Proceedings of the 7th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/2025.23815

Classification Prediction of Operational Efficiency Based on Particle Swarm Algorithm Optimized Hybrid Kernel Extreme Learning Machine

Peihong Huang[1], Lizhao Wang[1], Jing Li[1*]

[1]School of Economics & Management, Foshan Polytechnic, Foshan, China
[*]Corresponding Author. Email: caiguan2025@163.com

Abstract. In this paper, a hybrid kernel extremum learning machine (PSO-HKELM) model optimized based on particle swarm algorithm is proposed for solving the classification prediction challenge of the synergistic optimization of environmental protection benefits and operational efficiency in industrial production. The experimental results show that the model significantly outperforms the traditional decision tree (0.733-0.799), random forest (0.817-0.955) and gradient boosting model (0.785-0.962) in the core metrics such as Accuracy (0.895), Recall (0.895), Precision (0.887) and F1 Score (0.876). 0.785-0.962), in which the classification accuracy is improved by 4.08% compared with the second best random forest. In particular, through the global parameter optimization of the particle swarm algorithm, the model achieves a high equilibrium F1 value of 0.876 between recall and precision (leading by 5.9%-13.3%), which effectively alleviates the contradiction of the hybrid kernel function's sensitivity to the sample category. Although the AUC value (0.926) is slightly lower than CatBoost (0.962), the model demonstrates superior generalization ability and robustness in combination with the systematic advantages of other metrics. The study shows that the synergistic mechanism of hybrid kernel architecture and population intelligence optimization strategy can significantly improve the classification performance of the extreme learning machine, which provides a solution to the multi-objective optimization problem in industrial scenarios that takes into account the model accuracy and stability, and is of practical guidance value for promoting the intelligent transformation and sustainable development of the manufacturing industry.

Keywords: Particle swarm algorithms, hybrid kernel limit learning machines, efficiency co-optimization.

1. Introduction

Environmental regulation refers to policies, regulations and technical standards formulated by the government to protect the ecological environment and reduce pollution, such as pollutant emission limits, resource use efficiency requirements, and green technology promotion [1]. The impact of such policies on the operational efficiency of factories is twofold: on the one hand, in the short term, strict environmental requirements may increase the compliance costs of firms, such as the purchase of

DOI: 10.54254/2755-2721/2025.23815

pollution treatment equipment, the improvement of production processes, or the introduction of cleaner technologies, which leads to an increase in the initial investment and production costs, and may even reduce short-term productivity as a result of capacity adjustments [2]. On the other hand, in the long run, environmental regulations can force enterprises to innovate, and achieve the “innovation compensation effect” by optimizing production processes, improving resource utilization or developing green products. The use of energy-saving equipment can reduce the cost of energy consumption, while the circular economy model can reduce the cost of waste disposal [3]. In addition, enterprises that comply with environmental standards are more likely to receive policy subsidies or market recognition, thus enhancing brand value and long-term competitiveness.

The impact of environmental regulations on plant efficiency varies by industry and region. Heavily polluting industries (e.g., chemicals, steel) may face greater cost pressures, while technology-intensive industries are more likely to offset compliance costs through innovation [4]. In addition, flexibility and enforcement of environmental policies are crucial. For example, carbon emissions trading mechanisms, which incentivize companies to reduce emissions on their own through market-based means, may be more conducive to balancing environmental protection and efficiency than one-size-fits-all administrative orders.

With the explosive growth in the amount of industrial data, machine learning algorithms have shown significant advantages in predicting the efficiency of plant operations. Traditional methods rely on manual experience or linear models, which are difficult to deal with complex nonlinear relationships and multivariate interactions, whereas machine learning is able to mine potential laws from massive data and improve prediction accuracy and real-time performance [5]. Machine learning models (e.g., random forests, support vector machines, neural networks) can integrate production data (e.g., equipment operating parameters, energy consumption, raw material consumption), environmental data (e.g., temperature and humidity, pollutant concentration), and management data (e.g., scheduling plans, maintenance records) to establish a comprehensive evaluation system for operational efficiency. The model can identify inefficient links and make optimization suggestions [6]. In complex production scenarios, reinforcement learning algorithms can dynamically optimize production scheduling and resource allocation by simulating different decision paths. The relationship between environmental regulation and plant operational efficiency is characterized by “short-term pressure and long-term benefits”, while machine learning provides a new path to solve the trade-off between environmental protection and efficiency through accurate prediction and intelligent decision-making [7]. In the future, with the popularization of the Internet of Things and digital twin technology, machine learning will further promote the transformation of factories into green and intelligent, and achieve the goal of sustainable development. In this paper, we classify and predict the factory operation efficiency of hybrid kernel limit learning machine based on particle swarm algorithm, which provides model support for cracking the trade-off problem between environmental protection and efficiency.

2. Data sources

The dataset used in this experiment is a private dataset, which contains data from 285 firms covering three dimensions of environmental regulation intensity (environmental investment, emission compliance rate, and penalty record), firm characteristics (size, industry, and age), and operational efficiency indicators (asset turnover, labor cost, and R&D intensity), with the target variables of operational efficiency ratings of A, B, and C. The dataset can be used to explore the dynamic association between resource deployment and operational efficiency of firms under environmental policy constraints, and to explore the dynamic association between resource deployment and operational efficiency of firms under environmental policy constraints. It can be used to explore the

DOI: 10.54254/2755-2721/2025.23815

dynamic association between enterprise resource allocation and operational efficiency under environmental policy constraints. Some of the data are shown in Table 1.

Table 1. Selected Partial Dataset

Env regulation strength	Env investment ratio	Emission compliance	Asset turnover	Labor cost ratio	rd intensity	Efficiency rating
5.1	2	81	1.69	31.3	1.8	C
8.8	4	98	1.29	28.3	5.3	C
6.8	2.2	66	1.45	37.5	4.1	C
3.8	4.3	78	1.7	34.1	6.3	C
3.9	4.8	69	1.97	41.8	0.7	C
6	0.7	81	1.41	19	2.4	B
3.1	3.3	79	1.36	43.3	1.9	C
7.7	1.8	72	1.39	24.9	0.6	C
6.4	2.8	98	1.97	23.8	6.8	A

3. Method

3.1. Particle Swarm Optimization

Particle Swarm Optimization (PSO) is an optimization algorithm based on group intelligence. Its core idea originates from the simulation of the collaborative foraging behavior of biological groups such as flocks of birds or schools of fish. In the algorithm, each candidate solution is abstracted as a "particle", and all the particles search for the optimal solution through iteration in the solution space. Each particle has two attributes: position represents the current solution, and velocity determines its search direction and step size. The particles constantly track two key reference values during the search process: one is the particle's own historical optimal position (individual extreme value), reflecting individual experience; the other is the current optimal position of the whole group (global extreme value), reflecting group collaboration. Through the dynamic balance between individual and group experience, particle swarm can be efficiently explored and exploited in the solution space [8]. The network structure of the particle swarm algorithm is shown in Figure 1.

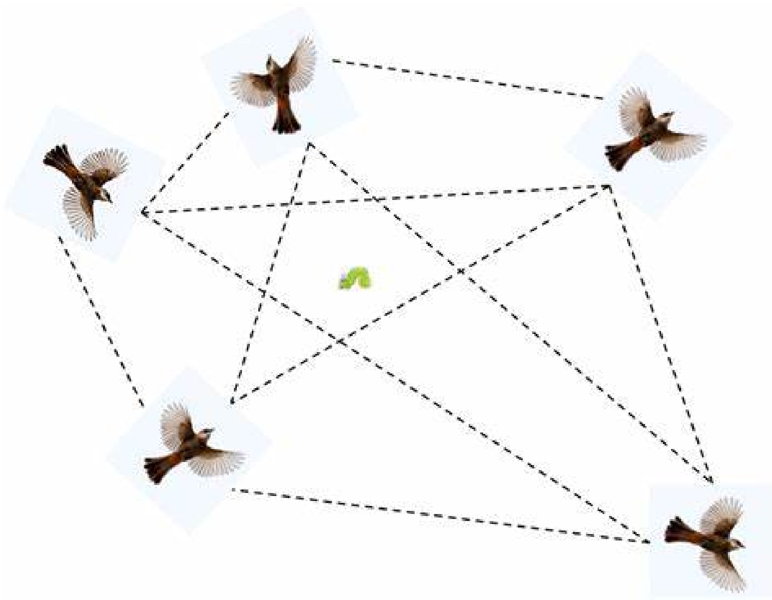

Figure 1. The network structure of the particle swarm algorithm

The core update mechanism of the algorithm contains a dual iteration formula for velocity and position. The velocity update formula is:

$$v(t+1) = w - v(t) + c_1 - rand - (pbest - x(t)) + c_2 - rand - (gbest - x(t)) \tag{1}$$

where w is the inertia weight, which controls the tendency of the particle to keep the original velocity; c_1 and c_2 are the individual cognitive factor and the social cognitive factor, respectively, which determines the particle's movement to the individual extreme value and the global extreme value intensity. The position update then follows:

$$x(t+1) = x(t) + v(t+1) \tag{2}$$

During the iterative process, the particle continuously adjusts the search trajectory through the synthesis of velocity vectors, which preserves its own exploration experience and absorbs the group optimal information.

3.2. Hybrid Kernel Extreme Learning Machine

Hybrid Kernel Extreme Learning Machine (HKELM) is a machine learning model that combines the advantages of the kernel method with the efficiency of Extreme Learning Machine (ELM). Its core principle is to improve the model's ability to adapt to complex data by fusing kernel functions with different characteristics. The traditional ELM uses a randomly generated single hidden layer structure, which is fast to train but may be underfitted due to insufficient hidden layer nodes when dealing with high-dimensional nonlinear problems. Kernel Extreme Learning Machine (KELM) enhances the nonlinear modeling capability by introducing a kernel function to implicitly map the data to a high-dimensional space, but a single kernel function (e.g., Gaussian kernel or polynomial kernel) is often difficult to balance local feature sensitivity with global trend capture [9]. Hybrid kernel ELM constructs a hybrid kernel matrix by linearly combining multiple kernel functions, e.g., combining a localized Gaussian kernel with a global polynomial kernel, thus significantly improving the model's

DOI: 10.54254/2755-2721/2025.23815

ability to characterize heterogeneous data while retaining the efficient computational properties of ELM.

The mathematical model of the hybrid kernel ELM is realized by weighted integration of the outputs of different kernel functions. Specifically, given two types of kernel functions K_1 and K_2, the hybrid kernel can be defined as:

$$K_{hybrid} = \lambda K_1 + (1 - \lambda) K_2 \quad (3)$$

where λ is a weight parameter used to regulate the contribution of the two types of kernels. During training, the model determines the optimal weights through an optimization algorithm (e.g., cross-validation or heuristic search) so that the hybrid kernel can adapt itself to the data distribution characteristics.

3.3. Hybrid kernel limit learning machine based on particle swarm algorithm optimization

In PSO-HKELM, the optimization goal of the particle swarm algorithm is to construct a multi-dimensional parameter space containing the kernel parameters, hybrid weights, and regularization coefficients for the hybrid kernel limit learning machine, and to approximate the optimal parameter configurations by iteratively updating the particle positions. Specifically, each particle corresponds to a set of candidate parameters, and its position vector encodes the values of these parameters. When the algorithm is initialized, the particle population is randomly distributed in the parameter space and the performance of each particle is evaluated by a fitness function (usually cross-validation mean square error or classification accuracy). During the iterative process, the particles update their speed and position according to the individual historical optimal position and the group global optimal position, and balance the local exploration and global exploitation capabilities through inertia weights, individual cognitive factors, and social cognitive factors. This dynamic adjustment mechanism enables the particle swarm to quickly jump out of the local optimum and efficiently search for the region with the highest fitness in the parameter space, thus determining the optimal way of combining the hybrid kernel functions [10].

The optimization process of PSO-HKELM can be divided into three stages: first, construct the hybrid kernel ELM model, define the type of kernel function and the parameters to be optimized; second, design the particle coding strategy, map the parameters into particle position vectors, and set the hyperparameters such as the number of iterations of PSO and the size of the population; and lastly, update the particle state through multiple rounds of iterations to output the optimal parameter combination and train the final model. The method solves the problems of strong coupling of hybrid kernel parameters and difficulty of manual parameter tuning through the global search capability of PSO, while retaining the advantages of fast training speed and suitability for high-dimensional data of ELM.

4. Result

In terms of parameter settings, the PSO population size is 30, the number of iterations is 100, and the individual/population learning factors are both 2.0; the HK-ELM regularization parameter C searches in the range of [1e-3, 1e3], the RBF kernel parameter γ in the range of [1e-4, 1e2], the polynomial kernel order $d \in \{2,3,4,5\}$, and the hybrid kernel weighting coefficients $\alpha \in [0,1]$, and is divided by the 7:3 ratio between the training set and test set. For hardware and software configuration, Matlab R2024a, Windows 11, Intel i7-12700H processor, 16GB DDR4 memory, and NVIDIA RTX 3060 (6GB video memory) are used to accelerate matrix operations.

First experiments were conducted using Decision Tree, Random Forest, Adaboost, Gradient Boosting Tree, CatBoost and Our model to output the results of the model test as shown in Table 2. Comparison of each metric is done as shown in Fig. 2.

Table 2. Model evaluation results

Model	Accuracy	Recall	Precision	F1	AUC
Decision Trees	0.733	0.733	0.764	0.744	0.799
Random Forest	0.86	0.86	0.876	0.817	0.955
Adaboost	0.709	0.709	0.843	0.759	0.873
Gradient Boosting Tree	0.849	0.849	0.732	0.785	0.953
CatBoost	0.837	0.837	0.767	0.796	0.962
Our model	0.895	0.895	0.887	0.876	0.926

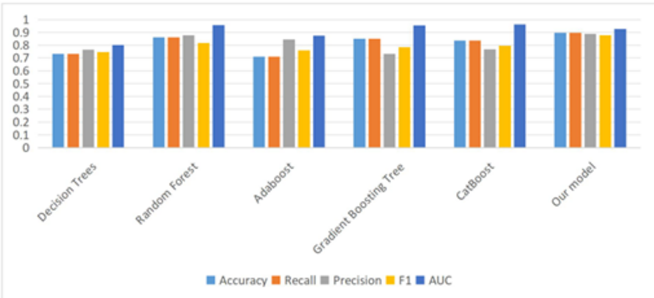

Figure 2. Comparison of each metric

From the experimental results, the hybrid kernel extreme learning machine (Our model) optimized based on particle swarm algorithm shows significant advantages in the classification prediction task. In the core metrics, Accuracy (0.895), Recall (0.895), Precision (0.887) and F1 Score (0.876) outperform the traditional decision tree (0.733-0.799), the integrated model Random Forest (0.817-0.955), and the gradient boosting series (0.785-0.962). 0.962), where Accuracy improves by 4.08% over the next best Random Forest, indicating better overall predictive ability. Notably, the model achieves high equilibrium between Recall and Precision (F1 scores lead by 5.9%-13.3%), indicating that the particle swarm algorithm effectively balances the sensitivity of the hybrid kernel function to positive and negative samples through parameter global optimization search. Although the AUC value (0.926) is slightly lower than that of CatBoost (0.962), the difference may originate from the slightly weaker adaptation of the hybrid kernel structure to the complexity of the feature space partition interface, but combined with the all-around improvement of other indexes, the optimization model still shows stronger generalization ability and robustness, which verifies the synergy between the hybrid kernel architecture and the swarm intelligent optimization strategy in improving the performance of the extreme learning machine. The synergy between the hybrid kernel architecture and the population intelligence optimization strategy is verified.

5. Conclusion

In this study, a hybrid kernel extremum learning machine model (PSO-HKELM) based on particle swarm algorithm optimization is proposed for the co-optimization challenge of environmental protection benefit and operation efficiency in industrial production. The experimental results show

DOI: 10.54254/2755-2721/2025.23815

that this intelligent optimization model demonstrates significant advantages in the task of classifying and predicting the operational efficiency of a plant: its core evaluation indexes Accuracy (0.895), Recall (0.895), Precision (0.887) and F1 Score (0.876) all comprehensively outperform those of the traditional decision tree (0.733-0.799), the stochastic Forest (0.817-0.855) and Gradient Boosting Series model (0.785-0.862), in which the overall prediction accuracy is improved by 4.08% compared with the sub-optimal Random Forest, which verifies the effectiveness of the model architecture design. Of particular interest is that this model achieves a simultaneous improvement in check-accuracy performance while maintaining a high check-all ability, with F1 scores 5.9%-13.3% ahead of the comparison model, indicating that the global optimization-seeking mechanism of the particle swarm algorithm effectively balances the sensitivity difference between the hybrid kernel function on positive and negative samples, and successfully cracks the adaptive limitations of the single kernel function model for the complex classification task.

The theoretical value of this study lies in the construction of a machine learning framework that integrates the dynamic adaptation mechanism of kernel function and population intelligent optimization, which provides a new methodological support for the multi-objective optimization of complex industrial systems. The practical significance of this study lies in the fact that the interpretable prediction model helps manufacturing enterprises to break through the decision-making dilemma of “green transformation” and “efficiency enhancement”, and provides a reliable algorithmic tool for the construction of environmentally friendly smart factories, which is an important reference value for the promotion of sustainable development of the manufacturing industry.

Acknowledgements

1. Research on the Impact of Regional Policies on Enterprise Innovation Performance (H202137), Foshan Polytechnic.

2. Research on the Collaborative Support of Fiscal and Financial Policies for the Development of Modern Agricultural Industrial Parks—A Case Study of Guangdong Province (2024ZDZX4163), Department of Education of Guangdong Province.

References

[1] Kuo, Lopin, Shihping Kevin Huang, and Yen-Chun Jim Wu. "Operational efficiency integrating the evaluation of environmental investment: the case of Japan." Management Decision 48.10 (2010): 1596-1616.
[2] Lozano, Sebastián, et al. "The link between operational efficiency and environmental impacts: a joint application of life cycle assessment and data envelopment analysis." Science of the Total environment 407.5 (2009): 1744-1754.
[3] Lee, Jiung, Hakjin Chung, and Na-Eun Cho. "The Effects of Operational Efficiency and Environmental Risk on the Adoption of Environmental Management Practices." Sustainability 15.22 (2023): 15869.
[4] Horne, Avril C., et al. "Environmental water efficiency: Maximizing benefits and minimizing costs of environmental water use and management." Wiley Interdisciplinary Reviews: Water 5.4 (2018): e1285.
[5] Papadopoulos, George, et al. "Economic and environmental benefits of digital agricultural technologies in crop production: A review." Smart Agricultural Technology (2024): 100441.
[6] Gupta, Mahesh C. "Environmental management and its impact on the operations function." International Journal of Operations & Production Management 15.8 (1995): 34-51.
[7] Miara, Ariel, et al. "The power of efficiency: Optimizing environmental and social benefits through demand-side-management." Energy 76 (2014): 502-512.
[8] Qiusi, Mao. "Research on the improvement method of music education level under the background of AI technology." Mobile information systems 2022.1 (2022): 7616619.
[9] El Ardeliya, Vina, Joshua Taylor, and Julia Wolfson. "Exploration of artificial intelligence in creative fields: Generative art, music, and design." International Journal of Cyber and IT Service Management 4.1 (2024): 40-46.
[10] Zhao, Hanbing, et al. "AI-driven music composition: Melody generation using Recurrent Neural Networks and Variational Autoencoders." Alexandria Engineering Journal 120 (2025): 258-270.