sustainability

MDPI

Article

Analysis of Renewable Energy Policies through Decision Trees

Dania Ortiz [1,*], Vera Migueis [2], Vitor Leal [3], Janelle Knox-Hayes [4] and Jungwoo Chun [5]

1 MIT Portugal Program, Faculty of Engineering, University of Porto (FEUP), 4200-465 Porto, Portugal
2 Institute for Systems and Computer Engineering, Technology and Science (INESCTEC), 4200-465 Porto, Portugal; vera.migueis@inesctec.pt
3 Department of Mechanical Engineering, Faculty of Engineering, University of Porto (FEUP), 4200-465 Porto, Portugal; vleal@fe.up.pt
4 Department of Urban Studies and Planning, Massachusetts Institute of Technology, 77 Massachusetts Avenue 9-424, Cambridge, MA 02139, USA; jankh@mit.edu
5 Department of Urban Studies and Planning, Massachusetts Institute of Technology, 105 Massachusetts Avenue 9-366, Cambridge, MA 02139, USA; jwchun@mit.edu
* Correspondence: up201608014@fe.up.pt

Abstract: This paper presents an alternative way of making predictions on the effectiveness and efficacy of Renewable Energy (RE) policies using Decision Trees (DT). As a data-driven process for decision-making, the analysis uses the Renewable Energy (RE) target achievement, predicting whether or not a RE target will likely be achieved (efficacy) and to what degree (effectiveness), depending on the different criteria, including geographical context, characterizing concerns, and policy characteristics. The results suggest different criteria that could help policymakers in designing policies with a higher propensity to achieve the desired goal. Using this tool, the policy decision-makers can better test/predict whether the target will be achieved and to what degree. The novelty in the present paper is the application of Machine Learning methods (through the Decision Trees) for energy policy analysis. Machine learning methodologies present an alternative way to pilot RE policies before spending lots of time, money, and other resources. We also find that using Machine Learning techniques underscores the importance of data availability. A general summary for policymakers has been included.

Keywords: energy policy; policy effectiveness; policy efficacy; decision trees; machine learning

Citation: Ortiz, D.; Migueis, V.; Leal, V.; Knox-Hayes, J.; Chun, J. Analysis of Renewable Energy Policies through Decision Trees. *Sustainability* **2022**, *14*, 7720. https://doi.org/10.3390/su14137720

Academic Editor: Valentina Palermo

Received: 3 May 2022
Accepted: 21 June 2022
Published: 24 June 2022

1. Introduction

The main motivation of this paper is to provide an alternative to elaborate predictions on the effectiveness and efficacy of Renewable Energy (RE) policies. Critical situations highlight the motivation of this analysis, like the non-achievement of the 2020 RE targets by the countries (e.g., France) [1]. RE policies were selected to be analyzed for their relevance towards energy efficiency improvement and direct impact on GHG emissions, but also due to their data availability for the conformation of the database.

The main contribution of this proposal is the provision of a data-driven solution that may support policymakers in designing policies with a higher propensity to achieve the desired goal (as when designing the policies, decision-makers can test/predict whether the target will be achieved or not or in what degree).

Machine Learning (ML) has been selected for the analysis. It includes different techniques, such as Support Vector Machine (SVM), Artificial Neural Networks (ANN), Random Forest (RF), or Decision Trees (DT) [2]. Regarding the application of these techniques across the energy field, there are two main ideas to be highlighted. First, ML has been mainly used to analyze patterns of energy features like energy consumption and prices, using the results to design alternatives and solutions that better respond to specified conditions (like the weather or the integration of RE resources) [3]. Second, ML techniques are generating new opportunities for innovative research in energy, especially connected with topics like economics and finance [4]. These include different applications like risk management, trading strategies, energy prices, demand forecasting, and analyzing macro/energy trends, among others.

A review of recent literature examining Machine Learning in the energy field (Appendix A) reveals the growing use of ML techniques, especially for the short-term predictions in the energy industry. The studies considered mainly use neural networks, support vector regression techniques, or Decision Trees, although other methodologies are being developed and applied.

The literature recognizes the potential of applying the ML methods to the tasks of predicting elements in different stages of the energy provision process, from utility scheduling [5] to Demand-Side-Management [6]. To our knowledge, this is one of the few analyses done for energy policy analysis.

Literature research suggests that there is a growing interest in applying Decision Trees for finding solutions to short-term utilizations, such as security dispatches [7], as there is for long-term ones as planning of energy storage systems [6]. The results indicate that the application of Decision Trees minimizes the risk of contingencies or structural problems, which can be translated into the potential for energy and economic savings [7,8].

Decision Trees (DT) have been successfully used in different energy systems applications. DT has an easy reasoning and reading process through what is known as induction rules [9]. The induction rules prevent the existence of a total black box that characterizes other types of techniques in which the internal logic is not that easy to be understood [10]. The rules have served as the base for developing real scheduling for optimizing electricity generation and gas production [5].

Problem statements at the strategical, tactical, and operational levels can be observed across the DT literature, for which different target users of the models have been observed, including energy aggregators, planners in utilities, or energy planners.

Some of the benefits of the application of Decision Trees may be obvious such as their ease of problem structuring and interpretability. However, the diversity of the challenges associated with their application (and of other ML techniques) in policymaking include the possible lack of understanding of a high share of public policy makers around their applicability to solving prediction and the lack of data availability [11]. In addition, poor communication between academics and researchers, and public policy decision-makers may raise a communication barrier [12].

Nevertheless, the potential benefits of the application of Decision Trees support the development of this analysis. Overall, this research brings up an easy-to-implement and interpret modeling technique application while simultaneously contributing to filling the gap in the literature on machine learning techniques for energy policy analysis.

This work contributes to proposing a novel orientation mechanism for policy analysis through Decision Trees. Predictive analytics has been applied to provide the alternative of anticipating RE share target achievement and its degree of fulfillment.

Such techniques support scientific research incorporation into the country's policy design, locating scientific progress as an essential key for development [13]; with this in mind, the conclusions of this work are presented as a summary for policymakers. The summary aims to leverage the key highlights of this work, serving as a reference point for evidence-based energy policy making.

2. Materials and Methods

Building a DT is mainly divided into two stages: (1) tree building and (2) testing [5]. In the building stage, the training data is partitioned using testing conditions. The first node (at the top of the tree) acts as the root, specifying the testing condition for the following branches, known as internal nodes. Each internal node can split into two or more subspaces [10]. The terminal nodes (located in an end node) assign the classifications and are also known as leaf nodes [14].

In the testing stage, the performance of the model is tested [15]. Different methods can be used for testing purposes (further explained in the validation section of this paper). An iterative process of identifying the best performing tree aims to increase the performance of the Decision Tree.

2.1. Algorithm

For DT, different types of algorithms are available; some of the most frequent ones include CART, C4.5, CHAID, and QUEST YAN YAN [16]. They all share the task of predicting the value of a dependent variable from a set of independent variables [3]. The criteria to decide the splits depends on the algorithm [10]. Each algorithm leads to a different way of building up the tree under three main elements [14]:

- Splitting decision: Selection of the independent variable with the strongest interaction with the dependent variable in every step of the tree.
- Stopping decision: The splitting process is repeated for all input variables until the defined tree levels are achieved.
- Assignment decision: The class assigned to each terminal node.

For this study, the CHAID (Chi-squared Automatic Interaction Detection) algorithm was selected. This algorithm selects the variables for partitioning using Chi-square tests.

2.2. Data

The final dataset used to construct the DT was self-elaborated. The dataset is composed of a sample of 292 observations. Each one reflects a set of RE policies (policy mixes) assessed in a specific year and country, including their respective dependent variable and independent variables. The variables composing the dataset have different origins. A variables directory is presented in Appendix B (Tables A1–A9).

All the observations used for this analysis refer to the years 2004 to 2018. A total of 89 policy programs were initially inputted. These policies were extracted from the MURE-ODYSEE database [17]. When considering each policy program by its year and country of implementation, 889 policy programs are presented. However, after integrating them as policy mixes per year per country, the final data set consisted of 292 observations. This conglomeration of the policy programs, i.e., policy mixes (done by grouping sets of policies by year and place of implementation), can assess the RE target share achievement (dependent variable). An individual assessment per policy program was not possible to be carried out; as a result (dependent variable) that can be directly assessed by a single policy was not identified in the literature.

An example of the RE policy programs integrated to generate the policy mixes used for the database is provided in Table 1.

Table 1. Example of the integration of Policy mixes.

Policy Mix No.	Year of Evaluation	Individual Policy	Policy Instruments *	Policy Subtypes
1	2004	Promotion of Electricity from Renewable Sources	Information/Education	Information campaigns
2	2005	Promotion of Electricity from Renewable Sources	Information/Education	Information campaigns
		Mobility management consulting and funding program	Financial	Grants/Subsidies for investments

* The policy instruments and the subtypes can be reviewed in Appendix B (Table A10).

Existing online policy databases were searched. The MURE-ODYSEE database was identified. It gathers policy programs under different energy topics, including RE policies. The policies included in this database represent policies of the European Union and Schengen space countries.

Other general data availability challenges were identified:

- Fragmented data: In the policymaking field, the availability of data can regard the disposal of information from different sources and in diverse formats [18].
- Unreachable data: Even if national or international efforts are being taken to organize data required for policy evaluation or other analysis, the reachability of the info may be stressed by the open access to the data or its availability in only the mother tongue languages.
- Disparities in data: The different purposes behind the data organization and sources can result in disparity of information and their formats from source to source [3].
- Basic or non-existence information registration: In some cases, it can be noted that there are no existing records on policy data or other cross-cutting topics that regard the policy evaluation process [19].

To address these general challenges, we encourage the governments to share data with relevant stakeholders; adopt a data openness approach around policies [17]; create the infrastructure required to accumulate and generate reliable data [18]. In addition, ML techniques can only be successfully employed with the recognition that efforts are required from the governments to systematically assess the policy outcomes [20].

2.3. The Variables

- Dependent Variable

Effectiveness has been selected as the core of the assessment of policy performance and it is defined as "the degree to which something is successful in producing a desired result" [21]. The proxy measure of success or efficacy of RE policies is the share of RE in the total energy consumption. The dependent variable of the effectiveness model is the degree of Renewable Energy share achieved by country by year [22] (second column of Table 2).

Table 2. Estimation of the average intended growth by year.

Country	Initial RE Share Values in 2004	RE Target Until 2020	Intended Growth	Average Intended Growth by Year (Threshold) *
EU	8.5	20	11.5	0.71875
Belgium	1.9	13	11.1	0.69375

* Average Intended growth by year (threshold) = (RE Target until 2020 − Initial RE share values in 2004)/number of years in the period of analysis.

The EU defined a target of 20% of RE share from the total energy consumption by 2020. In addition, member countries established binding national targets for raising their shares under the Renewable Energy Directive [23]. The targets vary significantly by country, reflecting the different stages each country is at for RE production and their resource availability to expand the usage of RE sources. Contrasting examples range from 10% in Malta to 49% in Sweden [24].

An alternative dependent variable has been considered. The focus on effectiveness (the degree of achievement in producing the desired result) has been changed to efficacy (whether or not the intended result is produced [25]). As a categorical variable, it is the product of the binarization of the achievement of the RE share target (effectiveness model dependent variable). The detailed procedure is described as follows:

- Step 1: The self-defined "RE Target until 2020" of every country [23] and the difference from the starting value "Initial RE share values in 2004" was used to calculate their expected increase: "Intended growth". Afterward, the "Intended growth" was divided by the number of years in the period of analysis (16 years, from 2004 to 2020), producing a value under the name of "average intended growth by year" (being the reference threshold). Examples of the calculations are introduced in Table 2.

- Step 2: The "Average Intended growth by year" (calculated threshold in Table 2) is compared to the RE share achieved each year (achieved difference in RE share). To do so, the "Achieved Difference in RE share" is calculated by subtracting the "Achieved RE share" value of the consecutive year from the year of evaluation (Consecutive year value—Year of evaluation value). An example of this estimation is presented in Table 3.
- Step 3: For the dependent variable binarization (under the name of "Efficacy by country"), if the "Achieved Difference in RE share" (step 2) is equal or lower than the "Average Intended growth by year" (threshold Table 3), the value of 0 (non-efficacious) is assigned. If the "Achieved Difference in RE share" is higher than the "Average Intended growth by year", the variable takes the value of 1 (efficacious). An example of this estimation is presented in Table 4 and a representation by country for the year 2010 can be seen in Figure 1.

Table 3. Estimation of the intended growth by year.

Year of Evaluation	Achieved RE Share	Achieved Difference in RE Share
2005	8.5	0.600 *
2006	9.1	

* Achieved Difference in RE share of 2005 = (RE share by year 2006) − (RE share by year 2005).

Table 4. Example of the estimation of the category of efficacy by country.

Year of Evaluation	RE Share by Year	Achieved Difference in RE Share	Average Intended Growth by Year	Efficacy by Country
2005	8.5	0.600	0.71875	0 Non-efficacious
2006	9.107	0.578	0.71875	0 Non-efficacious
2007	9.685	0.931	0.71875	1 Efficacious
2008	10.616	0.752	0.71875	1 Efficacious
2009	11.368	1.254	0.71875	1 Efficacious

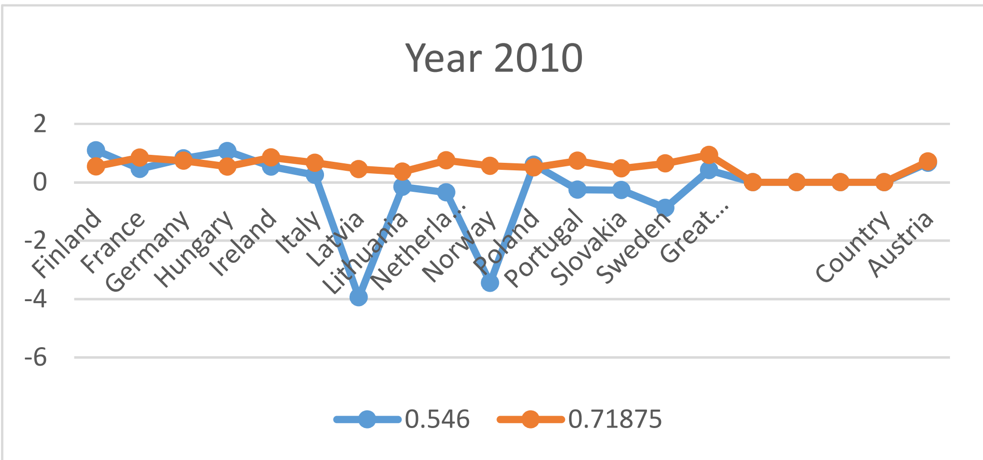


Figure 1. Comparison of the Achieved growth and Intended Growth for the year 2010.

- Independent Variables

As part of the data organization, it is important to define the attributes that will be tested in the model for prediction purposes [10]. Attributes that characterize energy policy effectiveness from different angles were identified. Through the incorporation of the effectiveness concerns, a multidisciplinary assessment has been done. The aim is to enrich the proposed data-driven analysis and perform an integral analysis.

The independent variables have been organized into three groups. A complete description of each variable is presented in Appendix B (Tables A1–A9), together with the possible values that each variable can take, their units, etc.

Country Characteristics include variables that regard sociodemographic and other contextual conditions of the country. To enable the model to be used for future prediction purposes, all the independent variables of this group report the data from one year before the evaluation year (year-1 format). The variables of this kind are introduced in Table 5 below.

Table 5. Country Contextual Characteristics variables.

Variables	
Country location	Population
EU member	Human Development Index (HDI)
Member of the Eurozone	Harmonised Index of Consumer Prices (HICP)
Population	Total people at risk of poverty or social exclusion by age and sex
Long-term interest rate	Gini coefficient
Cooling degree days	Average household size
Heating degree days	Air pollutants by source sector
Final energy consumption	Pollution, grime, or other environmental problems
Total energy supply by product	Exposure to air pollution
Renewables and biofuels	Environmental protection expenditure
Oil and petroleum products	Total general government expenditure
Non-renewable waste	Number of national civil servants in central public administration
Nuclear heat	Feed-in Tariff (FiT)-Small Hydro
Electricity	FiT-Biomass
Solid fossil fuels	FiT-Waste
Final energy supply	FiT-Waste
Expenditure on research and development	FiT-Geothermal
Gross Domestic Product (GDP) per capita	FiT-Marine

1. Policy Characteristics include the variables that are self-calculated by using other database variables. These variables provide radiography of the policy mix at the moment of performing the analysis. For example, this group includes the number of policies in the sample and the number of years of operation. This group also includes the amount, types, and subtype of instruments included in the sample. The different types and subtypes of policies that a policy mix can include are detailed in Appendix B (Table A10). The variables under this group are introduced in Table 6 below.
2. Policy Effectiveness Concerns, these variables are focused on assessing policy effectiveness. The literature on energy policy effectiveness is extensive [26]. These variables have been identified based on the systematic review of energy policy effectiveness concerns by Ortiz and Leal. We follow the concerns identified and categorized by Ortiz and Leal. The categorization is based on eight core groups. With further analysis of their database, additional concerns and metrics are identified. All the variables in this category have a year-1 format. An important observation is that from the eight condensed groups of concerns, it is only possible to find variables included in seven

groups; no data for the employment subcategory are identified. The variables under this group are introduced in Table 7 below.

Table 6. Policy Characteristics variables.

Variables
Other related policies
Average duration of the programs
Average years of activity
Median years of activity
Number of RE programs
Policy Type
Policy subtype

Table 7. Policy Effectiveness Concerns variables.

Category	Subcategory	Variable
Economic	1. Affordability	Electricity price for households
		Electricity prices for non-household consumers
		Gas Price for households
		Gas prices for non-household consumers
	2. Economic Competitiveness	Market share of the biggest competitor
		Energy productivity
		Energy intensity of the economy
		Energy intensity of the population
	3. Accessibility	Energy balance
		Energy import dependency
Environmental	4. Impact on climate-change	Total Greenhouse gas emissions
		Greenhouse gas emissions intensity of energy consumption
	5. Impact on the environment	Increase in renewable electricity capacity
		Energy efficiency
Social	6. Health	Population unable to keep home adequately warm by poverty status
Institutional	7. Governance effectiveness and efficacy	Expenditure on energy and fuels expenditure
		Environmental tax revenues (energy tax)

3. Results

The results of the two models (with the two dependent variables) are now presented, including the tree diagram, the Decision Tree results, and the classification rules.

The results for the continuous dependent variable "RE share" are presented first as the variable used for analyzing energy policy effectiveness and, therefore, the degree of RE share achievement. Secondly, the categorical dependent variable "Efficacious by country" results are presented.

3.1. Effectiveness Analysis

For the continuous variable used for the Energy Policy effectiveness model, the predicting performance has been estimated using the Coefficient of Determination or R-squared. It is estimated via the proportion of the variation in the dependent variable that is predictable from the independent variables [27]. It has been calculated using the formula:

$$R^2 = 1 - \frac{SSres}{SStot} = 1 - \frac{\sum i(Yi - Y\hat{\imath})^2}{\sum i(Yi - \gamma)^2}$$

SSres = Residual sum of squared errors; Y actual values − Y predicted values
SStot = Total sum of squared errors; Y actual values − the mean value
Yi = actual y values
Yî = predicted y values
y = baseline model the mean

The validation procedure adopted was the ten cross-fold validation. This means that the entire data was randomly partitioned into ten parts, nine parts were used for training the model, and the remaining part was used for testing. This process is iterated ten times.

The R-squared compares the predicted values and the observed ones. As seen in the calculation done below, with an R-squared value of 0.917, and as seen in Figure 2, the EP effectiveness model estimates 91 percent of the variation in the dependent variable based on the variation in the independent variables.

$$R^2 = 1 - \frac{5096.5}{61529} = 1 - 0.82 = 0.917$$

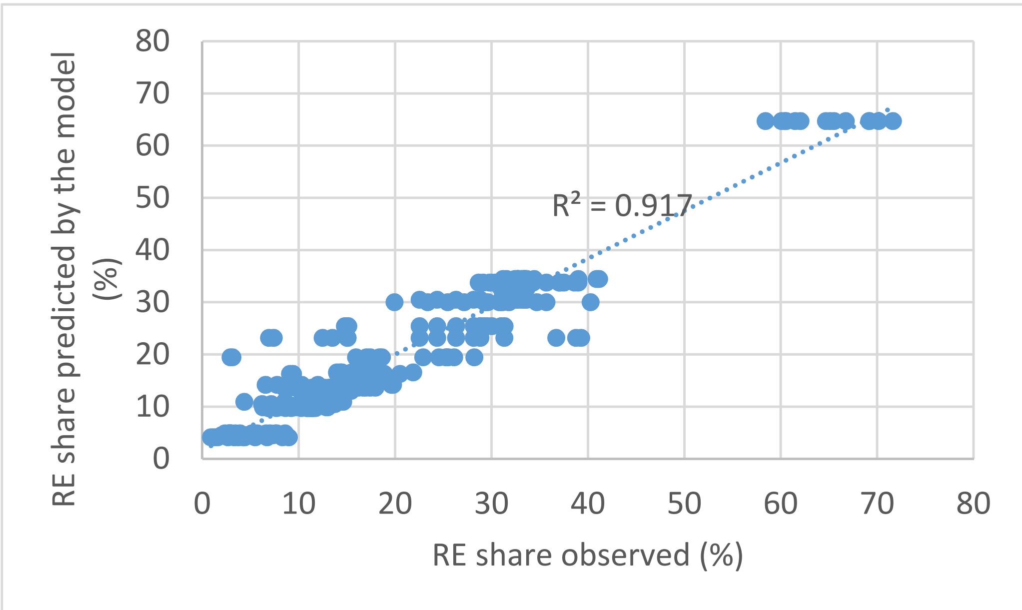


Figure 2. EP effectiveness model: predicted vs. observed.

3.1.1. Tree Diagram

The tree diagram is the graphic representation of the model. The technique for determining the optimal pathway in a Decision Tree is known as the "rollback method" [28]. To apply this method, the tree is analyzed from the bottom to the top (or right to the left in this case) by considering the later decisions first. The Decision Tree is presented in Figure 3.

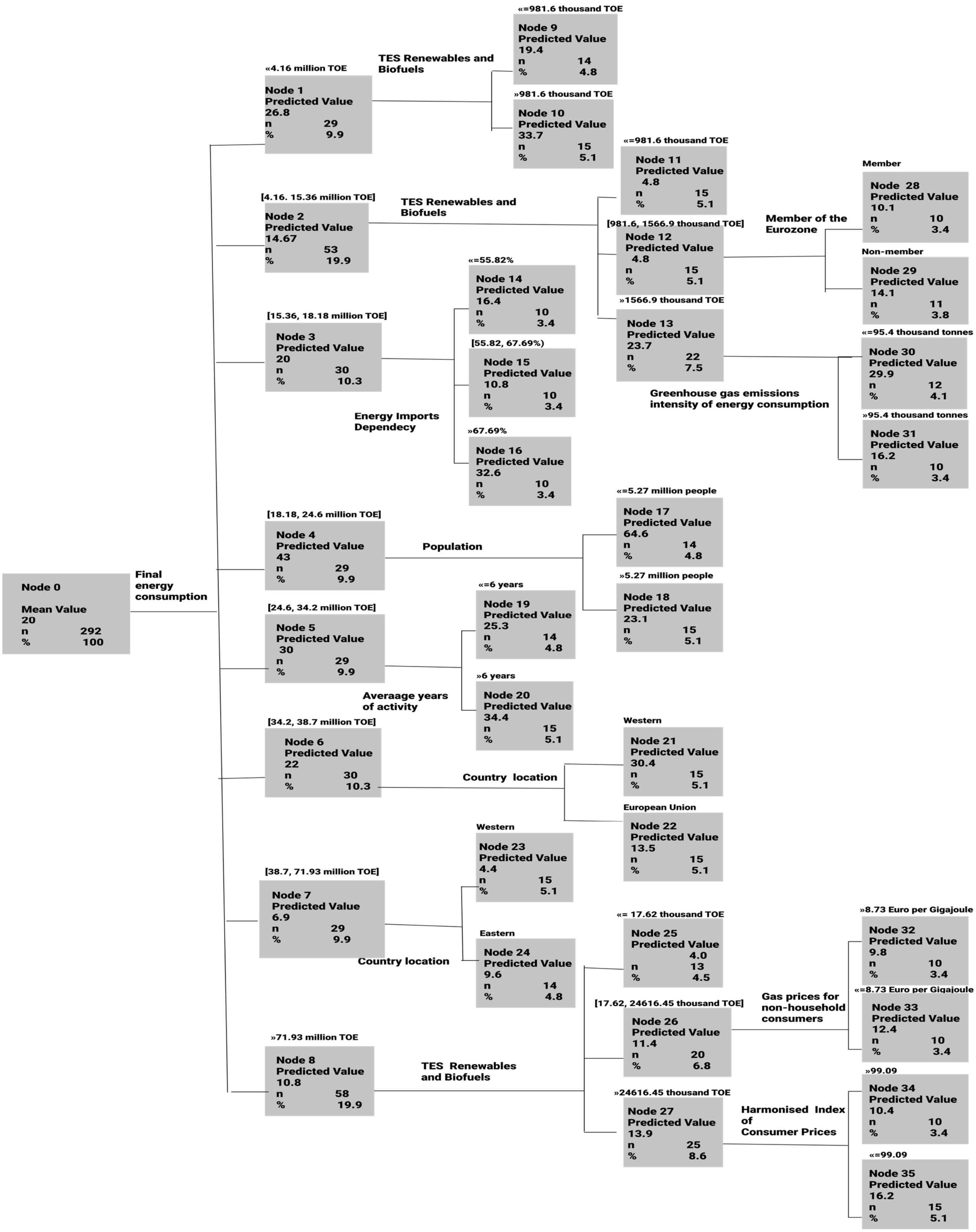


Figure 3. Decision Tree for energy policy effectiveness analysis.

From the Decision Tree above, we find that:

- *Final energy consumption* is the best discriminator of *Energy policy effectiveness*.
- The subsequent branch on the lower side of the DT selected the *Total energy supply by renewables and biofuels* as the second-best discriminator. The policy mixes included in the three consecutive nodes present higher mean values, as the value of the *total energy supply by renewables and biofuels* increases (with a 17,625.4 thousand TOE, the mean value of RE share is 4%, with TOE between 17,625.4 and 24,616.4 the mean value of the RE share is 11.4%, with TOE higher than 24,616.4 the mean value of the RE share is 13.9%). Two child nodes are derived from this branch (child nodes), for which the *Gas prices for non-household consumers* and the *Harmonised Index of Consumer Prices*, respectively, are the best next discriminators.
- For the next five branches above, a second-best predictor has been identified: *Final energy consumption*. If the *Final energy consumption* is between 38.7 and 71.9 million TOE, and the country is located in the Western area of Europe, the predicted mean value for the RE Target to be achieved is 4.4%, while if located in the Eastern the mean value is 9.6%.
- Similar to the previous observation, if the *Final energy consumption* is between 34.2 and 38.7 million TOEs, and the country is located in the Western part of Europe, the predicted mean value for the RE Target to be achieved is 30.4%, compared to the mean value of the European Union which is 13.5%.
- If the *Final energy consumption* is between 24.6 and 34.2 million TOE, and the *Average years of activity* of the policy mix is higher than 6 years, the mean value predicted to be achieved is 34.4%, compared to a 25.3% if lower than 6 years.
- If the *Final energy consumption* is between 18.1 and 24.6 million TOE, and the *Population* is higher than 5.2 million, the RE target mean value predicted to be achieved is 23.1%, compared to 64.6% if lower than the 5.2 million in the population.
- If the *Final energy consumption* is between 15.3 and 18.1 million TOE, and the *Energy import dependency* is lower than 55.8%, in between 55.8% and 67.6%, or higher than 67.6%, the mean values of the predicted RE target are respectively 16.4%, 10.8%, and 32.6%. As import dependency increases in countries with a final energy consumption between 15.3 and 18.1 million TOE, the share of renewable energy also increases.
- Whenever the *Final energy consumption* is in between 4.1 and 15.3 million TOE, the next best predictor is *the Total Energy Supply with Renewables and Biofuels*. At this level, an increase in RE share mean value is observed as the variable increases (from 981.6 to 1566.9 thousand TOE). *A* third level can be observed, in which if the country is a member of the Eurozone, the predicted RE target is 10.1% compared to 14.1% if not. If the *GHG emissions intensity of energy consumption* is lower than 95.4%, the predicted RE target value is 29.9%, falling to 16.2% if higher than 95.4%.
- Finally, the second predictor on the upper branch of the tree is *the Total Energy Supply with Renewables and Biofuels*. If its value is lower than 981.6 thousand TOE, the predicted RE Target to be achieved is 19.4%, compared to 33.7% if the predictor is higher than 981.6 thousand TOE.

The DT data has been organized in a table (Table 8) presenting the details for each node. It includes the number of observations (N) included in each node and the percentage of observations that it represents. The table also presents the predicted value of each node, being the policy mix prediction values. The table also includes the parent node, as the immediate ascendant node and the Primary Independent Variable, as the most influential attribute in determining the tree partition. The terminal nodes of the model have been highlighted in Table 8.

Table 8. Decision Tree results in a table format.

Node	N	Percent	Predicted Mean	Parent Node	Primary Independent Variable
0	292	100.0%	20.0		
1	29	9.9%	26.8	0	Final energy consumption
2	58	19.9%	14.7	0	Final energy consumption
3	30	10.3%	20.0	0	Final energy consumption
4	29	9.9%	43.2	0	Final energy consumption
5	29	9.9%	30.0	0	Final energy consumption
6	30	10.3%	22.0	0	Final energy consumption
7	29	9.9%	7.0	0	Final energy consumption
8	58	19.9%	10.8	0	Final energy consumption
9	**14**	**4.8%**	**19.4**	**1**	**TES renewables and biofuels**
10	**15**	**5.1%**	**33.7**	**1**	**TES renewables and biofuels**
11	**15**	**5.1%**	**4.8**	**2**	**TES renewables and biofuels**
12	21	7.2%	12.2	2	TES renewables and biofuels
13	22	7.5%	23.7	2	TES renewables and biofuels
14	**10**	**3.4%**	**16.5**	**3**	**Energy import dependency**
15	**10**	**3.4%**	**10.9**	**3**	**Energy import dependency**
16	**10**	**3.4%**	**32.6**	**3**	**Energy import dependency**
17	**14**	**4.8%**	**64.7**	**4**	**Population**
18	**15**	**5.1%**	**23.1**	**4**	**Population**
19	**14**	**4.8%**	**25.4**	**5**	**Average years of activity**
20	**15**	**5.1%**	**34.4**	**5**	**Average years of activity**
21	**15**	**5.1%**	**30.4**	**6**	**Country location**
22	**15**	**5.1%**	**13.6**	**6**	**Country location**
23	**15**	**5.1%**	**4.5**	**7**	**Country location**
24	**14**	**4.8%**	**9.7**	**7**	**Country location**
25	**13**	**4.5%**	**4.1**	**8**	**TES renewables and biofuels**
26	20	6.8%	11.4	8	TES renewables and biofuels
27	25	8.6%	13.9	8	TES renewables and biofuels
28	**10**	**3.4%**	**10.1**	**12**	**Member of the Eurozone**
29	**11**	**3.8%**	**14.1**	**12**	**Member of the Eurozone**
30	**12**	**4.1%**	**30.0**	**13**	**Greenhouse gas emissions intensity of energy consumption**
31	**10**	**3.4%**	**16.2**	**13**	**Greenhouse gas emissions intensity of energy consumption**
32	**10**	**3.4%**	**9.8**	**26**	**Gas prices for non-household consumers**
33	**10**	**3.4%**	**13.0**	**26**	**Gas prices for non-household consumers**
34	**10**	**3.4%**	**10.4**	**27**	**Harmonised index of consumer prices**
35	**15**	**5.1%**	**16.2**	**27**	**Harmonised index of consumer Pprices**

3.1.2. Classification Rules

The classification rules (also known as induction rules) reveal the paths followed by the tree for making predictions, and identifying routes with different degrees of RE target achievement. Via the classification rules, the replication of the tree is possible using modeling software (e.g., Microsoft Excel or Python).

The identification and application of the rules can simplify the definition of policy profiles, making it easier to determine the possible course of action for certain choices. The classification rules generated by the model are presented in Table 9.

Table 9. Classification Rules.

Node	Rules
Node 9	IF (Finalenergyconsumption $\leq$ 4.169) AND (Renewablesandbiofuels $\leq$ 981.668) THEN RE share = 19.40%.
Node 10	IF (Finalenergyconsumption $\leq$ 4.169) AND (Renewablesandbiofuels $\geq$ 981.668) THEN RE share = 33.72%.
Node 11	IF (Finalenergyconsumption $\geq$ 4.169 AND $\leq$ 15.364) AND (Renewablesandbiofuels $\leq$ 981.668) THEN RE share = 4.84%.
Node 28	IF (Finalenergyconsumption $\geq$ 4.169 AND $\leq$ 15.364) AND (Renewablesandbiofuels $\geq$ 981.668 AND $\leq$ 1566.912) AND (MemberoftheEurozone = 1) THEN RE share = 10.11%.
Node 29	IF (Finalenergyconsumption $\geq$ 4.169 AND $\leq$ 15.364) AND (Renewablesandbiofuels $\geq$ 981.668 AND $\leq$ 1566.912) AND (MemberoftheEurozone = 1) THEN RE share = 14.12%.
Node 30	IF(Finalenergyconsumption $\geq$ 4.169 AND $\leq$ 15.364) AND (Renewablesandbiofuels $\geq$ 1566.912) AND(Greenhousegasemissionsintensityofenergyconsumption $\leq$ 95.4) THEN RE share = 29.95%.
Node 31	IF (Finalenergyconsumption $\geq$ 4.169 AND $\leq$ 15.364) AND (Renewablesandbiofuels $\geq$ 1566.912) AND (Greenhousegasemissionsintensityofenergyconsumption $\geq$ 95.4) THEN RE share = 16.23%.
Node 14	IF (Finalenergyconsumption $\geq$ 15.364 AND $\leq$ 18.188) AND (Energyimportdependency $\leq$ 55.823) THEN RE share = 16.49%.
Node 15	IF (Finalenergyconsumption $\geq$ 15.364 AND $\leq$ 18.188) AND (Energyimportdependency $\geq$ 55.823 AND $\leq$ 67.691) THEN RE share = 10.88%.
Node 16	IF (Finalenergyconsumption $\geq$ 15.364 AND $\leq$ 18.188) AND (Energyimportdependency $\geq$ 67.691) THEN RE share = 32.63%.
Node 17	IF (Finalenergyconsumption $\geq$ 18.188 AND $\leq$ 24.680) AND (Population $\leq$ 5.276) THEN RE share = 64.67%.
Node 18	IF (Finalenergyconsumption $\geq$ 18.188 AND $\leq$ 24.680) AND (Population $\geq$ 5.276) THEN RE share = 23.13%.
Node 19	IF (Finalenergyconsumption $\geq$ 24.68 AND $\leq$ 34.24) AND (Averageyearsofactivity $\leq$ 6) THEN RE share = 25.37%.
Node 20	IF (Finalenergyconsumption $\geq$ 24.68 AND $\leq$ 34.24) AND (Averageyearsofactivity $\geq$ 6) THEN RE share = 34.40%.
Node 21	IF (Finalenergyconsumption $\geq$ 34.240 AND $\leq$ 38.742) AND (CountryLocation = "European Union") THEN RE share = 30.44%.
Node 22	IF (Finalenergyconsumption $\geq$ 34.24083 AND $\leq$ 38.742) AND (CountryLocation = "European Union") THEN = RE share = 13.59%.
Node 23	IF (Finalenergyconsumption $\geq$ 38.742 AND $\leq$ 71.933) AND (CountryLocation = "Eastern") THEN RE share = 4.46%.
Node 24	IF (Finalenergyconsumption $\geq$ 38.742 AND $\leq$ 71.933) AND (CountryLocation = "Eastern") THEN RE share = 9.68%.
Node 25	IF (Finalenergyconsumption $\geq$ 71.933) AND (Renewablesandbiofuels $\leq$ 17,625.451) THEN RE share = 4.08%.
Node 32	IF (Finalenergyconsumption $\geq$ 71.933) AND (Renewablesandbiofuels $\geq$ 17,625.451 AND $\leq$ 24,616.453) AND (Gaspricesfornonhouseholdconsumers $\leq$ 8.734) THEN RE share = 9.82%.
Node 33	IF (Finalenergyconsumption $\geq$ 71.933) AND (Renewablesandbiofuels $\geq$ 17,625.451 AND $\leq$ 24,616.453) AND (Gaspricesfornonhouseholdconsumers $\geq$ 8.734). THEN RE share = 12.98%.
Node 34	IF (Finalenergyconsumption $\geq$ 71.933) AND (Renewablesandbiofuels $\geq$ 24,616.453) AND (HarmonisedIndexofConsumerPrices $\leq$ 99.09) THEN RE share = 10.43%.
Node 35	IF (Finalenergyconsumption $\geq$ 71.933) AND (Renewablesandbiofuels $\geq$ 24,616.453) AND (HarmonisedIndexofConsumerPrices OR $\geq$ 99.09) THEN RE share = 16.21%.

Command Orders (IBM, 2021): AND = And, IF = If, OR = Other than, THEN = Compute.

3.2. *Efficacy Analysis*

For categorical dependent variables, the performance of the model results by the calculation of the accuracy of the model (overall right prediction of the positive and negative classes) [29]. It is calculated using the following formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

TP (True Positive): amount of right predictions for the positive class in the model.
TN (True Negative): amount of right predictions for the negative class in the model.
FP (False Positive): amount of wrong predictions for the positive class in the model.
FN (False Negative): amount of wrong predictions for the negative class in the model.

$$Accuracy = \frac{107 + 111}{107 + 111 + 35 + 39} = 0.746$$

The accuracy rate of 74.6% is expressed in the confusion matrix (Table 10). It reflects the overall percentage of correct classifications.

Table 10. Confusion Matrix.

Classification			
Observed	**Predicted**		
	Non-Efficacious	**Efficacious**	**Percent Correct**
Non-Efficacious	111	39	74.0%
Efficacious	35	107	75.4%
Overall Percentage	50.0%	50.0%	74.6%

With the results displayed in the Confusion Matrix, the Sensitivity (True positive rate) and Specificity (True negative rate) measures can be calculated. Sensitivity, also known as True Positive Rate, measures how well a test can identify true efficacious. Specificity, also known as True Negative Rate, measures how well a test can identify true non-efficacious [29]. For the efficiency analysis, Sensitivity refers to the rate of observations predicted as efficacious and Specificity as non-efficacious. They are calculated using the following formulas:

$$Sensitivity = TP/(TP + FN)$$

$$Specificity = TN/(TN + FP)$$

TP (True Positive): amount of right predictions for the positive class in the model.
FN (False Negative): amount of wrong predictions for the negative class in the model.
TN (True Negative): amount of right predictions for the negative class in the model.
FP (False Positive): amount of wrong predictions for the positive class in the model.
The Sensitivity and Specificity of the RE efficacy model are calculated below:

$$Sensitivity = 107/(107 + 35)\ observations = 0.754$$

$$Specificity = 111/(111 + 39)\ observations = 0.740$$

The results for the Energy Policy efficiency analysis are presented below, including the Decision Tree Diagram (Figure 4), the DT result in a table (Table 11), findings and Classification rules (Table 12).

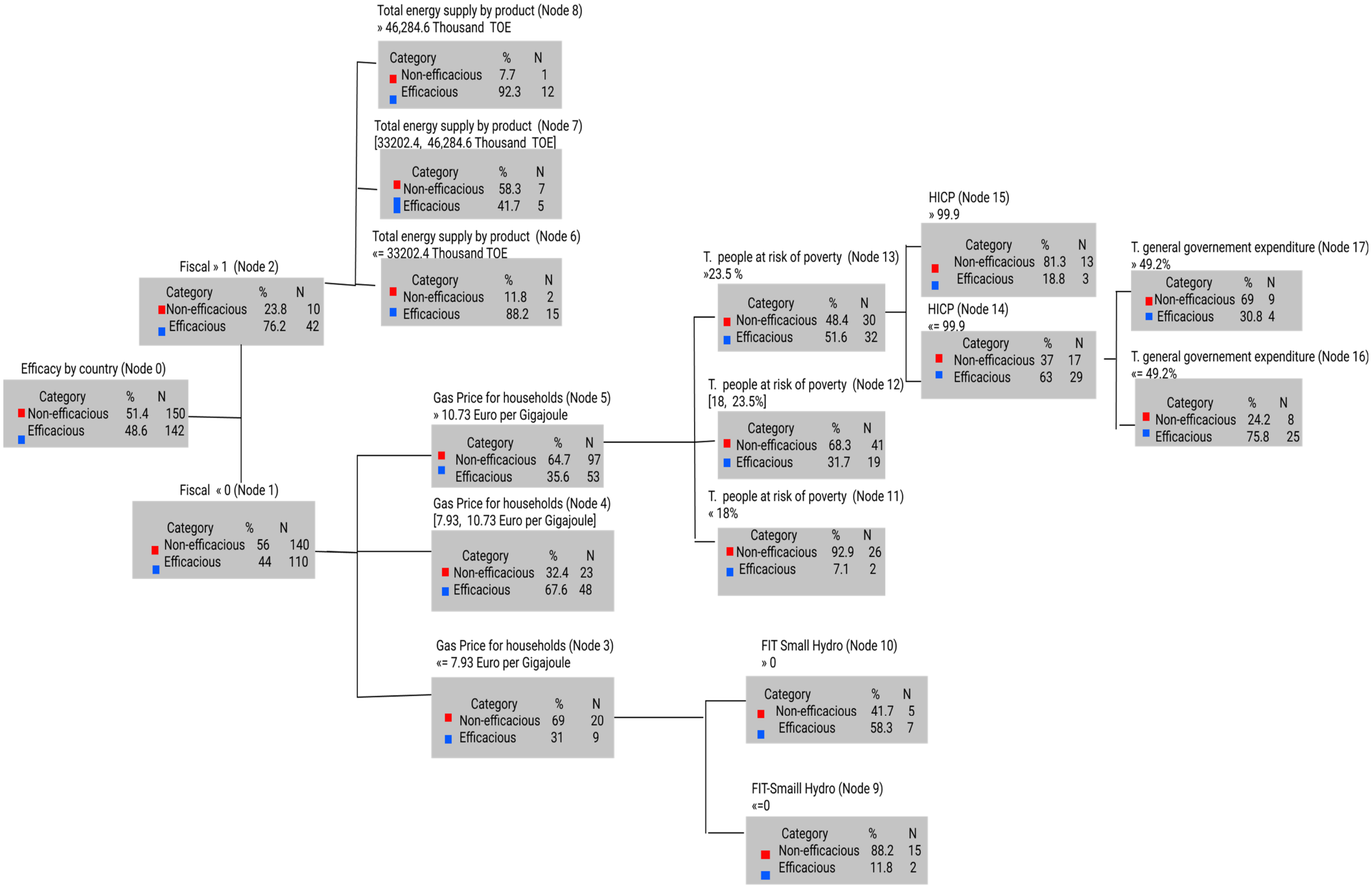


Figure 4. Energy Policy efficacy Decision Tree.

Table 11. Decision Tree results in a table format.

Node	Non-Efficacious		Efficacious		Total		Predicted Category	Parent Node	Primary Independent Variable
	Number of policies	Percentage of policies	Number of policies	Percentage of policies	Number of policies	Percentage of policies			Variable used to split the node
0	150	51.40%	142	48.60%	292	100.00%	Non-efficacious		
1	140	56.00%	110	44.00%	250	85.60%	Non-efficacious	0	Fiscal
2	10	23.80%	32	76.20%	42	14.40%	Efficacious	0	Fiscal
3	20	69.00%	9	31.00%	29	9.90%	Non-efficacious	1	Gas price for households
4	***23***	***32.40%***	***48***	***67.60%***	***71***	***24.30%***	***Efficacious***	***1***	**Gas price for households**
5	97	64.70%	53	35.30%	150	51.40%	Non-efficacious	1	Gas price for households
6	***2***	***11.80%***	***15***	***88.20%***	***17***	***5.80%***	***Efficacious***	***2***	**Total energy supply by product**
7	***7***	***58.30%***	***5***	***41.70%***	***12***	***4.10%***	***Non-efficacious***	***2***	**Total energy supply by product**
8	***1***	***7.70%***	***12***	***92.30%***	***13***	***4.50%***	***Efficacious***	***2***	**Total energy supply by product**
9	***15***	***88.20%***	***2***	***11.80%***	***17***	***5.80%***	***Non-efficacious***	***3***	**FiT-Small Hydro**
10	***5***	***41.70%***	***7***	***58.30%***	***12***	***4.10%***	***Efficacious***	***3***	**FiT-Small Hydro**
11	***26***	***92.90%***	***2***	***7.10%***	***28***	***9.60%***	***Non-efficacious***	***5***	**Total people at risk of poverty**
12	***41***	***68.30%***	***19***	***31.70%***	***60***	***20.50%***	***Non-efficacious***	***5***	**Total people at risk of poverty**
13	30	48.40%	32	51.60%	62	21.20%	Efficacious	5	Total people at risk of poverty
14	17	37.00%	29	63.00%	46	15.80%	Efficacious	13	Harmonised index of consumer Pprices
15	***13***	***81.30%***	***3***	***18.80%***	***16***	***5.50%***	***Non-efficacious***	***13***	**Harmonised index of consumer prices**
16	***8***	***24.20%***	***25***	***75.80%***	***33***	***11.30%***	***Efficacious***	***14***	**Total general government expenditure**
17	***9***	***69.20%***	***4***	***30.80%***	***13***	***4.50%***	***Non-efficacious***	***14***	**Total general government expenditure**

Table 12. Classification rules.

Node	Rules
Node 9	IF (FiscalTaxC.TaxI = 1) AND (Gaspricehousehold ≤ 7.931) AND (FiTSmallHydro ≤ 0) THEN Non-efficacious (Prob = 0.882)
Node 10	IF (FiscalTaxC.TaxI = 1) AND (Gaspricehousehold ≤ 7.931) AND (FiTSmallHydro ≥ 0)THEN Efficacious (Prob = 0.583)
Node 4	IF (FiscalTaxC.TaxI = 1) AND (Gaspricehousehold ≥ 7.931) AND (Gaspricehousehold ≤ 10.731) THEN Efficacious (Prob = 0.676)
Node 11	IF (FiscalTaxC.TaxI = 1) AND (Gaspricehousehold ≥ 10.731) AND (Peopleatriskofpovertyorsocialexclusionbyageandsex ≤ 18) THEN Non-Efficacious (Prob = 0.928).
Node 12	IF (FiscalTaxC.TaxI = 1) AND (Gaspricehousehold ≥ 10.731) AND (Peopleatriskofpovertyorsocialexclusionbyageandsex ≥ 18 AND ≤ 23.5) THEN Non-efficacious (Prob = 0.683).
Node 16	IF (FiscalTaxC.TaxI = 1) AND (Gaspricehousehold ≥ 10.731) AND OR (Peopleatriskofpovertyorsocialexclusionbyageandsex ≥ 23.5) AND (HICP ≤ 99.91) AND (Totalgeneralgovernmentexpenditure ≤ 49.2) THEN Efficacious (Prob = 0.757).
Node 17	IF (FiscalTaxC.TaxI = 1) AND (Gaspricehousehold ≥ 10.731) AND (Peopleatriskofpovertyorsocialexclusionbyageandsex ≥ 23.5)) AND (HICP ≤ 99.91) AND (Totalgeneralgovernmentexpenditure ≥ 49.2) THEN Non-efficacious (Prob = 0.692).
Node 15	IF (FiscalTaxC.TaxI = 1) AND (Gaspricehousehold ≥ 10.731) AND (Peopleatriskofpovertyorsocialexclusionbyageandsex ≥ 23.5) AND (HICP ≥ 99.91) THEN Non-efficacious (Prob = 0.812).
Node 6	IF (FiscalTaxC.TaxI = 1) AND (Totalenergysupplybyproduct ≤ 33,202.406) THEN Efficacious (Prob = 0.882).
Node 7	IF (FiscalTaxC.TaxI = 1) AND (Totalenergysupplybyproduct ≥ 33,202.406) AND (Totalenergysupplybyproduct ≤ 46,284.6) THEN Non-Efficacious (Prob = 0.583).
Node 8	IF (FiscalTaxC.TaxI = 1) AND (Totalenergysupplybyproduct ≥ 46,284.6) THEN Efficacious (Prob = 0.923).

From the Decision Tree we report that:

- *Fiscal* is the best discriminator of *Energy policy efficacy.*
- The subsequent branch on the upper side of the DT have selected the *Total energy supply by product* as the second-best discriminator (with a p-Value of 0.023). The policy mixes included in the three consecutive nodes, independently of the total energy supply by product, have a high average probability of achieving their national RE targets (therefore classified as efficacious). As there are no child nodes below this branch, the nodes are considered terminal nodes.
- On the bottom branch, the second-best predictor is the *Gas price for households*; only one out of the tree branches of it is a terminal node, in which if the *Gas price for households* is between 7.93 and 10.73 Euro/Gigajoules, there is a probability of 67.7% for the RE Target to be achieved (classified as efficacious and therefore a 32.3% probability for the target to not be achieved).
- If the *Gas price for households* is lower than 7.93 Euro/Gigajoules and there is a *Feed-In Tariff for Small Hydro* (FIT), there is a probability of 58.3% for the RE Target to be achieved. If this FIT is not present in the analyzed policy mix, the probability of being classified as efficacious falls to 11.8%.
- The third predictor on the left branch of the tree is *People at risk of poverty*. If its value is lower than 18% or between 18–23%, the probability is very low, with 7 and 31%, respectively, for the policy mixes to be efficacious. At this same level of the tree, if the *people at risk of poverty* is higher than 23%, and the *Harmonised Index of Consumer Prices* (*HICP)* (fourth predictor) is higher than 99.9, the policy mix is predicted to be efficacious with an 18.8% of probability (if it is lower than 99,9 the policy mix is predicted to be efficacious with a 63% of probability).
- Whenever the HICP is lower than 99.9, the next (and fifth) best predictor is the *Total general government expenditure*. Whenever it is lower than 49.2% of GDP, the probability of falling into the category of efficacious is 75.8%, and whenever higher than 49.2%, the probability is 30.8%.

Table 11 presents the results of the Decision Tree in a table format. The terminal nodes have been highlighted in bold and italic letters. The highest percentage of policies by category in each node has been underlined.

Classification Rules

The classification rules generated by the model are presented in Table 12.

3.3. Analysis with Other Dependent Variables

As part of the robustness validation of the models, other possible dependent variables were tested. The additional variables are variations of the initial one "RE share" (Eurostat, 2020). The descriptions are included in Table 13 for the continuous variables and Table 14 for the categorical variables that were tested. For the EP effectiveness model, two additional continuous dependent variables were tested, comparing their R-squared values. This process is documented in Table 14. RE share, as the selected variable (highlighted variable), has the highest R-squared compared to the additional tested variables.

Table 13. Continuous Dependent Variables.

Possible Dependent Variable	Dependent Variable Description	Type of Variable (Categorical/Continuous) & (Binary/Non-Binary)	R-Squared
1. RE share *	**Observed share of Renewable Energy of the total electricity consumption achieved in the year of measurement.**	**Continuous**	**0.917**
2. RE share variation (self-calculated)	Observed difference in RE share between consecutive years.	Continuous	0.039
3. RE share ratio variation (self-calculated)	Variation of the Ratio between the observed increase in RE share and the country's expected growth in a year.	Continuous	0.198

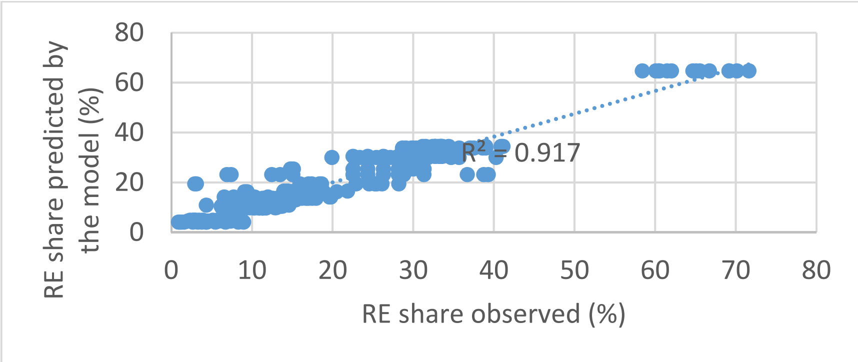
Source: * (Eurostat, 2020).

Table 14. Categorical Dependent Variables.

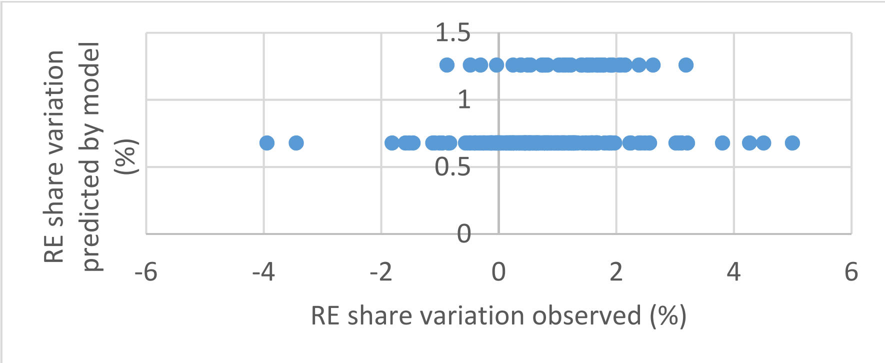
Possible Dependent Variable	Description	Type of Variable (Categorical/Continuous) & (Binary/Non-Binary)	Accuracy	Sensitivity	Specificity	True Positives	True Negatives
Efficacy by country (self-calculated)	Comparison of the observed RE share increase and the country expected growth in a year (threshold)	Categorical; binary	0.746	0.754	0.740	107	111
RE share Increase vs. Decrease (self-calculated)	Increase or decrease in the RE share compared to the RE share value of the consecutive year by country.	Categorical; binary	0.718	0.674	0.802	130	81
RE share Efficacy over 0.6% (self-calculated)	Comparison of the observed RE share increase and the EU expected growth of 0.6% per year. **	Categorical; binary	0.554	0.786	0.516	33	130
RE share Efficacy (over 1) (self-calculated)	Comparison of the observed RE share increase and the self-defined value of 1 (from the RE share mean value of 0.7 plus 1/3 of one standard deviation value). *	Categorical; binary	0.732	0.659	0.765	60	156

* Efficacious class is assigned if the RE share growth is higher than 1; otherwise, non-efficacious. ** Efficacious class is assigned if the RE share growth is higher than 0.6; otherwise, non-efficacious.

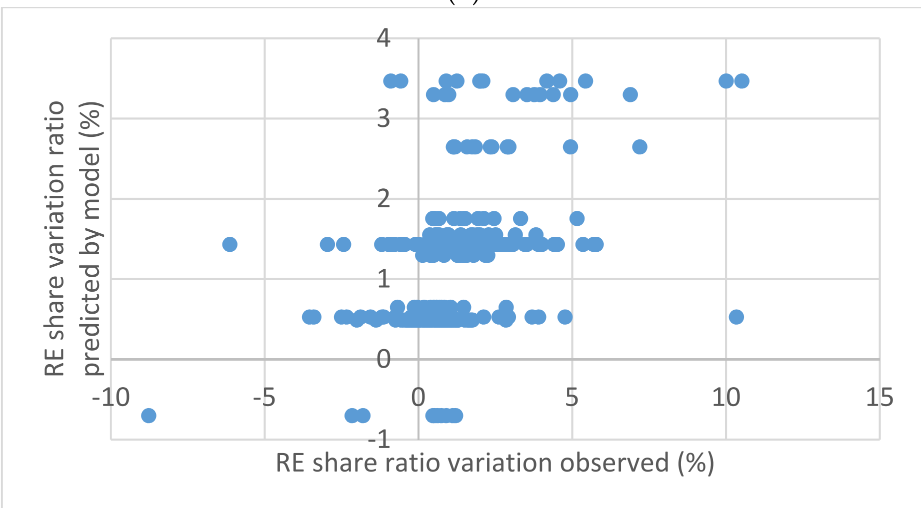
The results of the predicted and observed values by the models are presented graphically in Figure 5.



(**a**)



(**b**)



(**c**)

Figure 5. Comparison of possible dependent variables. (**a**) RE share (selected variable); (**b**) RE share variation; (**c**) RE share ratio variation.

For the categorical variables tested for the EP efficacy model, the Sensitivity and Specificity measures, as well as the accuracy, were estimated and compared. The comparisons are reported in Table 14.

Variables 1, 2, and 4 have similar accuracy values (with around 0.7). Nevertheless, when comparing their sensitivity and specificity values, differences can be seen in having a non-balanced value among Sensitivity and Specificity. Variable number 1, "Efficacy by country" (highlighted in the table), has more equilibrated values among its Sensitivity and Specificity values, meaning that the model will perform similarly in predicting both the efficacious policies and the non-efficacious ones. It also has the highest accuracy rate, and the higher its value, the better the model is [18].

The comparison of the accuracy values of the categorical variables tested are presented in Figure 6. The comparison of the sensitivity and specificity values of each variable can be seen in Figure 7.

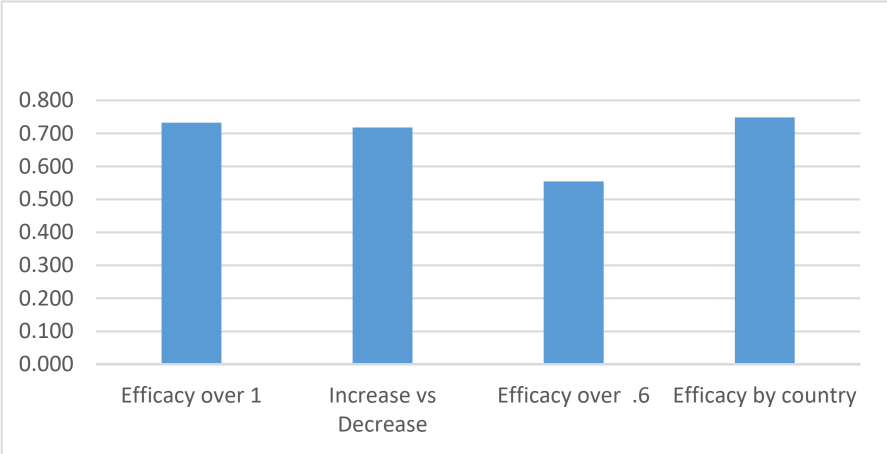


Figure 6. Comparison of the accuracy values of the categorical variables.

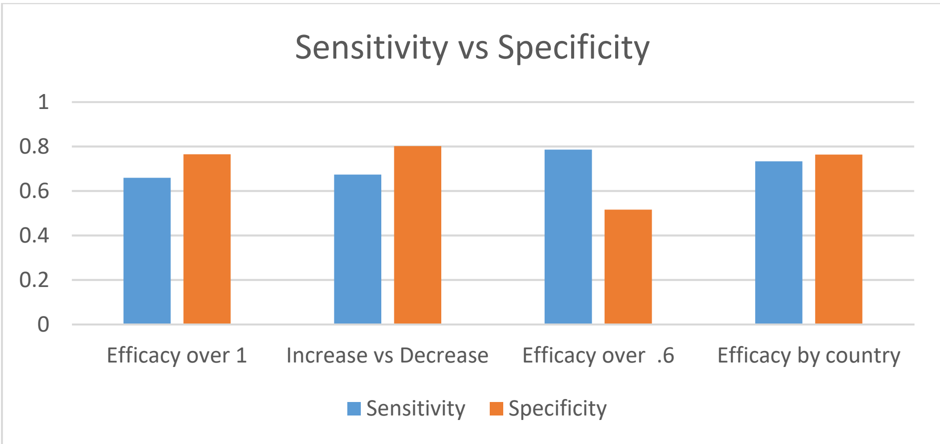


Figure 7. Comparison of the sensitivity and specificity values of the categorical variables.

4. Discussion

The study's objective has been to develop a support tool that provides information that could guide targeted interventions and policies. With the use of advanced data analytics, we aimed to determine whether there is a relationship between the different independent variables selected and the RE share and its target achievement.

Using Decision Trees to predict the degree of RE share and classify the policy mixes into efficacious and non-efficacious outcomes and a predicted RE share value, we present the utility of machine learning techniques and the application to energy policymaking. The methodology explored in this paper also aims to bring different substantive fields together in advancing data-driven policymaking. The results of this study are one example that illustrates the intersection of public policy and computational sciences. A wider application of these techniques in social sciences could enable better future policy design, capturing the real problems and needs on the ground.

The model proposed in this document is a data-driven model whose performance was validated using historical data. In the future, we plan to validate the model with new data that are gathered each year.

The results of the Decision Tree models have produced the stratification of the policy mixes according to their potential to promote RE share targets (effectiveness model) and the potential to display achievement or non-achievement (efficacy model). Each of the nodes of the trees has different profiles of predictors and, therefore, different RE share degrees (for the EP effectiveness analysis) and different achievement probability of achieving the efficacious or non-efficacious categories (for the EP efficiency analysis).

To the best of our knowledge, this study is one of the few studies in policymaking to incorporate ML methods to analyze energy policies. No renewable energy target and policy achievement analysis has been found to compare its results against the ones presented in this work. An ex-ante methodology for policy impact assessment is proposed to the European Commission [30], serving as a guideline for the member states supporting the objective achievement of climate change policies (energy policies included). It proposes comparing GHG levels ex-ante and ex-post of the implementation of the policies.

4.1. Effectiveness Findings

The findings of this study suggest some policy implications both in the European and global contexts with similar circumstances or policy environments. These results need to be seen with care, as the Decision Tree model detects associations, which do not always imply causation.

Policymakers must be selective when designing energy policy mixes, as small variations in predictors can bring very different outcomes. This predictive model aims to be an inspiration to maximize the feasibility of policy targets and limit policy target misalignment.

Energy Policy Efficacy implications:

- When the Population served by a policy mix is higher, the predicted RE share value is reduced by almost half (from 64.6% to 23.1%) (as seen in NODE 17 & 18). The model predicts more effective policy mixes in the context of smaller populations (within the range of 18.1 and 24.6 million TOE of final energy consumption).
- Considerable variation in the predicted RE share values can be seen based on the duration of how long the policy mixes have been in effect. For example, if the policy mixes have more than six years of average duration, compared to under six years, with a difference of more than 10% RE share (from 34.4% to 23.6%). The model predicts a more efficacious policy mix if the average duration surpasses six years (within the range of 24.6 and 34.2 million TOE of final energy consumption).
- As the amount of the Total Energy Supply by renewables and biofuels is increased due to policy mixes, the respective RE share values will also increase (as seen in NODE 9 to 13 and 25 to 27).
- The threshold determined by the model at 95.4 for the GHG emissions intensity of energy consumption denotes a decrease in RE share value (within the range of 4.169 and 15.364 million TOE of final energy consumption and higher than 1566.9 of TES Renewables and Biofuels thousand TOE) (as seen in NODE 30 & 31).

4.2. Efficacy Findings

- First, when a policy mix addresses FIT for Small Hydro, a considerable difference in the probability of target achievement or of being efficacious is reached (from 11.8% to 58.3%) (as seen in NODE 9 and 10).
- When a policy mix is designed to address a fiscal policy type, and with a high gas price for households (after 10.73 Gigajoule/euro), the higher the risk of poverty, the more efficacious a policy mix becomes (as seen in NODE 11 and 12).
- When the total general government expenditure addressed by a policy mix is higher than 49.2% (Percentage of GDP), the probability for policies to fall in the category of efficacious is reduced from 30.8 to 75.8% (as seen in NODE 16 and 17).
- If there is a fiscal policy included in the policy mix, independent of the amount of energy supply, all profiles will be in the trend of RE target achievement (efficacious) (70% average) (as seen in NODE 6, 7 and 8).

- The threshold determined by the model at 99.9% for the HICP value denotes an increase in the probability for the policy mixes to achieve their RE target if the threshold is surpassed (as seen in NODE 4).

Classification rules have been generated to estimate outcomes, highlighting the easy application of DT as prediction tools. Through the tree paths that are created, predictions of future trends can be made.

The main outcomes of the predictive models rely on estimations of the RE share and the target achievements before the proposed policies are implemented (in their design stage). Based on the anticipated likely degrees of policy effectiveness and efficacy (or success) of the policy mixes, the decision-makers can make informed policy decisions, considering the country's characteristics and circumstances.

The proposed predictive models allow policymakers at various scales and localities to not only improve the precision of the policies in achieving the desired outcomes but also to come to realize the generalizable and testable lessons over time. These lessons or trends could be shared across national borders. Data-driven machine learning methodologies present an alternative to save time, money, and resources required to see the real outcomes by the end of the policy life cycle. The issue of analyzing the efficacy of policy mixes after a long period alone can be debatable. Policymakers ought to experiment with these models as examples of "tools", allowing them to implement policies with a higher probability of being effective or efficacious.

The findings from this research highlight the importance of data availability. Without the data that can be used to illustrate the variables that feed into the models, the possibility of maximizing our ability to take action based on estimation is not possible. The shortage of data at the country level must be continued to be improved with more aggressive measures and transparency adopted by the governments. At the same time, the need for multidisciplinary methodologies that can systematically seek evidence-based approaches in policymaking needs more focused attention [31].

As the limitations of this analysis, it is recognized that the RE share is also influenced by the energy efficiency and other factors; and that different energy policies are available to promote energy efficiency (each policy subject involves different types of policy instruments and technical measures). The policy analysis here presented was applied only to a certain type of policy (Renewable Energy policies).

In addition, missing concerns to be included in further applications of this analysis include information on foreign business (as investment and trade) [32], political stability and the absence of violence [33], energy access, and employment [26].

Finally, all the 22 sample countries used in the analysis are located in the same continent (Europe). A situation that resulted mainly due to the data availability. Even in Europe, not all the countries were evaluated as not enough information for the assessed years was found.

5. Conclusions

A summary for Policy Makers on applying ML techniques for energy policy analysis has been organized and is now presented.

Machine Learning use. Machine Learning methods have been applied within the energy sector to estimate short- and long-term factors. Their usability is promoted by their capability to alleviate the uncertainty of controllable and non-controllable circumstances and provide estimations that lead to coping with a better use of resources.

Decision Trees. They are used to predict a dependent variable based on independent variables. They are characterized by their easiness of being interpreted. Better performing policies can be identified, leading to choices with a higher propensity for "success".

Supported information for Policy-Making. A support tool based on the examination of ex-post policy information is available to be used in an ex-ante form to adjust or design policies and their interaction (through policy mixes). By doing so, it is possible to provide information that could guide targeted interventions.

Predictive analytics to improve RE target achievement. Predictive analytics provides an alternative to anticipating RE share target achievement and its degree of achievement through the results of the effectiveness (degree of achievement) and efficacy (achievement or not) analysis.

Data Availability Challenges. Different challenges have been identified as key aspects to be considered for the fulfillment of the data requirements; actions must be taken by policy developers, institutions, and academics toward data registration and openness.

Conformation of a Database. Data combined with Machine Learning has a great potential to tackle complex problems. The database conformed for the analysis includes variables effectiveness assessment concerns, country and policy characteristics. The database can serve for further application of studies.

Findings. This paper aimed to determine whether there is a relationship between the different independent variables selected and the RE share and its target achievement. The findings aim to lead to further exploration of certain variables.

Machine Learning opportunities. A call for collaboration among experts from different fields of knowledge is made to continue energy policy analysis based on machine learning methods. The call is extended to energy policymakers and other stakeholders to raise problems on the topic. Computational sciences at the service of qualitative-based sciences ought to be supported.

Author Contributions: Conceptualization, V.M., V.L. and D.O.; Formal analysis, V.M. and D.O.; Funding acquisition, V.M., D.O., J.K.-H. and J.C.; Investigation, D.O.; Methodology, V.M. and D.O.; Supervision, V.M. and V.L.; Validation, V.M., J.K.-H. and J.C. and V.L. Visualization, D.O.; Writing—original draft, D.O. and V.M.; Writing—review & editing, V.M., J.K.-H. and J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the MIT Portugal Program.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The dataset was organized by the authors from different sources.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Examples of Machine Learning Applications in the Energy Field

Year	Authors	Application	Document Focus
2018	Marugán et al. [34]	Wind energy systems	NN applied in forecasting and predictions; design optimization; fault detection and diagnosis; and optimal control in wind energy systems
2018	Zaidi et al. [35]	Energy-water nexus	Description of different ML methodologies applied in the Energy-water nexus
2018	Zeng et al. [36]	Optimization of reactors	Development of a support vector regression for the creation of autonomous control for small reactors
2018	Zendehboudi et al. [37]	Oil and gas processes	Review on hybrid models with focus on applications in chemical, petroleum, and energy systems
2017	Anifowose et al. [38]	Oil and gas field exploration	Modeling based in ensemble learning paradigm for oil reservoir and other applications
2017	Voyant et al. [39]	Solar radiation forecasting	Development of hybrid models to use an ensemble forecast approach on solar radiation
2016	Heinermann and Kramer [40]	Wind power forecasting	Proposal of a prediction framework based on heterogeneous machine learning ensembles
2016	Fulford et al. [41]	Shale gas well diagnosis	Model application for shale gas well
2015	Gupta et al. [42]	Failure prediction	Development of a support vector machine model for a blackout prediction

Year	Authors	Application	Document Focus
2016	Costa et al. [43]	Security dispatch method for coupled natural gas and electric power networks	DT for preventing contingencies that may cause interruption in the power networks
2016	Ottesen [5]	Total cost minimization in energy systems for the prosumers' buildings	DT for the minimization of total costs
2016	Moutis et al. [6]	Energy storage planning and energy controlling	Application of a DT to improve storage planning
2017	Kamali et al. [7]	Prediction of the risk of a blackout in electric energy systems	Smart grid early warning system DT support (in several scenarios)
2018	Aguado et al. [8]	Railway electric energy systems optimal operation	DT applied for multiple railway electric energy systems and operation modes

Appendix B. Variables Directory

- Context Variables

Table A1. Country Characteristics (1).

Variable Name	Description	Example	Binary, Polynomial or Numeric	Units	Min	Max	Mean	Std. Deviation	Source
Country Location (2)	Geographical Country location in which the program has been implemented	Western	Polynomial	no units	n.a.	n.a.	n.a.	n.a.	(Unstat, 2020)
EU member	The country was a member of the EU at the moment of evaluation or not	1 (or 0)	Binary	no units	0	1	0.95	0.23	(European Comission)
Member of the Eurozone	The country had the Euro as currency at the moment or not	1 (or 0)	Binary	no units	0	1	0.47	0.5	(European Comission, 2020)
Population	Number of persons having their usual residence in a country	10 million	Numeric	million people	0.49	82.91	20.61	24.95	(Eurostat, 2020)
Cooling degree days	Weather-based technical indexes designed to describe the energy requirements of buildings in terms of cooling	4.94	Numeric	Numeric Value	0	320.38	54.48	71.73	(Eurostat, 2020)
Heating degree days	Weather-based technical indexes designed to describe the energy requirements of buildings in terms of heating	36	Numeric	Numeric Value	1054.83	6179.75	3329.1	1081.02	(Eurostat, 2020)

Table A1. *Cont.*

Variable Name	Description	Example	Binary, Polynomial or Numeric	Units	Min	Max	Mean	Std. Deviation	Source
Final energy consumption	Total final energy consumption	38,381 Million tonnes of oil equivalent (TOE)	Numeric	Million tonnes of oil equivalent (TOE)	2.77	225.35	49.99	58.49	(Eurostat, 2020)
Total energy supply by product	For the total of all energy products it is the total energy delivered/consumed in a country	33,218.11	Numeric	Thousand TOE	3948.82	335,474.27	73,683	88,544.81	(Eurostat, 2020)
Supply by Renewables and biofuels	Supply by renewables and biofuels	9789.91	Numeric	Thousand TOE	125.54	43,693.14	9627.3	9781.01	(Eurostat, 2020)
Supply by Oil and petroleum products	Supply by the Oil and petroleum products	11,698.52	Numeric	Thousand TOE	61.1	110,941.29	23,479	28,753.99	(Eurostat, 2020)
Supply by Non-renewable waste	Supply by the non-renewable waste	634.23	Numeric	Thousand TOE	0	4514.02	629.28	941.77	(Eurostat, 2020)
Supply by Nuclear Heat	Supply by the Nuclear Heat	0	Numeric	Thousand TOE	0	115,209	9470.7	23,010.26	(Eurostat, 2020)
Supply by Electricity	Supply by Electricity products	534.19	Numeric	Thousand TOE	−5777.3	3987.79	−17.6	1564.75	(Eurostat, 2020)
Supply by Solid Fossil Fuels	Supply by Solid Fossil Fuels	3205.97	Numeric	Thousand TOE	8.66	81,561.13	12,034	20,038.21	(Eurostat, 2020)
Final energy supply	Total final energy supply	58,031 (Million TOE)	Numeric	Million TOE	3.95	349.21	73.95	88.02	(Eurostat, 2020)
Expenditure on Research and Development	Percentage of government budget allocations for Research & Development	1.90%	Numeric	Percentage	0.35	2.13	1.28	0.44	(Eurostat, 2020)
Long-term interest rate	The amount a lender charges a borrower, being a percentage of the principal	4%	Numeric	Percentage	0.09	10.55	3.39	1.92	(Eurostat, 2020)
Air pollutants by source sector	Air pollutants emissions for the entire territory	240,593	Numeric	Tonnes	32,138	1,783,137	391,508	444,570.42	(Eurostat, 2020)
Pollution or other environmental problems	Environmental problems by degree of urbanization	14.1	Numeric	Percentage	5.2	43.6	18.33	7.5	(Eurostat, 2020)
GDP per capita	Country's GDP per capita	36,883.87 (U.S. dollars per capita)	Numeric	U.S. dollars per capita	4477.03	117,366.01	36,674	22,373.49	(Woeld Economic Outlook, 2020)

Table A1. *Cont.*

Variable Name	Description	Example	Binary, Polynomial or Numeric	Units	Min	Max	Mean	Std. Deviation	Source
Human Development Index	Average achievement of a long and healthy life, being knowledgeable and have a decent standard of living	0.86	Numeric	Numeric Value (Scale 0 to 1)	0.77	0.95	0.88	0.04	(UNDP, 2020)
Harmonised Index of Consumer Prices	Change over time of the prices of consumer goods and services acquired by households	80.6	Numeric	Index	63.06	108.05	93.72	8.38	(Eurostat, 2020)
Total People at risk of poverty	Indicator to monitor the EU 2030 target on poverty and social exclusion	0.14	Numeric	Percentage	12.2	61.3	23.31	7.94	(Eurostat, 2020)
Gini coefficient of equivalised disposable income	Indicator relating to poverty risk	0.43 (Scale 0 to 1)	Numeric	Numeric Value (Scale 0 to 1)	20.9	40.2	29.62	3.91	(Eurostat, 2020)
Average household size	Average household inhabitants	2.3 inhabitants	Numeric	Number of inhabitants by house	1.9	2.9	2.36	0.27	(Eurostat, 2020)
Exposure to air pollution by particulate matter	Measures the population weighted annual mean concentration of particulate matter	23.3	Numeric	Numeric Value	4.9	51.4	15.8	7.73	(Eurostat, 2020)
Environmental protection expenditure	General government expenditure by environmental protection	0.005	53.7	Euro per inhabitant	−0.3	1.7	0.72	0.29	(Eurostat, 2020)
Environmental tax revenues	Proportion of environmental tax revenues in total revenues from all taxes and social contributions	43.3	Numeric	Percentage	173.76	49,474	11,402	15,029.44	(Eurostat, 2020)
Remuneration of civil servants	Annual evolution in the remuneration of national civil servants working in central public administrations	102.5	Numeric	Index	73	148.7	103.21	5.91	(Eurostat, 2020)
Number of national civil servants	Number of national civil servants in central public administration	42,005	Numeric	Numeric Value	825	435,660	65,250	85,104.49	(Eurostat, 2020)

Table A1. *Cont.*

Variable Name	Description	Example	Binary, Polynomial or Numeric	Units	Min	Max	Mean	Std. Deviation	Source
FiT-Solar PV	Feed in Tariff for Solar technologies	0.6717	Numeric	US Dollars	0	0.83	0.15	0.22	(OECD, 2020)
FiT-Wind	Feed in Tariff for Wind technologies	0.097	Numeric	US Dollars	0	0.68	0.07	0.09	(OECD, 2020)
FiT-Small Hydro	Feed in Tariff for Hydro technologies	0.6717	Numeric	US Dollars	0	0.25	0.06	0.06	(OECD, 2020)
FiT-Biomass	Feed in Tariff for Biomass technologies	0.097	Numeric	US Dollars	0	0.21	0.06	0.07	(OECD, 2020)
FiT-Waste	Feed in Tariff for Waste technologies	0	Numeric	US Dollars	0	1.13	0.08	0.17	(OECD, 2020)
FiT-Geothermal	Feed in Tariff for Geothermal technologies	0.1477	Numeric	US Dollars	0	0.33	0.05	0.08	(OECD, 2020)
FiT-Marine	Feed in Tariff for Marine technologies	0.0964	Numeric	US Dollars	0	0.71	0.05	0.1	(OECD, 2020)
Total general government expenditure	Total General government expenditure	0.65	Numeric	Percentage of GDP	25.6	65.1	45.62	6.31	(Eurostat, 2020)

(1) Country Characteristics: The variables of this group have not been included for the EPEI application. (2) Country Location possible values: Western, Eastern, Northern, Southern or European Union (as the average values).

- Effectiveness Concerns

Table A2. Affordability Variables.

Variable Name	Description	Example	Binary, Polynomial or Numeric	Units	Min	Max	Mean	Std. Deviation	Source
Electricity price for households	Yearly average of the price paid for electricity in a country by households users	0.89 €/kWh	Numeric	Euros/ Kilowatt-hour/	0.06	0.2	0.12	0.03	(Eurostat, 2020)
Electricity prices for non-households	Yearly average of the price paid for electricity in a country by industrial users	0.81 €/kWh	Numeric	Euros/ Kilowatt-hour/	0.04	0.14	0.08	0.02	(Eurostat, 2020)
Gas Price for households	Yearly average of the price paid for gas in a country by household users	10.4 €/GJ	Numeric	Euros/ Gigajoules	3.67	21.05	11.8	3.11	(Eurostat, 2020)
Gas prices for non-households	Yearly average of the price paid for gas in a country by industrial users	7.3 €/GJ	Numeric	Euros/ Gigajoules	2.75	14.03	8.38	1.83	(Eurostat, 2020)

Table A3. Accessibility Variables.

Variable Name	Description	Example	Binary, Polynomial or Numeric	Units	Min	Max	Mean	Std. Deviation	Source
Energy Balance	Difference between the final energy consumption and final energy supply	19,650 Million tonnes of oil equivalent (TOE)	Numeric	Million tonnes of oil equivalent (TOE)	−0.18	123.86	23.96	30.53	Own calculated
Energy import dependency	Share of total energy needs of a country met by imports from other countries	50%	Numeric	Percentage	−702.61	97.51	19.15	139.86	(Eurostat, 2020)

Table A4. Economic Competitiveness Variables.

Variable Name	Description	Example	Binary, Polynomial or Numeric	Units	Min	Max	Mean	Std. Deviation	Source
Market share of the biggest competitor	Market share of the largest generator in the electricity market	55.30%	Numeric	Percentage	15.3	96.5	49.11	21.36	(Eurostat, 2020)
Energy Productivity	Indicator that measures the amount of economic output that is produced per unit of gross available energy	8.05	Numeric	Euro per Kilogram of Oil Equivalent (KGOE)	1.67	18.58	7.13	3.13	(Eurostat, 2020)
Energy Intensity of the economy	Primary energy consumption per unit of GDP	12.7	Numeric	Numeric Value	1.15	13.03	3.11	2.49	Own Calculated
Energy Intensity of the population	Primary energy consumption per inhabitant	1.04	Numeric	Numeric Value	0.04	7.29	1.48	1.66	Own Calculated

Table A5. Impact on the Environment Variables.

Variable Name	Description	Example	Binary, Polynomial or Numeric	Units	Min	Max	Mean	Std. Deviation	Source
Total Greenhouse gas emissions	Total GHG emissions in a country	81,987.78	Numeric	Thousand tonnes	6636.22	1,007,867	181,129	236,340.85	(Eurostat, 2020)
GHG emissions intensity of energy consumption	Ratio between energy-related GHG emissions and gross inland consumption of energy	105.4	Numeric	Numeric Value	68.5	124.9	91.61	8.88	(Eurostat, 2020)

Table A6. Impact on Climate Change Variables.

Variable Name	Description	Example	Binary, Polynomial or Numeric	Units	Min	Max	Mean	Std. Deviation	Source
Renewable Electricity capacity	Maximum net generating capacity of power plants and other installations that use renewable energy sources	303	Numeric	MW	−168.48	11,316.33	941.11	1807.05	(IRENA, 2020)
Energy Efficiency	Ratio of the total energy delivered by the sytem and the energy consumed by it	55.6 (Million tonnes of oil equivalent (TOE))	Numeric	Million TOE	4.27	332.75	70.96	85.12	(European Commission, 2020)

Table A7. Equity Variables.

Variable Name	Description	Example	Binary, Polynomial or Numeric	Units	Min	Max	Mean	Std. Deviation	Score
Population unable to keep home adequately warm	Population unable to keep home adequately warm by poverty status	0.03%	Numeric	Percentage	0.3	69.5	10.05	12.27	(Eurostat, 2020)

Table A8. Governance effectiveness and efficacy variables.

Variable Name	Description	Example	Binary, Polynomial or Numeric	Units	Min	Max	Mean	Std. Deviation	Score
Expenditure on energy and fuels expenditure	General government expenditure by fuels and energy	0.53	Numeric	Percentage of GDP	−0.5	1.8	0.23	0.28	(Eurostat, 2020)
Environmental tax revenues (energy tax) of GDP	Proportion of environmental tax revenues in Gross Domestic Product (GDP)	0.01	Numeric	Percentage of GDP	0.97	3.05	1.94	0.4	(Eurostat, 2020)

- Policy Characteristics

Table A9. Policy Characteristics (1).

Variable Name	Description	Example	Binary, Polynomial or Numeric	Units	Min	Max	Mean	Std. Deviation	Source
Other related policies	Amount of RE policies included part of the policy mix	2	Numeric	Numeric Value	0	136	25.61	21.25	Own Calculated
Average duration of the programs	Average of the years the programs were designed to be active	15 years	Numeric	no units	4	16	11.38	2.89	Own Calculated
Average years of activity	Average of the years the programs have been active	5	Numeric	no units	1	15	6.28	3.62	Own Calculated
Median years of activity	Median of the years the programs have been active	5.5	Numeric	no units	1	15	6.51	4.06	Own Calculated
Number of RE programs	Total amount of RE programs that are active simultaneously	3 policies active at the analyzed year	Numeric	no units	1	14	3.02	2.73	Own Calculated
Age of oldest program	Age of the oldest policy program considered	1 or 0	Binary	no units	4	141	14.41	13.19	Own Calculated
Policy Type (2)	General program type class	General Programme	Polynomial	no units	n.a.	n.a.	n.a.	n.a.	(Fraunhofer Institute, 2020)
Policy Sub type (3)	Detailed program type class	General programme renewables	Polynomial	no units	n.a.	n.a.	n.a.	n.a.	(Fraunhofer Institute, 2020)

(1) Policy Characteristics: The variables of this group have not been included for the EPEI application; (2) Legislative/Normative; Legislative/Informative; Financial; Fiscal/Tariffs; Information/Education; Co-operative Measures; Cross-cutting; (3) Mandatory Standards for Buildings; Regulation for Heating Systems; Other Regulation in the Field of Buildings; Mandatory Standards for Appliances; Mandatory labelling; Mandatory energy efficiency certificates; Mandatory audits; Grants/Subsidies for investments; Grants/Subsidies for audits; Loans/Others; VAT Reduction; Income tax reduction; Linear electricity tariffs; Voluntary labelling; Information campaigns; Detailed energy/electrical bill; Regional and local information centers; Voluntary/Negotiated agreements; Voluntary DSM measures of suppliers; Technology procurement; Eco-tax on electricity/energy; Eco-tax on CO_2-emissions.

Table A10. Policy instruments and subtypes.

Policy Type	Policy Subtype
Legislative/Normative	Mandatory Standards for Buildings Regulation for Heating Systems Other Regulation in the Field of Buildings Mandatory Standards for Appliances
Legislative/Informative	Mandatory labelling Mandatory energy efficiency certificates Mandatory audits
Financial	Grants/Subsidies for investments Grants/Subsidies for audits Loans/Others
Fiscal/Tariffs	VAT Reduction Income tax reduction Linear electricity tariffs
Information/Education	Voluntary labelling Information campaigns Detailed energy/electrical bill Regional and local information centres
Co-operative Measures	Voluntary/Negotiated agreements Voluntary DSM measures of suppliers Technology procurement
Cross-cutting	Eco-tax on electricity/energy Eco-tax on CO_2-emissions

(Boonekamp & Piccioni, 2014).

References

1. EU Overachieves 2020 Renewable Energy Target—Products Eurostat News—Eurostat. 2022. Available online: https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20220119-1 (accessed on 2 May 2022).
2. Nitze, I.; Schulthess, U.; Asche, H. Comparison of machine learning algorithms random forest, artificial neuronal network and support vector machine to maximum likelihood for supervised crop type classification. In Proceedings of the 4th Conference on GEographic Object-Based Image Analysis—GEOBIA 2012, Salzburg, Austria, 7–9 May 2012.
3. The Royal Society. *Machine Learning: The Power and Promise of Computers that Learn by Example*; The Royal Society: London, UK, 2017.
4. Ghoddusi, H.; Creamer, G.G.; Rafizadeh, N. Machine learning in energy economics and finance: A review. *Energy Econ.* **2019**, *81*, 709–727. [CrossRef]
5. Ottesen, S.Ø.; Tomasgard, A.; Fleten, S.-E. Prosumer bidding and scheduling in electricity markets. *Energy* **2016**, *2016*, 94. [CrossRef]
6. Moutis, P.; Skarvelis-Kazakos, S.; Brucoli, M. Decision tree aided planning and energy balancing of planned community microgrids. *Appl. Energy* **2016**, *161*, 197–205. [CrossRef]
7. Kamali, S.; Amraee, T. Blackout prediction in interconnected electric energy systems considering generation re-dispatch and energy curtailment. *Appl. Energy* **2017**, *187*, 50–61. [CrossRef]
8. Aguado, J.A.; Racero, A.J.S.; de la Torre, S. Optimal operation of electric railways with renewable energy and electric storage systems. *IEEE Trans. Smart Grid* **2018**, *2018*, 9. [CrossRef]
9. Mosavi, A.; Salimi, M.; Ardabili, S.F.; Rabczuk, T.; Shamshirband, S.; Varkonyi-Koczy, A.R. State of the art of machine learning models in energy systems, a systematic review. *Energies* **2019**, *12*, 1301. [CrossRef]
10. Rokach, L.; Maimon, O. Decision Trees. *Tel. Aviv. Univ.* **2005**, *1897*, 44–45. [CrossRef]
11. Steuer, F. *Machine Learning for Public Policy Making How to Use Data-Driven Predictive Modeling for the Social Good*; Institut Batcelona: Barcelona, Spain, 2018.
12. Rudin, C.; Wagstaff, K.L. Machine learning for science and society. *Mach Learn.* **2014**, *95*, 1–9. [CrossRef]
13. Bush, V. Science, the Endless Frontier. United States Government Printing Office. 1945. Available online: https://www.nsf.gov/about/history/EndlessFrontier_w.pdf?linkId=81927236 (accessed on 2 May 2022).
14. Baesens, B. Analytics in a Big Data World. In *Wiley & SAS Business Series*; John Wiley & Sons: Hoboken, NJ, USA, 2014.
15. Zhang, Z.; Zhao, Z.; Yeom, D. Decision Tree Algorithm-Based Model and Computer Simulation for Evaluating the Effectiveness of Physical Education in Universities. *Hinday Complex.* **2020**, *2020*, 1–11. [CrossRef]
16. Song, Y.Y.; Lu, Y. Decision tree methods: Applications for classification and prediction. *Shanghai Arch. Psychiatry* **2015**, *27*, 130–135. [CrossRef]
17. ADEME. Aides à la Rénovation | Agir Pour la Transition Ecologique | ADEME. 2020. Available online: https://particuliers.ademe.fr/finances/aides-la-renovation (accessed on 7 July 2020).
18. Buchanan, B.; Miller, T. *Machine Learning for Policymakers*; Harvard Kennedy School: Cambridge, MA, USA, 2017.

19. Kim, E.S.; Choi, Y.; Byun, J. Big Data Analytics in Government: Improving Decision Making for R&D Investment in Korean SMEs. *Sustainability* **2020**, *12*, 202. [CrossRef]
20. Developments for Sustainability. *Deliverable n D3. 2—RS 2b "Final Set of Sustainability Criteria and Indicators for Assessment of Electricity Supply Options"*; Paul Scherrer Institut: Villigen, Switzerland, 2008.
21. Cambridge Dictionary. Effectiveness. 2018. Available online: https://dictionary.cambridge.org/es-LA/dictionary/essential-british-english/effectiveness (accessed on 12 September 2018).
22. Eurostat. *Energy, Transport and Environment Indicators*, 2017th ed.; Eurostat Pocketbook; Eurostat: Luxemburg, 2020; Volume 25.
23. European Commission. *EUROPE 2020: A Strategy for Smart, Sustainable and Inclusive Growth*; European Commission: Brussels, Belgium, 2020.
24. European Commission. Commission Awards 2003 Prizes to the Best European Projects in Renewable Energy. Available online: https://ec.europa.eu/commission/presscorner/detail/en/IP_04_76/ (accessed on 9 September 2020).
25. Cambridge Dictionary. Efficacy. 2018. Available online: https://dictionary.cambridge.org/es-LA/dictionary/essential-british-english/efficacy (accessed on 12 September 2018).
26. Ortiz, D.; Leal, V. Energy Policy Concerns, Objectives and Indicators: A Review towards a Framework for Effectiveness Assessment. *Energies* **2020**, *13*, 6533. [CrossRef]
27. Wright, S. The Method of Path Coefficients. *Ann. Math. Stat.* **1934**, *5*, 161–215. [CrossRef]
28. Lucertini, G. *Evaluating Public Policies Normative Models Beyond Cost Benefit Analysis*; Universita degli Studi di Padova: Padova, Italy, 2012.
29. Rottmann, M.; Maag, K.; Chan, R.; Huger, F.; Schlicht, P.; Gottschalk, H. *Detection of False Positive and False Negative Samples in Semantic Segmentation*; IEEE: Washington, DC, USA, 2020.
30. Debnath, R.; Darby, S.; Bardhan, R.; Mohaddes, K.; Sunikka-blank, M. Grounded reality meets machine learning: A deep-narrative analysis framework for energy policy research. *Energy Res. Soc. Sci.* **2020**, *69*, 101704. [CrossRef]
31. AEA Group. *Quantification of the Effects on Greenhouse Gas Emissions of Policies and Measures Final Report a Final Report to the European Commission AEA Technology*; AEA Group: Didcot, UK, 2009.
32. Zakari, A.; Li, G.; Khan, I.; Jindal, A.; Tawiah, V.; Alvarado, R. Are abundant energy resources and Chinese business a solution to environmental prosperity in Africa? *Energy Policy* **2022**, *163*, 112829. [CrossRef]
33. Zakari, A.; Khan, I.; Tawiah, V.; Alvarado, R. Reviewing the ecological footprints of Africa top carbon consumer: A quantile on quantile analysis. *Int. J. Environ. Sci. Technol.* **2022**. [CrossRef]
34. Marugán, A.P.; Márquez, F.P.; Perez, J.M.; Ruiz-Hernández, D. A survey of artificial neural network in wind energy systems. *Appl. Energy* **2018**, *228*, 1822–1836. [CrossRef]
35. Zaidi, S.M.A.; Chandola, V.; Allen, M.R.; Sanyal, J.; Stewart, R.N.; Bhaduri, B.L.; McManamay, R.A. Machine learning for energy-water nexus: Challenges and opportunities. *Big Earth Data* **2018**, *2*, 228–267. [CrossRef]
36. Zeng, Y.; Liu, J.; Sun, K.; Hu, L.-W. Machine learning based system performance prediction model for reactor control. *Ann. Nucl. Energy* **2018**, *113*, 270–278. [CrossRef]
37. Zendehboudi, S.; Rezaei, N.; Lohi, A. Applications of hybrid models in chemical, petroleum, and energy systems: A systematic review. *Appl. Energy* **2018**, *228*, 2539–2566. [CrossRef]
38. Anifowose, F.A.; Labadin, J.; Abdulraheem, A. Ensemble machine learning: An untapped modeling paradigm for petroleum reservoir characterization. *J. Pet. Sci. Eng.* **2017**, *151*, 480–487. [CrossRef]
39. Voyant, C.; Notton, G.; Kalogirou, S.; Nivet, M.-L.; Paoli, C.; Motte, F.; Fouilloy, A. Machine learning methods for solar radiation forecasting: A review. *Renew. Energy* **2017**, *105*, 569–582. [CrossRef]
40. Heinermann, J.; Kramer, O. Machine learning ensembles for wind power prediction. *Renew. Energy* **2016**, *89*, 671–679. [CrossRef]
41. Fulford, D.S.; Bowie, B.; Berry, M.E.; Bowen, B.; Turk, D.W. Machine learning as a reliable technology for evaluating time/rate performance of unconventional wells. *SPE Econ. Manag.* **2016**, *8*, 23–39. [CrossRef]
42. Gupta, S.; Kambli, R.; Wagh, S.; Kazi, F. Support-vector-machine-based proactive cascade prediction in smart grid using probabilistic framework. *IEEE Trans. Ind. Electron.* **2015**, *62*, 2478–2486. [CrossRef]
43. Costa, D.C.; Nunes, M.V.; Vieira, J.P.; Bezerra, U.H. Decision tree-based security dispatch application in integrated electric power and natural-gas networks. *Electr. Power Syst. Res.* **2016**, *2016*, 141. [CrossRef]