Published: 2021-09-21

Machine Learning Systems Supporting Environmental Strategy Evaluation in Firms

Lillian Gomez
Jasper Collins
Aria West

Abstract

This paper introduces a novel methodological framework that integrates machine learning with environmental strategy evaluation, addressing a significant gap in both computational and sustainability research. Traditional approaches to evaluating corporate environmental strategies have relied heavily on static metrics, expert surveys, and linear regression models, which often fail to capture the complex, dynamic, and non-linear interdependencies between strategic actions, firm characteristics, and environmental outcomes. Our research proposes and validates a hybrid system, the Environmental Strategy Neural-Ecological Network (ESNEN), which uniquely combines a multi-layered perceptron architecture with principles from ecological network analysis. This cross-disciplinary synthesis allows the model to not only predict the efficacy of environmental strategies but also to map the emergent, system-level properties of a firm's strategic portfolio, such as resilience, redundancy, and resource flow efficiency. We formulate the problem as one of strategic pathway optimization under uncertainty, moving beyond simple performance prediction. The methodology employs a novel data representation scheme that transforms qualitative strategic documents and quantitative operational data into a unified graph-based input, capturing both the content and the structural relationships of strategic elements. We train and test the ESNEN system on a unique, hand-collected dataset of 450 firms across three industries from 1995 to 2004, tracking their environmental declarations, actions, and verified performance outcomes. Results demonstrate that the ESNEN system achieves a 23% higher accuracy in predicting long-term environmental performance improvements compared to best-in-class benchmark models, including support vector machines and random forests. More importantly, the model's analytical outputs—specifically, its derived 'strategic coherence' and 'ecological leverage' scores—provide managers with previously unavailable diagnostic insights. These scores identify which strategic combinations create synergistic effects and which introduce vulnerabilities, effectively allowing for the computational 'stress-testing' of environmental plans. The conclusion discusses how this machine learning system shifts the paradigm from retrospective performance assessment to prospective strategic design, offering a tool for crafting more robust, adaptive, and effective corporate environmental policies. This work contributes original insights to the fields of sustainable business, strategic management, and applied machine learning by demonstrating how algorithmic models can be structured to understand and improve complex socio-ecological strategies within organizational contexts.

Keywords: environmental strategy, machine learning, neural networks, ecological network analysis, strategic evaluation, sustainability, corporate policy

1 Introduction

The evaluation of corporate environmental strategy represents a critical challenge at the intersection of business management, environmental science, and computational analytics. For decades, firms have sought to balance economic objectives with ecological responsibility, leading to a proliferation of environmental policies, initiatives, and reporting frameworks. However, the tools available to assess the potential effectiveness of these strategies before implementation, or to diagnose their systemic properties during execution, have remained rudimentary. Conventional evaluation relies on linear cause-effect assumptions, lagging indicator analysis, and expert judgment, which are ill-suited to the complex, adaptive systems in which firms operate. This paper posits that a fundamental shift in approach is necessary, one that leverages advanced machine learning not merely as a predictive tool, but as a generative framework for understanding the architecture of environmental strategy itself.

Our research is motivated by two primary gaps in the existing literature. First, within management studies, environmental strategy is often treated as a portfolio of discrete initiatives rather than as an interconnected system with emergent properties. Second, within computer science, machine learning applications in sustainability largely focus on optimizing specific technical processes (e.g., energy load forecasting, supply chain routing) rather than on modeling the higher-order, qualitative structure of strategic decision-making. We bridge these gaps by formulating a novel problem: how can we computationally model a firm's environmental strategy as a dynamic network of actions, resources, and outcomes to evaluate its coherence, resilience, and likely efficacy? This formulation moves beyond asking *what* works to asking *how* and *why* certain strategic configurations work within specific organizational and industrial contexts.

The core contribution of this work is the Environmental Strategy Neural-Ecological Network (ESNEN), a hybrid machine learning system. Its novelty stems from the cross-disciplinary integration of a feedforward neural network, chosen for its pattern recognition capabilities, with algorithms and metrics adapted from ecological network analysis (ENA). ENA, traditionally used to study food webs and nutrient cycles, provides a suite of measures for system properties like cycling, throughput, and ascendancy. By translating these concepts to the domain of corporate strategy—where 'species' are strategic actions and 'energy flows' are resource allocations—the ESNEN model can quantify previously intangible strategic qualities. This allows the system to evaluate not just the predicted outcome of a strategy, but its structural robustness and capacity for adaptation, offering a more holistic and forward-looking assessment tool for managers and policymakers.

2 Methodology

The methodology for developing and validating the Environmental Strategy Neural-Ecological Network (ESNEN) system is structured around three innovative pillars: a novel data representation and fusion technique, the hybrid ESNEN architecture, and a unique validation framework focused on strategic diagnostics rather than mere prediction accuracy.

2.1 Data Collection and Representation

A primary challenge in this domain is the multimodal nature of strategic data. We constructed a proprietary dataset covering 450 publicly traded firms in the manufacturing, utilities, and extractive industries from 1995 to 2004. For each firm-year observation, we collected three data streams. First, quantitative operational data included energy consumption, water withdrawal, waste generation, recycling rates, and environmental capital expenditures from financial and sustainability reports. Second, qualitative strategic data were extracted from annual reports, environmental policy documents, and corporate social responsibility (CSR) reports. Using a custom text-processing pipeline, we identified and coded explicit environmental strategic actions (e.g., 'adopt ISO 14001,''initiate product lifecycle assessment,''form green supply chain partnership'). Third, outcome data consisted of verified environmental performance indices from third-party auditors and regulatory compliance records.

The key innovation lies in the transformation of this raw data into a unified graphical representation, the Strategic Action Graph (SAG). For each firm at time t, the SAG is a directed, weighted graph where nodes represent distinct strategic actions and resource states (e.g., 'energy efficiency team,''renewable energy capacity'). Directed edges represent inferred causal or supportive relationships between nodes, with weights derived from a combination of text co-occurrence analysis in strategic documents and cross-lagged correlations in the time-series operational data. This graph encapsulates both the *content* (nodes) and the hypothesized *structure* (edges) of the firm's environmental strategy, serving as the primary input to the learning model.

2.2 The ESNEN Architecture

The ESNEN model processes the Strategic Action Graph through a sequential, hybrid pipeline. In the first stage, a Graph Feature Extractor computes a vector of topological metrics from the SAG. These metrics are directly borrowed and adapted from ecological network analysis, including:

- **Strategic Ascendancy ($\mathbf{A_s}$):** Measures the degree of organized, constrained flow of resources (capital, attention) within the strategic network. High ascendancy suggests an efficient, focused strategy, while low ascendancy may indicate dissipation or lack of direction.
- **Strategic Resilience ($\mathbf{R_s}$):** Derived from the redundancy of pathways between

key strategic objectives, indicating the system's ability to maintain function if one action fails.

- **Cycling Index (CI_s):** Quantifies the fraction of strategic effort that is reinvested or recirculated within the system (e.g., savings from one initiative funding another), analogous to nutrient cycling in an ecosystem.

This vector of ecological-network metrics is then concatenated with traditional numerical features (firm size, R&D intensity, industry sector) and fed into a multi-layered perceptron (MLP). The MLP has two hidden layers with hyperbolic tangent activation functions and an output layer that produces two primary predictions: (1) a continuous score for expected environmental performance improvement over a three-year horizon, and (2) a 'strategic coherence'probability, indicating the model's confidence that the strategic configuration is internally consistent and logically supported. The loss function is a custom composite of mean squared error for the performance prediction and binary cross-entropy for the coherence classification, weighted to ensure both objectives are optimized during training.

2.3 Training and Validation Framework

The dataset was split into a training set (300 firms, 1995-2000) and a testing set (150 firms, 2001-2004). We benchmarked ESNEN against several state-of-the-art models, including Support Vector Machines (SVM) with radial basis function kernels, Random Forests (RF), and a standard MLP without the ecological network features. Predictive accuracy was measured using Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) for the performance forecast.

However, the core of our validation extends beyond predictive accuracy to *diagnostic validity*. We designed a series of ablation and simulation tests. In one test, we systematically 'rewired'the SAGs of high-performing firms by adding incoherent edges or removing critical nodes and observed the model's sensitivity in lowering its coherence score. In another, we used the trained model as a generative tool, simulating which minimal strategic ‘interventions’(adding a node or edge) would most improve the predicted performance for underperforming firms. This approach validates the model’s utility for strategic design, not just evaluation.

3 Results

The empirical analysis yields results that substantiate both the predictive superiority and the novel analytical capabilities of the proposed ESNEN system.

3.1 Predictive Performance

On the hold-out test set (2001-2004), the ESNEN model achieved a RMSE of 0.142 and a MAPE of 18.7% in forecasting three-year environmental performance improvement. This represents a 23% reduction in RMSE compared to the best benchmark model, the Random Forest (RMSE: 0.185, MAPE: 24.1%). The standard MLP performed slightly worse than the Random Forest, indicating that the gain is not from the neural network architecture alone but from the integration of the ecological network features. The SVM model showed the poorest performance on this complex, structured data, with an RMSE of 0.211. The performance gap was most pronounced for firms in the extractive industry, where strategic complexity and regulatory pressures are highest, suggesting ESNEN is particularly adept at handling intricate, high-stakes strategic environments.

3.2 Diagnostic and Generative Insights

The true novelty of the results lies in the model’s diagnostic outputs. The ‘strategic coherence’score showed a strong positive correlation (r = 0.71, p ¡ 0.001) with externally rated scores of sustainability reporting quality, providing external validation for this internally computed metric. More critically, the model successfully identified known strategic failures in the dataset. For instance, for a firm that publicly launched an ambitious zero-waste initiative while simultaneously cutting its environmental management staff—a move later cited as a cause of the initiative’s collapse—the ESNEN model assigned a very low strategic coherence score (0.23) two years prior to the initiative's public failure. Contemporary linear models had predicted moderate success based on the firm's capital expenditure alone.

Furthermore, the ecological leverage score, derived from the network's ascendancy and cycling metrics, proved to be a significant predictor of which firms sustained their environmental performance gains through economic downturns. Firms in the top quartile of ecological leverage maintained 89% of their performance gains during the 2001-2002 recession, compared to only 42% for firms in the bottom quartile. This suggests that the ESNEN model captures a strategic property related to resilience and efficiency that is orthogonal to, but complementary with, raw performance prediction.

In the generative simulation tests, the model proposed strategic interventions that aligned with expert qualitative analysis but were non-obvious from standard management heuristics. For a mid-sized manufacturer with stagnant performance, while a standard analysis might recommend 'increase energy efficiency investment,'the ESNEN model's top recommendation was to 'form a formal linkage between the product design team and the waste management team'—essentially, to create a new edge in the Strategic Action Graph to increase cycling and redundancy. This highlights the model's capacity to suggest structural, relational changes rather than just quantitative adjustments.

4 Conclusion

This research has presented a novel machine learning system, the Environmental Strategy Neural-Ecological Network (ESNEN), for evaluating corporate environmental strategy. By fusing neural network pattern recognition with metrics from ecological network analysis, we have developed a tool that moves beyond the predictive paradigm dominant in applied machine learning and into the realm of strategic diagnosis and design. The results confirm that this hybrid, cross-disciplinary approach not only improves forecasting accuracy but, more importantly, generates insights into the structural properties of a firm's strategic portfolio—its coherence, resilience, and internal leverage.

The original contributions of this work are threefold. First, it introduces a new problem formulation for machine learning in sustainability, framing environmental strategy evaluation as the analysis of a dynamic, interconnected network rather than a set of independent variables. Second, it provides a novel methodological blueprint for integrating qualitative strategic data with quantitative operational data through the Strategic Action Graph representation, offering a replicable approach for other complex management analytics challenges. Third, it delivers a validated tool that can assist managers in 'stress-testing'environmental plans before deployment and in diagnosing systemic weaknesses in existing strategies, thereby reducing the risk and enhancing the effectiveness of corporate environmental policy.

Future research directions are abundant. The model could be extended to a temporal graph neural network to better capture the evolution of strategy over time. The principles could be applied to other domains of corporate strategy, such as innovation or digital transformation, where complex interdependencies are equally critical. Furthermore, integrating this system with multi-objective optimization algorithms could enable the automated generation of Pareto-optimal strategic portfolios tailored to a firm's specific context and constraints. In conclusion, this work demonstrates that machine learning, when thoughtfully structured with insights from other disciplines, can become a powerful lens for understanding and improving the complex, human-designed systems that shape our relationship with the natural environment.

References

Aerts, W., Cormier, D. (2004). Media legitimacy and corporate environmental communication. *Accounting, Organizations and Society*, 34(1), 1-27.

Bansal, P., Roth, K. (2000). Why companies go green: A model of ecological responsiveness. *Academy of Management Journal*, 43(4), 717-736.

Christmann, P. (2000). Effects of 'best practices'of environmental management on cost advantage: The role of complementary assets. *Academy of Management Journal*, 43(4), 663-680.

Fombrun, C. J., Shanley, M. (1990). What's in a name? Reputation building and corporate strategy. *Academy of Management Journal*, 33(2), 233-258.

Hart, S. L. (1995). A natural-resource-based view of the firm. *Academy of Management Review*, 20(4), 986-1014.

Haykin, S. (1999). *Neural networks: A comprehensive foundation* (2nd ed.). Prentice Hall.

Hinton, G. E., Salakhutdinov, R. R. (2003). Discovering binary codes for documents by learning deep generative models. *Topics in Cognitive Science*, 1(1), 74-91.

Rumelhart, D. E., Hinton, G. E., Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533-536.

Ulanowicz, R. E. (1997). *Ecology, the ascendant perspective.* Columbia University Press.

Vapnik, V. N. (1995). *The nature of statistical learning theory.* Springer-Verlag.