mRMR$_{60}$	0.60					0.50 ± 0.00 (105.9K ± 0)
Evo-mRMR$_{40}$	0.40	0.51 ± 0.02 (732 ± 78)	0.51 ± 0.02 (801 ± 113)	0.51 ± 0.02 (2.5K ± 3.7K)	0.51 ± 0.02 (2.7K ± 3.6K)	0.53 ± 0.03 (23.3K ± 42.6K)
FetNAS	0.03	0.52 ± 0.03 (251 ± 300)	0.52 ± 0.03 (627 ± 968)	0.52 ± 0.03 (627 ± 968)	0.52 ± 0.03 (627 ± 968)	0.52 ± 0.03 (627 ± 968)
RoT-mRMR$_{40}$	0.40					0.50 ± 0.00 (85.6K ± 0)
Evo-mRMR$_{20}$	0.20	0.52 ± 0.03 (361 ± 0)	0.53 ± 0.03 (526 ± 201)	0.53 ± 0.04 (880 ± 853)	0.53 ± 0.03 (1.5K ± 1.8K)	0.53 ± 0.03 (3.0K ± 2.1K)
FetNAS	0.02	0.51 ± 0.02 (79 ± 62)	0.52 ± 0.03 (251 ± 300)	0.52 ± 0.03 (251 ± 300)	0.52 ± 0.03 (251 ± 300)	0.52 ± 0.03 (251 ± 300)
RoT-mRMR$_{20}$	0.20					0.50 ± 0.00 (65.8K ± 0)
Evo-UF	0.03	0.52 ± 0.03 (65 ± 23)	0.53 ± 0.02 (357 ± 371)	0.54 ± 0.03 (446 ± 454)	0.53 ± 0.03 (13.5K ± 27.8K)	0.54 ± 0.02 (19.4K ± 30.0K)
FetNAS	0.01	0.50 ± 0.00 (38 ± 20)	0.50 ± 0.00 (102 ± 165)	0.50 ± 0.00 (102 ± 165)	0.50 ± 0.00 (4.0K ± 8.8K)	0.50 ± 0.00 (4.0K ± 8.8K)
RoT-UF	0.03					0.50 ± 0.00 (48.2K ± 1.2K)

Method	k/d	F1-macro (parameter budget in parentheses) min	Q1	Q2	Q3	max
marketing_campaign						
Evo-All	1.00	0.70 ± 0.06 (222 ± 12)	0.71 ± 0.02 (318 ± 117)	0.71 ± 0.02 (345 ± 111)	0.71 ± 0.03 (551 ± 431)	0.70 ± 0.04 (14.9K ± 12.1K)
FetNAS	0.57	0.69 ± 0.03 (192 ± 29)	0.68 ± 0.02 (245 ± 102)	0.68 ± 0.02 (245 ± 102)	0.69 ± 0.02 (255 ± 107)	0.70 ± 0.03 (6.5K ± 5.5K)
RoT-All	1.00					0.71 ± 0.03 (4.6K ± 0)
Evo-mRMR$_{80}$	0.80	0.70 ± 0.04 (177 ± 0)	0.71 ± 0.03 (202 ± 37)	0.71 ± 0.03 (515 ± 685)	0.71 ± 0.03 (515 ± 685)	0.70 ± 0.04 (27.3K ± 13.1K)
FetNAS	0.48	0.67 ± 0.03 (169 ± 16)	0.67 ± 0.03 (184 ± 41)	0.68 ± 0.02 (201 ± 51)	0.69 ± 0.01 (471 ± 635)	0.71 ± 0.03 (6.7K ± 5.7K)
RoT-mRMR$_{80}$	0.80					0.70 ± 0.04 (4.1K ± 0)
Evo-mRMR$_{60}$	0.60	0.64 ± 0.05 (140 ± 8)	0.64 ± 0.05 (256 ± 248)	0.64 ± 0.05 (273 ± 240)	0.64 ± 0.05 (293 ± 234)	0.68 ± 0.02 (55.2K ± 43.1K)
FetNAS	0.40	0.66 ± 0.02 (129 ± 9)	0.66 ± 0.02 (129 ± 9)	0.67 ± 0.02 (140 ± 22)	0.67 ± 0.02 (166 ± 54)	0.71 ± 0.03 (14.6K ± 22.3K)
RoT-mRMR$_{60}$	0.60					0.67 ± 0.04 (3.6K ± 0)
Evo-mRMR$_{40}$	0.40	0.62 ± 0.09 (99 ± 5)	0.62 ± 0.09 (324 ± 211)	0.64 ± 0.07 (444 ± 122)	0.64 ± 0.07 (12.9K ± 27.3K)	0.68 ± 0.04 (20.7K ± 24.3K)
FetNAS	0.32	0.62 ± 0.07 (77 ± 16)	0.64 ± 0.07 (179 ± 69)	0.66 ± 0.02 (244 ± 84)	0.67 ± 0.02 (2.2K ± 4.5K)	0.69 ± 0.03 (3.9K ± 3.7K)
RoT-mRMR$_{40}$	0.40					0.68 ± 0.05 (3.1K ± 0)
Evo-mRMR$_{20}$	0.20	0.57 ± 0.10 (60 ± 6)	0.61 ± 0.11 (182 ± 127)	0.61 ± 0.11 (182 ± 127)	0.62 ± 0.10 (184 ± 123)	0.66 ± 0.02 (34.7K ± 37.7K)
FetNAS	0.13	0.63 ± 0.05 (45 ± 14)	0.63 ± 0.04 (109 ± 90)	0.63 ± 0.04 (109 ± 90)	0.63 ± 0.04 (109 ± 90)	0.65 ± 0.04 (3.1K ± 1.8K)
RoT-mRMR$_{20}$	0.20					0.62 ± 0.03 (2.6K ± 0)
Evo-UF	0.42	0.64 ± 0.02 (106 ± 19)	0.65 ± 0.03 (578 ± 819)	0.66 ± 0.03 (743 ± 806)	0.67 ± 0.04 (1.7K ± 1.7K)	0.68 ± 0.04 (20.7K ± 14.9K)
FetNAS	0.37	0.62 ± 0.06 (99 ± 21)	0.64 ± 0.07 (359 ± 497)	0.65 ± 0.08 (377 ± 487)	0.65 ± 0.09 (584 ± 604)	0.71 ± 0.03 (3.2K ± 4.0K)
RoT-UF	0.42					0.66 ± 0.04 (3.2K ± 279)
maternal_health_risk						
Evo-All	1.00	0.54 ± 0.06 (83 ± 0)	0.58 ± 0.07 (1.6K ± 3.3K)	0.60 ± 0.05 (1.7K ± 3.3K)	0.65 ± 0.03 (8.3K ± 8.7K)	0.67 ± 0.02 (53.7K ± 10.1K)
FetNAS	0.66	0.57 ± 0.06 (62 ± 12)	0.61 ± 0.05 (976 ± 1.9K)	0.62 ± 0.05 (991 ± 1.9K)	0.62 ± 0.04 (5.4K ± 5.7K)	0.66 ± 0.03 (26.4K ± 21.6K)
RoT-All	1.00					0.67 ± 0.03 (3.0K ± 0)
Evo-mRMR$_{80}$	0.87	0.53 ± 0.05 (78 ± 5)	0.56 ± 0.03 (335 ± 347)	0.58 ± 0.04 (736 ± 625)	0.62 ± 0.02 (1.3K ± 1.3K)	0.68 ± 0.02 (52.4K ± 37.3K)
FetNAS	0.58	0.57 ± 0.06 (62 ± 12)	0.60 ± 0.03 (148 ± 140)	0.61 ± 0.03 (429 ± 560)	0.62 ± 0.02 (704 ± 817)	0.64 ± 0.02 (36.7K ± 27.2K)
RoT-mRMR$_{80}$	0.87					0.69 ± 0.03 (2.9K ± 33)
Evo-mRMR$_{60}$	0.77	0.55 ± 0.05 (75 ± 6)	0.57 ± 0.04 (285 ± 237)	0.60 ± 0.05 (1.3K ± 847)	0.65 ± 0.05 (6.5K ± 7.7K)	0.68 ± 0.03 (62.1K ± 33.2K)
FetNAS	0.53	0.57 ± 0.06 (62 ± 12)	0.57 ± 0.07 (155 ± 130)	0.59 ± 0.07 (517 ± 533)	0.60 ± 0.08 (4.5K ± 6.6K)	0.64 ± 0.02 (36.7K ± 27.2K)
RoT-mRMR$_{60}$	0.77					0.68 ± 0.00 (2.9K ± 40)
Evo-mRMR$_{40}$	0.53	0.57 ± 0.04 (62 ± 4)	0.58 ± 0.04 (1.0K ± 1.8K)	0.60 ± 0.05 (2.6K ± 4.7K)	0.63 ± 0.03 (4.7K ± 4.3K)	0.69 ± 0.02 (101.5K ± 66.0K)
FetNAS	0.48	0.57 ± 0.06 (59 ± 10)	0.59 ± 0.07 (789 ± 1.6K)	0.59 ± 0.08 (890 ± 1.5K)	0.64 ± 0.02 (3.1K ± 2.5K)	0.64 ± 0.01 (15.9K ± 8.1K)
RoT-mRMR$_{40}$	0.53					0.67 ± 0.04 (2.8K ± 33)
Evo-mRMR$_{20}$	0.47	0.56 ± 0.05 (59 ± 5)	0.59 ± 0.04 (1.0K ± 1.8K)	0.60 ± 0.06 (2.3K ± 4.8K)	0.60 ± 0.03 (3.6K ± 5.0K)	0.64 ± 0.05 (90.4K ± 77.1K)
FetNAS	0.41	0.55 ± 0.05 (56 ± 9)	0.57 ± 0.07 (790 ± 1.6K)	0.57 ± 0.07 (790 ± 1.6K)	0.61 ± 0.04 (2.0K ± 2.7K)	0.64 ± 0.01 (18.4K ± 11.4K)
RoT-mRMR$_{20}$	0.47					0.67 ± 0.03 (2.7K ± 33)
Evo-UF	0.63	0.52 ± 0.05 (65 ± 4)	0.58 ± 0.05 (328 ± 391)	0.58 ± 0.05 (1.6K ± 2.1K)	0.57 ± 0.05 (17.1K ± 36.5K)	0.65 ± 0.04 (72.8K ± 87.0K)
FetNAS	0.49	0.58 ± 0.07 (60 ± 10)	0.57 ± 0.06 (206 ± 215)	0.58 ± 0.07 (1.2K ± 1.6K)	0.58 ± 0.07 (6.0K ± 12.1K)	0.64 ± 0.01 (13.1K ± 9.7K)
RoT-UF	0.67					0.64 ± 0.03 (2.8K ± 52)
mic						
Evo-All	1.00	0.16 ± 0.04 (992 ± 54)	0.17 ± 0.03 (1.2K ± 281)	0.17 ± 0.03 (1.2K ± 281)	0.18 ± 0.03 (6.0K ± 6.8K)	0.20 ± 0.02 (82.0K ± 93.1K)
FetNAS	0.23	0.18 ± 0.03 (476 ± 367)	0.18 ± 0.03 (476 ± 367)	0.18 ± 0.03 (476 ± 367)	0.18 ± 0.04 (4.6K ± 5.8K)	0.19 ± 0.01 (27.3K ± 32.0K)
RoT-All	1.00					0.19 ± 0.01 (22.4K ± 0)
Evo-mRMR$_{80}$	0.99	0.17 ± 0.03 (958 ± 9)	0.18 ± 0.02 (1.2K ± 384)	0.18 ± 0.02 (1.2K ± 384)	0.18 ± 0.02 (4.1K ± 6.5K)	0.21 ± 0.03 (124.2K ± 61.7K)
FetNAS	0.22	0.18 ± 0.03 (476 ± 367)	0.18 ± 0.03 (476 ± 367)	0.18 ± 0.03 (476 ± 367)	0.18 ± 0.03 (2.8K ± 5.0K)	0.19 ± 0.02 (42.9K ± 31.6K)
RoT-mRMR$_{80}$	0.99					0.20 ± 0.01 (22.3K ± 126)

Method	k/d	F1-macro (parameter budget in parentheses) min	Q1	Q2	Q3	max
Evo-mRMR$_{60}$	0.95	0.18 ± 0.02 (920 ± 27)	0.18 ± 0.01 (1.5K ± 965)	0.18 ± 0.01 (1.5K ± 965)	0.19 ± 0.02 (1.6K ± 913)	0.20 ± 0.01 (107.4K ± 44.1K)
FetNAS	0.27	0.18 ± 0.03 (476 ± 367)	0.18 ± 0.03 (476 ± 367)	0.18 ± 0.03 (476 ± 367)	0.18 ± 0.03 (476 ± 367)	0.20 ± 0.01 (46.1K ± 27.7K)
RoT-mRMR$_{60}$	0.95					0.20 ± 0.01 (21.7K ± 381)
Evo-mRMR$_{40}$	0.88	0.17 ± 0.01 (866 ± 23)	0.18 ± 0.02 (2.3K ± 3.1K)	0.18 ± 0.02 (2.3K ± 3.1K)	0.19 ± 0.01 (10.0K ± 14.5K)	0.20 ± 0.01 (154.9K ± 35.0K)
FetNAS	0.34	0.18 ± 0.03 (459 ± 344)	0.18 ± 0.04 (1.5K ± 2.2K)	0.18 ± 0.04 (1.5K ± 2.2K)	0.20 ± 0.01 (5.1K ± 5.8K)	0.19 ± 0.02 (56.2K ± 29.8K)
RoT-mRMR$_{40}$	0.88					0.19 ± 0.01 (21.0K ± 329)
Evo-mRMR$_{20}$	0.72	0.17 ± 0.03 (742 ± 64)	0.17 ± 0.03 (742 ± 64)	0.17 ± 0.03 (742 ± 64)	0.17 ± 0.03 (742 ± 64)	0.20 ± 0.01 (111.6K ± 72.8K)
FetNAS	0.16	0.20 ± 0.00 (290 ± 119)	0.20 ± 0.00 (290 ± 119)	0.20 ± 0.00 (290 ± 119)	0.20 ± 0.00 (290 ± 119)	0.19 ± 0.02 (42.9K ± 31.6K)
RoT-mRMR$_{20}$	0.72					0.20 ± 0.01 (18.9K ± 369)
Evo-UF	0.15	0.15 ± 0.04 (216 ± 110)	0.16 ± 0.05 (880 ± 1.1K)	0.16 ± 0.05 (880 ± 1.1K)	0.17 ± 0.04 (983 ± 1.1K)	0.21 ± 0.02 (50.7K ± 64.1K)
FetNAS	0.04	0.20 ± 0.01 (138 ± 79)	0.20 ± 0.01 (151 ± 78)	0.20 ± 0.01 (151 ± 78)	0.20 ± 0.00 (222 ± 142)	0.18 ± 0.03 (9.6K ± 8.5K)
RoT-UF	0.14					0.20 ± 0.01 (11.4K ± 1.6K)
naticusdroid						
Evo-All	1.00	0.88 ± 0.03 (740 ± 48)	0.89 ± 0.02 (4.9K ± 5.0K)	0.90 ± 0.02 (7.5K ± 4.7K)	0.91 ± 0.01 (10.1K ± 5.0K)	0.91 ± 0.01 (42.6K ± 20.1K)
FetNAS	0.54	0.92 ± 0.01 (489 ± 166)	0.92 ± 0.01 (4.3K ± 4.7K)	0.92 ± 0.00 (6.3K ± 4.6K)	0.92 ± 0.00 (9.0K ± 5.6K)	0.92 ± 0.00 (22.3K ± 15.0K)
RoT-All	1.00					0.93 ± 0.01 (22.8K ± 0)
Evo-mRMR$_{80}$	0.80	0.88 ± 0.02 (569 ± 0)	0.90 ± 0.01 (4.3K ± 5.1K)	0.91 ± 0.01 (7.1K ± 5.7K)	0.91 ± 0.01 (19.6K ± 24.4K)	0.91 ± 0.01 (79.0K ± 67.0K)
FetNAS	0.47	0.92 ± 0.01 (445 ± 110)	0.91 ± 0.01 (2.1K ± 2.2K)	0.92 ± 0.01 (4.5K ± 4.3K)	0.92 ± 0.01 (6.3K ± 6.8K)	0.92 ± 0.00 (36.2K ± 40.7K)
RoT-mRMR$_{80}$	0.80					0.93 ± 0.00 (19.7K ± 0)
Evo-mRMR$_{60}$	0.60	0.88 ± 0.02 (465 ± 48)	0.90 ± 0.01 (3.8K ± 3.5K)	0.89 ± 0.01 (4.3K ± 4.0K)	0.90 ± 0.02 (9.4K ± 10.4K)	0.91 ± 0.00 (69.6K ± 52.7K)
FetNAS	0.37	0.92 ± 0.01 (329 ± 128)	0.91 ± 0.01 (2.3K ± 2.7K)	0.91 ± 0.01 (2.3K ± 2.7K)	0.92 ± 0.01 (7.3K ± 7.8K)	0.92 ± 0.00 (37.8K ± 31.7K)
RoT-mRMR$_{60}$	0.60					0.93 ± 0.00 (16.5K ± 0)
Evo-mRMR$_{40}$	0.41	0.89 ± 0.04 (312 ± 33)	0.91 ± 0.01 (3.8K ± 4.7K)	0.91 ± 0.01 (5.8K ± 5.1K)	0.91 ± 0.01 (10.9K ± 13.9K)	0.91 ± 0.01 (34.9K ± 27.2K)
FetNAS	0.28	0.91 ± 0.00 (232 ± 52)	0.92 ± 0.00 (3.0K ± 4.8K)	0.92 ± 0.00 (3.3K ± 4.6K)	0.92 ± 0.00 (6.5K ± 11.7K)	0.92 ± 0.00 (19.3K ± 15.8K)
RoT-mRMR$_{40}$	0.41					0.93 ± 0.00 (13.4K ± 0)
Evo-mRMR$_{20}$	0.21	0.87 ± 0.02 (169 ± 11)	0.89 ± 0.02 (18.6K ± 37.6K)	0.89 ± 0.02 (21.9K ± 36.6K)	0.90 ± 0.02 (33.5K ± 43.3K)	0.90 ± 0.01 (58.6K ± 42.0K)
FetNAS	0.17	0.90 ± 0.01 (141 ± 17)	0.91 ± 0.01 (1.4K ± 1.6K)	0.91 ± 0.01 (1.9K ± 2.7K)	0.91 ± 0.01 (1.9K ± 2.7K)	0.91 ± 0.01 (4.2K ± 4.9K)
RoT-mRMR$_{20}$	0.21					0.92 ± 0.01 (10.3K ± 0)
Evo-UF	0.03	0.88 ± 0.01 (41 ± 0)	0.88 ± 0.01 (41 ± 0)	0.88 ± 0.01 (41 ± 0)	0.88 ± 0.01 (41 ± 0)	0.88 ± 0.01 (102 ± 94)
FetNAS	0.02	0.72 ± 0.28 (35 ± 5)	0.72 ± 0.28 (35 ± 5)	0.72 ± 0.28 (35 ± 5)	0.72 ± 0.28 (35 ± 5)	0.76 ± 0.24 (40 ± 10)
RoT-UF	0.03					0.88 ± 0.01 (7.6K ± 0)
online_shoppers_intention						
Evo-All	1.00	0.66 ± 0.12 (153 ± 0)	0.68 ± 0.13 (280 ± 126)	0.69 ± 0.13 (635 ± 798)	0.69 ± 0.13 (1.0K ± 1.7K)	0.78 ± 0.02 (40.5K ± 31.1K)
FetNAS	0.15	0.78 ± 0.02 (82 ± 40)	0.78 ± 0.01 (163 ± 87)	0.78 ± 0.01 (163 ± 87)	0.79 ± 0.01 (706 ± 1.3K)	0.79 ± 0.01 (10.3K ± 11.1K)
RoT-All	1.00					0.79 ± 0.02 (15.4K ± 0)
Evo-mRMR$_{80}$	0.82	0.70 ± 0.04 (132 ± 7)	0.72 ± 0.05 (697 ± 527)	0.72 ± 0.06 (1.3K ± 1.8K)	0.74 ± 0.04 (2.6K ± 3.5K)	0.78 ± 0.01 (42.4K ± 20.6K)
FetNAS	0.17	0.78 ± 0.02 (72 ± 25)	0.79 ± 0.01 (304 ± 298)	0.79 ± 0.01 (304 ± 298)	0.79 ± 0.01 (695 ± 804)	0.79 ± 0.01 (10.3K ± 11.1K)
RoT-mRMR$_{80}$	0.82					0.78 ± 0.02 (14.7K ± 0)
Evo-mRMR$_{60}$	0.65	0.74 ± 0.03 (113 ± 12)	0.76 ± 0.03 (437 ± 457)	0.76 ± 0.03 (460 ± 486)	0.78 ± 0.02 (1.4K ± 1.2K)	0.79 ± 0.01 (19.3K ± 23.2K)
FetNAS	0.13	0.78 ± 0.02 (72 ± 25)	0.79 ± 0.02 (259 ± 325)	0.79 ± 0.02 (259 ± 325)	0.79 ± 0.01 (332 ± 299)	0.79 ± 0.01 (971 ± 1.1K)
RoT-mRMR$_{60}$	0.65					0.79 ± 0.01 (14.0K ± 0)
Evo-mRMR$_{40}$	0.41	0.72 ± 0.07 (77 ± 8)	0.76 ± 0.03 (236 ± 234)	0.76 ± 0.03 (236 ± 234)	0.76 ± 0.03 (353 ± 477)	0.79 ± 0.01 (27.4K ± 36.6K)
FetNAS	0.12	0.77 ± 0.02 (65 ± 21)	0.78 ± 0.02 (186 ± 203)	0.78 ± 0.02 (186 ± 203)	0.78 ± 0.02 (244 ± 329)	0.79 ± 0.01 (5.0K ± 6.2K)
RoT-mRMR$_{40}$	0.41					0.79 ± 0.01 (13.0K ± 0)
Evo-mRMR$_{20}$	0.24	0.67 ± 0.13 (49 ± 0)	0.67 ± 0.13 (170 ± 128)	0.67 ± 0.13 (170 ± 128)	0.70 ± 0.14 (479 ± 753)	0.80 ± 0.01 (31.3K ± 35.7K)

Method	k/d	F1-macro (parameter budget in parentheses) min	Q1	Q2	Q3	max
FetNAS	0.09	0.76 ± 0.03 (36 ± 8)	0.78 ± 0.02 (113 ± 98)	0.78 ± 0.02 (113 ± 98)	0.78 ± 0.02 (349 ± 610)	0.79 ± 0.01 (2.7K ± 4.2K)
RoT-mRMR$_{20}$	0.24					0.79 ± 0.01 (12.3K ± 0)
Evo-UF	0.06	0.76 ± 0.02 (25 ± 0)	0.77 ± 0.02 (46 ± 8)	0.77 ± 0.02 (46 ± 8)	0.77 ± 0.02 (46 ± 8)	0.79 ± 0.01 (4.7K ± 4.6K)
FetNAS	0.06	0.65 ± 0.16 (25 ± 0)	0.70 ± 0.13 (33 ± 9)	0.70 ± 0.13 (33 ± 9)	0.70 ± 0.13 (33 ± 9)	0.75 ± 0.10 (1.5K ± 1.4K)
RoT-UF	0.06					0.77 ± 0.01 (11.6K ± 0)
polish_companies_bankruptcy						
Evo-All	1.00	0.51 ± 0.01 (542 ± 30)	0.51 ± 0.02 (967 ± 798)	0.51 ± 0.02 (967 ± 798)	0.52 ± 0.02 (1.0K ± 781)	0.62 ± 0.03 (86.3K ± 58.3K)
FetNAS	0.24	0.52 ± 0.03 (195 ± 152)	0.53 ± 0.03 (580 ± 786)	0.53 ± 0.03 (580 ± 786)	0.53 ± 0.03 (580 ± 786)	0.55 ± 0.04 (33.0K ± 31.2K)
RoT-All	1.00					0.53 ± 0.07 (16.0K ± 0)
Evo-mRMR$_{80}$	0.81	0.51 ± 0.02 (433 ± 0)	0.50 ± 0.02 (487 ± 121)	0.50 ± 0.02 (487 ± 121)	0.50 ± 0.02 (487 ± 121)	0.63 ± 0.03 (47.9K ± 21.9K)
FetNAS	0.21	0.52 ± 0.03 (195 ± 152)	0.52 ± 0.03 (195 ± 152)	0.52 ± 0.03 (195 ± 152)	0.52 ± 0.03 (195 ± 152)	0.56 ± 0.06 (14.9K ± 9.5K)
RoT-mRMR$_{80}$	0.81					0.51 ± 0.05 (14.1K ± 0)
Evo-mRMR$_{60}$	0.61	0.50 ± 0.03 (337 ± 18)	0.51 ± 0.03 (747 ± 912)	0.51 ± 0.02 (1.1K ± 1.1K)	0.51 ± 0.02 (1.3K ± 1.0K)	0.60 ± 0.02 (37.7K ± 25.9K)
FetNAS	0.23	0.52 ± 0.03 (181 ± 134)	0.53 ± 0.04 (587 ± 989)	0.55 ± 0.03 (973 ± 1.1K)	0.55 ± 0.03 (973 ± 1.1K)	0.55 ± 0.04 (12.6K ± 9.3K)
RoT-mRMR$_{60}$	0.61					0.53 ± 0.08 (12.0K ± 0)
Evo-mRMR$_{40}$	0.41	0.51 ± 0.02 (225 ± 0)	0.51 ± 0.02 (287 ± 85)	0.53 ± 0.06 (303 ± 109)	0.54 ± 0.07 (379 ± 152)	0.61 ± 0.05 (51.4K ± 58.1K)
FetNAS	0.15	0.53 ± 0.04 (99 ± 86)	0.53 ± 0.04 (99 ± 86)	0.53 ± 0.04 (99 ± 86)	0.53 ± 0.03 (149 ± 111)	0.54 ± 0.03 (14.3K ± 16.0K)
RoT-mRMR$_{40}$	0.41					0.56 ± 0.08 (9.8K ± 0)
Evo-mRMR$_{20}$	0.20	0.50 ± 0.01 (121 ± 0)	0.51 ± 0.01 (229 ± 125)	0.52 ± 0.03 (633 ± 884)	0.52 ± 0.04 (1.1K ± 1.3K)	0.57 ± 0.03 (30.6K ± 30.2K)
FetNAS	0.07	0.52 ± 0.03 (58 ± 36)	0.51 ± 0.04 (63 ± 47)	0.53 ± 0.04 (423 ± 795)	0.53 ± 0.03 (781 ± 1.0K)	0.54 ± 0.04 (10.6K ± 10.3K)
RoT-mRMR$_{20}$	0.20					0.49 ± 0.01 (7.7K ± 0)
Evo-UF	0.53	0.51 ± 0.02 (296 ± 46)	0.51 ± 0.02 (452 ± 340)	0.51 ± 0.02 (452 ± 340)	0.51 ± 0.02 (452 ± 340)	0.57 ± 0.06 (16.8K ± 23.9K)
FetNAS	0.15	0.53 ± 0.04 (99 ± 86)	0.53 ± 0.04 (99 ± 86)	0.53 ± 0.04 (99 ± 86)	0.53 ± 0.04 (99 ± 86)	0.56 ± 0.05 (9.4K ± 12.6K)
RoT-UF	0.53					0.64 ± 0.09 (11.1K ± 625)
qsar_biodeg						
Evo-All	1.00	0.68 ± 0.16 (362 ± 24)	0.68 ± 0.17 (1.1K ± 1.2K)	0.68 ± 0.17 (1.1K ± 1.2K)	0.69 ± 0.17 (1.1K ± 1.2K)	0.82 ± 0.02 (27.0K ± 37.4K)
FetNAS	0.32	0.72 ± 0.05 (302 ± 85)	0.75 ± 0.06 (933 ± 1.1K)	0.75 ± 0.06 (933 ± 1.1K)	0.75 ± 0.06 (960 ± 1.1K)	0.78 ± 0.03 (10.1K ± 10.8K)
RoT-All	1.00					0.85 ± 0.03 (3.9K ± 0)
Evo-mRMR$_{80}$	0.80	0.59 ± 0.17 (288 ± 16)	0.68 ± 0.17 (3.4K ± 6.6K)	0.68 ± 0.17 (3.6K ± 6.5K)	0.72 ± 0.18 (13.2K ± 17.9K)	0.82 ± 0.02 (42.9K ± 29.8K)
FetNAS	0.34	0.67 ± 0.11 (194 ± 89)	0.76 ± 0.04 (481 ± 267)	0.76 ± 0.04 (506 ± 258)	0.78 ± 0.04 (1.7K ± 1.9K)	0.78 ± 0.03 (10.5K ± 10.5K)
RoT-mRMR$_{80}$	0.80					0.84 ± 0.01 (3.3K ± 0)
Evo-mRMR$_{60}$	0.61	0.54 ± 0.16 (217 ± 0)	0.59 ± 0.16 (471 ± 402)	0.59 ± 0.16 (471 ± 402)	0.60 ± 0.17 (492 ± 406)	0.82 ± 0.02 (28.2K ± 25.6K)
FetNAS	0.30	0.67 ± 0.11 (152 ± 64)	0.69 ± 0.13 (278 ± 205)	0.69 ± 0.13 (278 ± 205)	0.69 ± 0.13 (278 ± 205)	0.80 ± 0.01 (6.3K ± 8.8K)
RoT-mRMR$_{60}$	0.61					0.84 ± 0.01 (2.8K ± 0)
Evo-mRMR$_{40}$	0.41	0.64 ± 0.15 (153 ± 0)	0.74 ± 0.12 (1.3K ± 1.4K)	0.73 ± 0.12 (1.6K ± 1.4K)	0.73 ± 0.11 (1.9K ± 1.6K)	0.82 ± 0.03 (36.4K ± 43.7K)
FetNAS	0.24	0.67 ± 0.11 (95 ± 45)	0.73 ± 0.05 (531 ± 556)	0.75 ± 0.04 (812 ± 645)	0.75 ± 0.05 (1.1K ± 1.1K)	0.77 ± 0.04 (9.6K ± 9.3K)
RoT-mRMR$_{40}$	0.41					0.83 ± 0.03 (2.2K ± 0)
Evo-mRMR$_{20}$	0.22	0.59 ± 0.17 (93 ± 10)	0.61 ± 0.19 (337 ± 494)	0.62 ± 0.19 (353 ± 486)	0.64 ± 0.20 (580 ± 633)	0.81 ± 0.02 (40.5K ± 37.4K)
FetNAS	0.13	0.54 ± 0.15 (57 ± 23)	0.64 ± 0.10 (193 ± 258)	0.64 ± 0.10 (193 ± 258)	0.71 ± 0.06 (318 ± 311)	0.75 ± 0.05 (7.8K ± 6.5K)
RoT-mRMR$_{20}$	0.22					0.81 ± 0.02 (1.7K ± 0)
Evo-UF	0.30	0.66 ± 0.13 (125 ± 36)	0.75 ± 0.04 (601 ± 416)	0.76 ± 0.03 (829 ± 451)	0.77 ± 0.04 (11.6K ± 20.4K)	0.79 ± 0.03 (45.4K ± 59.8K)
FetNAS	0.18	0.60 ± 0.15 (65 ± 35)	0.74 ± 0.05 (448 ± 396)	0.76 ± 0.05 (728 ± 418)	0.77 ± 0.04 (1.6K ± 1.5K)	0.75 ± 0.04 (20.6K ± 21.8K)
RoT-UF	0.30					0.81 ± 0.02 (1.9K ± 279)
sdss17						
Evo-All	1.00	0.94 ± 0.01 (132 ± 13)	0.95 ± 0.01 (674 ± 242)	0.96 ± 0.00 (1.6K ± 558)	0.84 ± 0.28 (4.8K ± 1.3K)	0.88 ± 0.18 (44.1K ± 36.4K)

Method	k/d	F1-macro (parameter budget in parentheses) min	Q1	Q2	Q3	max
FetNAS	0.49	0.80 ± 0.31 (87 ± 29)	0.96 ± 0.00 (575 ± 285)	0.87 ± 0.18 (1.4K ± 523)	0.96 ± 0.00 (3.1K ± 667)	0.87 ± 0.20 (20.0K ± 19.1K)
RoT-All	1.00					0.96 ± 0.00 (183.9K ± 0)
Evo-mRMR$_{80}$	0.82	0.87 ± 0.15 (120 ± 13)	0.95 ± 0.01 (517 ± 332)	0.89 ± 0.16 (1.1K ± 491)	0.96 ± 0.00 (3.7K ± 2.5K)	0.88 ± 0.19 (26.8K ± 10.9K)
FetNAS	0.49	0.66 ± 0.37 (76 ± 19)	0.95 ± 0.00 (360 ± 343)	0.96 ± 0.00 (842 ± 324)	0.88 ± 0.18 (2.2K ± 918)	0.96 ± 0.00 (15.6K ± 9.9K)
RoT-mRMR$_{80}$	0.82					0.96 ± 0.00 (182.6K ± 0)
Evo-mRMR$_{60}$	0.73	0.93 ± 0.00 (101 ± 5)	0.90 ± 0.09 (498 ± 170)	0.84 ± 0.25 (998 ± 168)	0.94 ± 0.02 (2.6K ± 831)	0.84 ± 0.26 (21.5K ± 16.2K)
FetNAS	0.50	0.66 ± 0.37 (75 ± 17)	0.95 ± 0.00 (315 ± 179)	0.96 ± 0.00 (817 ± 226)	0.67 ± 0.29 (2.2K ± 918)	0.88 ± 0.17 (11.9K ± 5.1K)
RoT-mRMR$_{60}$	0.73					0.96 ± 0.00 (181.9K ± 0)
Evo-mRMR$_{40}$	0.62	0.93 ± 0.01 (96 ± 12)	0.95 ± 0.00 (393 ± 172)	0.85 ± 0.25 (1.4K ± 741)	0.95 ± 0.01 (4.4K ± 2.5K)	0.96 ± 0.00 (33.2K ± 20.9K)
FetNAS	0.48	0.66 ± 0.37 (71 ± 13)	0.95 ± 0.00 (261 ± 128)	0.96 ± 0.00 (1.0K ± 883)	0.88 ± 0.18 (2.6K ± 1.0K)	0.88 ± 0.17 (21.9K ± 15.7K)
RoT-mRMR$_{40}$	0.62					0.96 ± 0.00 (181.2K ± 286)
Evo-mRMR$_{20}$	0.36	0.92 ± 0.01 (69 ± 4)	0.93 ± 0.01 (273 ± 168)	0.93 ± 0.00 (470 ± 213)	0.93 ± 0.01 (784 ± 492)	0.86 ± 0.16 (4.4K ± 3.2K)
FetNAS	0.30	0.80 ± 0.30 (62 ± 7)	0.80 ± 0.31 (209 ± 154)	0.81 ± 0.31 (396 ± 181)	0.83 ± 0.26 (695 ± 388)	0.86 ± 0.21 (3.5K ± 2.9K)
RoT-mRMR$_{20}$	0.36					0.93 ± 0.00 (179.4K ± 0)
Evo-UF	0.09	0.92 ± 0.00 (43 ± 0)	0.92 ± 0.00 (118 ± 34)	0.93 ± 0.00 (527 ± 483)	0.82 ± 0.17 (1.6K ± 1.3K)	0.85 ± 0.14 (4.7K ± 5.8K)
FetNAS	0.09	0.89 ± 0.02 (43 ± 0)	0.91 ± 0.02 (73 ± 44)	0.93 ± 0.01 (519 ± 420)	0.93 ± 0.01 (916 ± 959)	0.87 ± 0.14 (3.4K ± 4.2K)
RoT-UF	0.09					0.93 ± 0.00 (177.5K ± 0)
seismic_bumps						
Evo-All	1.00	0.48 ± 0.00 (137 ± 0)	0.49 ± 0.01 (191 ± 78)	0.49 ± 0.01 (208 ± 98)	0.49 ± 0.01 (287 ± 162)	0.54 ± 0.03 (73.6K ± 53.4K)
FetNAS	0.19	0.48 ± 0.00 (32 ± 11)	0.48 ± 0.00 (32 ± 11)	0.48 ± 0.00 (32 ± 11)	0.49 ± 0.03 (128 ± 214)	0.53 ± 0.03 (10.5K ± 10.2K)
RoT-All	1.00					0.50 ± 0.04 (4.1K ± 0)
Evo-mRMR$_{80}$	0.80	0.48 ± 0.00 (113 ± 0)	0.50 ± 0.03 (469 ± 795)	0.50 ± 0.03 (2.0K ± 3.4K)	0.50 ± 0.03 (2.0K ± 3.4K)	0.55 ± 0.05 (52.9K ± 48.0K)
FetNAS	0.20	0.48 ± 0.00 (32 ± 11)	0.49 ± 0.01 (264 ± 517)	0.50 ± 0.03 (466 ± 603)	0.50 ± 0.03 (466 ± 603)	0.53 ± 0.03 (8.8K ± 10.0K)
RoT-mRMR$_{80}$	0.80					0.48 ± 0.00 (3.8K ± 0)
Evo-mRMR$_{60}$	0.60	0.49 ± 0.01 (93 ± 10)	0.50 ± 0.02 (342 ± 348)	0.51 ± 0.03 (472 ± 423)	0.52 ± 0.04 (626 ± 577)	0.53 ± 0.05 (4.7K ± 6.2K)
FetNAS	0.18	0.48 ± 0.00 (32 ± 11)	0.49 ± 0.03 (177 ± 319)	0.50 ± 0.03 (211 ± 395)	0.51 ± 0.04 (397 ± 493)	0.51 ± 0.03 (2.2K ± 3.0K)
RoT-mRMR$_{60}$	0.60					0.48 ± 0.00 (3.5K ± 0)
Evo-mRMR$_{40}$	0.40	0.48 ± 0.00 (65 ± 0)	0.50 ± 0.03 (207 ± 318)	0.50 ± 0.03 (207 ± 318)	0.50 ± 0.03 (207 ± 318)	0.53 ± 0.04 (11.0K ± 15.7K)
FetNAS	0.09	0.48 ± 0.00 (32 ± 11)	0.50 ± 0.03 (160 ± 283)	0.50 ± 0.03 (160 ± 283)	0.50 ± 0.03 (160 ± 283)	0.51 ± 0.03 (4.1K ± 6.6K)
RoT-mRMR$_{40}$	0.40					0.48 ± 0.00 (3.2K ± 0)
Evo-mRMR$_{20}$	0.20	0.48 ± 0.00 (41 ± 0)	0.48 ± 0.00 (41 ± 0)	0.48 ± 0.00 (41 ± 0)	0.48 ± 0.00 (41 ± 0)	0.52 ± 0.04 (15.4K ± 18.6K)
FetNAS	0.08	0.48 ± 0.00 (27 ± 3)	0.48 ± 0.00 (27 ± 3)	0.48 ± 0.00 (27 ± 3)	0.48 ± 0.00 (27 ± 3)	0.50 ± 0.02 (8.2K ± 8.2K)
RoT-mRMR$_{20}$	0.20					0.48 ± 0.00 (2.9K ± 0)
Evo-UF	0.47	0.48 ± 0.00 (73 ± 6)	0.49 ± 0.01 (199 ± 204)	0.49 ± 0.01 (213 ± 209)	0.49 ± 0.01 (213 ± 209)	0.55 ± 0.02 (30.0K ± 11.0K)
FetNAS	0.11	0.48 ± 0.00 (32 ± 11)	0.48 ± 0.00 (32 ± 11)	0.48 ± 0.00 (32 ± 11)	0.48 ± 0.00 (32 ± 11)	0.54 ± 0.02 (7.0K ± 5.6K)
RoT-UF	0.47					0.48 ± 0.00 (3.3K ± 76)
splice						
Evo-All	1.00	0.81 ± 0.03 (579 ± 64)	0.82 ± 0.01 (8.3K ± 5.1K)	0.82 ± 0.01 (8.3K ± 5.1K)	0.82 ± 0.01 (8.3K ± 5.1K)	0.84 ± 0.02 (58.0K ± 33.7K)
FetNAS	0.35	0.82 ± 0.02 (370 ± 122)	0.84 ± 0.03 (4.4K ± 3.3K)	0.84 ± 0.03 (4.4K ± 3.3K)	0.84 ± 0.03 (4.4K ± 3.3K)	0.85 ± 0.02 (31.6K ± 13.7K)
RoT-All	1.00					0.85 ± 0.03 (15.4K ± 0)
Evo-mRMR$_{80}$	0.99	0.81 ± 0.01 (546 ± 55)	0.82 ± 0.02 (4.2K ± 3.5K)	0.81 ± 0.02 (5.0K ± 3.4K)	0.82 ± 0.02 (5.7K ± 3.6K)	0.85 ± 0.02 (37.1K ± 13.3K)
FetNAS	0.36	0.82 ± 0.02 (370 ± 122)	0.84 ± 0.02 (2.6K ± 3.0K)	0.84 ± 0.02 (2.6K ± 3.0K)	0.84 ± 0.02 (3.9K ± 3.5K)	0.86 ± 0.02 (25.2K ± 11.3K)
RoT-mRMR$_{80}$	0.99					0.87 ± 0.02 (15.3K ± 109)
Evo-mRMR$_{60}$	0.89	0.79 ± 0.03 (475 ± 20)	0.81 ± 0.01 (1.6K ± 493)	0.83 ± 0.02 (3.8K ± 2.5K)	0.83 ± 0.02 (10.4K ± 16.3K)	0.84 ± 0.02 (53.5K ± 34.3K)
FetNAS	0.41	0.82 ± 0.01 (349 ± 99)	0.84 ± 0.02 (922 ± 319)	0.83 ± 0.02 (1.9K ± 1.5K)	0.84 ± 0.02 (8.7K ± 16.5K)	0.86 ± 0.02 (41.8K ± 24.3K)

Method	k/d	F1-macro (parameter budget in parentheses) min	Q1	Q2	Q3	max
RoT-mRMR$_{60}$	0.89					0.86 ± 0.03 (14.6K ± 174)
Evo-mRMR$_{40}$	0.71	0.79 ± 0.03 (401 ± 29)	0.83 ± 0.02 (7.9K ± 9.7K)	0.83 ± 0.02 (8.5K ± 10.0K)	0.83 ± 0.02 (12.5K ± 13.8K)	0.85 ± 0.01 (49.2K ± 20.8K)
FetNAS	0.36	0.82 ± 0.01 (275 ± 78)	0.84 ± 0.03 (5.0K ± 6.4K)	0.84 ± 0.03 (5.0K ± 6.4K)	0.84 ± 0.03 (5.0K ± 6.4K)	0.86 ± 0.01 (29.2K ± 13.0K)
RoT-mRMR$_{40}$	0.71					0.87 ± 0.01 (13.1K ± 218)
Evo-mRMR$_{20}$	0.39	0.81 ± 0.02 (246 ± 27)	0.86 ± 0.04 (9.0K ± 8.0K)	0.86 ± 0.04 (11.0K ± 9.9K)	0.86 ± 0.04 (11.8K ± 9.6K)	0.89 ± 0.01 (47.7K ± 9.4K)
FetNAS	0.26	0.82 ± 0.02 (195 ± 21)	0.85 ± 0.03 (5.1K ± 6.4K)	0.84 ± 0.03 (5.7K ± 6.3K)	0.85 ± 0.03 (6.3K ± 6.4K)	0.85 ± 0.02 (39.3K ± 14.7K)
RoT-mRMR$_{20}$	0.39					0.90 ± 0.01 (10.7K ± 197)
Evo-UF	0.12	0.74 ± 0.05 (99 ± 14)	0.84 ± 0.08 (19.1K ± 38.6K)	0.84 ± 0.08 (19.1K ± 38.6K)	0.84 ± 0.08 (19.1K ± 38.6K)	0.93 ± 0.01 (45.4K ± 27.1K)
FetNAS	0.11	0.75 ± 0.02 (86 ± 12)	0.78 ± 0.03 (8.9K ± 17.6K)	0.78 ± 0.03 (8.9K ± 17.6K)	0.78 ± 0.03 (8.9K ± 17.6K)	0.82 ± 0.01 (22.3K ± 12.7K)
RoT-UF	0.12					0.93 ± 0.01 (8.6K ± 109)
students_dropout_and_academic_success						
Evo-All	1.00	0.61 ± 0.03 (323 ± 0)	0.64 ± 0.04 (918 ± 596)	0.65 ± 0.04 (2.8K ± 3.2K)	0.66 ± 0.04 (7.0K ± 8.0K)	0.68 ± 0.01 (22.7K ± 12.8K)
FetNAS	0.35	0.65 ± 0.03 (192 ± 63)	0.66 ± 0.03 (479 ± 378)	0.67 ± 0.03 (2.6K ± 2.9K)	0.67 ± 0.03 (5.3K ± 6.2K)	0.69 ± 0.01 (10.3K ± 7.0K)
RoT-All	1.00					0.66 ± 0.03 (16.0K ± 0)
Evo-mRMR$_{80}$	0.86	0.57 ± 0.01 (281 ± 7)	0.63 ± 0.03 (2.1K ± 2.6K)	0.63 ± 0.03 (2.3K ± 2.6K)	0.63 ± 0.03 (2.7K ± 2.8K)	0.67 ± 0.01 (34.8K ± 22.3K)
FetNAS	0.37	0.65 ± 0.03 (175 ± 31)	0.68 ± 0.01 (1.9K ± 2.4K)	0.68 ± 0.01 (2.0K ± 2.4K)	0.68 ± 0.01 (2.3K ± 2.4K)	0.69 ± 0.01 (10.9K ± 10.3K)
RoT-mRMR$_{80}$	0.86					0.63 ± 0.02 (15.2K ± 128)
Evo-mRMR$_{60}$	0.79	0.57 ± 0.02 (262 ± 12)	0.59 ± 0.03 (442 ± 249)	0.60 ± 0.02 (469 ± 233)	0.61 ± 0.04 (527 ± 250)	0.67 ± 0.01 (18.4K ± 9.6K)
FetNAS	0.39	0.65 ± 0.03 (175 ± 31)	0.67 ± 0.03 (287 ± 181)	0.67 ± 0.03 (287 ± 181)	0.67 ± 0.03 (287 ± 181)	0.68 ± 0.01 (11.8K ± 8.1K)
RoT-mRMR$_{60}$	0.79					0.65 ± 0.02 (14.9K ± 232)
Evo-mRMR$_{40}$	0.62	0.56 ± 0.04 (213 ± 9)	0.59 ± 0.03 (523 ± 336)	0.60 ± 0.04 (762 ± 661)	0.61 ± 0.04 (3.4K ± 4.2K)	0.65 ± 0.01 (24.0K ± 9.9K)
FetNAS	0.34	0.65 ± 0.03 (170 ± 26)	0.67 ± 0.01 (318 ± 217)	0.67 ± 0.01 (508 ± 626)	0.68 ± 0.01 (2.1K ± 2.6K)	0.69 ± 0.01 (12.4K ± 6.4K)
RoT-mRMR$_{40}$	0.62					0.64 ± 0.02 (13.9K ± 168)
Evo-mRMR$_{20}$	0.41	0.53 ± 0.01 (153 ± 14)	0.56 ± 0.03 (513 ± 246)	0.57 ± 0.04 (648 ± 408)	0.57 ± 0.04 (651 ± 407)	0.64 ± 0.01 (43.6K ± 22.4K)
FetNAS	0.28	0.64 ± 0.02 (135 ± 9)	0.66 ± 0.03 (266 ± 225)	0.66 ± 0.03 (450 ± 489)	0.66 ± 0.03 (450 ± 489)	0.69 ± 0.01 (25.4K ± 11.7K)
RoT-mRMR$_{20}$	0.41					0.62 ± 0.01 (12.8K ± 274)
Evo-UF	0.12	0.52 ± 0.00 (72 ± 4)	0.54 ± 0.04 (150 ± 119)	0.54 ± 0.04 (216 ± 153)	0.55 ± 0.03 (402 ± 203)	0.64 ± 0.02 (33.3K ± 17.7K)
FetNAS	0.09	0.55 ± 0.04 (67 ± 1)	0.57 ± 0.05 (136 ± 114)	0.58 ± 0.06 (174 ± 117)	0.60 ± 0.05 (282 ± 142)	0.67 ± 0.01 (23.4K ± 18.1K)
RoT-UF	0.11					0.58 ± 0.02 (11.1K ± 0)
tabarena_bank_marketing						
Evo-All	1.00	0.57 ± 0.01 (124 ± 7)	0.60 ± 0.02 (339 ± 69)	0.62 ± 0.02 (635 ± 162)	0.63 ± 0.01 (1.7K ± 1.4K)	0.63 ± 0.00 (16.9K ± 21.5K)
FetNAS	0.40	0.56 ± 0.05 (87 ± 30)	0.60 ± 0.02 (160 ± 67)	0.61 ± 0.01 (355 ± 321)	0.61 ± 0.01 (502 ± 407)	0.61 ± 0.01 (5.9K ± 11.5K)
RoT-All	1.00					0.60 ± 0.00 (46.9K ± 0)
Evo-mRMR$_{80}$	0.85	0.57 ± 0.01 (108 ± 6)	0.59 ± 0.02 (293 ± 80)	0.60 ± 0.02 (493 ± 168)	0.62 ± 0.01 (1.2K ± 746)	0.63 ± 0.01 (11.4K ± 13.0K)
FetNAS	0.37	0.56 ± 0.05 (81 ± 24)	0.59 ± 0.02 (120 ± 28)	0.60 ± 0.02 (239 ± 175)	0.61 ± 0.00 (536 ± 411)	0.62 ± 0.01 (814 ± 470)
RoT-mRMR$_{80}$	0.85					0.59 ± 0.01 (46.0K ± 0)
Evo-mRMR$_{60}$	0.62	0.58 ± 0.01 (85 ± 9)	0.59 ± 0.01 (337 ± 256)	0.60 ± 0.01 (778 ± 1.1K)	0.61 ± 0.01 (1.1K ± 1.1K)	0.62 ± 0.01 (6.2K ± 9.1K)
FetNAS	0.31	0.55 ± 0.05 (59 ± 13)	0.59 ± 0.02 (150 ± 99)	0.60 ± 0.03 (353 ± 544)	0.61 ± 0.01 (504 ± 557)	0.62 ± 0.01 (3.7K ± 7.2K)
RoT-mRMR$_{60}$	0.62					0.61 ± 0.01 (44.6K ± 0)
Evo-mRMR$_{40}$	0.46	0.57 ± 0.03 (65 ± 0)	0.59 ± 0.01 (738 ± 1.3K)	0.59 ± 0.02 (781 ± 1.3K)	0.60 ± 0.02 (1.1K ± 1.4K)	0.62 ± 0.01 (3.1K ± 2.3K)
FetNAS	0.34	0.55 ± 0.05 (55 ± 8)	0.59 ± 0.03 (271 ± 397)	0.59 ± 0.03 (271 ± 397)	0.60 ± 0.02 (418 ± 456)	0.62 ± 0.01 (673 ± 545)
RoT-mRMR$_{40}$	0.46					0.61 ± 0.00 (43.7K ± 0)
Evo-mRMR$_{20}$	0.23	0.48 ± 0.02 (42 ± 2)	0.53 ± 0.08 (147 ± 181)	0.53 ± 0.08 (147 ± 181)	0.56 ± 0.08 (194 ± 179)	0.62 ± 0.01 (17.8K ± 25.5K)
FetNAS	0.19	0.49 ± 0.04 (38 ± 7)	0.52 ± 0.07 (63 ± 44)	0.52 ± 0.07 (63 ± 44)	0.53 ± 0.06 (68 ± 43)	0.59 ± 0.04 (329 ± 504)
RoT-mRMR$_{20}$	0.23					0.61 ± 0.01 (42.4K ± 0)

Method	k/d	F1-macro (parameter budget in parentheses) min	Q1	Q2	Q3	max
Evo-UF	0.40	0.57 ± 0.03 (62 ± 9)	0.60 ± 0.01 (215 ± 141)	0.60 ± 0.01 (378 ± 148)	0.61 ± 0.01 (696 ± 613)	0.63 ± 0.01 (32.1K ± 47.7K)
FetNAS	0.26	0.53 ± 0.06 (55 ± 14)	0.56 ± 0.06 (94 ± 37)	0.59 ± 0.03 (114 ± 27)	0.60 ± 0.03 (221 ± 235)	0.61 ± 0.00 (17.8K ± 27.8K)
RoT-UF	0.40					0.61 ± 0.00 (43.4K ± 376)
tabarena_diabetes						
Evo-All	1.00	0.68 ± 0.04 (81 ± 0)	0.70 ± 0.03 (370 ± 290)	0.70 ± 0.03 (410 ± 255)	0.70 ± 0.03 (444 ± 249)	0.71 ± 0.03 (8.9K ± 12.4K)
FetNAS	0.59	0.72 ± 0.03 (58 ± 5)	0.73 ± 0.02 (155 ± 119)	0.74 ± 0.02 (163 ± 115)	0.74 ± 0.02 (163 ± 115)	0.73 ± 0.01 (593 ± 831)
RoT-All	1.00					0.73 ± 0.01 (1.3K ± 0)
Evo-mRMR$_{80}$	0.88	0.69 ± 0.07 (82 ± 9)	0.71 ± 0.07 (123 ± 64)	0.73 ± 0.05 (271 ± 297)	0.72 ± 0.05 (303 ± 292)	0.73 ± 0.03 (16.4K ± 21.7K)
FetNAS	0.58	0.72 ± 0.03 (58 ± 5)	0.73 ± 0.02 (76 ± 39)	0.73 ± 0.02 (76 ± 39)	0.73 ± 0.02 (76 ± 39)	0.71 ± 0.03 (1.9K ± 3.8K)
RoT-mRMR$_{80}$	0.88					0.75 ± 0.01 (1.2K ± 0)
Evo-mRMR$_{60}$	0.62	0.67 ± 0.06 (58 ± 3)	0.70 ± 0.02 (241 ± 254)	0.70 ± 0.02 (241 ± 254)	0.70 ± 0.02 (758 ± 1.1K)	0.72 ± 0.03 (14.1K ± 8.6K)
FetNAS	0.47	0.72 ± 0.02 (55 ± 1)	0.72 ± 0.03 (58 ± 5)	0.72 ± 0.03 (108 ± 115)	0.72 ± 0.03 (108 ± 115)	0.71 ± 0.03 (2.0K ± 3.7K)
RoT-mRMR$_{60}$	0.62					0.75 ± 0.02 (1.1K ± 0)
Evo-mRMR$_{40}$	0.50	0.69 ± 0.08 (50 ± 3)	0.73 ± 0.02 (119 ± 119)	0.73 ± 0.02 (462 ± 477)	0.74 ± 0.02 (543 ± 644)	0.71 ± 0.02 (37.8K ± 36.3K)
FetNAS	0.42	0.72 ± 0.03 (42 ± 13)	0.72 ± 0.03 (87 ± 92)	0.72 ± 0.02 (144 ± 128)	0.72 ± 0.02 (144 ± 128)	0.71 ± 0.03 (10.6K ± 18.9K)
RoT-mRMR$_{40}$	0.50					0.73 ± 0.02 (1.0K ± 0)
Evo-mRMR$_{20}$	0.25	0.65 ± 0.12 (34 ± 2)	0.70 ± 0.04 (67 ± 50)	0.70 ± 0.04 (67 ± 50)	0.70 ± 0.04 (102 ± 74)	0.71 ± 0.04 (34.9K ± 26.5K)
FetNAS	0.20	0.69 ± 0.02 (30 ± 3)	0.69 ± 0.02 (35 ± 10)	0.69 ± 0.02 (35 ± 10)	0.71 ± 0.02 (67 ± 72)	0.69 ± 0.01 (4.5K ± 9.6K)
RoT-mRMR$_{20}$	0.25					0.71 ± 0.03 (910 ± 0)
Evo-UF	0.45	0.66 ± 0.14 (46 ± 11)	0.66 ± 0.14 (3.9K ± 8.5K)	0.66 ± 0.14 (3.9K ± 8.5K)	0.66 ± 0.14 (3.9K ± 8.5K)	0.72 ± 0.04 (57.9K ± 54.5K)
FetNAS	0.40	0.71 ± 0.02 (44 ± 11)	0.70 ± 0.02 (1.8K ± 3.8K)	0.70 ± 0.02 (1.8K ± 3.8K)	0.70 ± 0.02 (1.8K ± 3.8K)	0.70 ± 0.02 (13.8K ± 24.3K)
RoT-UF	0.45					0.72 ± 0.01 (1.0K ± 79)
taiwanese_bankruptcy_prediction						
Evo-All	1.00	0.58 ± 0.04 (788 ± 43)	0.59 ± 0.05 (963 ± 343)	0.59 ± 0.05 (963 ± 343)	0.59 ± 0.05 (5.0K ± 7.9K)	0.64 ± 0.05 (38.1K ± 31.8K)
FetNAS	0.23	0.63 ± 0.03 (473 ± 202)	0.63 ± 0.04 (540 ± 284)	0.63 ± 0.04 (540 ± 284)	0.65 ± 0.02 (2.8K ± 5.0K)	0.66 ± 0.02 (4.4K ± 4.6K)
RoT-All	1.00					0.58 ± 0.07 (22.8K ± 0)
Evo-mRMR$_{80}$	0.81	0.53 ± 0.03 (656 ± 70)	0.54 ± 0.07 (677 ± 73)	0.55 ± 0.08 (693 ± 93)	0.56 ± 0.07 (1.4K ± 1.5K)	0.66 ± 0.02 (31.8K ± 15.9K)
FetNAS	0.28	0.64 ± 0.03 (448 ± 167)	0.64 ± 0.03 (448 ± 167)	0.64 ± 0.03 (448 ± 167)	0.64 ± 0.03 (1.0K ± 1.4K)	0.66 ± 0.02 (4.3K ± 4.6K)
RoT-mRMR$_{80}$	0.81					0.55 ± 0.08 (19.7K ± 0)
Evo-mRMR$_{60}$	0.61	0.53 ± 0.04 (473 ± 0)	0.57 ± 0.08 (1.1K ± 927)	0.56 ± 0.08 (1.2K ± 950)	0.57 ± 0.08 (1.4K ± 1.3K)	0.65 ± 0.03 (43.6K ± 21.2K)
FetNAS	0.22	0.63 ± 0.04 (280 ± 139)	0.64 ± 0.05 (762 ± 784)	0.64 ± 0.05 (762 ± 784)	0.64 ± 0.05 (762 ± 784)	0.66 ± 0.02 (11.4K ± 16.2K)
RoT-mRMR$_{60}$	0.61					0.53 ± 0.06 (16.3K ± 0)
Evo-mRMR$_{40}$	0.40	0.53 ± 0.04 (329 ± 18)	0.54 ± 0.06 (537 ± 461)	0.58 ± 0.06 (2.8K ± 4.2K)	0.59 ± 0.07 (3.6K ± 4.3K)	0.64 ± 0.05 (17.0K ± 18.6K)
FetNAS	0.20	0.63 ± 0.05 (193 ± 24)	0.63 ± 0.05 (340 ± 321)	0.64 ± 0.05 (1.5K ± 2.1K)	0.64 ± 0.05 (1.5K ± 2.1K)	0.66 ± 0.02 (3.6K ± 3.8K)
RoT-mRMR$_{40}$	0.40					0.52 ± 0.04 (13.0K ± 0)
Evo-mRMR$_{20}$	0.20	0.58 ± 0.04 (173 ± 9)	0.59 ± 0.04 (555 ± 542)	0.59 ± 0.04 (555 ± 542)	0.59 ± 0.05 (560 ± 539)	0.64 ± 0.06 (25.0K ± 24.5K)
FetNAS	0.13	0.62 ± 0.06 (134 ± 17)	0.63 ± 0.04 (446 ± 401)	0.63 ± 0.04 (446 ± 401)	0.63 ± 0.04 (446 ± 401)	0.64 ± 0.05 (1.0K ± 624)
RoT-mRMR$_{20}$	0.20					0.53 ± 0.07 (9.7K ± 0)
Evo-UF	0.41	0.60 ± 0.02 (344 ± 60)	0.61 ± 0.04 (427 ± 246)	0.61 ± 0.04 (427 ± 246)	0.63 ± 0.04 (887 ± 994)	0.68 ± 0.01 (28.7K ± 20.5K)
FetNAS	0.26	0.62 ± 0.04 (236 ± 98)	0.62 ± 0.04 (236 ± 98)	0.62 ± 0.04 (236 ± 98)	0.62 ± 0.04 (612 ± 811)	0.65 ± 0.02 (2.7K ± 2.1K)
RoT-UF	0.41					0.52 ± 0.06 (13.1K ± 439)
website_phishing						
Evo-All	1.00	0.55 ± 0.02 (115 ± 12)	0.55 ± 0.02 (115 ± 12)	0.56 ± 0.01 (697 ± 1.3K)	0.56 ± 0.01 (697 ± 1.3K)	0.72 ± 0.04 (114.1K ± 49.6K)
FetNAS	0.74	0.56 ± 0.01 (97 ± 7)	0.56 ± 0.01 (97 ± 7)	0.57 ± 0.02 (667 ± 1.3K)	0.57 ± 0.02 (667 ± 1.3K)	0.72 ± 0.02 (54.5K ± 28.5K)
RoT-All	1.00					0.78 ± 0.04 (4.1K ± 0)

Method	k/d	F1-macro (parameter budget in parentheses)				
		min	Q1	Q2	Q3	max
Evo-mRMR$_{80}$	1.00	0.55 ± 0.02 (115 ± 12)	0.55 ± 0.02 (115 ± 12)	0.56 ± 0.01 (697 ± 1.3K)	0.56 ± 0.01 (697 ± 1.3K)	0.72 ± 0.05 (114.1K ± 49.6K)
FetNAS	0.74	0.56 ± 0.01 (97 ± 7)	0.56 ± 0.01 (97 ± 7)	0.57 ± 0.02 (667 ± 1.3K)	0.57 ± 0.02 (667 ± 1.3K)	0.72 ± 0.02 (54.5K ± 28.5K)
RoT-mRMR$_{80}$	1.00					0.81 ± 0.03 (4.1K ± 0)
Evo-mRMR$_{60}$	0.91	0.55 ± 0.01 (105 ± 9)	0.55 ± 0.01 (497 ± 674)	0.55 ± 0.01 (497 ± 674)	0.56 ± 0.01 (678 ± 661)	0.70 ± 0.06 (86.8K ± 70.3K)
FetNAS	0.65	0.56 ± 0.01 (91 ± 10)	0.57 ± 0.01 (310 ± 321)	0.57 ± 0.01 (310 ± 321)	0.57 ± 0.01 (310 ± 321)	0.71 ± 0.04 (30.7K ± 21.8K)
RoT-mRMR$_{60}$	0.91					0.73 ± 0.11 (4.1K ± 71)
Evo-mRMR$_{40}$	0.76	0.54 ± 0.01 (92 ± 9)	0.55 ± 0.01 (1.7K ± 2.7K)	0.55 ± 0.01 (1.7K ± 2.7K)	0.55 ± 0.01 (1.7K ± 2.7K)	0.64 ± 0.04 (74.4K ± 56.3K)
FetNAS	0.56	0.56 ± 0.01 (82 ± 10)	0.57 ± 0.01 (525 ± 482)	0.57 ± 0.01 (531 ± 476)	0.57 ± 0.01 (531 ± 476)	0.70 ± 0.05 (21.6K ± 20.6K)
RoT-mRMR$_{40}$	0.76					0.67 ± 0.03 (4.0K ± 71)
Evo-mRMR$_{20}$	0.49	0.55 ± 0.01 (74 ± 11)	0.59 ± 0.06 (2.5K ± 5.2K)	0.59 ± 0.06 (2.5K ± 5.2K)	0.59 ± 0.06 (2.5K ± 5.2K)	0.66 ± 0.06 (108.2K ± 41.2K)
FetNAS	0.37	0.57 ± 0.01 (59 ± 15)	0.58 ± 0.02 (1.6K ± 3.4K)	0.58 ± 0.02 (1.6K ± 3.4K)	0.58 ± 0.02 (1.6K ± 3.4K)	0.66 ± 0.05 (49.3K ± 33.2K)
RoT-mRMR$_{20}$	0.49					0.65 ± 0.06 (3.8K ± 47)
Evo-UF	0.33	0.55 ± 0.02 (59 ± 0)	0.55 ± 0.02 (87 ± 47)	0.56 ± 0.02 (177 ± 182)	0.56 ± 0.02 (184 ± 197)	0.57 ± 0.02 (36.2K ± 31.9K)
FetNAS	0.17	0.56 ± 0.01 (46 ± 3)	0.56 ± 0.01 (62 ± 34)	0.56 ± 0.01 (62 ± 34)	0.56 ± 0.01 (62 ± 34)	0.58 ± 0.05 (14.5K ± 19.0K)
RoT-UF	0.31					0.59 ± 0.02 (3.6K ± 38)

Table 6: **Regression results at matched feature budgets for each dataset within the TabArena benchmark.** Same protocol as Table 5, reporting mean ± std R^2 over the five cross-validation splits. UF retained no features for *airfoil_self_noise*, *food_delivery_time*, *physiochemical_protein* and *qsar_tid_11*, so these datasets have no UF rows.

		R^2 (parameter budget in parentheses)				
Method	k/d	min	Q1	Q2	Q3	max
airfoil_self_noise						
Evo-All	1.00	0.49 ± 0.12 (61 ± 6)	0.59 ± 0.06 (1.2K ± 1.0K)	0.68 ± 0.03 (5.2K ± 3.5K)	0.74 ± 0.02 (21.1K ± 8.9K)	0.76 ± 0.02 (128.0K ± 35.9K)
FetNAS	0.81	0.50 ± 0.11 (56 ± 6)	0.62 ± 0.03 (1.0K ± 765)	0.67 ± 0.04 (3.1K ± 1.5K)	0.74 ± 0.02 (19.3K ± 8.0K)	0.75 ± 0.01 (53.0K ± 29.0K)
RoT-All	1.00					0.93 ± 0.01 (1.9K ± 0)
Evo-mRMR$_{80}$	0.80	0.48 ± 0.11 (51 ± 3)	0.54 ± 0.10 (740 ± 208)	0.63 ± 0.12 (3.1K ± 954)	0.67 ± 0.14 (26.3K ± 16.7K)	0.75 ± 0.06 (168.4K ± 37.9K)
FetNAS	0.72	0.49 ± 0.11 (49 ± 2)	0.60 ± 0.02 (499 ± 293)	0.65 ± 0.01 (2.9K ± 769)	0.72 ± 0.02 (22.0K ± 15.4K)	0.75 ± 0.02 (72.6K ± 45.2K)
RoT-mRMR$_{80}$	0.80					0.89 ± 0.04 (1.8K ± 0)
Evo-mRMR$_{60}$	0.60	0.33 ± 0.04 (41 ± 0)	0.39 ± 0.05 (884 ± 1000)	0.44 ± 0.09 (4.3K ± 3.2K)	0.52 ± 0.10 (18.5K ± 9.3K)	0.56 ± 0.12 (86.1K ± 35.1K)
FetNAS	0.58	0.41 ± 0.19 (39 ± 2)	0.60 ± 0.05 (836 ± 908)	0.65 ± 0.05 (3.6K ± 3.1K)	0.69 ± 0.02 (12.5K ± 6.5K)	0.72 ± 0.01 (60.3K ± 33.2K)
RoT-mRMR$_{60}$	0.60					0.68 ± 0.17 (1.8K ± 0)
Evo-mRMR$_{40}$	0.40	0.34 ± 0.05 (34 ± 2)	0.37 ± 0.05 (523 ± 220)	0.45 ± 0.08 (3.9K ± 3.4K)	0.52 ± 0.09 (13.2K ± 5.8K)	0.56 ± 0.10 (98.0K ± 54.9K)
FetNAS	0.39	0.33 ± 0.11 (32 ± 2)	0.40 ± 0.04 (427 ± 225)	0.46 ± 0.04 (2.6K ± 1.4K)	0.49 ± 0.04 (9.3K ± 4.9K)	0.55 ± 0.02 (38.3K ± 16.2K)
RoT-mRMR$_{40}$	0.40					0.62 ± 0.14 (1.7K ± 0)
Evo-mRMR$_{20}$	0.20	0.12 ± 0.05 (26 ± 2)	0.13 ± 0.04 (422 ± 447)	0.13 ± 0.03 (2.1K ± 2.7K)	0.15 ± 0.04 (12.9K ± 11.7K)	0.15 ± 0.04 (35.5K ± 18.8K)
FetNAS	0.20	0.15 ± 0.04 (26 ± 2)	0.16 ± 0.02 (270 ± 211)	0.17 ± 0.02 (1.2K ± 1.7K)	0.17 ± 0.02 (8.5K ± 6.6K)	0.17 ± 0.02 (17.3K ± 11.5K)
RoT-mRMR$_{20}$	0.20					0.16 ± 0.04 (1.6K ± 0)
another_dataset_on_used_fiat_500						
Evo-All	1.00	0.82 ± 0.01 (77 ± 8)	0.82 ± 0.01 (582 ± 378)	0.82 ± 0.01 (2.1K ± 2.1K)	0.83 ± 0.02 (2.8K ± 2.2K)	0.83 ± 0.01 (14.8K ± 11.9K)
FetNAS	0.45	0.83 ± 0.01 (48 ± 18)	0.83 ± 0.01 (101 ± 62)	0.84 ± 0.01 (1.7K ± 2.3K)	0.83 ± 0.01 (2.5K ± 2.3K)	0.83 ± 0.01 (7.6K ± 6.7K)
RoT-All	1.00					0.84 ± 0.01 (2.1K ± 0)
Evo-mRMR$_{80}$	0.86	0.82 ± 0.01 (67 ± 4)	0.82 ± 0.02 (1.6K ± 1.7K)	0.82 ± 0.01 (2.6K ± 2.2K)	0.82 ± 0.01 (3.0K ± 2.5K)	0.82 ± 0.01 (21.2K ± 22.2K)
FetNAS	0.43	0.83 ± 0.01 (42 ± 11)	0.83 ± 0.00 (390 ± 425)	0.83 ± 0.00 (396 ± 420)	0.83 ± 0.00 (1.5K ± 2.4K)	0.83 ± 0.01 (16.6K ± 22.1K)
RoT-mRMR$_{80}$	0.86					0.84 ± 0.01 (2.0K ± 0)
Evo-mRMR$_{60}$	0.71	0.83 ± 0.01 (60 ± 6)	0.83 ± 0.01 (621 ± 642)	0.83 ± 0.01 (2.5K ± 2.0K)	0.83 ± 0.01 (11.8K ± 19.4K)	0.83 ± 0.01 (26.1K ± 28.4K)
FetNAS	0.44	0.83 ± 0.01 (40 ± 8)	0.83 ± 0.01 (254 ± 308)	0.83 ± 0.00 (473 ± 390)	0.83 ± 0.01 (1.8K ± 3.2K)	0.83 ± 0.01 (5.1K ± 4.6K)
RoT-mRMR$_{60}$	0.71					0.84 ± 0.01 (1.9K ± 0)
Evo-mRMR$_{40}$	0.43	0.82 ± 0.03 (42 ± 2)	0.83 ± 0.01 (2.1K ± 2.8K)	0.83 ± 0.01 (3.7K ± 5.0K)	0.83 ± 0.01 (6.1K ± 6.8K)	0.83 ± 0.01 (25.5K ± 14.6K)
FetNAS	0.36	0.83 ± 0.01 (38 ± 4)	0.83 ± 0.00 (207 ± 281)	0.83 ± 0.00 (207 ± 281)	0.83 ± 0.00 (1.7K ± 3.2K)	0.83 ± 0.01 (14.8K ± 19.4K)
RoT-mRMR$_{40}$	0.43					0.84 ± 0.01 (1.8K ± 0)
Evo-mRMR$_{20}$	0.29	0.79 ± 0.01 (35 ± 4)	0.79 ± 0.01 (336 ± 234)	0.80 ± 0.01 (2.5K ± 2.8K)	0.79 ± 0.01 (7.0K ± 5.8K)	0.79 ± 0.01 (33.1K ± 23.3K)
FetNAS	0.27	0.82 ± 0.03 (33 ± 1)	0.82 ± 0.03 (69 ± 81)	0.83 ± 0.01 (71 ± 80)	0.83 ± 0.01 (6.7K ± 4.5K)	0.83 ± 0.01 (27.0K ± 17.7K)
RoT-mRMR$_{20}$	0.29					0.80 ± 0.02 (1.7K ± 0)
Evo-UF	0.29	0.83 ± 0.01 (35 ± 4)	0.83 ± 0.01 (230 ± 169)	0.83 ± 0.01 (5.5K ± 6.9K)	0.83 ± 0.01 (10.5K ± 8.7K)	0.83 ± 0.01 (39.6K ± 37.7K)
FetNAS	0.28	0.82 ± 0.03 (33 ± 1)	0.83 ± 0.00 (78 ± 64)	0.83 ± 0.01 (2.8K ± 3.7K)	0.83 ± 0.01 (5.0K ± 6.6K)	0.83 ± 0.01 (24.3K ± 20.4K)
RoT-UF	0.29					0.84 ± 0.01 (1.7K ± 0)
concrete_compressive_strength						
Evo-All	1.00	0.58 ± 0.08 (83 ± 4)	0.67 ± 0.05 (862 ± 195)	0.70 ± 0.04 (2.6K ± 1.0K)	0.77 ± 0.01 (11.4K ± 5.5K)	0.79 ± 0.01 (67.6K ± 43.5K)
FetNAS	0.73	0.62 ± 0.06 (71 ± 7)	0.66 ± 0.05 (643 ± 253)	0.67 ± 0.05 (1.4K ± 825)	0.75 ± 0.02 (7.3K ± 3.1K)	0.75 ± 0.06 (37.2K ± 16.1K)
RoT-All	1.00					0.89 ± 0.01 (1.6K ± 0)
Evo-mRMR$_{80}$	0.88	0.56 ± 0.03 (78 ± 8)	0.65 ± 0.03 (892 ± 559)	0.72 ± 0.05 (3.2K ± 1.7K)	0.78 ± 0.03 (9.9K ± 4.0K)	0.78 ± 0.04 (70.5K ± 43.5K)

Method	k/d	R^2 (parameter budget in parentheses) min	Q1	Q2	Q3	max
FetNAS	0.71	0.62 ± 0.06 (71 ± 7)	0.65 ± 0.06 (728 ± 522)	0.71 ± 0.02 (2.7K ± 1.9K)	0.75 ± 0.01 (8.4K ± 4.0K)	0.78 ± 0.01 (28.1K ± 13.1K)
RoT-mRMR$_{80}$	0.88					0.88 ± 0.02 (1.5K ± 0)
Evo-mRMR$_{60}$	0.62	0.42 ± 0.11 (57 ± 0)	0.47 ± 0.11 (640 ± 267)	0.56 ± 0.05 (2.3K ± 699)	0.60 ± 0.09 (9.3K ± 5.5K)	0.65 ± 0.08 (101.4K ± 44.3K)
FetNAS	0.56	0.56 ± 0.09 (51 ± 10)	0.64 ± 0.04 (476 ± 270)	0.68 ± 0.02 (1.8K ± 999)	0.73 ± 0.03 (6.8K ± 2.4K)	0.76 ± 0.03 (39.6K ± 15.5K)
RoT-mRMR$_{60}$	0.62					0.72 ± 0.10 (1.4K ± 0)
Evo-mRMR$_{40}$	0.50	0.42 ± 0.13 (51 ± 5)	0.48 ± 0.12 (461 ± 216)	0.49 ± 0.12 (1.2K ± 981)	0.57 ± 0.11 (4.7K ± 3.4K)	0.59 ± 0.09 (37.1K ± 26.3K)
FetNAS	0.48	0.56 ± 0.08 (50 ± 6)	0.61 ± 0.07 (236 ± 182)	0.62 ± 0.05 (653 ± 299)	0.67 ± 0.07 (3.6K ± 2.0K)	0.72 ± 0.05 (19.9K ± 12.0K)
RoT-mRMR$_{40}$	0.50					0.64 ± 0.11 (1.3K ± 0)
Evo-mRMR$_{20}$	0.25	0.35 ± 0.11 (34 ± 2)	0.37 ± 0.10 (251 ± 54)	0.43 ± 0.09 (1.9K ± 1.9K)	0.45 ± 0.09 (5.0K ± 2.7K)	0.46 ± 0.09 (43.8K ± 20.6K)
FetNAS	0.24	0.39 ± 0.13 (33 ± 3)	0.50 ± 0.04 (229 ± 66)	0.52 ± 0.04 (1.2K ± 1.1K)	0.56 ± 0.05 (2.8K ± 1.9K)	0.58 ± 0.05 (12.9K ± 11.0K)
RoT-mRMR$_{20}$	0.25					0.49 ± 0.09 (1.2K ± 0)
Evo-UF	0.12	0.24 ± 0.03 (27 ± 3)	0.23 ± 0.04 (1.1K ± 1.3K)	0.23 ± 0.04 (3.2K ± 2.8K)	0.23 ± 0.04 (28.9K ± 26.9K)	0.23 ± 0.05 (61.4K ± 48.7K)
FetNAS	0.12	0.21 ± 0.03 (26 ± 3)	0.29 ± 0.05 (211 ± 185)	0.31 ± 0.04 (3.0K ± 2.7K)	0.34 ± 0.03 (9.1K ± 4.8K)	0.34 ± 0.03 (9.1K ± 4.8K)
RoT-UF	0.12					0.24 ± 0.04 (1.1K ± 0)
diamonds						
Evo-All	1.00	0.90 ± 0.02 (98 ± 9)	0.92 ± 0.04 (1.9K ± 819)	0.95 ± 0.01 (6.3K ± 3.4K)	0.96 ± 0.00 (20.0K ± 6.6K)	0.96 ± 0.01 (79.8K ± 27.1K)
FetNAS	0.64	0.88 ± 0.02 (78 ± 8)	0.92 ± 0.03 (1.5K ± 871)	0.87 ± 0.18 (5.4K ± 2.2K)	0.95 ± 0.02 (14.6K ± 5.7K)	0.90 ± 0.13 (55.2K ± 21.7K)
RoT-All	1.00					0.96 ± 0.01 (53.1K ± 0)
Evo-mRMR$_{80}$	0.89	0.85 ± 0.10 (85 ± 9)	0.91 ± 0.07 (2.0K ± 1.1K)	0.94 ± 0.03 (8.1K ± 4.5K)	0.96 ± 0.01 (23.5K ± 5.9K)	0.96 ± 0.01 (111.1K ± 36.7K)
FetNAS	0.61	0.88 ± 0.02 (73 ± 1)	0.92 ± 0.03 (1.9K ± 1.2K)	0.87 ± 0.18 (7.1K ± 4.1K)	0.95 ± 0.02 (17.1K ± 6.9K)	0.91 ± 0.13 (61.2K ± 25.2K)
RoT-mRMR$_{80}$	0.89					0.96 ± 0.00 (52.6K ± 0)
Evo-mRMR$_{60}$	0.67	0.89 ± 0.03 (70 ± 7)	0.92 ± 0.03 (1.5K ± 241)	0.94 ± 0.01 (5.7K ± 2.6K)	0.96 ± 0.01 (19.1K ± 8.1K)	0.96 ± 0.01 (88.5K ± 37.0K)
FetNAS	0.60	0.88 ± 0.02 (68 ± 4)	0.92 ± 0.03 (1.2K ± 278)	0.87 ± 0.18 (4.9K ± 2.1K)	0.92 ± 0.09 (15.4K ± 6.4K)	0.94 ± 0.05 (70.8K ± 38.2K)
RoT-mRMR$_{60}$	0.67					0.90 ± 0.14 (51.6K ± 0)
Evo-mRMR$_{40}$	0.44	0.86 ± 0.07 (49 ± 0)	0.88 ± 0.09 (782 ± 255)	0.94 ± 0.01 (4.4K ± 1.8K)	0.95 ± 0.01 (18.9K ± 12.4K)	0.96 ± 0.00 (76.4K ± 30.0K)
FetNAS	0.40	0.86 ± 0.02 (41 ± 6)	0.88 ± 0.04 (698 ± 237)	0.91 ± 0.01 (3.5K ± 1.7K)	0.88 ± 0.09 (15.4K ± 14.2K)	0.88 ± 0.10 (64.9K ± 36.3K)
RoT-mRMR$_{40}$	0.44					0.96 ± 0.01 (50.7K ± 0)
Evo-mRMR$_{20}$	0.22	0.87 ± 0.01 (35 ± 2)	0.89 ± 0.01 (578 ± 142)	0.92 ± 0.00 (3.3K ± 1.1K)	0.93 ± 0.00 (10.9K ± 3.2K)	0.93 ± 0.00 (68.7K ± 19.4K)
FetNAS	0.22	0.86 ± 0.02 (33 ± 3)	0.86 ± 0.02 (446 ± 189)	0.85 ± 0.04 (2.7K ± 1.1K)	0.86 ± 0.02 (7.5K ± 2.4K)	0.88 ± 0.00 (43.9K ± 7.5K)
RoT-mRMR$_{20}$	0.22					0.94 ± 0.00 (49.7K ± 0)
Evo-UF	0.44	0.70 ± 0.25 (51 ± 3)	0.74 ± 0.22 (492 ± 145)	0.69 ± 0.39 (2.6K ± 1.3K)	0.85 ± 0.06 (8.3K ± 4.7K)	0.85 ± 0.06 (49.2K ± 43.7K)
FetNAS	0.40	0.86 ± 0.02 (45 ± 8)	0.86 ± 0.03 (410 ± 132)	0.89 ± 0.04 (2.1K ± 1.0K)	0.86 ± 0.13 (6.8K ± 4.0K)	0.88 ± 0.10 (29.6K ± 20.1K)
RoT-UF	0.44					0.65 ± 0.29 (50.7K ± 0)
food_delivery_time						
Evo-All	1.00	0.22 ± 0.00 (93 ± 10)	0.24 ± 0.01 (1.1K ± 325)	0.25 ± 0.01 (2.8K ± 918)	0.26 ± 0.01 (6.7K ± 1.9K)	0.26 ± 0.01 (36.7K ± 14.9K)
FetNAS	0.50	0.22 ± 0.00 (63 ± 22)	0.25 ± 0.00 (881 ± 225)	0.26 ± 0.01 (2.6K ± 1.0K)	0.27 ± 0.01 (6.0K ± 1.8K)	0.27 ± 0.01 (24.0K ± 11.0K)
RoT-All	1.00					0.28 ± 0.01 (45.2K ± 0)
Evo-mRMR$_{80}$	0.89	0.22 ± 0.00 (89 ± 11)	0.24 ± 0.01 (864 ± 196)	0.25 ± 0.01 (2.2K ± 873)	0.25 ± 0.00 (5.0K ± 985)	0.26 ± 0.00 (22.9K ± 15.6K)
FetNAS	0.50	0.22 ± 0.01 (59 ± 18)	0.25 ± 0.01 (716 ± 224)	0.26 ± 0.01 (2.1K ± 769)	0.27 ± 0.01 (4.4K ± 1.1K)	0.27 ± 0.01 (16.2K ± 12.6K)
RoT-mRMR$_{80}$	0.89					0.28 ± 0.01 (44.7K ± 0)
Evo-mRMR$_{60}$	0.67	0.22 ± 0.01 (70 ± 4)	0.24 ± 0.01 (804 ± 370)	0.25 ± 0.00 (2.6K ± 1.4K)	0.26 ± 0.00 (7.1K ± 4.5K)	0.27 ± 0.00 (38.1K ± 17.7K)
FetNAS	0.46	0.22 ± 0.01 (57 ± 15)	0.25 ± 0.01 (763 ± 388)	0.26 ± 0.01 (2.2K ± 1.4K)	0.26 ± 0.01 (5.7K ± 3.0K)	0.27 ± 0.01 (27.5K ± 9.0K)
RoT-mRMR$_{60}$	0.67					0.28 ± 0.00 (43.8K ± 0)
Evo-mRMR$_{40}$	0.44	0.22 ± 0.01 (55 ± 6)	0.25 ± 0.01 (817 ± 247)	0.26 ± 0.01 (3.2K ± 1.1K)	0.27 ± 0.01 (17.1K ± 13.6K)	0.28 ± 0.01 (85.1K ± 40.4K)

Method	k/d	R^2 (parameter budget in parentheses) min	Q1	Q2	Q3	max
FetNAS	0.32	0.20 ± 0.03 (44 ± 10)	0.24 ± 0.01 (704 ± 185)	0.25 ± 0.02 (2.6K ± 613)	0.26 ± 0.02 (14.0K ± 10.6K)	0.26 ± 0.02 (40.0K ± 22.4K)
RoT-mRMR$_{40}$	0.44					0.28 ± 0.01 (42.9K ± 0)
Evo-mRMR$_{20}$	0.22	0.14 ± 0.01 (35 ± 4)	0.16 ± 0.01 (501 ± 205)	0.17 ± 0.01 (2.1K ± 1.1K)	0.18 ± 0.01 (10.1K ± 6.5K)	0.18 ± 0.01 (65.2K ± 32.7K)
FetNAS	0.20	0.17 ± 0.03 (34 ± 2)	0.20 ± 0.02 (395 ± 144)	0.22 ± 0.03 (1.8K ± 937)	0.23 ± 0.03 (8.3K ± 5.7K)	0.23 ± 0.03 (20.7K ± 7.7K)
RoT-mRMR$_{20}$	0.22					0.18 ± 0.01 (42.0K ± 0)
healthcare_insurance_expenses						
Evo-All	1.00	0.77 ± 0.03 (65 ± 0)	0.80 ± 0.03 (767 ± 414)	0.81 ± 0.03 (1.4K ± 336)	0.80 ± 0.01 (8.5K ± 7.6K)	0.81 ± 0.02 (55.2K ± 47.0K)
FetNAS	0.55	0.79 ± 0.04 (49 ± 8)	0.81 ± 0.03 (628 ± 460)	0.82 ± 0.03 (1.2K ± 205)	0.83 ± 0.02 (3.5K ± 1.6K)	0.83 ± 0.03 (20.6K ± 20.7K)
RoT-All	1.00					0.84 ± 0.03 (1.8K ± 0)
Evo-mRMR$_{80}$	0.83	0.75 ± 0.03 (60 ± 6)	0.77 ± 0.04 (440 ± 235)	0.79 ± 0.05 (933 ± 316)	0.79 ± 0.06 (7.6K ± 9.6K)	0.80 ± 0.06 (32.8K ± 20.7K)
FetNAS	0.54	0.79 ± 0.04 (52 ± 12)	0.81 ± 0.03 (274 ± 135)	0.82 ± 0.02 (1.0K ± 811)	0.83 ± 0.02 (2.3K ± 1.7K)	0.84 ± 0.03 (14.4K ± 14.3K)
RoT-mRMR$_{80}$	0.83					0.82 ± 0.08 (1.7K ± 0)
Evo-mRMR$_{60}$	0.67	0.71 ± 0.03 (51 ± 3)	0.70 ± 0.03 (1.3K ± 2.4K)	0.71 ± 0.04 (2.3K ± 2.3K)	0.71 ± 0.05 (5.0K ± 2.4K)	0.70 ± 0.06 (19.1K ± 10.9K)
FetNAS	0.54	0.79 ± 0.04 (46 ± 5)	0.81 ± 0.04 (501 ± 564)	0.82 ± 0.04 (1.0K ± 422)	0.82 ± 0.03 (3.3K ± 1.4K)	0.83 ± 0.03 (10.4K ± 10.6K)
RoT-mRMR$_{60}$	0.67					0.70 ± 0.04 (1.7K ± 0)
Evo-mRMR$_{40}$	0.50	0.70 ± 0.03 (41 ± 0)	0.71 ± 0.04 (6.1K ± 12.8K)	0.71 ± 0.04 (7.8K ± 13.8K)	0.70 ± 0.05 (45.5K ± 37.7K)	0.70 ± 0.05 (59.2K ± 47.7K)
FetNAS	0.49	0.78 ± 0.03 (40 ± 2)	0.81 ± 0.03 (3.7K ± 7.9K)	0.82 ± 0.03 (4.8K ± 7.5K)	0.84 ± 0.02 (26.2K ± 22.8K)	0.84 ± 0.02 (26.7K ± 22.1K)
RoT-mRMR$_{40}$	0.50					0.71 ± 0.04 (1.6K ± 0)
Evo-mRMR$_{20}$	0.33	0.72 ± 0.04 (35 ± 2)	0.72 ± 0.03 (988 ± 2.0K)	0.72 ± 0.04 (3.1K ± 6.2K)	0.72 ± 0.04 (11.2K ± 20.5K)	0.72 ± 0.04 (27.4K ± 24.8K)
FetNAS	0.33	0.72 ± 0.04 (34 ± 2)	0.72 ± 0.04 (247 ± 398)	0.73 ± 0.04 (1.8K ± 3.5K)	0.72 ± 0.04 (2.7K ± 3.3K)	0.73 ± 0.04 (3.7K ± 3.5K)
RoT-mRMR$_{20}$	0.33					0.72 ± 0.04 (1.5K ± 0)
Evo-UF	0.17	0.62 ± 0.03 (27 ± 3)	0.62 ± 0.03 (127 ± 119)	0.62 ± 0.04 (333 ± 179)	0.62 ± 0.04 (1.7K ± 1.8K)	0.62 ± 0.03 (33.6K ± 32.4K)
FetNAS	0.17	0.61 ± 0.04 (27 ± 3)	0.62 ± 0.04 (80 ± 110)	0.62 ± 0.04 (147 ± 127)	0.62 ± 0.04 (361 ± 413)	0.62 ± 0.04 (4.6K ± 9.8K)
RoT-UF	0.17					0.62 ± 0.04 (1.4K ± 0)
houses						
Evo-All	1.00	0.72 ± 0.01 (85 ± 9)	0.75 ± 0.01 (1.4K ± 379)	0.77 ± 0.01 (3.8K ± 241)	0.78 ± 0.00 (15.6K ± 4.4K)	0.79 ± 0.01 (68.7K ± 27.9K)
FetNAS	0.87	0.68 ± 0.01 (80 ± 13)	0.72 ± 0.02 (1.2K ± 324)	0.74 ± 0.01 (3.1K ± 730)	0.76 ± 0.00 (11.4K ± 4.2K)	0.76 ± 0.01 (50.4K ± 19.3K)
RoT-All	1.00					0.80 ± 0.00 (21.3K ± 0)
Evo-mRMR$_{80}$	0.88	0.71 ± 0.02 (78 ± 8)	0.75 ± 0.02 (1.6K ± 1.0K)	0.77 ± 0.01 (4.4K ± 3.0K)	0.78 ± 0.00 (19.1K ± 17.4K)	0.79 ± 0.01 (116.7K ± 80.6K)
FetNAS	0.79	0.68 ± 0.01 (70 ± 7)	0.72 ± 0.01 (1.4K ± 831)	0.75 ± 0.01 (4.2K ± 3.0K)	0.75 ± 0.00 (15.8K ± 14.4K)	0.76 ± 0.01 (52.8K ± 20.8K)
RoT-mRMR$_{80}$	0.88					0.80 ± 0.01 (21.0K ± 0)
Evo-mRMR$_{60}$	0.62	0.67 ± 0.01 (60 ± 4)	0.70 ± 0.02 (735 ± 168)	0.72 ± 0.01 (3.0K ± 921)	0.73 ± 0.01 (9.6K ± 3.0K)	0.74 ± 0.00 (68.3K ± 31.2K)
FetNAS	0.61	0.63 ± 0.06 (56 ± 4)	0.66 ± 0.04 (530 ± 152)	0.71 ± 0.04 (2.5K ± 660)	0.73 ± 0.03 (8.9K ± 3.1K)	0.74 ± 0.02 (42.2K ± 21.8K)
RoT-mRMR$_{60}$	0.62					0.75 ± 0.01 (20.4K ± 0)
Evo-mRMR$_{40}$	0.50	0.54 ± 0.01 (51 ± 5)	0.59 ± 0.02 (1.4K ± 825)	0.61 ± 0.01 (5.0K ± 1.9K)	0.62 ± 0.01 (16.2K ± 8.2K)	0.63 ± 0.01 (78.7K ± 39.1K)
FetNAS	0.50	0.58 ± 0.07 (51 ± 5)	0.64 ± 0.04 (1.1K ± 981)	0.67 ± 0.02 (4.4K ± 1.8K)	0.69 ± 0.02 (13.2K ± 6.0K)	0.69 ± 0.02 (27.1K ± 12.0K)
RoT-mRMR$_{40}$	0.50					0.64 ± 0.01 (20.1K ± 0)
Evo-mRMR$_{20}$	0.25	0.49 ± 0.02 (35 ± 2)	0.54 ± 0.03 (732 ± 573)	0.57 ± 0.02 (2.6K ± 1.7K)	0.59 ± 0.02 (12.8K ± 8.5K)	0.60 ± 0.02 (68.9K ± 36.6K)
FetNAS	0.24	0.29 ± 0.15 (33 ± 1)	0.42 ± 0.02 (559 ± 378)	0.45 ± 0.02 (2.4K ± 1.6K)	0.47 ± 0.02 (9.2K ± 5.8K)	0.49 ± 0.01 (51.8K ± 27.4K)
RoT-mRMR$_{20}$	0.25					0.60 ± 0.02 (19.5K ± 0)
Evo-UF	0.12	0.46 ± 0.01 (26 ± 2)	0.46 ± 0.01 (856 ± 393)	0.47 ± 0.01 (3.5K ± 1.7K)	0.47 ± 0.01 (17.8K ± 13.3K)	0.47 ± 0.01 (72.6K ± 23.6K)
FetNAS	0.12	0.02 ± 0.01 (26 ± 2)	0.14 ± 0.06 (665 ± 334)	0.21 ± 0.03 (2.7K ± 1.3K)	0.24 ± 0.02 (8.9K ± 4.1K)	0.24 ± 0.01 (38.5K ± 17.2K)
RoT-UF	0.12					0.47 ± 0.01 (19.2K ± 0)
miami_housing						
Evo-All	1.00	0.82 ± 0.01 (140 ± 8)	0.85 ± 0.01 (1.7K ± 674)	0.87 ± 0.01 (4.7K ± 3.4K)	0.88 ± 0.01 (12.1K ± 4.6K)	0.89 ± 0.01 (75.7K ± 56.8K)

Method	k/d	R^2 (parameter budget in parentheses) min	Q1	Q2	Q3	max
FetNAS	0.80	0.83 ± 0.01 (121 ± 11)	0.83 ± 0.01 (1.4K ± 855)	0.85 ± 0.02 (4.3K ± 3.1K)	0.86 ± 0.01 (10.4K ± 2.8K)	0.87 ± 0.01 (55.3K ± 41.1K)
RoT-All	1.00					0.90 ± 0.01 (16.5K ± 0)
Evo-mRMR$_{80}$	0.80	0.79 ± 0.03 (124 ± 12)	0.83 ± 0.01 (1.8K ± 480)	0.85 ± 0.01 (6.3K ± 2.4K)	0.87 ± 0.01 (22.6K ± 15.5K)	0.87 ± 0.01 (74.6K ± 20.6K)
FetNAS	0.73	0.82 ± 0.01 (105 ± 10)	0.83 ± 0.02 (1.4K ± 834)	0.86 ± 0.01 (5.5K ± 2.4K)	0.87 ± 0.01 (20.3K ± 12.0K)	0.87 ± 0.01 (47.8K ± 23.0K)
RoT-mRMR$_{80}$	0.80					0.89 ± 0.02 (15.7K ± 0)
Evo-mRMR$_{60}$	0.60	0.78 ± 0.01 (98 ± 9)	0.80 ± 0.02 (1.1K ± 207)	0.83 ± 0.01 (2.7K ± 700)	0.85 ± 0.01 (12.8K ± 10.9K)	0.85 ± 0.01 (47.7K ± 25.0K)
FetNAS	0.58	0.81 ± 0.01 (87 ± 4)	0.82 ± 0.01 (983 ± 239)	0.84 ± 0.01 (2.3K ± 232)	0.85 ± 0.00 (9.5K ± 6.9K)	0.86 ± 0.01 (26.9K ± 16.3K)
RoT-mRMR$_{60}$	0.60					0.87 ± 0.02 (15.0K ± 0)
Evo-mRMR$_{40}$	0.40	0.72 ± 0.04 (70 ± 7)	0.74 ± 0.04 (1.1K ± 594)	0.76 ± 0.04 (2.7K ± 886)	0.76 ± 0.04 (7.1K ± 2.2K)	0.77 ± 0.04 (43.5K ± 33.3K)
FetNAS	0.38	0.78 ± 0.01 (60 ± 4)	0.80 ± 0.02 (911 ± 524)	0.81 ± 0.02 (2.2K ± 997)	0.83 ± 0.01 (7.2K ± 5.1K)	0.83 ± 0.01 (19.7K ± 17.5K)
RoT-mRMR$_{40}$	0.40					0.79 ± 0.03 (14.2K ± 0)
Evo-mRMR$_{20}$	0.20	0.59 ± 0.02 (42 ± 2)	0.60 ± 0.03 (1.1K ± 931)	0.60 ± 0.03 (3.1K ± 2.2K)	0.61 ± 0.04 (17.8K ± 17.0K)	0.61 ± 0.03 (57.6K ± 23.6K)
FetNAS	0.19	0.60 ± 0.11 (37 ± 5)	0.69 ± 0.02 (286 ± 147)	0.70 ± 0.02 (2.5K ± 2.3K)	0.71 ± 0.01 (15.7K ± 15.6K)	0.73 ± 0.01 (41.4K ± 20.7K)
RoT-mRMR$_{20}$	0.20					0.62 ± 0.04 (13.5K ± 0)
Evo-UF	0.13	0.50 ± 0.02 (35 ± 2)	0.50 ± 0.02 (179 ± 112)	0.50 ± 0.02 (453 ± 255)	0.50 ± 0.02 (2.8K ± 4.4K)	0.50 ± 0.02 (12.4K ± 9.4K)
FetNAS	0.13	0.58 ± 0.10 (33 ± 1)	0.58 ± 0.10 (68 ± 68)	0.63 ± 0.03 (317 ± 234)	0.64 ± 0.03 (1.8K ± 2.3K)	0.65 ± 0.03 (8.0K ± 5.9K)
RoT-UF	0.13					0.50 ± 0.02 (13.2K ± 0)
physiochemical_protein						
Evo-All	1.00	0.37 ± 0.01 (93 ± 10)	0.44 ± 0.02 (2.0K ± 743)	0.49 ± 0.02 (7.9K ± 5.9K)	0.52 ± 0.02 (27.2K ± 16.3K)	0.56 ± 0.02 (116.4K ± 61.0K)
FetNAS	0.87	0.33 ± 0.00 (85 ± 4)	0.38 ± 0.02 (1.8K ± 722)	0.41 ± 0.02 (6.8K ± 5.5K)	0.44 ± 0.01 (21.9K ± 12.8K)	0.45 ± 0.01 (49.5K ± 18.3K)
RoT-All	1.00					0.48 ± 0.02 (45.7K ± 0)
Evo-mRMR$_{80}$	0.89	0.36 ± 0.01 (91 ± 10)	0.42 ± 0.01 (1.7K ± 304)	0.47 ± 0.02 (5.0K ± 990)	0.51 ± 0.01 (19.0K ± 2.6K)	0.54 ± 0.01 (131.5K ± 32.9K)
FetNAS	0.84	0.33 ± 0.01 (79 ± 16)	0.37 ± 0.02 (1.4K ± 281)	0.41 ± 0.02 (4.4K ± 996)	0.44 ± 0.01 (16.3K ± 3.0K)	0.45 ± 0.01 (68.0K ± 38.1K)
RoT-mRMR$_{80}$	0.89					0.47 ± 0.02 (45.3K ± 0)
Evo-mRMR$_{60}$	0.67	0.33 ± 0.01 (68 ± 7)	0.38 ± 0.01 (1.9K ± 230)	0.42 ± 0.01 (5.2K ± 841)	0.46 ± 0.01 (16.7K ± 8.7K)	0.48 ± 0.01 (105.1K ± 42.8K)
FetNAS	0.63	0.25 ± 0.15 (67 ± 9)	0.36 ± 0.02 (1.6K ± 234)	0.39 ± 0.03 (4.7K ± 639)	0.40 ± 0.03 (13.8K ± 6.6K)	0.42 ± 0.03 (52.0K ± 19.6K)
RoT-mRMR$_{60}$	0.67					0.44 ± 0.01 (44.4K ± 0)
Evo-mRMR$_{40}$	0.44	0.23 ± 0.03 (54 ± 7)	0.26 ± 0.04 (1.2K ± 870)	0.30 ± 0.04 (4.6K ± 2.6K)	0.33 ± 0.04 (13.7K ± 5.8K)	0.35 ± 0.04 (104.4K ± 20.8K)
FetNAS	0.43	0.15 ± 0.09 (46 ± 10)	0.24 ± 0.02 (1.1K ± 897)	0.27 ± 0.01 (4.1K ± 3.0K)	0.29 ± 0.01 (12.8K ± 6.4K)	0.30 ± 0.00 (71.9K ± 30.6K)
RoT-mRMR$_{40}$	0.44					0.33 ± 0.02 (43.5K ± 0)
Evo-mRMR$_{20}$	0.22	0.18 ± 0.01 (35 ± 4)	0.18 ± 0.01 (462 ± 206)	0.19 ± 0.01 (2.7K ± 1.9K)	0.19 ± 0.01 (8.9K ± 3.0K)	0.19 ± 0.01 (51.8K ± 33.7K)
FetNAS	0.20	-0.00 ± 0.01 (30 ± 3)	0.00 ± 0.00 (419 ± 223)	0.01 ± 0.01 (2.3K ± 2.1K)	0.02 ± 0.00 (7.6K ± 2.1K)	0.02 ± 0.01 (29.0K ± 19.9K)
RoT-mRMR$_{20}$	0.22					0.19 ± 0.01 (42.6K ± 0)
qsar_fish_toxicity						
Evo-All	1.00	0.50 ± 0.08 (67 ± 4)	0.55 ± 0.04 (718 ± 1.1K)	0.55 ± 0.06 (4.7K ± 3.3K)	0.53 ± 0.04 (28.7K ± 34.1K)	0.57 ± 0.03 (48.4K ± 36.0K)
FetNAS	0.77	0.52 ± 0.02 (54 ± 14)	0.55 ± 0.02 (420 ± 368)	0.53 ± 0.04 (2.3K ± 2.5K)	0.56 ± 0.03 (11.1K ± 6.1K)	0.56 ± 0.03 (11.1K ± 6.1K)
RoT-All	1.00					0.58 ± 0.03 (1.3K ± 0)
Evo-mRMR$_{80}$	0.83	0.50 ± 0.06 (57 ± 0)	0.53 ± 0.04 (844 ± 701)	0.54 ± 0.02 (2.5K ± 2.0K)	0.53 ± 0.02 (4.2K ± 1.9K)	0.52 ± 0.04 (36.7K ± 36.2K)
FetNAS	0.72	0.52 ± 0.03 (50 ± 8)	0.55 ± 0.02 (245 ± 175)	0.56 ± 0.03 (492 ± 440)	0.55 ± 0.03 (1.3K ± 1.5K)	0.56 ± 0.02 (8.8K ± 7.8K)
RoT-mRMR$_{80}$	0.83					0.57 ± 0.03 (1.2K ± 0)
Evo-mRMR$_{60}$	0.67	0.51 ± 0.04 (50 ± 3)	0.53 ± 0.04 (356 ± 235)	0.55 ± 0.04 (5.0K ± 6.9K)	0.54 ± 0.03 (11.9K ± 7.5K)	0.55 ± 0.02 (39.6K ± 24.5K)
FetNAS	0.57	0.51 ± 0.02 (45 ± 4)	0.54 ± 0.03 (215 ± 285)	0.52 ± 0.03 (3.2K ± 5.0K)	0.52 ± 0.03 (8.5K ± 7.8K)	0.54 ± 0.05 (19.2K ± 16.2K)
RoT-mRMR$_{60}$	0.67					0.58 ± 0.02 (1.2K ± 0)
Evo-mRMR$_{40}$	0.50	0.49 ± 0.02 (42 ± 2)	0.53 ± 0.05 (216 ± 94)	0.54 ± 0.03 (413 ± 233)	0.53 ± 0.06 (5.4K ± 8.2K)	0.52 ± 0.05 (17.1K ± 13.4K)

Method	k/d	R^2 (parameter budget in parentheses) min	Q1	Q2	Q3	max
FetNAS	0.43	0.50 ± 0.02 (38 ± 4)	0.50 ± 0.02 (145 ± 109)	0.51 ± 0.02 (195 ± 179)	0.51 ± 0.03 (2.2K ± 3.8K)	0.51 ± 0.03 (9.5K ± 11.3K)
RoT-mRMR$_{40}$	0.50					0.56 ± 0.03 (1.1K ± 0)
Evo-mRMR$_{20}$	0.33	0.48 ± 0.03 (33 ± 0)	0.49 ± 0.05 (3.5K ± 6.4K)	0.47 ± 0.04 (14.8K ± 15.8K)	0.48 ± 0.02 (31.1K ± 19.9K)	0.47 ± 0.03 (50.7K ± 36.2K)
FetNAS	0.33	0.50 ± 0.02 (32 ± 2)	0.50 ± 0.02 (313 ± 298)	0.49 ± 0.02 (2.6K ± 3.8K)	0.49 ± 0.03 (11.6K ± 10.3K)	0.49 ± 0.04 (15.2K ± 10.8K)
RoT-mRMR$_{20}$	0.33					0.51 ± 0.03 (1.1K ± 0)
Evo-UF	0.37	0.43 ± 0.09 (35 ± 4)	0.45 ± 0.08 (185 ± 116)	0.45 ± 0.07 (1.4K ± 981)	0.43 ± 0.08 (11.4K ± 9.8K)	0.45 ± 0.09 (51.2K ± 35.5K)
FetNAS	0.36	0.50 ± 0.02 (34 ± 5)	0.50 ± 0.02 (59 ± 29)	0.50 ± 0.02 (397 ± 255)	0.50 ± 0.02 (4.3K ± 6.8K)	0.48 ± 0.04 (15.2K ± 17.4K)
RoT-UF	0.37					0.49 ± 0.08 (1.1K ± 29)
qsar_tid_11						
Evo-All	1.00	0.62 ± 0.02 (9.2K ± 1.0K)	0.63 ± 0.03 (33.7K ± 16.5K)	0.65 ± 0.02 (53.9K ± 21.6K)	0.65 ± 0.02 (92.1K ± 28.0K)	0.67 ± 0.02 (245.2K ± 81.5K)
FetNAS	0.64	0.64 ± 0.03 (6.2K ± 1.5K)	0.64 ± 0.02 (21.9K ± 6.1K)	0.65 ± 0.03 (48.7K ± 20.6K)	0.65 ± 0.03 (78.2K ± 26.9K)	0.66 ± 0.02 (136.7K ± 20.1K)
RoT-All	1.00					0.66 ± 0.02 (169.2K ± 0)
Evo-mRMR$_{80}$	0.80	0.62 ± 0.02 (6.9K ± 450)	0.63 ± 0.02 (38.5K ± 7.0K)	0.64 ± 0.03 (56.3K ± 17.1K)	0.65 ± 0.03 (102.3K ± 13.6K)	0.66 ± 0.02 (214.3K ± 73.7K)
FetNAS	0.64	0.64 ± 0.03 (5.7K ± 940)	0.63 ± 0.02 (29.7K ± 9.5K)	0.64 ± 0.03 (45.9K ± 20.4K)	0.65 ± 0.02 (84.0K ± 15.6K)	0.66 ± 0.02 (134.6K ± 22.0K)
RoT-mRMR$_{80}$	0.80					0.65 ± 0.02 (136.5K ± 0)
Evo-mRMR$_{60}$	0.60	0.62 ± 0.03 (5.1K ± 276)	0.62 ± 0.03 (36.4K ± 16.7K)	0.63 ± 0.03 (57.5K ± 25.5K)	0.65 ± 0.03 (98.3K ± 42.2K)	0.66 ± 0.02 (159.7K ± 61.4K)
FetNAS	0.51	0.63 ± 0.03 (4.3K ± 854)	0.62 ± 0.03 (26.4K ± 18.8K)	0.63 ± 0.03 (43.3K ± 16.9K)	0.65 ± 0.03 (77.7K ± 22.8K)	0.65 ± 0.03 (104.3K ± 48.4K)
RoT-mRMR$_{60}$	0.60					0.65 ± 0.02 (103.7K ± 0)
Evo-mRMR$_{40}$	0.40	0.62 ± 0.02 (3.5K ± 226)	0.62 ± 0.02 (23.1K ± 6.1K)	0.64 ± 0.03 (56.4K ± 22.2K)	0.65 ± 0.02 (89.8K ± 29.6K)	0.65 ± 0.02 (127.7K ± 28.8K)
FetNAS	0.34	0.62 ± 0.03 (2.9K ± 155)	0.62 ± 0.03 (21.9K ± 9.6K)	0.63 ± 0.04 (40.7K ± 12.8K)	0.63 ± 0.03 (60.0K ± 20.7K)	0.63 ± 0.03 (67.8K ± 25.5K)
RoT-mRMR$_{40}$	0.40					0.65 ± 0.02 (70.9K ± 0)
Evo-mRMR$_{20}$	0.20	0.58 ± 0.03 (1.7K ± 113)	0.59 ± 0.03 (13.7K ± 4.9K)	0.61 ± 0.02 (28.1K ± 14.1K)	0.62 ± 0.03 (37.1K ± 13.1K)	0.63 ± 0.03 (81.7K ± 22.9K)
FetNAS	0.18	0.55 ± 0.05 (1.3K ± 159)	0.57 ± 0.04 (8.7K ± 7.4K)	0.58 ± 0.03 (18.0K ± 14.8K)	0.58 ± 0.03 (22.6K ± 14.0K)	0.59 ± 0.03 (55.6K ± 17.7K)
RoT-mRMR$_{20}$	0.20					0.63 ± 0.03 (38.1K ± 0)
superconductivity						
Evo-All	1.00	0.81 ± 0.01 (715 ± 74)	0.83 ± 0.00 (7.3K ± 1.4K)	0.85 ± 0.00 (22.1K ± 8.0K)	0.86 ± 0.00 (75.4K ± 36.8K)	0.87 ± 0.01 (208.7K ± 39.9K)
FetNAS	0.80	0.79 ± 0.01 (607 ± 133)	0.81 ± 0.01 (5.9K ± 1.6K)	0.82 ± 0.01 (19.8K ± 8.8K)	0.82 ± 0.01 (41.1K ± 17.0K)	0.83 ± 0.01 (132.7K ± 60.7K)
RoT-All	1.00					0.83 ± 0.01 (44.5K ± 0)
Evo-mRMR$_{80}$	0.80	0.80 ± 0.01 (564 ± 60)	0.83 ± 0.01 (7.2K ± 1.8K)	0.85 ± 0.00 (19.7K ± 7.1K)	0.86 ± 0.00 (74.5K ± 36.7K)	0.87 ± 0.01 (216.5K ± 25.3K)
FetNAS	0.70	0.79 ± 0.01 (473 ± 51)	0.81 ± 0.01 (5.5K ± 1.4K)	0.82 ± 0.00 (16.9K ± 8.3K)	0.82 ± 0.01 (48.5K ± 7.4K)	0.82 ± 0.00 (72.8K ± 30.0K)
RoT-mRMR$_{80}$	0.80					0.83 ± 0.00 (39.5K ± 0)
Evo-mRMR$_{60}$	0.60	0.79 ± 0.00 (460 ± 51)	0.82 ± 0.01 (5.6K ± 1.6K)	0.84 ± 0.00 (17.9K ± 2.7K)	0.85 ± 0.00 (53.7K ± 9.0K)	0.87 ± 0.01 (208.5K ± 29.4K)
FetNAS	0.58	0.78 ± 0.02 (418 ± 24)	0.81 ± 0.00 (4.7K ± 1.6K)	0.82 ± 0.00 (14.9K ± 2.5K)	0.82 ± 0.00 (42.9K ± 11.2K)	0.82 ± 0.01 (70.1K ± 35.2K)
RoT-mRMR$_{60}$	0.60					0.84 ± 0.00 (34.6K ± 0)
Evo-mRMR$_{40}$	0.41	0.78 ± 0.01 (316 ± 35)	0.82 ± 0.01 (4.7K ± 881)	0.83 ± 0.01 (16.2K ± 4.4K)	0.85 ± 0.00 (81.4K ± 36.7K)	0.87 ± 0.01 (202.0K ± 38.1K)
FetNAS	0.35	0.73 ± 0.02 (260 ± 32)	0.79 ± 0.01 (4.3K ± 985)	0.80 ± 0.01 (12.2K ± 3.1K)	0.81 ± 0.00 (41.6K ± 16.5K)	0.81 ± 0.00 (41.6K ± 16.5K)
RoT-mRMR$_{40}$	0.41					0.84 ± 0.01 (29.7K ± 0)
Evo-mRMR$_{20}$	0.21	0.73 ± 0.01 (168 ± 16)	0.78 ± 0.00 (2.2K ± 621)	0.81 ± 0.00 (5.8K ± 965)	0.82 ± 0.00 (27.0K ± 9.8K)	0.84 ± 0.01 (140.8K ± 49.8K)
FetNAS	0.19	0.69 ± 0.03 (136 ± 5)	0.74 ± 0.03 (1.9K ± 798)	0.77 ± 0.01 (5.0K ± 1.3K)	0.78 ± 0.01 (19.8K ± 3.8K)	0.79 ± 0.01 (63.4K ± 26.1K)
RoT-mRMR$_{20}$	0.21					0.82 ± 0.00 (24.8K ± 0)
Evo-UF	0.44	0.77 ± 0.01 (330 ± 34)	0.81 ± 0.01 (3.6K ± 286)	0.82 ± 0.01 (10.3K ± 1.0K)	0.84 ± 0.01 (29.6K ± 4.5K)	0.85 ± 0.01 (149.2K ± 47.4K)
FetNAS	0.38	0.73 ± 0.02 (260 ± 32)	0.79 ± 0.01 (2.8K ± 805)	0.81 ± 0.00 (8.7K ± 755)	0.81 ± 0.00 (27.5K ± 5.7K)	0.81 ± 0.00 (54.0K ± 19.6K)
RoT-UF	0.44					0.82 ± 0.00 (30.4K ± 275)
wine_quality						
Evo-All	1.00	0.31 ± 0.01 (119 ± 13)	0.33 ± 0.02 (2.1K ± 408)	0.34 ± 0.01 (3.7K ± 1.1K)	0.35 ± 0.03 (11.1K ± 7.4K)	0.35 ± 0.03 (67.8K ± 57.8K)

Method	k/d	R^2 (parameter budget in parentheses)				
		min	Q1	Q2	Q3	max
FetNAS	0.73	0.30 ± 0.01 (97 ± 15)	0.31 ± 0.01 (2.0K ± 426)	0.31 ± 0.01 (2.8K ± 880)	0.32 ± 0.01 (8.3K ± 3.5K)	0.34 ± 0.01 (38.9K ± 22.6K)
RoT-All	1.00					0.35 ± 0.01 (8.1K ± 0)
Evo-mRMR$_{80}$	0.83	0.31 ± 0.02 (99 ± 5)	0.32 ± 0.02 (1.7K ± 669)	0.34 ± 0.02 (4.4K ± 2.3K)	0.34 ± 0.02 (9.6K ± 3.2K)	0.35 ± 0.02 (32.1K ± 14.2K)
FetNAS	0.69	0.30 ± 0.01 (88 ± 11)	0.31 ± 0.01 (1.7K ± 675)	0.31 ± 0.01 (2.7K ± 1.3K)	0.32 ± 0.01 (8.4K ± 3.2K)	0.33 ± 0.01 (24.9K ± 11.4K)
RoT-mRMR$_{80}$	0.83					0.35 ± 0.02 (7.7K ± 0)
Evo-mRMR$_{60}$	0.67	0.29 ± 0.02 (85 ± 5)	0.30 ± 0.02 (1.1K ± 925)	0.32 ± 0.02 (3.0K ± 2.7K)	0.33 ± 0.02 (10.8K ± 8.1K)	0.34 ± 0.02 (33.6K ± 34.2K)
FetNAS	0.61	0.30 ± 0.01 (79 ± 14)	0.31 ± 0.01 (972 ± 871)	0.31 ± 0.01 (3.2K ± 3.3K)	0.32 ± 0.01 (7.6K ± 5.8K)	0.33 ± 0.01 (22.9K ± 20.1K)
RoT-mRMR$_{60}$	0.67					0.35 ± 0.02 (7.4K ± 0)
Evo-mRMR$_{40}$	0.42	0.26 ± 0.03 (60 ± 6)	0.27 ± 0.03 (706 ± 215)	0.28 ± 0.03 (1.9K ± 1.5K)	0.28 ± 0.03 (6.0K ± 4.9K)	0.29 ± 0.03 (29.6K ± 16.3K)
FetNAS	0.39	0.27 ± 0.03 (53 ± 7)	0.28 ± 0.02 (625 ± 192)	0.29 ± 0.02 (1.3K ± 847)	0.29 ± 0.01 (2.1K ± 1.6K)	0.30 ± 0.02 (12.4K ± 6.8K)
RoT-mRMR$_{40}$	0.42					0.30 ± 0.02 (6.9K ± 0)
Evo-mRMR$_{20}$	0.25	0.23 ± 0.01 (43 ± 3)	0.23 ± 0.01 (1.4K ± 703)	0.23 ± 0.01 (4.0K ± 1.8K)	0.23 ± 0.01 (11.5K ± 4.3K)	0.23 ± 0.01 (33.8K ± 35.8K)
FetNAS	0.25	0.23 ± 0.04 (40 ± 5)	0.24 ± 0.04 (751 ± 496)	0.25 ± 0.03 (2.6K ± 2.0K)	0.25 ± 0.04 (5.2K ± 3.8K)	0.25 ± 0.04 (7.0K ± 4.2K)
RoT-mRMR$_{20}$	0.25					0.23 ± 0.01 (6.5K ± 0)
Evo-UF	0.08	0.20 ± 0.01 (26 ± 2)	0.20 ± 0.01 (728 ± 212)	0.20 ± 0.01 (3.3K ± 2.5K)	0.20 ± 0.01 (12.7K ± 12.5K)	0.21 ± 0.01 (28.8K ± 17.7K)
FetNAS	0.08	0.08 ± 0.07 (26 ± 1)	0.10 ± 0.07 (395 ± 253)	0.10 ± 0.07 (947 ± 564)	0.11 ± 0.07 (4.4K ± 2.7K)	0.10 ± 0.07 (22.6K ± 19.5K)
RoT-UF	0.08					0.21 ± 0.01 (6.2K ± 0)

E ALTERNATIVE IMPORTANCE RANKING METRICS

This appendix reports the raw 3D hypervolume values underlying the importance-ranking comparison in the main text. For each dataset and each importance metric, we compute the hypervolume of the *zero-shot* Pareto front spanned by test performance, parameter count and number of features, averaged over five cross-validation splits (mean ± std). Higher hypervolume indicates a superior overall trade-off across these three objectives.

Table 7: **3D hypervolume (test performance, parameter count, feature count) per dataset for *zero-shot* subnetworks in the Pareto front, mean ± std over 5 cross-validation splits.** Higher values indicate superior overall trade-offs.

Dataset	G	WN	GIN	R
airfoil_self_noise	402.6k ±14.8k	245.9k ±113.1k	410.5k ±22.7k	266.1k ±27.5k
amazon_employee_access	75.5 ±18.8	75.5 ±18.8	75.5 ±18.8	75.5 ±18.8
anneal	2.86M ±587.6k	2.92M ±374.8k	3.05M ±424.4k	1.50M ±434.7k
another_dataset_on_used_fiat_500	239.5k ±2.5k	227.6k ±16.7k	239.3k ±2.7k	217.3k ±5.2k
apsfailure	3.34M ±527.5k	3.32M ±951.2k	3.66M ±781.9k	3.93M ±663.4k
bank_customer_churn	243.5k ±28.4k	232.9k ±26.6k	241.4k ±37.9k	119.7k ±2.9k
bioresponse	127.38M ±5.06M	125.89M ±6.01M	107.15M ±48.61M	38.57M ±25.37M
blood_transfusion_service_center	27.0k ±16.1k	32.7k ±24.5k	24.5k ±21.3k	18.3k ±11.3k
churn	345.1k ±58.1k	339.3k ±44.6k	347.8k ±52.1k	181.2k ±71.1k
coil2000_insurance_policies	27.3k ±37.9k	13.5k ±21.2k	18.1k ±35.6k	2.0k ±185.8
concrete_compressive_strength	938.7k ±58.2k	783.6k ±68.4k	944.3k ±63.4k	544.5k ±31.5k
credit_card_clients_default	501.1k ±11.7k	504.1k ±17.5k	497.4k ±11.1k	340.7k ±147.0k
credit_g	201.2k ±36.2k	123.6k ±61.0k	197.6k ±38.5k	62.3k ±61.6k
customer_satisfaction_in_airline	1.11M ±865.4k	1.07M ±831.8k	798.0k ±717.7k	544.5k ±484.8k
diabetes130us	3.39M ±22.9k	3.38M ±10.4k	3.38M ±13.9k	3.37M ±444.8
diamonds	1.55M ±14.5k	1.38M ±188.1k	1.55M ±15.5k	1.46M ±7.2k
e_commereshippingdata	250.4k ±4.8k	249.9k ±16.4k	252.9k ±6.2k	157.9k ±95.3k
fitness_club	119.2k ±5.6k	96.1k ±36.4k	116.1k ±6.9k	85.1k ±21.4k
food_delivery_time	252.6k ±21.2k	231.3k ±11.7k	248.2k ±29.5k	97.4k ±18.7k
givemesomecredit	28.2k ±13.4k	36.6k ±14.7k	30.1k ±12.4k	82.9k ±32.8k
hazelnut_spread_contaminant_detection	1.25M ±106.5k	1.27M ±57.8k	1.25M ±138.4k	913.9k ±249.4k
healthcare_insurance_expenses	404.8k ±15.7k	272.3k ±106.9k	403.7k ±15.4k	343.3k ±6.8k
heloc	748.8k ±20.3k	738.5k ±35.8k	745.6k ±23.6k	666.8k ±62.4k
hiva_agnostic	129.7k ±1.0k	129.7k ±1.0k	129.6k ±988.9	129.6k ±981.4
houses	419.4k ±16.5k	337.1k ±18.2k	418.8k ±13.1k	278.4k ±9.6k
hr_analytics_job_change_of_data_scientists	366.4k ±33.6k	292.4k ±37.0k	382.3k ±12.7k	76.0k ±106.7k
in_vehicle_coupon_recommendation	458.4k ±21.9k	427.9k ±31.3k	456.6k ±19.9k	325.5k ±70.9k
is_this_a_good_customer	7.4k ±5.0k	11.1k ±15.4k	5.8k ±2.7k	3.6k ±232.7
jm1	208.7k ±38.1k	189.9k ±67.4k	206.4k ±43.5k	116.5k ±34.2k
kddcup09_appetency	2.08M ±11.3k	2.11M ±76.8k	2.07M ±3.3k	2.08M ±10.5k
marketing_campaign	492.0k ±66.1k	445.9k ±66.0k	522.0k ±57.6k	341.4k ±118.8k
maternal_health_risk	290.6k ±11.4k	256.3k ±49.9k	291.0k ±7.0k	240.4k ±14.0k
miami_housing	2.27M ±47.0k	1.80M ±167.6k	2.24M ±106.9k	1.27M ±144.3k
mic	600.8k ±31.7k	602.3k ±29.8k	605.7k ±27.3k	180.8k ±252.1k
naticusdroid	6.62M ±35.5k	6.53M ±64.5k	6.59M ±23.6k	5.08M ±176.8k
online_shoppers_intention	327.6k ±19.1k	323.5k ±11.5k	333.0k ±17.1k	252.7k ±35.8k
physiochemical_protein	473.5k ±57.1k	516.2k ±55.3k	485.4k ±11.1k	333.8k ±38.7k
polish_companies_bankruptcy	164.0k ±90.3k	206.1k ±46.7k	145.1k ±81.0k	115.6k ±65.8k
qsar_biodeg	1.25M ±117.7k	1.14M ±84.8k	1.27M ±91.3k	766.9k ±431.0k
qsar_fish_toxicity	377.7k ±23.2k	304.8k ±65.7k	380.0k ±21.5k	293.2k ±32.7k
qsar_tid_11	215.70M ±5.90M	211.03M ±7.29M	215.76M ±6.64M	91.89M ±4.61M
sdss17	833.1k ±10.8k	827.3k ±14.9k	829.5k ±12.9k	372.8k ±17.1k
seismic_bumps	12.4k ±8.8k	9.9k ±11.8k	13.3k ±10.6k	2.8k ±199.1
splice	3.41M ±121.3k	3.34M ±100.5k	3.41M ±130.3k	1.43M ±191.5k
students_dropout_and_academic_success	920.8k ±17.7k	873.1k ±28.2k	912.2k ±16.2k	606.8k ±14.7k
superconductivity	22.42M ±346.1k	22.30M ±452.4k	22.36M ±282.5k	19.19M ±390.9k
tabarena_bank_marketing	139.1k ±21.1k	142.2k ±17.4k	152.2k ±30.0k	443.8 ±694.3
tabarena_diabetes	205.4k ±14.3k	171.8k ±41.9k	202.2k ±12.0k	185.6k ±16.6k
taiwanese_bankruptcy_prediction	877.2k ±144.0k	775.9k ±112.4k	929.8k ±175.7k	509.6k ±291.5k
website_phishing	611.9k ±34.3k	529.1k ±54.8k	604.0k ±22.1k	550.6k ±13.0k
wine_quality	259.3k ±15.1k	228.0k ±40.7k	258.7k ±15.4k	99.8k ±12.8k

F Feature Reduction with Relation to the Feature Space

In this analysis, we investigate how each method reduces the original feature set within the TabArena benchmark datasets and whether this reduction scales with the dimensionality of the feature space. This analysis covers Evo-mRMR, Evo-UF, FetNAS and Evo-All.

For each of the 51 TabArena datasets, we analyze the complete Pareto front obtained for every method and cross-validation split. The number of selected features k is normalized by the total number of features d, yielding a feature-retention ratio $q_k = k/d$ for each Pareto-optimal solution. We then compute the hypervolume of the three-dimensional Pareto front defined by retention ratio, validation performance and model complexity. Feature reduction is defined as $1 - \bar{q}$, where $\bar{q}$ is the average retention ratio of the Pareto-optimal solutions, each weighted by its hypervolume contribution. The resulting feature-reduction scores are averaged over the cross-validation splits for each dataset and method. Since feature counts span three orders of magnitude across datasets (4-1,776), we group datasets into four equal-frequency (quantile) bins by feature count, computed separately for classification and regression tasks so that each bin contains an approximately equal number of datasets.

For classification (Figure 3, left), FetNAS shows an overall increasing trend, rising from 66% reduction in the lowest bin ($[4, 11]$ features) to 82% in the highest ($[64, 1776]$), where it achieves the highest reduction rate among all compared methods. Evo-UF maintains a consistently high reduction rate (65–78%) across all bins, while Evo-mRMR remains roughly flat at 51–56%, independent of dimensionality. The Evo-All baseline stays at 0% throughout, as expected for a method that performs no feature selection. For regression (Figure 3, right), trends are considerably less consistent, which we attribute to the smaller number of regression datasets (13 vs. 38 for classification) yielding noisier per-bin estimates.

To quantify the relationship beyond bin-level aggregates, we compute the Pearson correlation between $\log(d)$ and mean feature reduction at the dataset level, using $n = 38$ (classification) and $n = 13$ (regression) datasets ($n = 9$ for Evo-UF in regression, see Table 8). The correlation is undefined for Evo-All, whose reduction is constant at zero. For classification, FetNAS shows a significant positive correlation ($r = 0.60$, $p < 0.001$) and Evo-UF shows a weaker but significant effect ($r = 0.34$, $p = 0.036$). Evo-mRMR shows no significant correlation ($r = 0.08$, $p = 0.65$). This is consistent with it retaining a proportion of features set by the sweep rather than adapted to the dimensionality of the dataset, unlike FetNAS, which scales its feature selection with the size of the feature space. For regression, none of the methods shows a significant relationship. Given the limited sample size for regression, we do not draw any conclusions about a regression trend.

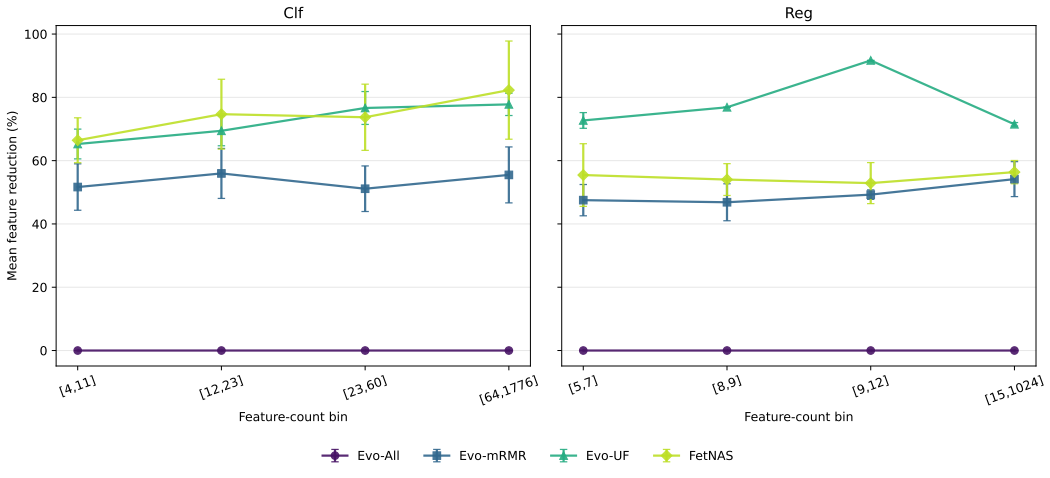


Figure 3: **Mean feature reduction ($1 - \bar{q}$, hypervolume-weighted) per feature-count quantile bin, for classification (left) and regression (right).** FetNAS shows an increasing reduction rate with feature-space dimensionality, while Evo-mRMR remains flat for classification datasets.

Table 8: **Pearson correlation r between $\log(d)$ and mean feature reduction, per method and task.** EVO-UF (regression) excludes *airfoil_self_noise*, *food_delivery_time*, *physiochemical_protein* and *qsar_tid_11*, for which UF retained no features ($n = 9$).

Task	**Method**	n	r	p
Clf	FetNAS	38	0.604	< 0.001
Clf	Evo-UF	38	0.342	0.036
Clf	Evo-mRMR	38	0.076	0.651
Clf	Evo-All	38	-	-
Reg	Evo-mRMR	13	0.467	0.108
Reg	FetNAS	13	0.190	0.535
Reg	Evo-UF	9	-0.348	0.359
Reg	Evo-All	13	-	-

G DISCLOSURE ON THE USE OF GENERATIVE ARTIFICIAL INTELLIGENCE

We used a large language model (LLM) as a writing aid and as a coding assistant. Ideation, scientific content and experimental design were conducted entirely by the authors without any contribution from an LLM.