Measuring the Trade-off Between Model Complexity and Ecological

Ms. Abitha S
Assistant Professor, Department of computer Application
Thiruthangal Nadar College, Chennai

Mr. Krishna Karthik
Assistant Professor, Department of computer Application
Thiruthangal Nadar College, Chennai

CJ. Preethi
Assistant Professor, Department of computer Application
Thiruthangal Nadar College, Chennai

OPEN ACCESS

Volume: 13

Special Issue: 3

Month: February

Year: 2026

P-ISSN: 2321-788X

E-ISSN: 2582-0397

Citation:
S, Abitha, et al. "Measuring the Trade-off Between Model Complexity and Ecological." *Shanlax International Journal of Arts, Science and Humanities,* vol. 13, no. 3, 2026, pp. 414–18.

DOI:
https://doiorg/10.34293/sijash.v13iS3-i2-Feb.10313

Abstract

The current obsession with massive neural architectures has led to a crisis of resource waste in the tech sector. Our study investigates a more sustainable path by testing whether high-performance results truly require high energy hardware. We put "slim" algorithms—specifically K-Nearest Neighbors and Naive Bayes—up against heavy-duty Deep Learning systems to see if we can achieve similar outcomes with a fraction of the electricity. Our data indicates that for many standard classification problems, the industry's reliance on power-heavy GPUs is misplaced. By switching to CPU-centered statistical models, we found that it is possible to slash electricity requirements by a factor of ten, while only experiencing a tiny change in output quality. We argue for a fundamental change in how the industry judges "success." By introducing Resource-Weighted Accuracy (RWA) as a primary metric, we can move toward a future where a model's efficiency is considered just as important as its precision.

Keywords: Lean Computing, Ecological ML, Algorithmic Parsimony, Resource-Weighted Accuracy (RWA), and Carbon-Efficient Classification

Introduction

The current era of Machine Learning is marked by a "growth-centric" bias, whereby a record-breaking performance is equivalent to the complexity of the architecture and the usage of high-wattage GPU hardware. Although the revolutionary breakthroughs made by the application of deep learning in the domain of computer vision are undeniable, the copious environmental and economic costs it has exacted are now becoming increasingly hard to justify.

Our research advocates a new paradigm based on the tenets of Computational Minimalism. It disputes the need for complex architectures to attain efficient classification. Instead, we reveal the efficacy of "lean" methods like KNN and Naïve Bayes, which are extremely efficacious yet underappreciated methods that may be applied using normal hardware with a negligible energy impact.

Quantifying Efficiency: An empirical comparison between "parsimonious" statistical methods and deep learning baselines, with a focus on the relationship between classification accuracy and energy costs.

The RWA Framework: An explanation of the development process behind the Resource-Weighted Accuracy metric, whereby energy costs are a non-secondary consideration in the assessment of performance.

Sustainable AI Deployment: Evidence-based reasoning to support the assertion that accurate classification is feasible using normal hardware.

Literature Review

The major achievement in this domain of deep learning has resulted in a significant impact on the field of predictive analytics, ensuring high accuracy levels in a number of application domains. However, the computational costs associated with the application of these large-scale neural networks have been increasingly identified as a major drawback in this domain by a number of researchers. This has resulted in the birth of a new branch in the domain of artificial intelligence, referred to as Green AI, focusing on the importance of using efficiency as a measure to evaluate the performance of artificial intelligence-based systems.

From the existing body of literature, it is identified that a number of scenarios have resulted in major success in the context of applying deep learning-based systems in large-scale and unstructured data scenarios. However, in the context of applying these systems in structured and moderately scaled scenarios, it is identified that the difference in the performances between machine learning and deep learning-based systems is less prominent. The Naïve Bayes algorithm seems to be a promising algorithm in this context, as it is identified to be trained at a faster rate. Additionally, satisfactory results have been achieved in a number of scenarios by applying the K Nearest Neighbour algorithm.

With the rise in sustainability issues, a number of recent studies have identified the need to consider a number of efficiency-based factors in this context. Though these studies are identified to be promising, a detailed evaluation comparing the performances of lightweight statistical-based classification scenarios with deep learning-based scenarios from a sustainability perspective has yet to be conducted. Additionally, a composite metric to facilitate this process has yet to be identified.

In order to add a new dimension to the existing body of literature in this domain of sustainability-based evaluation, a detailed evaluation comparing the performances of classical statistical-based classification scenarios with a deep learning-based scenario using an efficiency-based evaluation process is conducted in this study. Additionally, a novel metric referred to as Resource-Weighted Accuracy is proposed in this study.

Methodology

Dataset

Experiments were conducted on standard classification datasets obtained from publicly available repositories. The datasets represent structured classification problems of moderate size, suitable for evaluating both classical and neural approaches. Each dataset was divided into training and testing subsets using an 80:20 split.

Models Evaluated

The study compares three models:

- Lightweight classifiers
- Naïve Bayes
- K-Nearest Neighbour (KNN)
- Deep learning baseline
- Multi-Layer Perceptron (MLP)

Default hyperparameters were initially used, followed by minimal tuning to ensure fair comparison.

Evaluation Metrics

Model performance was evaluated using:

- Accuracy
- Precision
- Recall
- F1-score

Computational efficiency was measured using:

- Training time
- Prediction time
- Memory usage (where applicable)

Resource-Weighted Accuracy (RWA)

To jointly evaluate performance and efficiency, this study proposes Resource-Weighted Accuracy (RWA): RWA= Accuracy/ Normalized computational cost

where computational cost is derived from training time and resource dependency. Higher RWA values indicate better efficiency–performance balance.

Experimental Setup

Feature Engineering & Dimensionality Alignment

In order to establish a fair baseline for all learners to be compared against, a cleaning and transformation pipeline was established on the raw data. To deal with missing data, a statistical imputation mechanism was used to ensure feature space integrity. To negate the effects of "magnitude bias" commonly found in distance-based learning algorithms such as KNN, Linear Range Normalization between 0 and 1 was used on all continuous features. To deal with discrete features, a Discrete Integer Mapping mechanism was used. To ensure the validity and accuracy of this experiment, a Hold-out Validation Strategy was used to randomly set aside a portion of 20% of the data to be used in a final audit.

Algorithmic Topology & Hyper-parameter Control

In order to ensure the efficacy of the base architectures, minimal hyper-parameter tuning was used to ensure maximum inherent efficiency in the base architectures.

Instance-Based Learning: To optimize the KNN learner implementation, a heuristic search algorithm to optimize the coefficient of proximity k to maximize local density sensitivity.

Probabilistic Modeling: To optimize the Naive Bayes learner implementation, a Gaussian Kernel is used to assume a normal probability density function on continuous features.

Connectionist Reference: In the context of the proposed MLP learner, a shallow MLP was used for the baseline implementation. By capping the number of hidden layers, we proposed a 'frugal' version of deep learning, which reduces the high energy costs associated with deeper and over-parameterized networks, yet still captures non linear patterns.

Instrumentation of Resource Consumption

A significant innovation in the proposed workflow is the evaluation of the Hardware Latency and Power Proxies. In comparison to other studies, which only consider error rates, we incorporated a high-resolution timestamping mechanism to obtain the following values:Stochastic Training Durations - the time to converge weights or fit the model.Inference Throughput - the millisecond expense per prediction cycle.While the statistical models are constrained to a standard x86-based CPU, the MLP was also evaluated against a variety of both CPU and GPU backends. This allowed us to compute the 'Computational Premium' - the exact energy expense to transition from a lightweight statistical approach to a tensor-based neural approach. This was then combined to compute the Resource-Weighted Accuracy (RWA) score to determine the overall ecological benefit of the respective models.

Deployment Considerations

Deployment Environment

The viability of the models in a real-world environment has been considered by viewing the models in the context of a resource-constrained computing environment. For example, models like the classifier for Naïve Bayes and K-NN are efficient in a normal CPU environment and do not demand special accelerators. This makes the models viable for environments like schools and small businesses where the necessary high-end GPU accelerators are not present.

On the other hand, models like the neural networks demand a significant speed and scalability boost to operate efficiently in a GPU environment. However, for a low throughput or mid-scaled classifier model, the necessary hardware costs might not be justified if other models are performing at the same level of efficiency.

Implementation Workflow

The workflow for the proposed method, if implemented, would involve the following steps:

Data Acquisition and Preprocessing: In this step, the data received would be cleaned and normalized according to the model's training.

Model Selection Using RWA: Using the Resource Weighted Accuracy method, the models would be scored, and the most efficient model would be selected.

Model Serialization: After the most efficient model has been selected, it would be serialized in line with the standard methods.

Inference Service: After serialization, the model would be used in its capacity in a production environment.

Performance Monitoring: Periodic checks would be made to ensure the quality and efficiency of the model's predictions and computation are at desired levels.

Sustainability Implications

Looking at the sustainability implications, the findings of the study indicate that many of the classification activities could be carried out in a computationally efficient manner, which would ensure the system is not hardware-dependent, making it a more efficient system, especially in environments where the use of GPUs would be unfavorable.

The use of the Resource Weighted Accuracy method would ensure that organizations and companies are able to make the most out of the use of AI models, making it a more sustainable approach to the use of AI.

Results and Discussion

The results show that lightweight classifiers have what it takes to compete in structured classification tasks. Naive Bayes was impressive for its training speed and moderate changes in accuracy. KNN performed well but at the cost of reduced speed for each prediction.

The accuracy was better for the MLP classifier but at the cost of training time and resources. When the RWA metric was used for evaluation, the lightweight classifiers performed better than the deep learning classifier. The results show that when choosing a classifier, accuracy and efficiency should be considered.

Conclusion

The paper aims to investigate the relationship between model complexity and ecological cost by comparing classical statistical machine learning with deep learning. The results show that lightweight machine learning competes well in structured classification tasks.

The RWA metric was found to be useful in evaluating machine learning efficiency. The paper also highlights the importance of frugal machine learning in sustainable ICT.

Future directions for this paper involve evaluating the cost of training for each model using the proposed method and conducting large-scale evaluations.

References

1. D. Patterson, J. Gonzalez, Q. Le, et al., "Carbon emissions and large neural network training," arXiv preprint arXiv:2104.10350, 2021.
2. R. Schwartz, J. Dodge, N. A. Smith, and O. Etzioni, "Green AI," Communications of the ACM, vol. 63, no. 12, pp. 54–63, 2020.
3. J. Brownlee, Machine Learning Mastery with Python. Melbourne, Australia: Machine Learning Mastery, 2016.
4. T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning. New York, NY, USA: Springer, 2009.
5. C. M. Bishop, Pattern Recognition and Machine Learning. New York, NY, USA: Springer, 2006. [6] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.
6. P. Domingos and M. Pazzani, "On the optimality of the simple Bayesian classifier under zero-one loss," Machine Learning, vol. 29, no. 2–3, pp. 103–130, 1997.
7. T. Cover and P. Hart, "Nearest neighbor pattern classification," IEEE Transactions on Information Theory, vol 13.