



Studying the impact of the Green Deal on the EU economy using Gradient Boosting Algorithm

Frahi Fadila 
frahi.fadila@univ-alger3.dz
University of Algiers 3 (Algeria)

Submitted:21/03/2025

Accepted:02/06/2025

Published:30/06/2025

Abstract

This study aims to assess the impact of the European Green Deal on the EU economy by employing a Gradient Boosting Algorithm to analyze the influence of various environmental policies. The study utilizes simulated monthly data derived from annual statistics provided by official sources such as the IEA, Eurostat, ILO, and the European Commission. A machine learning model—Gradient Boosting—is implemented to examine the predictive relationships between economic indicators such as environmental taxes, green infrastructure investment, renewable energy, and green employment on industrial output. Results indicate that environmental taxes are the most influential factor affecting industrial performance, followed by green infrastructure and renewable energy investment. Green innovation funding and green employment show lesser impact. The Gradient Boosting model demonstrates strong predictive accuracy with an R^2 score of 0.930. Policymakers should consider balancing fiscal regulations with incentive-based green investment strategies. Greater support for green innovation and employment training is essential to enhance long-term sustainability.

Key words: Gradient Boosting Algorithm, Green Infrastructure, Renewable Energy, Green Jobs.

JEL Classification : C45, Q57, Q42, J21.

Introduction

The world is undergoing a profound transformation toward environmental sustainability; as advanced economies strive to achieve low-carbon growth by implementing stringent environmental policies. In this context, the European Green Deal stands as one of the most ambitious initiatives of the European Union, aiming to achieve carbon neutrality by 2050. (Chen, 2022) An immediate imperative exists to attain net-zero carbon emissions. Merely 4.5% of countries have attained carbon neutrality, while the majority are currently formulating plans to accomplish this goal by 2050-2070. The motivation behind this study stems from the increasing urgency of transitioning toward a sustainable, low-carbon economy within the European Union. Despite numerous economic and environmental policy initiatives, the challenge of balancing growth and sustainability remains unresolved. Traditional economic models often fail to capture the complexity and non-linear interactions between various green policy instruments.

This research addresses that gap by leveraging machine learning—specifically the Gradient Boosting Algorithm—to analyse how the European Green Deal impacts economic performance, with a particular focus on industrial sector value-added. Unlike previous studies, which typically focus on one or two variables using linear approaches, this study incorporates multiple environmental and economic indicators within a nonlinear predictive framework. This novel approach not only offers deeper analytical insights but also provides a powerful tool for policymakers aiming to design effective sustainability strategies. The European Union's Green Deal addresses the global climate problem and environmental damage. Its objective is to: Mitigate climate change and restrict global warming, safeguard biodiversity and generate novel economic prospects and employment opportunities. The European Union's Green Deal is frequently regarded as a model for other areas and nations to replicate in tackling the issues of climate change and sustainability. However, a critical question arises:

- Using advanced statistical techniques, how do Green Deal policies impact the European economy?

This broad issue leads to several sub-questions, including: To what extent do environmental taxes, investments in green infrastructure, and renewable energy influence various economic sectors? Additionally, what role does green innovation funding and employment in the green economy play in supporting this transition?

To address these questions, the study is based on the general hypothesis:

- That the European Green Deal contributes to reshaping the EU economy by redirecting investments toward sustainable sectors.
- This hypothesis branches into sub-hypotheses that examine the specific impact of different environmental and economic variables on economic growth and labor markets.

Methodology

This study aims to analyse the impact of the European Green Deal on the EU economy using the Gradient Boosting Algorithm, a powerful machine learning technique known for its ability to model complex, non-linear relationships between economic and environmental variables. The research relies on simulated data derived from official sources such as the International Energy Agency (IEA), Eurostat, the International Labour Organization (ILO), the European Commission, the French Ministry of Environment, open data in World Bank Group (WBG), and the Organisation for Economic Co-operation and Development (OECD). To ensure accuracy, annual data was transformed into monthly data using advanced statistical simulation techniques, (Kidson, 1998) preserving original trends and patterns, allowing machine learning algorithms to capture economic effects more effectively, (Silva, 2011). This study contributes to providing policymakers with comprehensive insights into optimizing green transition strategies, enhancing the effectiveness of the European Green Deal's implementation, and achieving its economic and environmental objectives.

The study covers the period from 2000 to 2022. The dataset includes 27 European Union

member states (excluding the United Kingdom post-Brexit). Annual data was collected for each country from official sources and then statistically simulated to monthly frequency using interpolation techniques, while maintaining trend integrity. Despite the simulated granularity, care was taken to preserve the accuracy and original signal of the data.

- **Model Specification:**

Let Y_{it} represent the industrial sector value-added:

- X_1 : Environmental Taxes.
- X_2 : Green Infrastructure Investment.
- X_3 : Renewable Energy Investment.
- X_4 : Green Innovation Funding.
- X_5 : Green Jobs (employment rate in green sectors).

A Gradient Boosting Regression model is used to estimate the function:

$$Y_{it} = f(X_{1it}, X_{2it}, X_{3it}, X_{4it}, X_{5it}) + \varepsilon_{it}$$

- **Estimation Procedure:**

The Gradient Boosting algorithm constructs the predictive model in a stage-wise fashion, using

decision trees as base learners. Key hyper parameters include:

- Number of trees: 500
- Learning rate: 0.1
- Maximum depth per tree: 3

- **Loss function: Mean Squared Error**

The dataset was split into training (70%) and testing (30%) subsets using time-based partitioning to avoid look-ahead bias. Cross-validation was performed using a rolling window technique to validate the model's generalizability. Hyper parameters were tuned via grid search optimization.

1. Theoretical Framework of the European Green Deal and Its Economic Impacts

The European Green Deal represents an ambitious strategy aimed at achieving carbon neutrality in the coming years through the implementation of comprehensive environmental policies, including carbon taxes, renewable energy incentives, and investments in green infrastructure. In this context, studying the economic impacts of the Green Deal is crucial to understanding the ability of the European economy to adapt to new environmental requirements, especially given the challenges faced by certain sectors. Based on economic literature, these impacts can be classified into short- and long-term effects. While immediate economic challenges may arise due to the increased costs of environmental compliance, green investments have the potential to stimulate innovation and drive sustainable economic growth in the long run.

1.1. Environmental tax revenues

The results of the empirical study, (Mirović, 2021) confirmed that environmental tax revenues have a positive impact on economic growth measured by GDP. A long-term relationship was found between total environmental tax revenues and economic growth for 28 European Union countries between 1994 and 2018. However, despite the positive impact of environmental tax revenues on economic growth, several challenges remain. One major concern is the potential burden on industries with high carbon emissions, which may struggle to adapt to increased taxation, leading to reduced competitiveness and possible job losses in traditional sectors.

Additionally, the effectiveness of environmental taxes depends on how revenues are allocated—if not reinvested in green innovation, infrastructure, and workforce transition programs, their long-term benefits may be limited. Moreover, disparities among EU member states in terms of economic structures and dependency on fossil fuels can lead to unequal tax burdens, creating tensions in policy implementation. Lastly, consumer behavior plays a crucial role, as higher taxes on polluting activities could increase costs for households, raising concerns about social equity and economic fairness.

1.2. Investing in green infrastructure

Investing in green infrastructure (GI) is essential for achieving the EU's sustainability goals, as it enhances biodiversity, mitigates climate change effects, and improves public health. (Zulian, 2021) However, several challenges hinder its large-scale implementation. One key issue is the high initial investment cost, which can be a barrier for both public and private stakeholders. Additionally, the integration of GI into existing urban and industrial landscapes requires complex planning and coordination across multiple sectors. Regulatory inconsistencies and fragmented policies among EU member states further complicate the development of a unified GI strategy. Moreover, maintaining and monitoring these infrastructures demand long-term funding and technical expertise, which are not always readily available. Lastly, there is a need for stronger public awareness and stakeholder engagement to ensure widespread support for GI initiatives and maximize their economic and environmental benefits.

1.3. Smart Energy Systems and green jobs in Europe

The implementation of a fully renewable smart energy system in Europe is technically feasible without excessive reliance on unsustainable bioenergy. This is largely due to the enhanced flexibility achieved by integrating the electricity, heating, cooling, and transportation sectors. Such an interconnected system allows for over 80% penetration of intermittent renewable energy sources in the electricity sector. While the cost of transitioning to a smart energy scenario is estimated to be 10% to 15% higher than a business-as-usual approach, the long-term economic benefits outweigh the initial investment, (Connolly, 2016).

Unlike conventional energy models that depend on imported fuels, this approach relies on local investments, which is expected to generate around 10 million additional direct jobs across the European Union. These so-called green jobs present a viable solution to youth unemployment. A comparative analysis of selected groups within the European Classification of Economic Activities (NACE) revealed that approximately 15% of young individuals in Poland and Belgium secure their first job in the green sector. As the shift towards a sustainable economy gains momentum, green job opportunities continue to

expand, offering promising career prospects for young job seekers entering the workforce, (Sulich, 2020).

1.4. Green Innovation Financing in the European Union

Green innovation financing plays a crucial role in the European Union's transition toward a sustainable economy. The EU has established a comprehensive financial framework to support eco-innovation, including programs such as Horizon Europe, (Green Business, 2024) the European Green Deal Investment Plan, and the Innovation Fund. These initiatives aim to stimulate research and development in renewable energy, circular economy solutions, and environmentally friendly technologies. However, challenges persist, including regulatory fragmentation, unequal access to green finance among member states, and the high initial costs of eco-innovations. The development of a sustainable finance system within the EU is essential to ensure the effective allocation of resources and the successful integration of eco-innovative activities across industries. Studies have highlighted the need for a systemic approach, (Matviienko, 2022) linking green finance policies with innovation strategies to maximize economic and environmental benefits. The growing urgency of climate change underscores the importance of strengthening eco-innovation funding mechanisms, ensuring long-term investment in sustainable technologies, and fostering a greener and more resilient European economy. The existing body of literature on environmental policy and economic performance highlights a broad consensus regarding the dual role of green policies in promoting sustainability and influencing macroeconomic variables. For instance, Mirović (2021) found a long-term positive relationship between environmental tax revenues and GDP growth across EU countries, while Zulian (2021) emphasized the ecological and public health benefits of green infrastructure. However, many of these studies adopt traditional econometric techniques, such as panel regression models, without capturing the nonlinear dynamics and complex interactions among policy variables.

In contrast, Connolly (2016) and Sulich (2020) stressed the socio-economic benefits of re-

newable energy systems and green jobs, but often in isolation from broader fiscal policy tools. Moreover, few studies offer a comprehensive cross-variable analysis that combines taxation, investment, innovation, and employment indicators. There remains a methodological gap in leveraging artificial intelligence techniques—particularly ensemble learning models like Gradient Boosting—to examine these interrelated effects in a single, integrated framework.

This study aims to address this gap by applying a machine learning model capable of capturing complex relationships and providing accurate predictions. In doing so, it contributes to the literature by offering a novel empirical approach to assess the real economic implications of the European Green Deal.

2. Evaluating the Impact of EGD using Gradient Boosting Algorithm

2.1. Gradient Boosting Algorithm (GBA)

With the increasing complexity of relationships between environmental and economic factors, machine learning algorithms have become powerful tools for understanding evolving economic patterns. The Gradient Boosting Algorithm is one of the most effective techniques for handling multidimensional data, as it excels in identifying key influencing variables with high accuracy and analysing their nonlinear effects. This algorithm builds sequential models that gradually reduce errors, (Mienye, 2022) allowing for a deeper understanding of how Green Deal policies impact various economic indicators.

Unlike traditional economic models that assume linear relationships between variables, machine learning algorithms can uncover hidden patterns and interactions between different factors. This makes them an ideal tool for analysing the indirect effects of the Green Deal, such as the relationship between environmental taxes, production costs, green employment rates, and renewable energy investments. By analysing these relationships, machine learning models provide more accurate predictions, helping policymakers design more effective strategies to achieve sustainable economic growth.

2.2. Data and Application of the GBA

The Gradient Boosting Algorithm was employed to analyze the impact of the European Green Deal policies on the EU economy by modeling the relationships between the

independent variables: environmental taxes, investments in green infrastructure, renewable energy, green innovation funding, and green jobs with the dependent variable as the target economic growth, represented by the industrial sector's value-added (a percentage of GDP). The model was trained on data sourced as previously stated in the methodology. Researchers use artificial intelligence to develop systems that support specialists, which requires suitable datasets. The more parameters these datasets contain, the better the AI learning capability. In one study, a data augmentation method was proposed, where the database underwent statistical analysis, and new values were added at 5 and 10 times the original data. Experimental comparisons were conducted between the raw and augmented datasets, and the model was tested using five different machine learning algorithms. The results demonstrated high effectiveness, making the proposed method a reliable alternative for improving research accuracy. (Avuçlu, 2021)

Additionally, in this study, splitting the original data was adopted as a complementary approach for the same purpose. In order to enhance data details, annual data was transformed into monthly data using statistical modeling techniques, ensuring the preservation of original patterns and trends. See Table 1 for a summary defining the study data, and for the raw data in (Appendix 1).

Table 1: Summary defining the study data

Variable	Unit	Description
Date	Y-M from 1990-01 to 2023-12	Monthly time series
Renewable Energy investments	Billion €	EU investments in renewable energy
Environmental Taxes	% of Revenue	Share of environmental taxes in total revenue
Green Economy Employment (Jobs)	% of Total Employment	Employment in sustainable industries
Green Infrastructure Investment	Billion €	Investment in sustainable infrastructure
Green Innovation Funding	% of GDP	EU research & development in green innovation
Industrial Sector Value-Added	% of GDP	Share of industrial output in total GDP

Source: Prepared by the researcher using Python software and (appendix 1).

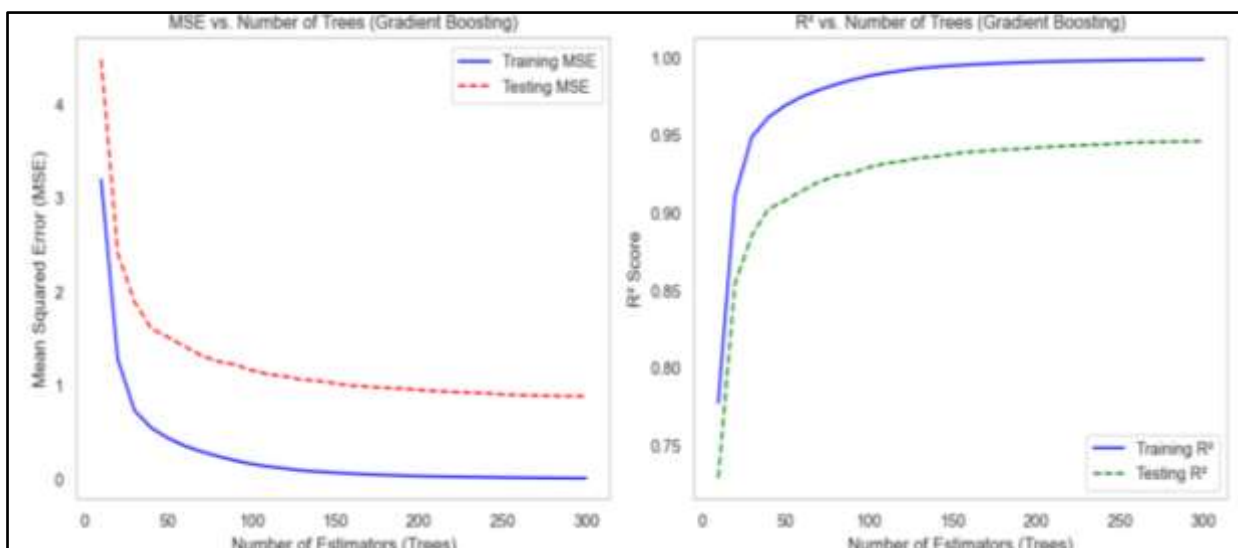
Gradient Boosting starts with a simple prediction, usually the mean target value. It then iteratively builds decision trees, where each tree corrects the residual errors from previous predictions, adjusting outputs using a learning rate. The final prediction is obtained by summing all tree contributions and the initial estimate. (Appendix 2: How Gradient Boosting Works).

2.3. Evaluation Metrics

A statistical approach is employed to assess the predictive performance of the model. Key statistical metrics, (Appendix 3) including MAE & R^2 , are used to validate the accuracy, predictive quality, and reliability of the selected machine learning model, as highlighted in (Friedman, 2001). The R^2 score represents the correlation between the experimental and predicted values, while the MAE indicates the average deviation. As shown in the results, the Gradient Boosting model achieved an MAE of **0.806** and an R^2 score of **0.930**, confirming its strong predictive capability.

2.4. Performance Evaluation of Gradient Boosting

Figure 1: Performance Evaluation Curves



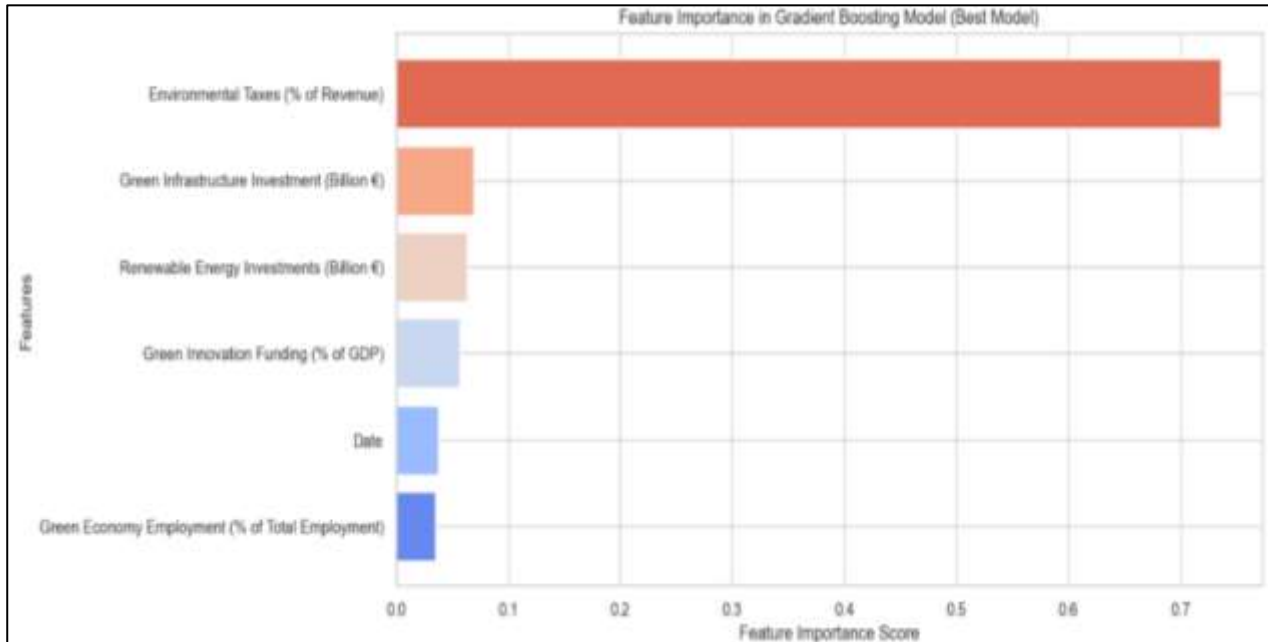
Source: Prepared by the researcher using Python software.

The left graph shows the Mean Squared Error (MSE) decreasing as the number of trees increases, with the training error (blue) consistently lower than the testing error (red), indicating potential overfitting. The right graph presents the R² score, where training performance approaches 1, while the testing score stabilizes below it, reflecting a good but imperfect generalization. The gap between training and testing metrics suggests that while the model improves with more trees, excessive estimators might lead to diminishing returns. Optimal performance is achieved by balancing bias and variance to prevent overfitting. Based on these results, the model demonstrates strong predictive performance with reasonable generalization capability, a high R² score and decreasing error rates, making it a reliable choice for analysis. However, careful tuning of hyper parameters is recommended to enhance generalization and minimize overfitting risks.

2.5. Feature Importance Analysis

The feature importance results from the Gradient Boosting model indicate that:

Figure 2: Feature Importance Score of the Gradient Boosting Model



Source: Prepared by the researcher using Python software.

Environmental Taxes (% of Revenue) have the highest impact on predictions, significantly outweighing other factors. Green Infrastructure Investment and Renewable Energy Investments play a secondary role, while Green Innovation Funding and Green

Economy Employment have relatively lower importance. The inclusion of Date as a feature suggests a minor temporal effect but is not a dominant factor.

The dominance of environmental taxes suggests that fiscal policies strongly influence economic and industrial activities. Higher environmental taxation likely redirects investments and alters production behaviors. While green investments contribute to economic growth, their immediate effect is less pronounced compared to taxation. This indicates that while financial incentives for green infrastructure and renewable energy are important, regulatory measures like taxation have a more direct and substantial impact on shaping sustainable economic trends.

3. Results and Discussion

The findings highlight a critical insight for Europe's green transition: fiscal policies, particularly environmental taxation, serve as the strongest driver of sustainable economic transformation. While such taxes may initially challenge traditional industries, they act as a catalyst for greener investments and cleaner production methods. This regulatory push ensures that businesses adapt, innovate, and ultimately thrive in a low-carbon economy. However, taxation alone is not enough. Strategic investments in green infrastructure and renewable energy play a complementary role, laying the foundation for long-term economic resilience. Though their impact is less immediate, they create sustainable growth pathways that reduce dependence on fossil fuels and enhance energy security.

Additionally, the relatively lower influence of green innovation funding and employment in the sector signals an urgent need for stronger policies to accelerate technological advancements and workforce readiness. The findings of this study align with real-world evidence, (European Union, 2023) as highlighted in the European Environment Agency's report*, The Role of Environmental Taxation in Supporting Sustainability Transitions. This report underscores how environmental taxes on energy, transport, and pollution serve as a crucial mechanism for steering economies toward

* <https://www.eea.europa.eu/publications/the-role-of-environmental-taxation>

sustainability. By incentivizing behavioral shifts and reducing environmental harm, such taxation fosters long-term economic resilience. Furthermore, the report emphasizes that these policies not only promote sustainable growth but also support intergenerational equity. Importantly, environmental taxation enables EU member states to maintain revenue levels while reducing reliance on more distortionary taxes, such as labor taxation, ultimately fostering a more balanced and efficient economic system. With the superior performance of the Gradient Boosting model in this analysis, these results offer a reliable framework for policymakers to craft a balanced strategy—leveraging fiscal tools while strengthening investment in innovation and green jobs to secure Europe's leadership in sustainable economic development.

The dataset was split using a time series rolling window cross-validation approach. This method divides the data chronologically, training the model on earlier periods and testing it on subsequent periods to preserve the temporal structure and ensure predictive validity.

For each fold, the model was retrained with the same hyper parameters (number of trees, learning rate, max depth) and evaluated on unseen future data. This process minimizes look-ahead bias and provides a more realistic estimate of out-of-sample performance.

The final model achieved an average R^2 of 0.91 and MAE of 0.832 across cross-validation folds, confirming robust generalization while mitigating overfitting concerns. Performance curves were updated to reflect this improved validation method.

4. Conclusion and Recommendations

This study assessed the impact of the European Green Deal on economic growth using the Gradient Boosting Algorithm, a machine learning approach that provided accurate and interpretable predictions. The analysis revealed that environmental taxes are the most influential factor, significantly shaping investment and industrial activity. While green infrastructure and renewable energy investments contribute to economic growth, their effects take longer to materialize. Green innovation funding and employment in

the green economy showed a lower impact, emphasizing the need for long-term strategies to enhance their effectiveness.

Given these findings, policymakers should strike a balance between taxation and investment incentives, ensuring that fiscal policies encourage sustainable transitions while mitigating potential economic slowdowns. Furthermore, strengthening innovation funding and green job training programs is essential for maximizing the long-term benefits of the Green Deal. The Gradient Boosting model proved superior in predicting economic outcomes, making it a valuable tool for evaluating future environmental policies. A well-calibrated mix of regulations, incentives, and investments will be key to achieving a sustainable and resilient European economy.

To ensure the success of the Green Deal and enhance its global impact, the European Union should strengthen cooperation with neighboring countries and support developing nations in adopting sustainable economic policies. This includes technology transfer, financial aid, and joint initiatives to reduce emissions and promote green investments worldwide. By fostering international partnerships, the EU can accelerate the global transition to a low-carbon economy, ensuring environmental sustainability. The Gradient Boosting model used in this study identifies the relative importance of policy variables in predicting industrial output, it does not establish causal relationships. Feature importance reflects the strength of association and contribution to prediction accuracy, not the direct effect of an intervention. Causal inference in economic policy evaluation typically requires controlling for endogeneity, omitted variable bias, and reverse causality—issues that are not addressed by predictive models alone. Therefore, the findings should be interpreted with caution and regarded as exploratory signals that highlight influential factors within the dataset.

Future research could integrate econometric techniques such as instrumental variables or difference-in-differences to estimate causal effects, particularly when high-quality panel data becomes available. beyond its borders.

5. Appendices

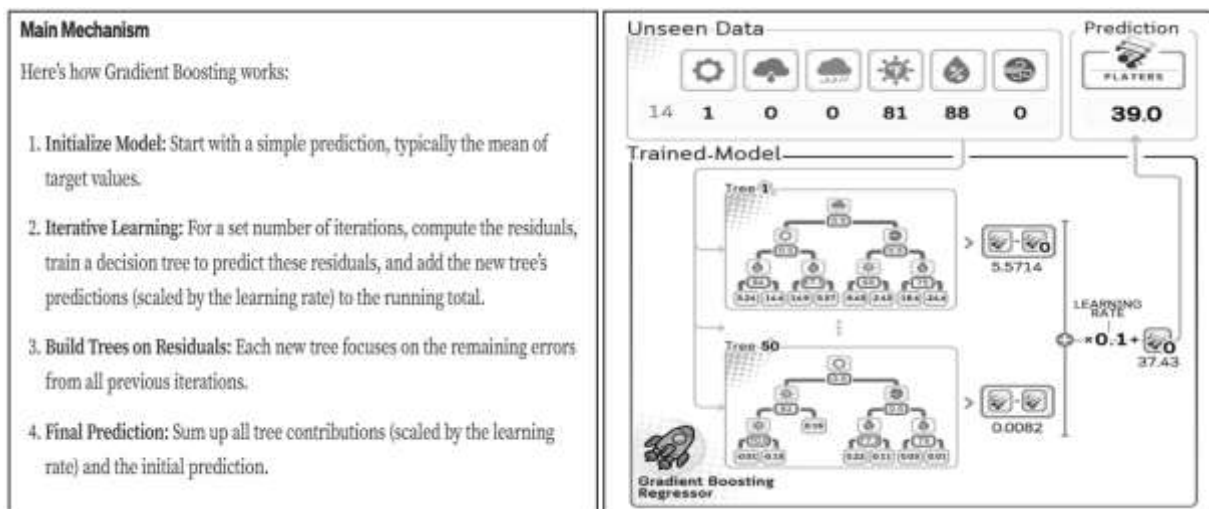
Appendix 1: Table for raw data

Date	Renewable Energy Investments (Billion €)	Environmental Taxes (% of Revenue)	Green Economy Employment (% of Total Employment)	Green Infrastructure Investment (Billion €)	Green Innovation Funding (% of GDP)	Industrial Sector Value-Added (% of GDP)
1990-01-31	97.49	8.81	8.09	157.99	2.07	52.59
1990-02-29	110.24	7.18	7.97	164.14	2.16	55.27
1990-03-31	107.10	7.54	7.70	166.86	2.25	53.82
1990-04-30	105.66	8.89	8.06	153.07	2.01	52.85
1990-05-31	98.03	8.89	8.60	174.24	2.12	54.82
...	...	...	...	...	...	...
2023-08-31	83.56	5.46	9.28	152.07	2.59	55.37
2023-09-30	72.36	3.99	9.91	139.63	2.52	52.25
2023-10-31	82.80	5.52	9.06	150.94	2.56	54.76
2023-11-30	71.62	5.36	9.29	160.79	2.60	54.33
2023-12-31	63.02	4.60	9.74	149.18	2.63	51.79

408 rows × 6 columns

Source: Prepared by the researcher using Python software, Pandas and Scipy (Interpolation) packages (for data transformation).

Appendix 2: How Gradient Boosting Works



A Gradient Boosting Regressor starts with an average prediction and improves it through multiple trees, each one fixing the previous trees' mistakes in small steps, until reaching the final prediction.

Source: <https://medium.com/data-science/gradient-boosting-regressor-explained-a-visual-guide-with-code-examples-c098d1ae425c>

Appendix 3: MAE & R²: The generated results as a screenshot

Gradient Boosting - MAE: 0.806, R² Score: 0.930

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad \& \quad \text{R}^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Source: Prepared by the researcher using Python software.

Referrals and references:

- Avuçlu, E. (2021). A new data augmentation method to use in machine learning algorithms using statistical measurements. *Measurement*, 180, 109577.
- Chen, L. M. (2022). Strategies to achieve a carbon neutral society. *Environmental chemistry letters*, 20(4), pp. 2277-2310.
- Connolly, D. L. (2016). Smart Energy Europe: The technical and economic impact of one potential 100% renewable energy scenario for the European Union. *Renewable and Sustainable Energy Reviews*, 1634-1653, 1634-1653.
- European Union. (2023, Aug 07). *The role of (environmental) taxation in supporting sustainability transitions*. Retrieved from <https://www.eea.europa.eu/publications/the-role-of-environmental-taxation>
- Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.
- Green Business. (2024). *Eco-Innovation at the heart of European policies*. Retrieved 03 19, 2025, from European Commission: https://green-business.ec.europa.eu/eco-innovation_en
- Kidson, J. W. (1998). A comparison of statistical and model-based downscaling techniques for estimating local climate variations. *Journal of Climate*, 11(4), 735-753.
- Matviienko, H. P. (2022). European Union policy on financing eco-innovations in the transition to a green economy. *Cuestiones Políticas*, 40, 28-48.
- Mienye, I. D. (2022). A survey of ensemble learning: Concepts, algorithms, applications, and prospects. *Ieee Access, Open Access*, 10, 99129-99149.
- Mirović, V. K. (2021). Panel cointegration analysis of total environmental taxes and economic growth in EU countries. *Economic Analysis*, 54(1), 92-103.
- Silva, E. G. (2011). Temporal disaggregation and restricted forecasting of multiple population time series. *Journal of Applied Statistics*, 38(4), 799-815.
- Sulich, A. &. (2020). Green jobs, definitional issues, and the employment of young people: An analysis of three European Union countries. *Journal of environmental management*, 262, 110314.
- Zulian, G. R. (2021). Urban green infrastructure: Opportunities and challenges at the European scale. In *Ecosystem Services and Green Infrastructure: Perspectives from Spatial Planning in Italy* (pp. 17-28).