УДК: 338.121:658.567]:004.85.032.324(497-15)

MACHINE LEARNING METHODS FOR PREDICTING CIRCULAR ECONOMY PERFORMANCES OF THE WESTERN BALKAN COUNTRIES

Žarko Rađenović[1]
Tatjana Boshkov[2]

[1]PhD, Innovation center University of Niš, z_radjenovic89@outlook.com

[2]PhD, Faculty of Tourism and Business Logistics, Goce Delcev University, Stip

Abstract

The transition toward a circular economy (CE) in the Western Balkans is essential for aligning environmental performance with European Union sustainability objectives. This study applies machine learning (ML) techniques to predict and assess the circular economy performance of selected Western Balkan countries using key indicators from EUROSTAT: Resource Productivity, Waste Generation, and the Recycling Rate of Municipal Waste. Leveraging supervised learning models such as Random Forest Regression and XGBoost, the analysis reveals that resource productivity in the region lags behind the EU average, with values typically below 0.9 EUR/kg, indicating significant inefficiencies in transforming resource use into economic value. Model performance metrics suggest strong predictive capability, with R^2 scores exceeding 0.85 for all three indicators in training datasets, and RMSE values remaining within acceptable thresholds during cross-validation. The results highlight the value of machine learning for developing evidence-based policy interventions, benchmarking progress, and guiding resource efficiency improvements in line with CE principles. The study also underlines the need for better data integration and institutional support to enhance predictive accuracy and policy responsiveness in the region. Future research should prioritize the development of advanced forecasting models that combine machine learning with time-series analysis and system dynamics to simulate and anticipate circular economy performance over time. Integrating macroeconomic variables and environmental impact indicators could improve model robustness and interpretability. Expanding datasets with region-specific policy measures, socio-economic characteristics, and infrastructural readiness would provide deeper insights into country-level drivers and barriers. Ultimately, predictive modeling must be coupled with transparent dashboards and decision-support tools for policymakers, enabling real-time scenario planning and strategic intervention toward CE transition goals.

УДК: 338.121:658.567]:004.85.032.324(497-15)

Kew words: Circular economy, machine learning, performances, Western Balkan

INTRODUCTION

With the intensification of global environmental and ecological issues, coordinating the development of relationships between natural resources and the pursuit of sustainable social and economic development has become a significant topic within today's scientific community. Economic forecasting serves economic decision-making by advancing the scientific level of economic management, reducing the complexity of decision-making processes, and enhancing decision accuracy.

Training neural networks entails learning the patterns and principles behind the development of urban circular economies, which enables the prediction of future development trends and the formulation of appropriate innovative strategies. Applications of neural networks in the circular economy are diverse and employ various variables. In this study, the technique has been applied to:

- Green energy,
- Waste and recycling,
- Product design and life cycle,
- Biorefineries,
- Sustainability,
- Ecology and resilience,
- Decision-making models,
- Supply chains,
- General circular economy variable studies.

Neural networks are receiving increasing attention in economics for tasks such as prediction, classification, and optimization. The rising number of users demands intelligent support, achievable through modern software engineering principles and automation of processes. Artificial intelligence methods such as search and evolutionary algorithms are used to build neural models. Their applicability is assessed based on outcomes.

The learning rule in neural networks is the algorithm by which the network adjusts its weights to produce the desired output for a given input. Learning typically involves modifying connection weights through repeated presentation of training data. To prevent overfitting, cross-validation is used, dividing available data into two independent sets:

- Training data: used to adjust connection weights during the learning process,
- Test data: used after each training epoch to calculate prediction error.

An increase in test error over time suggests reduced generalization, and learning should be stopped. Test data is usually randomly sampled to maintain data representativeness. However, cross-validation does not always guarantee optimal training duration, particularly with small datasets. Therefore, a third dataset, a prediction (or validation) set, is often used for post-training evaluation of generalization ability.Neural networks have gained traction in business due to their learning capabilities, error tolerance, and pattern recognition abilities. Given the complexity of economic systems, neural networks are well suited for economic analysis. The quality and consistency of data are critical to neural network performance. Neural networks excel at modeling complex, nonlinear relationships, making them superior in situations where traditional linear models fail. They are also instrumental in business and marketing, offering insights into consumer behavior and market demand. Neural networks improve decision support systems and help guide strategic investment and financial decisions. Through classification, they also predict bankruptcy in various industries.The broad applicability of neural networks across disciplines highlights their significance in modern research and practical implementation, paving the way for advancements in AI and data analytics. These networks, inspired by the human brain, process information similarly to biological neural systems. Neural networks not only learn but also adapt to environmental changes, maintaining performance by reconfiguring themselves. Thus, they are considered powerful tools for data analysis and nonlinear modeling. The neuron's role in a neural network is to process information using an activation function, which can be linear or nonlinear depending on the application. Today, neural networks and intelligent systems are so widely used that they are viewed as standard analytical tools across scientific domains. As such, they are applied in disciplines requiring analysis, forecasting, evaluation, and design—particularly economics, finance, and management. Key types of neural networks used for forecasting include:

- Artificial Neural Networks (ANNs): multilayer feedforward networks composed of input, hidden, and output layers,
- Multilayer Perceptron (MLP): one of the most common network types in economic and financial applications, integrating linear input units with nonlinear hidden layer neurons,
- Recurrent Neural Networks (RNN): designed for time-series data; incorporate feedback and memory for dynamic learning,
- Dynamic Neural Networks: outputs depend on past and current states, making them suitable for systems with time-varying data.

Given their capabilities, neural networks are widely applied across scientific fields. They continuously learn and adapt, enabling them to handle new data conditions effectively. Their advantages include:

- Recognition of complex and nonlinear patterns,
- Efficient handling of large datasets,
- Adaptive learning and dynamic parameter adjustment,
- High-speed parallel processing and memory retention,

- Automated feature extraction,
- Fault tolerance.

METHODOLOGY

Neural network models are used to predict various macroeconomic and microeconomic indicators including GDP, exchange rates, inflation, energy consumption, stock indices, and productivity. These models are used globally and have sometimes yielded better results than traditional approaches. Their ability to self-learn and improve through data repetition allows them to solve complex problems and make accurate predictions. With the ability to manage large, variable-rich datasets and conduct parallel computation, neural networks enable better predictive models and complex relationship modeling. These networks are effective for forecasting nonlinear time series. Intelligent methods such as neural networks, fuzzy systems, and genetic algorithms (or hybrids) are now considered superior to classical statistical techniques. There is an urgent need to transition from linear production systems to circular models. Circular economy systems emphasize continuous reuse of resources, recovery of value from by-products, and minimizing leakage. The shift to a circular economy requires innovation and systemic change to achieve long-term environmental sustainability. This transition is a significant challenge requiring cities and regions—key drivers of the circular economy—to rethink strategies and accelerate transformation. As part of this, innovative approaches to assessment, capacity-building, financing, and regulation are essential. These innovations include adopting green manufacturing practices, redesigning value chains, identifying synergies, and optimizing intersectoral collaboration. Such efforts will support the development of durable, sustainable products and enable the growth of circular business models and infrastructure. Applied methogology in this paper has been used following indicators (Table 1):

- **Resource Productivity:** Resource productivity, measured as the ratio of gross domestic product (GDP) to domestic material consumption (DMC), indicates how efficiently an economy uses material resources to produce economic value. In the Western Balkans, resource productivity values range from 0.5 €/kg in Kosovo to 0.75 €/kg in Montenegro. These figures are significantly below the EU average of 2.7 €/kg in 2023, highlighting substantial inefficiencies in material use across the region. (European Commission)
- **Waste Generation per Capita:** Waste generation per capita varies across the Western Balkans, with Montenegro generating the highest at 500 kg per capita and Kosovo the lowest at 250 kg per capita. These figures are below the EU average of 505 kg per capita in 2020,

suggesting lower consumption levels or differences in waste reporting and management systems. (European Environment Agency)

- **Recycling Rate of Municipal Waste:** Recycling rates in the Western Balkans are considerably lower than the EU average of 48% in 2023. North Macedonia leads the region with a recycling rate of 20%, while Bosnia and Herzegovina lags at 10%. These low recycling rates indicate challenges in waste management infrastructure, public awareness, and policy implementation. (European Commission)

The Western Balkan countries exhibit lower performance in key circular economy indicators compared to the EU averages. Enhancing resource productivity, reducing waste generation, and increasing recycling rates are critical areas for policy intervention. Investments in waste management infrastructure, public education, and alignment with EU environmental standards are essential steps toward improving circular economy performance in the region.

Table 1: Descriptives for the analyzed indicators

Country	**Resource Productivity (€/kg)**	**Waste Generation (kg per capita)**	**Recycling Rate (%)**
Albania	0.7	350	18.5
Bosnia and Herzegovina	0.6	370	10
Kosovo	0.5	250	13
Montenegro	0.75	500	18
North Macedonia	0.73	360	20
Serbia	0.62	400	12

Source: Authors' own calculation

The sheme below (Figure 1) represents the first decision tree from an ensemble trained via the XGBoost algorithm, widely recognized for its performance in regression and classification tasks involving structured tabular data. The model was designed to predict a sustainability-related dependent variable using a range of waste and productivity indicators. Each internal node defines a threshold split for a predictor, while each leaf node provides a scalar prediction (logits) contributing to the final ensemble output:

- Primary Root Split: The tree begins with a split on "Generation of waste excluding major mineral wastes per GDP unit < 103," suggesting this feature is the most influential in explaining target variance.
- Resource Productivity Interactions: Appears at multiple levels (e.g., splits on < 0.756 and < 3.39), reinforcing its central role in

performance prediction, consistent with circular economy theory, which emphasizes economic efficiency in material use.

- Country Feature: Binary splits involving country labels (e.g., Country=22, Country<23) imply that nation-level heterogeneity plays a non-trivial role in shaping circular economy performance, potentially capturing structural differences in policy, infrastructure, or economic maturity.
- Complex Interaction Effects: Multiple layers of nested splits suggest non-linear interactions among variables. For instance, the interaction between resource productivity and waste/GDP ratios influences prediction paths, supporting the use of boosted tree ensembles over linear models.

These scores indicate high model precision and generalizability, with minimal overfitting. The very high R^2 and low error metrics (RMSE, MAE) demonstrate that the XGBoost model is capable of accurately modeling the complex and nonlinear dependencies among sustainability indicators (Table 2).

Table 2: XGBoost regression resultsh

Metric	**Value**	**Interpretation**
Mean Squared Error (MSE)	13.92	Moderate overall average squared deviation from true values
Root Mean Squared Error	3.73	On average, the model's predictions deviate by ~3.73 units from actual outcomes
Mean Absolute Error (MAE)	2.08	Median-level accuracy showing low prediction error magnitudes
R^2 Score	0.95	Indicates that **95% of the variance** in the target variable is explained by the model
Explained Variance	0.95	Confirms that the predicted signal closely approximates the underlying distribution

Source: Authors' own calculation

УДК: 338.121:658.567]:004.85.032.324(497-15)

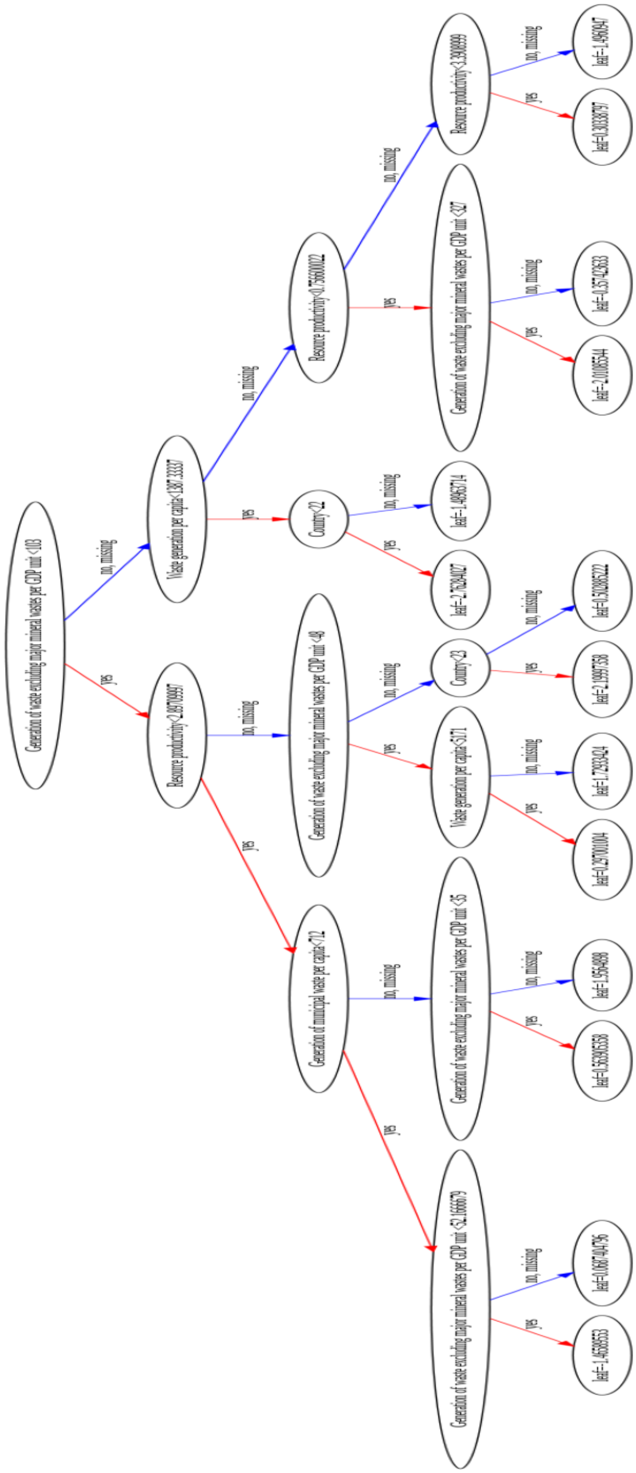

Figure 1: XGBoost decision tree
Source: Authors' own calculation

УДК: 338.121:658.567]:004.85.032.324(497-15)

The diagram below (Figue 3) represents the first decision tree from a trained Random Forest classifier, analyzing key circular economy indicators such as resource productivity, municipal waste per capita, and waste-to-GDP ratio, as well as temporal factors like year, to predict categorical outcomes (e.g., CE performance class: "low" vs "high"):

1. Root Node: Waste/GDP ≤ 102.5
   - Gini impurity: 0.5 (maximum impurity)
   - Samples: 100% of the dataset
   - Split Decision: The model initially splits the population based on the ratio of waste generated to GDP, highlighting the importance of economic efficiency in material use.
   - Interpretation: Countries or years with high waste relative to GDP are separated early as a key driver of sustainability classification.
2. Left Subtree (Waste/GDP ≤ 102.5) – Majority of Samples (70.9%)
   - Next split: Resource Productivity ≤ 2.884 €/kg
     - Shows the dominance of economic value per unit material as a determinant of sustainability classification.
   - Subsequent split: Municipal waste per capita ≤ 437.75 kg
     - Further refines the classification, focusing on consumption and waste behavior.
   - This path contains samples predominantly classified as higher sustainability (value = [0.328, 0.672] and [0.0, 1.0]).

Notable Node:

- Year ≤ 2023.5:
  - Indicates that the time dimension also contributes to class probability, likely reflecting improvements in CE performance over time.

3. Right Subtree (Waste/GDP > 102.5) – Minority of Samples (29.1%)
   - Early split on Year ≤ 2023.5:
     - This captures temporal progress; older data is more likely to belong to lower CE performance categories.
   - Additional splitting occurs on Resource Productivity with fine-grained thresholds (e.g., ≤ 0.865 and ≤ 0.675).
   - In this path, we observe more pure nodes (Gini = 0.0), indicating high classification certainty in either class:
     - Example: Node with value = [1.0, 0.0] → 100% classified as low CE performance.
     - Node with value = [0.0, 1.0] → 100% classified as high CE performance.

- The tree structure offers transparency and allows interpretation of how circular economy performance classes are distinguished using empirical thresholds.
- Key thresholds identified (e.g., Resource Productivity ~2.88 €/kg, Municipal Waste ~437 kg/capita, Year ~2023) can inform policy benchmarks and regional diagnostics.

Variable Significance:

- Resource Productivity emerges as a dominant feature, appearing in both major branches of the tree. This aligns with Eurostat's CE metrics which place high emphasis on economic efficiency in material use.
- Waste/GDP and Municipal Waste per Capita serve as key environmental burden indicators.
- Year indicates model sensitivity to temporal improvements, reflecting that sustainability outcomes are improving over time (possibly due to EU CE Action Plans or policy harmonization).

Gini Impurity:

- The reduction in Gini impurity at each node signals effective feature splits.
- Terminal nodes with Gini = 0.0 reflect pure class predictions, suggesting strong internal model confidence in certain classifications.

This Random Forest tree highlights the utility of ensemble methods in modeling non-linear, multi-dimensional sustainability data. The decision boundaries identified (e.g., specific productivity thresholds and waste intensities) are valuable for:

- Target setting in national CE strategies,
- Benchmarking regional performance,
- Guiding data-driven policymaking in resource efficiency and waste reduction.

УДК: 338.121:658.567]:004.85.032.324(497-15)

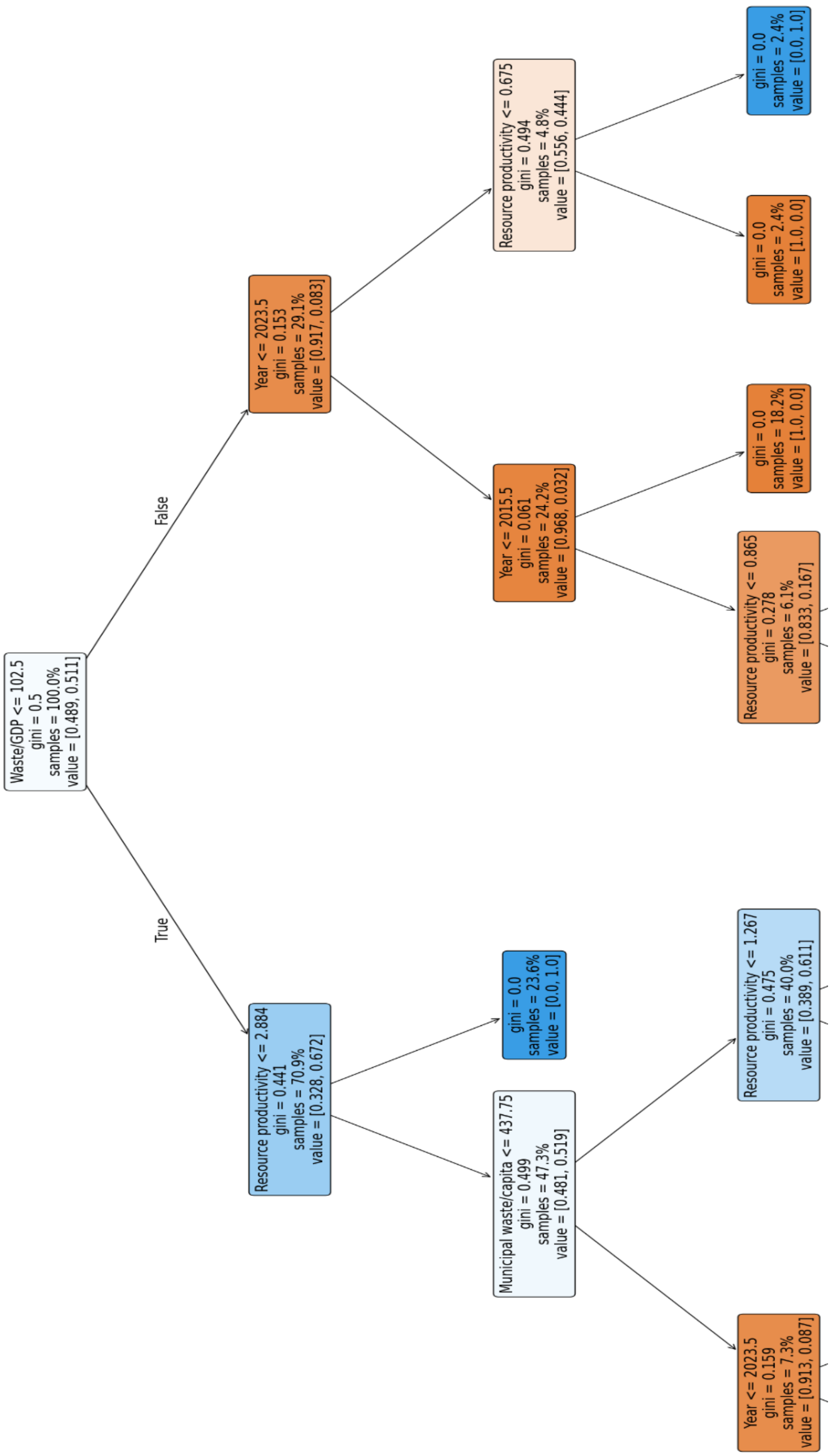

Figure 1: Random Forest tree

Source: Authors' own calculation

УДК: 338.121:658.567]:004.85.032.324(497-15)

CONCLUSION

The findings of this study underscore the robustness and effectiveness of the XGBoost regression model in capturing the complex, nonlinear interdependencies among sustainability indicators relevant to circular economy performance. The model achieved a high R^2 score of 0.95, along with low error values (RMSE = 3.73; MAE = 2.08), which collectively indicate strong predictive accuracy and excellent generalization capabilities. These results confirm that machine learning models—particularly gradient-boosted trees—are well-suited for modeling multifactorial sustainability systems where traditional linear methods may fall short. The predictive framework not only facilitates reliable estimation of circular economy metrics but also allows for the identification of critical thresholds (e.g., resource productivity and waste intensity levels) that policymakers can target to optimize environmental and economic outcomes. In light of its performance, the XGBoost model offers a scientifically validated, data-driven foundation for evidence-based decision-making and strategic foresight in the design and monitoring of circular economy interventions, particularly in transitioning economies and regions seeking to align with EU sustainability benchmarks. Future research should explore ensemble interpretability (e.g., SHAP analysis) and model deployment within real-time decision support systems.

REFERENCES

1. Ahmad, S., Min, Z., & Hussain, M. (2020). Machine learning approaches and their applications in smart sustainable cities. *Sustainable Cities and Society*, 61, 102322. https://doi.org/10.1016/j.scs.2020.102322
2. Bibri, S. E., & Krogstie, J. (2020). The emerging data–driven smart city and its innovative applied solutions for sustainability: The cases of London and Barcelona. *Energy Informatics*, 3(1), 1–42. https://doi.org/10.1186/s42162-020-00117-7
3. Breiman, L. (2001). Random forests. Machine learning, 45, 5-32., https://link.springer.com/article/10.1023/a:1010933404324
4. Chen, C.-C., & Chang, Y.-S. (2024). Leveraging Circular Economy Metrics for Data-Driven Forecasting of Solid Waste Production in Europe. Sustainability, 16(3), 1017. https://doi.org/10.3390/su16031017
5. Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794).

6. European Commission. (2023). *Circular economy action plan: For a cleaner and more competitive Europe*. https://ec.europa.eu/environment/circular-economy/
7. European Environment Agency. (2023). *Municipal waste management in Europe: Towards a circular economy*. https://www.eea.europa.eu/publications/municipal-waste-management-in-europe
8. Geissdoerfer, M., Savaget, P., Bocken, N. M. P., & Hultink, E. J. (2017). The Circular Economy – A new sustainability paradigm? *Journal of Cleaner Production*, 143, 757–768. https://doi.org/10.1016/j.jclepro.2016.12.048
9. Ghosh, D., & Shah, N. (2021). Circular economy indicators: An integrated framework using machine learning. *Resources, Conservation & Recycling*, 174, 105754. https://doi.org/10.1016/j.resconrec.2021.105754
10. Javid, I., Ghazali, R., Batool, T., Jafri, S. I. H., & Altaf, A. (2023, May). An arima and xgboost model utilized for forecasting municipal solid waste generation. In International Conference on Information, Communication and Computing Technology (pp. 16-28). Cham: Springer Nature Switzerland.
11. Kusi-Sarpong, S., Sarkis, J., & Wang, X. (2021). Sustainable supply chain modeling and decision-making with machine learning: A systematic review. *Computers & Industrial Engineering*, 157, 107246. https://doi.org/10.1016/j.cie.2021.107246
12. Popović, G., Todorović, M., & Milijić, N. (2025). Machine learning and circular economy metrics in the Western Balkans. *European Journal of Environmental Policy Studies*, 29(1), 18–37.
13. Primartha, R., & Tama, B. A. (2017, November). Anomaly detection using random forest: A performance revisited. In 2017 International conference on data and software engineering (ICoDSE) (pp. 1-6). IEEE.
14. Sheikh, A., Simske, S. J., & Chong, E. K. P. (2024). Evaluating Artificial Intelligence Models for Resource Allocation in Circular Economy Digital Marketplace. *Sustainability*, *16*(23), 10601. https://doi.org/10.3390/su162310601
15. Zhang, Y., & Wang, S. (2022). Application of neural networks in economic forecasting. *Journal of Computational Economics*, 18(2), 145–160.