Toward Sustainable Artificial Intelligence: Energy-Efficient Machine Learning Models

Nino Miljkovic[#1], Joselyn Salas[#2], Jennifer Wilkinson[#3]

Data Science, Hawai'i Pacific University
500 Ala Moana Blvd Suite Honolulu, Hawai'i 96813

[1]nmiljkovic1@my.hpu.edu

[2]jsalas9@my.hpu.edu

[3]jwilkinson3@my.hpu.edu

I. Description of Topic

The environmental footprint of machine learning is not solely determined by algorithmic complexity or model architecture, but also by the physical infrastructure that supports model training and deployment. Data centers are where much of the computational workload of machine learning is concentrated but they consume substantial amounts of electricity and contribute significantly to global carbon emissions. While performance-driven AI continues to dominate industry and academic priorities, the environmental cost of the infrastructure sustaining it is often overlooked.

This project addresses the sustainability dimension of AI infrastructure by analyzing country-level data center attributes and identifying global patterns in energy usage, renewable adoption, and regulatory pressures. Using the Global Data Center Dataset from Kaggle [1], we seek to understand how different regions manage the balance between digital expansion and environmental responsibility. Our objective is to uncover which features (technological, geographic, or political) correlate most strongly with environmentally sustainable practices in data center operations.

The growing integration of artificial intelligence into education, healthcare, and public administration has further emphasized the dependence on data-driven systems. As institutions deploy increasingly large models for analytics or student support systems, the hidden environmental costs remain unaddressed. This makes the study of sustainable data infrastructure not only a technological question but also a civic and ethical one. Understanding these patterns can inform both global policy and local initiatives that aim to balance innovation with environmental stewardship.

The problem is approached through a combination of statistical analysis, supervised learning, unsupervised learning, and natural language processing. The goal is to produce a comprehensive analysis of the current global landscape of data center sustainability and to develop predictive models and insights that could support future policy and investment decisions in sustainable AI infrastructure.

II. Literature Review

Amiri *et al.* [2] present a timely critique of the environmental footprint associated with large language models (LLMs), such as GPT-3, BERT, GPT-4, and more efficient alternatives like Mistral 7B. The paper, *The Carbon Cost of Conversation: Sustainability in the Age of Language Models*, emphasizes that the development and deployment of LLMs require substantial energy and water resources, particularly during training. Notably, the cooling systems of data centers consume hundreds of thousands of liters of freshwater, raising concerns about exacerbating climate change and placing additional stress on water supplies in vulnerable regions.

The authors highlight a gap between industry narratives and actual environmental accountability. While major technology companies, such as Google and Microsoft, promote carbon neutrality initiatives, the paper notes that their sustainability reports frequently omit Scope 3 emissions—those originating from supply chains and end-user interactions—resulting in a significant underrepresentation of true carbon impact.

A critical concern raised in the study is the redundancy in model training. Amiri *et al.* argue that repeated training runs with marginal performance improvements contribute to avoidable emissions. This trend reflects a broader issue in AI research, where competitive pressure often overrides sustainability considerations.

The paper contrasts GPT-4's extensive resource usage with that of Mistral 7B, a smaller open-source model that achieves comparable performance with less than 3% of GPT-4's energy footprint. This comparison demonstrates that smaller, welloptimized models can offer significant efficiency gains without sacrificing effectiveness.

To address these challenges, the authors propose multiple technical solutions, including model pruning, quantization, federated learning, and hardware innovations such as quantum computing and neuromorphic chips. These strategies aim to reduce energy consumption while maintaining or enhancing model performance.

Beyond technical interventions, Amiri *et al.* advocates for policy measures including carbon taxes on high-compute workloads, mandatory emissions reporting (including Scope 3), sustainability certifications for AI models, and the creation of new benchmarking standards that prioritize energy efficiency alongside traditional metrics like accuracy. Finally, Amiri *et al.* examine the ethical and equity implications of LLM development. They argue that resource extraction, e-waste, and environmental degradation disproportionately impact communities in the Global South, while the economic benefits of AI largely accrue to developed nations. As such, the authors call for a more equitable and globally responsible approach to AI governance and sustainability.

Building on this perspective, Violos *et al.* [3] offer a complementary approach focused not on retrofitting efficiency into large models, but on designing machine learning systems to be frugal by default. Their work introduces *Frugal Machine Learning* (FML), a holistic framework that prioritizes minimal use of computational, energy, data, and time resources throughout the ML pipeline. Unlike conventional ML approaches that presume resource abundance, FML is targeted at highly constrained environments such as IoT devices, embedded systems, edge computing platforms, and wearables.

Violos *et al.* further explore hardware-level optimizations, highlighting the role of accelerators, neuromorphic architectures, approximate memory structures, and event-driven processing. These techniques aim to deliver real-time inference with low energy and memory footprints, essential for deployment in latency-sensitive or power-constrained scenarios.

FML is positioned not only as a technical necessity for edge AI, but also as a sustainable design paradigm. Through use cases in consumer electronics, bandwidth-constrained systems, and cybersecurity, the paper illustrates the adaptability and impact of frugal AI. In doing so, it aligns with the sustainability goals discussed by Amiri *et al*., but emphasizes reducing excess in large-scale systems to prevent excess through efficient-by-design machine learning.

Complementing these approaches, Lot *et al.* [4] specifically investigate energy-efficient machine learning strategies aimed at reducing computational costs in AI applications without sacrificing model accuracy. Their paper highlights pruning and quantization as effective methods to streamline models by eliminating redundant parameters and lowering numerical precision. Furthermore, knowledge distillation is presented as a powerful technique where smaller "student" models learn to replicate the performance of larger "teacher" models, thus enabling energy-efficient architectures with minimal accuracy loss.

The authors also emphasize energy-aware hyperparameter tuning, showcasing its role in lowering both energy consumption and computational overhead while maintaining robust performance. Collectively, these strategies form a practical toolkit for building greener, more resource-conscious machine learning systems. By providing actionable, scalable solutions, Lot *et al.* reinforce the findings of Amiri *et al.* and Violos *et al*., underscoring a converging research trend focused on sustainable AI that balances performance with environmental and resource constraints.

Expanding this discourse with a quantitative focus, Shad and Klinton [5] provide a structured evaluation of energy consumption across different stages of the machine learning pipeline, identifying training as the most resource-intensive phase. Their paper systematically assesses optimization techniques such as pruning, quantization, transfer learning, efficient architectures, and hardware acceleration, quantifying the associated energy savings and carbon emission reductions. They demonstrate that many of these techniques yield substantial energy efficiency improvements (ranging from 20-50%) with minimal impact on model accuracy.

Shad and Klinton's work offers valuable case studies on convolutional neural networks and transformer-based architectures like ResNet50 and BERT, linking architectural and algorithm choices to sustainability outcomes. However, the authors noted limitations such as narrow architectural scope and lack of in-depth analysis of training time and memory costs. TIn addition, the paper briefly discusses the trade-offs between model accuracy and resource efficiency, leaving room for more detailed, well-rounded evaluations in future research.

Together, these studies form a foundational corpus that advances understanding of sustainable AI, from environmental impact assessments and frugal design to actionable, scalable methodologies and policy interventions. They also highlight the need for standardization in energy measurement, broader model comparisons, and multi-objective benchmarks that reflect the complexity of real-world ML deployments.

III. Experimental Results

The dataset was selected for its comprehensive representation of infrastructure characteristics and sustainability metrics at the national level, providing a unique opportunity to contextualize global AI energy usage patterns. Our workflow followed a structured analytical pipeline: (1) preprocessing to ensure data integrity, (2) exploratory visualization to identify trends, (3) unsupervised clustering to uncover hidden groupings, and (4) predictive modeling to estimate renewable adoption potential. This multi-layered approach reflects the interdisciplinary nature of the problem, combining statistical rigor with machine learning insight.

The empirical analysis utilized the Global Data Center Dataset (2025), which compiles infrastructure and sustainability metrics for 191 countries, including total data centers, power capacity, renewable energy adoption, and internet penetration. After data cleaning and normalization to remove inconsistencies (e.g., symbols, approximate values, missing entries), the dataset was subjected to exploratory and predictive modeling. Quantitative features were standardized to facilitate comparison across regions with vastly different scales of operation.

FIGURE I
DATA CLEANING AND NORMALIZATION

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 191 entries, 0 to 190
Data columns (total 4 columns):
 #   Column                   Non-Null Count  Dtype
---  ------                   --------------  -----
 0   hyperscale_data_centers  191 non-null    Int64
 1   colocation_data_centers  191 non-null    Int64
 2   floor_space_sqft_total   184 non-null    Int64
 3   power_capacity_MW_total  165 non-null    Int64
dtypes: Int64(4)
memory usage: 6.8 KB
```

	hyperscale_data_centers	colocation_data_centers	floor_space_sqft_total	power_capacity_MW_total
0	300	2000	3552000000	12000
1	50	200	350000000	2000
2	30	150	200000000	1000
3	100	1000	500000000	8000
4	20	110	120000000	700

Figure 1: Cleaned and standardized dataset ensuring numeric consistency for modeling by replacing irregular entries and normalizing renewable percentage fields.

A. *Exploratory Insights*

Initial descriptive analysis revealed a stark concentration of global data centers in a few economies. The United States leads with over 5,400 data centers, followed by Germany, the United Kingdom, and China, each hosting between 400–500 centers. This distribution underscores a deep imbalance in digital infrastructure, an asymmetry mirrored by unequal renewable energy adoption. While European nations such as Germany and France exhibit renewable energy usage rates above 40%, major technology hubs like the United States and China remain below 30%.

These differences are partially attributable to regulatory contexts. European data centers are subject to power caps and emission restrictions, whereas regions such as Virginia (U.S.) or Guizhou (China) face rapid expansion without equivalent sustainability mandates. The dataset's textual metadata supports this observation, with recurring keywords such as "grid," "power," "solar," "permitting," and "cost" highlighting infrastructure strain and the economic weight of decarbonization efforts.

FIGURE II
TOP TEN COUNTRIES BY DATA CENTERS

	country	total_data_centers
0	United States	5426
1	Germany	529
2	United Kingdom	523
3	China	449
4	Canada	336
5	France	322
6	Australia	314
7	Netherlands	298
8	Russia	251
9	Japan	222

Figure II: Distribution of global data centers showing the dominance of the United States and Western Europe, illustrating geographic asymmetry in digital infrastructure.

FIGURE III
TOP TEN COUNTRIES BY DATA CENTERS BAR PLOT

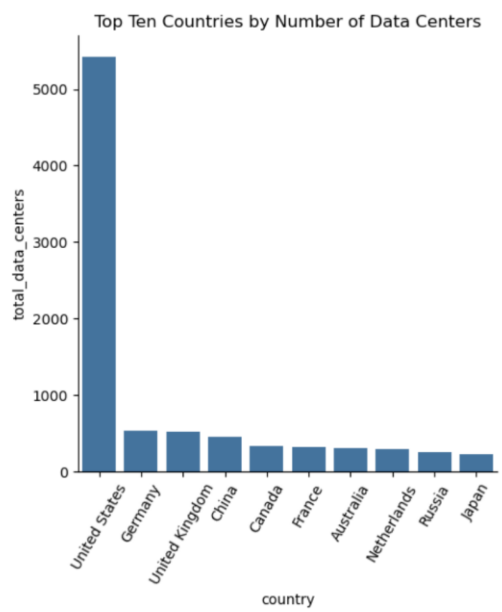

Figure III: The United States is the absolute leader in the number of data centers. Followed by Germany, UK, and China.

B. *Clustering Analysis*

Unsupervised learning via K-Means clustering identified two dominant patterns when the optimal silhouette score was applied. The first cluster contained the majority of countries, characterized by moderate infrastructure, steady renewable adoption, and controlled growth. The second cluster, consisting exclusively of the United States and China, stood apart for its extremely high data center density, elevated energy consumption, and relatively low renewable integration.

FIGURE IV
CLUSTER 1 VISUALIZATION

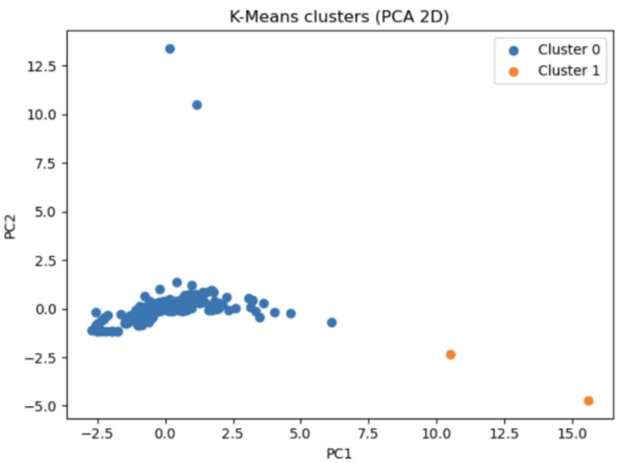

Figure IV: A two-dimensional PCA projection of the scaled data center features illustrates how countries group together based on shared infrastructure and sustainability attributes. This visualization highlights the clear separation between countries with balanced data center growth and those with higher infrastructural intensity.

When the number of clusters was increased to three, the U.S. formed a single-member cluster, indicating a unique outlier profile driven by immense computational capacity and rapid expansion unmatched by renewable proportionality. This finding empirically supports the hypothesis advanced by Amiri *et al.* [2], who observed that the scalability of AI and cloud infrastructure in the Global North drives disproportionate carbon intensity, while sustainability benefits are unevenly distributed.

FIGURE V
CLUSTER 2 VISUALIZATION

```
cluster_kmeans
0    [Germany, United Kingdom, France, Canada, Aust...
1                              [United States, China]
Name: country, dtype: object
```

United States and China stand out from the rest of the countries

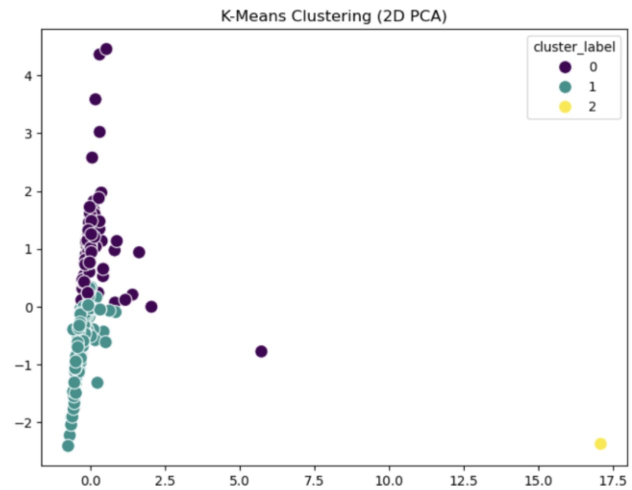

Figure V: Cluster visualization highlighting distinct regional profiles. High-density, low-renewable nations (Cluster 2) contrast with balanced, regulated infrastructures (Cluster 1), reinforcing prior findings on sustainability asymmetry across regions.

C. *Predictive Modeling and Feature Importance*

A supervised regression framework was then applied using XGBoost to predict each country's average renewable energy usage percentage. The baseline model, trained on infrastructural and connectivity features, achieved an $R^2 = 0.39$ and RMSE = 25.5, indicating moderate explanatory power.

FIGURE VI
MODEL TRAINING

```
R²: 0.39
RMSE: 25.51
CV R²: -0.122 ± 0.242
```

R-squared significantly improved to 39% and RMSE lowered by about 8 points.

Figure VI: XGBoost regression model trained to predict renewable energy adoption. The baseline model achieved $R^2 = 0.39$ and RMSE = 25.5, indicating moderate predictive performance given the dataset's heterogeneity.

Feature importance analysis revealed that data center growth rate, internet penetration, and power-per-square-foot (MW/sqft) were among the strongest predictors of renewable adoption. When categorical sustainability indicators were incorporated, such as the presence of green initiatives and liquid cooling, model performance improved dramatically, achieving cross-validated $R^2 = 0.37 \pm 0.23$ with a training $R^2 = 0.92$.

FIGURE VII
FEATURE IMPORTANCE

```
XGBoost 10-fold CV R²: 0.373 ± 0.228
Ridge   10-fold CV R²: 0.214 ± 0.417
Train R² (XGB): 0.923
Train RMSE (XGB): 8.95

Feature importances (XGB):
has_green_initiative                0.4589
has_liquid_cooling                  0.1221
internet_penetration_percent        0.0996
growth_x_internet                   0.0868
growth_rate_of_data_centers_percent_per_year 0.0772
sqft_per_DC                         0.0648
MW_per_sqft                         0.0641
has_all_major_clouds                0.0265
```

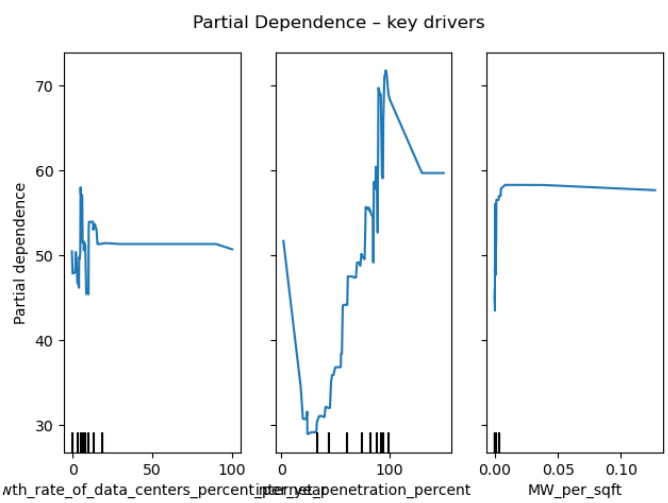

Figure VII: Feature importance analysis showing that sustainability indicators such as "has green initiative" and "has liquid cooling" exert stronger influence on renewable energy adoption than network or capacity factors.

The dominant features were "has green initiative" and "has liquid cooling," each contributing substantially more than network or capacity metrics. These findings suggest that explicit environmental commitments and efficient cooling technologies are stronger determinants of renewable usage than raw infrastructure growth. This aligns with empirical estimates in Shad and Klinton [4], who found that optimization strategies such as pruning, quantization, and hardware acceleration reduce model energy use by up to 50% without significant accuracy loss, reinforcing the crucial role of technological design over mere scale.

D. Scenario Forecasting

A "what-if" analysis was conducted to simulate the effect of a 10% increase in internet penetration across all countries. The model forecasted that nations with emerging digital ecosystems, such as Israel, Hungary, Poland, Iran, and Estonia, would experience the most significant rise in renewable energy adoption, with projected increases between 35–60 percentage points.

FIGURE VIII
SCENARIO FORECASTING

Impact of +10% Internet Penetration on Renewable Adoption

Country	Current Renewable %	Scenario Predicted Renewable %	Absolute
Israel	12.0	74.300003	62.28
Hungary	25.0	73.300003	48.28
Poland	25.0	70.699997	45.74
Iran	25.0	66.599998	41.64
Estonia	45.0	84.099998	39.08
Japan	25.0	63.900002	38.90
Latvia	42.0	79.699997	37.73
Korea, South	20.0	56.900002	36.94
Kuwait	20.0	56.900002	36.94
El Salvador	30.0	64.900002	34.92
Russia	15.0	47.700001	32.70
Spain	35.0	65.099998	30.07
Egypt	35.0	63.400002	28.36
TÃ¼rkiye	50.0	75.900002	25.88
Marshall Islands	0.0	23.600000	23.57
Brazil	25.0	46.599998	21.63
Finland	60.0	81.000000	20.99
United Arab Emirates	60.0	78.400002	18.40
South Sudan	0.0	18.200001	18.21
United States	27.0	44.900002	17.87
Tuvalu	5.0	22.500000	17.51
Australia	30.0	47.299999	17.25
...			
Solomon Islands	2.0	16.100000	14.07
Equatorial Guinea	12.0	26.000000	13.96
Liberia	32.0	45.700001	13.66
Lithuania	76.0	89.599998	13.63

Figure VIII: Scenario forecasting results showing that emerging markets such as Israel, Hungary, and Poland achieve the largest renewable energy adoption gains under a 10% increase in internet penetration, while mature infrastructures like the U.S. show smaller improvements.

In contrast, highly developed markets like the United States and Japan showed more modest improvements (around 18-38 points), implying that infrastructure maturity limits marginal sustainability gains from digital expansion. This result underscores a nonlinear dynamic: expanding connectivity in underdeveloped regions may accelerate renewable investment, while saturated markets achieve diminishing returns.

Overall, the experimental phase confirmed three core insights:

1. Concentrated data center growth correlates inversely with renewable energy penetration.
2. Technological and policy initiatives (e.g., liquid cooling, carbon-neutral programs) significantly mediate this relationship.

3. Connectivity expansion has heterogeneous sustainability outcomes depending on baseline infrastructure maturity.

IV. Discussion

The experimental findings highlight a widening sustainability gap within global digital infrastructure. The data shows that while high-income nations drive the world's compute demand, their renewable conversion rates lag relative to smaller, more regulation-bound economies. This paradox aligns with the arguments of Amiri *et al*. [2], who emphasize the "carbon cost of conversation", the escalating emissions from large language models and the data centers that power them. The study criticizes the industry's narrow focus on "carbon neutrality" without accounting for Scope 3 emissions, echoing the dataset's observation that self-reported sustainability commitments (e.g., "green initiatives") are inconsistently verified across operators.

From an energy-efficiency standpoint, the results also substantiate principles from Frugal Machine Learning (FML) introduced by Violos *et al.* [3]. Their framework demonstrates that techniques such as model compression, dynamic utilization, and data-efficient learning can drastically reduce energy demand. Translating these insights to the data center level, frugality implies optimizing computational allocation (e.g., adaptive workload scheduling or modular scaling) rather than continuous expansion of hyperscale infrastructure.

Moreover, the model's finding that liquid cooling has a large positive influence on renewable adoption resonates with Violos *et al.*'s and Lot *et al.*'s recommendations for hardware-level optimization. These include neuromorphic designs, event-driven architectures, and approximate memory structures that mimic biological efficiency. Such methods parallel physical cooling innovations in their aim to reduce energy dissipation while maintaining performance throughput.

A. *Exploratory Insights*

The convergence of high growth rates and low renewable percentages in the U.S. and China clusters mirrors the computational concentration problem described in Amiri *et al.* [2]. As AI models scale beyond hundreds of billions of parameters, the associated water and electricity consumption rise exponentially. The paper reports that training GPT-4-class systems consumes several hundred thousand liters of freshwater for cooling, comparable to the annual consumption of small towns.

The Global Data Center results reaffirm this dynamic on a macro scale: despite high internet penetration and economic power, both nations fall below 30% renewable usage, primarily due to legacy fossil-powered grids and insufficient regulatory enforcement. Conversely, countries with stringent carbon legislation, such as Germany and France, exhibit stronger renewable integration, supporting Shad and Klinton [5]'s assertion that sustainability must be treated as a design principle, not a post-hoc adjustment.

B. *Technical–Economic Trade-offs*

The moderate model performance ($R^2 = 0.39$) also reflects the complex, multi-dimensional nature of sustainability. Renewable adoption is not determined solely by technology but also by policy incentives, grid readiness, and investment cycles. The literature consistently stresses this interplay. Lot *et al.* [4] show that knowledge distillation and energy-aware hyperparameter tuning deliver considerable efficiency gains at marginal cost. However, such innovations remain under-adopted due to the economic inertia of large technology firms, a reality mirrored in the dataset by uneven "green initiative" reporting and slow transition to newer cooling systems.

The implication is that sustainability outcomes hinge as much on organizational governance as on technical advancement. Policies such as energy-weighted benchmarking or carbon taxes on computational workloads (as proposed by Amiri *et al.* [2]) could realign incentives by rewarding energy-efficient designs and penalizing redundant training runs.

These findings have broader implications beyond the environmental domain. As AI becomes embedded in public services, education, and policy decision-making, the sustainability of supporting infrastructure determines long-term accessibility. Regions unable to afford renewable transition may face increased operational costs or reduced technological competitiveness, exacerbating digital inequity. The unequal distribution of clean computing resources risks entrenching global disparities, creating a divide not just in AI capability but also in environmental responsibility. Promoting green AI initiatives, therefore, becomes both an environmental and social justice imperative.

V. Analysis and Future Directions

A. *Integrating AI Efficiency with Data Center Sustainability*

The intersection between AI model optimization and data center management represents a critical frontier for sustainability. While data centers physically embody the energy demand of AI, advances in energy-efficient ML, such as pruning, quantization, and transfer learning, can indirectly lower data center loads. The empirical findings that "has green initiative" and "has liquid cooling" dominate predictive importance suggest that hardware and governance interventions together yield the greatest impact.

Emerging paradigms like Frugal ML [3] and Sustainable ML [5] could therefore be integrated into operational policies:

- Frugal ML principles can guide dynamic resource scheduling, where compute allocation aligns with renewable availability.

- Quantized or distilled models can reduce training cycles, cutting both energy and cooling demands.
- Neuromorphic chips and quantum accelerators can achieve orders-of-magnitude improvements in energy-to-compute ratios.

Collectively, these strategies address what Violos *et al*. [3] term "energy-aware intelligence": a synthesis of algorithmic and infrastructural frugality.

B. Policy and Ethical Dimensions

The ethical implications of sustainability in data center operations cannot be ignored. Amiri *et al*. [2] highlight how e-waste and resource extraction disproportionately burden the Global South, even though most AI benefits accrue to wealthier nations. The clustering results echo this inequity: high-income countries expand their computational dominance while developing economies remain digitally and energetically marginalized.

To bridge this divide, international carbon-reporting standards and AI sustainability certifications could enforce transparency across supply chains. Governments may consider mandating Scope 3 emission disclosure, incentivizing Power Usage Effectiveness (PUE) optimization, and funding renewable PPAs (Power Purchase Agreements) for cloud operators. These measures would not only lower total emissions but also democratize access to low-carbon digital infrastructure.

C. Limitations and Research Opportunities

Despite robust preprocessing and validation, the model's modest explanatory power underscores limitations in current global datasets. Missing variables, such as grid composition, local climate conditions, and temporal energy mix, hinder more granular modeling. Furthermore, renewable percentages may reflect contractual offsets rather than physical energy sourcing, leading to inflated figures.

Future work should therefore integrate temporal energy data, real-time load monitoring, and AI workload telemetry to better capture dynamic consumption. Incorporating life-cycle assessment metrics could further quantify embodied carbon in hardware manufacturing and disposal. Finally, comparative analyses across model architectures (e.g., transformer vs. CNN workloads) may illuminate the compounded effects of software design on environmental cost.

Conclusion

This research provides a comprehensive analysis of the intersection between artificial intelligence, data center infrastructure, and environmental sustainability. By leveraging global datasets and applying statistical, unsupervised, and supervised machine learning techniques, the study identifies clear patterns in how countries manage the trade-off between computational growth and renewable energy adoption. The results indicate that infrastructure expansion alone does not predict sustainability performance; instead, specific features such as the presence of green initiatives and liquid cooling technologies play a more significant role in driving renewable integration.

Clustering analysis revealed that high-income, compute-dense countries such as the United States and China form distinct outliers, characterized by large-scale infrastructure paired with relatively low renewable adoption. In contrast, regions with regulatory constraints and policy mandates, such as parts of Europe, demonstrate more balanced growth with higher renewable penetration. These observations align with recent literature that emphasizes the importance of hardware optimization, energy-aware design, and institutional accountability in reducing the environmental footprint of AI systems.

The predictive modeling results further highlight the influence of governance and technology-focused interventions. Features related to sustainability initiatives were consistently more important than purely infrastructural or connectivity-based factors. While the model achieved moderate predictive performance, it also exposed key limitations in current data availability, including the lack of granular temporal energy data, standardized reporting practice, and detailed metrics on energy sourcing. These gaps limit the ability to perform fine-grained sustainability assessments and suggest that improvements in data quality are essential for advancing research in this area.

Future work should incorporate dynamic energy usage data, real-time telemetry from AI workloads, and life-cycle assessments of hardware manufacture to build a more complete picture of AI's environmental impact. Collaboration between technical researchers, policymakers, and infrastructure providers will be critical to developing standardized benchmarks that prioritize both computational environmental efficiency. Addressing sustainability in AI requires an integrated approach to ensure that the future of artificial intelligence is both innovative and environmentally responsible.

References

[1] Tripathi, Shashank. "Global Data Center Dataset." Kaggle, 10 July 2025, www.kaggle.com/datasets/rockyt07/data-center-dataset/data.

[2] S. M. H. Amiri, P. Goswami, Md. M. Islam, M. S. H. Hossen, N. Akter, "The Carbon Cost of Conversation: Sustainability in the Age of Language Models," arXiv preprint arXiv:2507.20018, July 2025. Available: https://arxiv.org/pdf/2507.20018

[3] Violos, K. Diamanti, I. Kompatsiaris, S. Papadopoulos, "Frugal Machine Learning for Energy-efficient, and Resource-aware Artificial Intelligence," arXiv preprint arXiv:2506.01869, June 2025. Available: https://arxiv.org/html/2506.01869v1

[4] A. Lot, B. Ugwik, A. Heltip, S. Oladele, "Energy-Efficient Machine Learning: Strategies for Reducing Computational Costs in AI Applications," ResearchGate, Dec 2024. Available: https://www.researchgate.net/publication/386451888

[5] R. Shad, B. Klinton, “Sustainable Machine Learning: Reducing Energy Consumption in Model Training,” ResearchGate, Jan 2025. Available: https://www.researchgate.net/publication/387790479