sustainability

MDPI

Article

Incremental Green Investment Rule Induction Using Intelligent Rough Sets from an Energy Perspective

Chun-Che Huang [1], Wen-Yau Liang [2,*], Horng-Fu Chuang [3], Tzu-Liang (Bill) Tseng [4] and Yi-Chun Shen [1]

[1] Department of Information Management, National Chi Nan University, 1 University Road, Nantou 545301, Taiwan
[2] Department of Information Management, National Changhua University of Education, 1 Jin-De Road, Changhua 500207, Taiwan
[3] Department of Accounting Information, Da-Yeh University, No. 168, University Rd., Changhua 515006, Taiwan
[4] Department of Industrial Manufacturing and Systems Engineering, The University of Texas at El Paso, 500 West University Avenue, El Paso, TX 79968, USA
* Correspondence: wyliang@cc.ncue.edu.tw

Abstract: The United Nations Sustainable Development Goals (SDGs) agenda has stated the importance of green investment. Energy-related green investment involves intricate economic behavior and ecological objectives. Green investment definitely requires agile decisions, e.g., rule-based decisions, to respond to changes outside the country. The identification of significant rules with numerous result features and the assurance of the stability and robustness of the rules in decision-making are crucial for green energy investment. The rough set (RS) methodology works well for processing qualitative data that are difficult to examine with traditional statistical methods in order to induce decision rules. The RS methodology starts with the analysis of the limits of discernibility of a subset of objects belonging to the domain to induce rules. However, traditional RS methods cannot incrementally generate rules with outcome features when new objects are added, which frequently occurs in green energy investment with the inclusion of big data. In this paper, an intelligent RS approach is proposed. This approach effectively identifies the rules that either stay the same or are altered based on four classified cases after a new object is introduced; it is novel because it can deal with a complicated investment environment by imposing multiple outcome features, specifically when it is required to flexibly extract new decision rules via adding new data sets.

Keywords: green energy; green investment; incremental technique; rough set; rule induction; multiple outcomes; big data

Citation: Huang, C.-C.; Liang, W.-Y.; Chuang, H.-F.; Tseng, T.-L.; Shen, Y.-C. Incremental Green Investment Rule Induction Using Intelligent Rough Sets from an Energy Perspective. *Sustainability* **2024**, *16*, 3655. https://doi.org/10.3390/su16093655

Academic Editor: Bruce Morley

Received: 25 February 2024
Revised: 18 April 2024
Accepted: 24 April 2024
Published: 26 April 2024

1. Introduction

As society becomes more urbanized, environmental issues are receiving more and more attention. Specifically, the ensuing introduction of dangerous and poisonous chemicals in the environment has sparked worries around the world [1]. Shen et al. investigated the effects of investing in green energy and discovered that it can improve the quality and sustainability of the environment [2]. According to different perspectives, green energy investment is also known as environmentally friendly investment; socially or sustainably responsible investment; environmental, social, and governance investing (ESG); or responsible investing (RI) [3]. Briefly, an environmentally friendly investment makes a positive impact on the environment and focuses on reducing negative environmental effects. Renewable energy sources and energy-efficient technologies are examples of environmentally friendly investment. Socially responsible investing requires that business not only focus on their financial earnings but also put efforts on social and environmental impact, striking a balance between investors' profit and ethical considerations. ESG investment focuses on three aspects, environmental, social, and governance, aiming to consider investment with regard to its impact on society and the environment. Responsible investing is a strategy focusing on profits for both investors and society while making investment decisions.

Investment in green energy has become crucial for economies and society to achieve sustainable development in the modern era, as environmental challenges continue to rise [4]. The role of green investment has been emphasized in the Sustainable Development Goals (SDGs) agenda of the United Nations as well [5]. However, investment data may be obtained from online surveys and involve qualitative data. The rough set (RS) methodology works well for processing qualitative data that are hard to examine with traditional statistical methods.

The rough set theory (RST) is an important theoretical foundation for attribute reduction problems [6]. RST is a useful mathematical technique for handling inconsistent and ambiguous data according to Pawlak [7]. Most academics gradually became aware of rough set theory due to its benefits in handling imprecise and partial data. The main goal of rough set attribute reduction is to reduce the number of unneeded attributes in the attribute set gradually in order to expedite the processing of data [8]. As rough sets are successfully applied in a variety of disciplines, new research findings on the theory are being discovered continuously [8], such as in project management [9], dynamical systems [10], food safety [11], hyperspectral band [12], linear regression model [13], mobile phones [14], decision support [15], multiset-valued information system, MSVIS [16], and occupational risk prediction [17].

Previous RS methods used for energy include energy usage [18–20], hybrid electric vehicles, HEV [21], and water quality [22]. However, traditional RS studies of energy have not considered multiple outcome features. In addition, old-fashioned rough set methods cannot generate rules that have a preference order, meaning that they cannot generate more meaningful and general rules. In order to solve this problem, multiple outcomes rule extraction (MORE) and multiple outcomes reduct generation (MORG) algorithms based on rough sets are required, specifically in dynamic big data sets [23].

Green investment research is increasingly incorporating big data analysis [24]. Big data is prone to the same types of biases as is data analysis utilizing different data sources since it neither modifies the nature of knowledge nor applies fundamentally novel techniques for extracting knowledge from the data. On the other hand, big data has the ability to increase the pace and volume of knowledge creation. In the big data world, investment data may be added because of multiple decision-making factors related to return on investment, with attention being paid to environmental, social, and governance (ESG) values in investing [25,26]. Consequently, the evaluation criteria have an impact on decision-making when further facts are added. The need to create incremental systems for managing databases and rule sets arises from the assessment criteria's costly computational complexity if rule extraction procedures have to be executed repeatedly when new data are added [27]. Therefore, an intelligent incremental RS-based approach with multiple outcome features for rule extraction is focused on in this study on green energy investment in order to extract decision rules, reduce time complexity, and reduce the data or feature space.

The research problem based on the gap in the existing literature is that in the agile environment, rule induction must be flexible and efficient in order to assist decision-making since data are changed dynamically and since the nature of multiple outcomes need to be focused on. In the case that new objects are added in, traditional RS approaches have to compute the whole database again, and it is time-consuming. These are the emergent problems to be solved. In this paper, an incremental rough-set-based approach that considers multiple outcome features for rule extraction, IMORE (incremental multiple outcomes rule extraction) is developed to overcome these problems. This study aims to extend the traditional rough set theory to deal with multiple outcomes in a situation where incremental data sets are added. The proposed approach is applied to a green energy investment case to solve a dynamic database problem. This study is novel since it provides a quick method for the extraction of decision rules that are important to investors and financial advisors. When there are multiple types of customers and different types of decision-making in an information system, the proposed approach can flexibly generate stable and effective rules. This contributes to understanding the significance of making green energy expenditures within enterprises and to identifying the factors that influence these decisions in the big data world [28].

2. Literature Review

This section surveys the literature on green investment to identify the specific research questions in Section 2.1. Consequently, in Section 2.2, the rule induction approaches based on the rough set theory are reviewed to contribute to the understanding of the existing gap and how the research problem can be solved. Section 2.3 summarizes the findings based on the literature review.

2.1. Green Investment in the Big Data Era

Climate change stands out as a crucial environmental concern, sparking widespread discussions in both academic circles and practical spheres. The detrimental impacts of greenhouse gases, landfill pollutants, and water contamination have resulted in significant environmental deterioration, compelling urgent interventions from both industrial leaders and governmental authorities. For example, policymakers can reduce emission activities based on investing in green technology [29]. Green energy investment is specifically that made to reduce industrial emissions and lessen environmental contamination [30]. Green energy investment is an economic behavior and complex management process, as it is not easy to achieve both ecological goals and economic benefits. Green energy investment is a novel resource allocation strategy that allocates limited resources toward the advancement of renewable resources and green technology [31]; however, only when green technology is profitable would firms be pushed to make green investments [32]. The interest in green energy investment among academics and practitioners has increased due to its promising impact on environmental improvements that support the ecological environment [33].

Previous studies have formulated green investment as a problem of multiple-criteria decision-making (MCDM), for example in [34–36]. Based on sub-discipline of operations research, MCDM explicitly evaluates multiple conflicting criteria in decision-making. In this way, the solutions to green investment can be explored and solved. However, the MCDM methods might have some limitations, such subjectivity, trade-offs and conflicting objectives, sensitivity to methodology, and lack of transparency [37]. Specifically, complexity is increased in cases where the data are more diversified and enlarged, which makes traditional MCDM methods difficult to apply in the big data era, and we know that green investment research is increasingly incorporating big data analysis [24]. For example, Wang et al. found out that big data technology makes it feasible to analyze enormous and complex data sets, giving businesses the ability to identify business innovation opportunities and improve their environmentally friendly operations [38]. Moreover, outcomes of big data analysis can facilitate access to financial resources for enterprises and stimulate environmental investment. McAfee et al. highlighted significant publications, academics, research findings, and areas of future study for green finance and energy policy, mentioning the possibility of future research involving financial technology, big data, and blockchain integration [39].

Big data is merely the continuation of data-driven studies or business analytics, which have long been utilized in academia and industry to provide market information and intelligence. Big data refers to the availability of much greater, frequently enormous volumes of data (volume) from a variety of different data sources (variety) continuously and frequently in real time (velocity). That is, in such a big data world, market change may occur often and time by time in green energy investment [40]. In addition, energy-related green investment is going through a significant paradigm shift that is moving it away from traditional management and toward an agile environment [41]. An agile company must be able to react swiftly to changes in the market, and swift decision-making support methods are essential [41,42]. However, since data are periodically altered dynamically, obtaining relevant, consistent, and up-to-date information across a large firm is a difficult and time-consuming procedure. Given this, rule induction must be flexible and efficient in order to assist decision-making because data are changed dynamically. (i.e., add-in and remove-out) at the implementation stage. To explore green energy investment decision rules, research topics in the green investment domain are various, including the quantitative approach [4,43], rule induction [44], data mining, and knowledge discovery [45,46]. The application of mathematical functions to model the quantitative links between incidents in green investment data has been the main focus of these investigations. Other examples of quantitative approaches focus on the effects of renewable energy investment resources and green finance on economic performance [46], business performance [47], and the effects of green banking and green investment on firm value [48].

2.2. Rule Induction Based on Rough Sets

Quantitative approaches in green investment analysis have their limitations in terms of their validity, the researcher's role, the academic impact, and the data-gathering process [49]. In addition, the green investment sector often produces a large amount of qualitative data related to online surveys [50,51], case studies and empirical research approaches in e-commerce applications [52,53], and in-depth interviews [38,54]. Since the data are qualitative, it is challenging to analyze them using conventional statistical methods [55]. Qualitative models are complicated systems at a higher degree of abstraction, enabling the modeling of systems that are too complicated to be represented with traditional methods [56]. The rough set (RS) approach functions well in processing qualitative data to induce decision rules. It has been used as a method for uncovering logical patterns concealed in large amounts of data. A rough set technique can produce decision rules by extracting valuable information from a collection of mixed data [57].

The RS approach derives decision rules by an inductive process using a decision table with rows representing objects and columns containing characteristics or criteria. The minimal subset of characteristics is generated in order to provide the same classification of universe elements as the entire collection of features. This is referred to as a reduct. However, the majority of rough set techniques disregard the problem of multiple result features, so they are not efficient and the nature of the rules is not easily understood [58]. For instance, the standard reduct generation process needs to be carried out twice for each outcome in a data set containing three characteristics and two outcomes. Furthermore, a decision rule constantly has a single potential result. The similarities between the rules in these two results are challenging to comprehend. For example, the outcome (1, 2) is different from (1, 3), but both should be included in (1, *).

In addition, there are relatively few statistics, funds, and qualitative details available for green investment initiatives, and there are not many defined guidelines for compiling these data [59,60]. The rough set (RS) approach also functions well in processing incomplete data to induce decision rules. As RS theory research's central issue is attribute reduction, it has garnered considerable attention since it can efficiently lower the dimension of data and produce reduction results with a clear semantic interpretation [61].

Specifically, due the dynamic structure in which data are added according to the big data era, an incremental rough-set-based approach to induce rules that considers multiple outcome features is desirable when addressing problems in the green energy investment field. Such an incremental technique allows new data to be added without re-implementing the algorithm in a dynamic database. In general, the incremental attribute reduction approach is used as the attribute reduction procedure for dynamic data sets [62]. The incremental attribute reduction method has garnered a great deal of interest because it may efficiently use the reduction results that have already been achieved, saving a significant amount of time and space [8]. There have been numerous studies on incremental rough set theory, e.g., in feature selection [63–66], incremental approximation calculation [67], incremental information [68], rule discovery [69], and case-based reasoning [70], where the incremental technique allows new data to be added without re-implementing the algorithm in a dynamic database. Nevertheless, the majority of rough set approaches fail to take into consideration the problem of many result attributes, making them ineffective and ill-suited to understanding the nature of the rules in the big data era. In addition, these studies of incremental rough set approaches do not use data sets with multiple outcome features. Therefore, this paper proposes a novel incremental rough-set-based approach to deal with multiple outcome features and decision rules with multiple outcomes which often arise in green energy investment.

2.3. Summary

The findings of the literature review are summarized as follows:

1. Green investment research is increasingly incorporating big data analysis. Energy-related green investment is going through a significant paradigm shift toward an agile environment that makes decisions more often via rules that are quick to respond to the changes in the market. Rule induction must be flexible and efficient to assist decision-making because data are changed dynamically.
2. According to agile decision-making, data are used to induce decision rules first. However, a decision rule may not consist of a single outcome. Therefore, if there are multiple decisions, multiple outcome rule extraction (MORE) is required to generate stable decision rules that traditional RS approaches cannot.
3. Most current rough set approaches do not consider the issue of utilizing a dynamic database. Sometimes, the rules generated by the rough set approach fail to predict newly entered objects because of non-deterministic rules.
4. The existing algorithms of the rough set have the ability to generate a set of classification rules efficiently, but they cannot generate rules incrementally when new objects are added. In practical applications, the number of recorders in the database is often increased dynamically. Thus, if a new object is added in, the traditional RS approaches have to compute the whole database again. This process consumes a huge amount of computation time and memory space.

Therefore, the solution approach, an incremental multiple outcome rule extraction algorithm, is proposed next.

3. The Solution Approach

In this paper, decision rules are extracted using a multiple-outcome table that depicts the link between condition attributes and choice outcomes. Table 1 displays the various result sets of the relevant tuples as O_j, and the element (e_{ij}) indicates the value of the feature (F_m) that an object (tuple) (X_i) possesses. This table also includes the number of values. The proposed incremental solution approach is developed to induce rules with multiple outcomes, whereas traditional approaches only consider a single outcome and data set without change, e.g., new data add-in.

Table 1. Fundamental structure of a multiple-outcome table.

Condition Features (F_j)					Decision Outcomes (O_p)				
Object (X_i)	**F_1**	**F_2**	...	**F_j**	**Outcome(O_i)**	**O_1**	**O_2**	...	**O_p**
1	e_{11}	e_{12}	...	e_{1j}	1	o_{11}	o_{12}	...	o_{1p}
2	e_{21}	e_{22}	...	e_{2j}	2	o_{21}	o_{22}	...	o_{2p}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n	e_{n1}	e_{n2}	...	e_{nj}	n	o_{q1}	o_{q2}	...	o_{qp}

3.1. The Proposed Incremental Approach

Note that in the proposed approach, a new data set often refers to an object. Since most databases contain ever-changing objects that are frequently added, removed, or modified, an effective approach for practical applications must make it easy to change the objects [71,72]. After an object is introduced, the rules either stay the same or change. The regulation might be altered as a result of being updated or created. The four types of results are classified in Table 2.

Table 2. The results of a new object being added in.

	Case I	**Case II**	**Case III**	**Case IV**
Is any original rule transformed?	No	Yes	No	Yes
Is any new rule generated?	No	No	Yes	Yes

Case I. Original rules are not transformed because the new data set does not conflict with the original rule set.
Case II. Original rules are transformed because the new data set conflicts with the original rule set. However, new rules are not generated.
Case III. New rules are generated by the new data set, but the original rules are not transformed.
Case IV. New rules are generated by the new data set. The original rules are transformed because the new data set conflicts with the original rule set.

The proposed IMORE algorithm uses the multiple outcome rule extract (MORE) algorithm and the multiple outcomes reduct generation (MORG) algorithm [23]. A portion of the original rule sets are modified in order to update the rules. Next, the IMORE procedure is presented with the following notations:

- T_i: original data set i;
- N_j: new data set j;
- R_k: original reduct set k;
- R_{new}: new reduct set;
- R_{merge}: reduct that has to merge;
- R_{add}: newly generated reduct;
- R_{tra}: reduct that has to transform;
- R_{tmp}: a temporary set to place the reduct;
- F_a: feature set;
- O_b: outcome set;
- m: the number of original data set;
- p: the number of feature s;
- q: the number of outcomes;
- r: the number of reduct data sets;
- t: the number of R_{tmp} with t = 0;
- f: feature value with f = 0;
- x: constant with x = false.

IMORE procedure:

```
Step 0. Add the new data set
Step 1. Check if new data sets are identical to the original data set.
         For i = 1 to m
For a = 1 to p
         If Nj(Fa) conflicts with Ti(Fa)
For b = 1 to q
If Nj(Ob) conflicts with Ti(Ob)
          Go to Step 2 //Case I
          Else go to Step 3 //Case II
End if
            Else x = true
End if
End for
If x = true
                      Go to Step 4 //Case III
End if
          End for
Step 2. For Case II, obtain reducts
Step 3. For Case III, obtain reducts
Step 4. For Case III, obtain reducts
Step 5. For Case III, obtain reducts
Step 6. Rule extraction
Step 6.1. Compute the number of identified values from features of each reduct rule
            (including new reducts, conflicting reducts, and merged reducts)
       For a = 1 to s
For k = 1 to r
If Rk(Fa) != "*"
f++
End if
End for
End for
          Step 6.2. Compute the SI of all reducts
          Step 6.3. If the objects have more than one reduct
The reduct with the maximal SI from each object (or each merge objects) is the final rule
                  of this object.
End if
Step 6.4. If the reduct and the sum of feature identical values are the same
Select first one
End if
Step 6.5. Update the final rule.
END
```

The details for each case to obtain reducts are as follows:

The column "number of objects" from the primary rules, which is identical to the new data set, is merged for Case I if the original object data set and the new object set are the same. The object number is the only element altered in the final rule data, and the original rule sets remain unchanged.

Case I: Data change, but the original rules do not change.

Step 2.1. add T_i and T_{m+1} to R_{merge}
Step 2.2. $R_{new} = \{ R_k + R_{merge} \}$
Step 2.3. Update the final rule

In Case II, a conflict will occur when the features of the incremental object set and the original data set are the same, yet the outcomes are not identical. In these instances, the original object set and the new data set merge together, and the conflict between the original reductions and the new data set is assessed. The reducts of the conflicting objects will be computed again if there is a conflict between the initial reducts and the updated data set. The SI and the number of identified values for every feature row are computed. The identical reducts are combined into a new reduct rule if there are any identical reducts in the multiple-outcomes table. The maximal SI reduct is chosen if the table contains no identical reducts. The initial reduction of an object is selected if there is not likewise a result of this kind. Next, the final rule sets are revised.

Case II: The features of the new data set are identical to the features of the original data set, but the outcomes are not identical.

Step 3.1. Add the new data set into the raw data set
Step 3.2. Merge the reduct of the new data set
reduct T_i and $T_{m+1\ \text{and}}$ then add to R_{merge}
Step 3.3. Re-compute the reduct of the object that generates the rule
For $k = 1$ to r
For $a = 1$ to p
If $Rk(Fa)$ dominated to R_{merge} (Fa)
Reduct Tk to R_{tra}
End if
End for
End for
Step 3.4. $R_{new} = \{ R_k + R_{merge} + R_{tra} \}$
Step 3.5. Merge the identical reducts
If $\{ R_{merge} + R_{tra} \}$ has identical reducts to R_k
Merge the identical reducts with a new reduct
End if
Step 3.6. Go to Step 6

This condition will be categorized as Case III or IV if the decision outcomes and condition attributes of a new data collection differ from those of the original data sets. The original reductions and the new data set are options. Next, the object that contradicts both the new reduct's condition features and the condition features of the original reduct is added to a buffer, known as a temp. The object in the temp is removed if the decision outcomes it contains are the same as those of the new data set. When the temperature is finally measured, Case III is identified if it is null.

Case III: Data are changed, and new reducts are generated

Step 4.1. Add the new data set into the raw data set
Step 4.2. Reduct N_j to R_{add}
Step 4.3. $R_{new} = \{ R_k + R_{add} \}$
Step 4.4. Find the transformed reducts in new reduct
For $k = 1$ to r
If $Rk(Fa)$ conflicted with T_{m+1} (Fa)
add Rk to R_{tmp}
$t{+}{+}$
End if
End for
For $tmp = 1$ to t
If $Rtmp(Ob)$ conflicts with T_{m+1} (Ob)
Delete $Rtmp$
End if
End for
If $Rtmp! = \varnothing$
Go to Step 5 //Case IV
End if
Step 4.5. Merge the identical reducts
If $\{ R_{add} \}$ has identical reducts in R_k
Merge the identical reducts with a new reduct
End if
Step 4.6. Go to Step 6

Case IV is similar to Case III, yet it is more complex. First, the reduct is detected and recalculated in the temp. The strength index (SI), which was developed to identify meaningful reducts, is then computed based on the condition aspects of each reduct and the number of identified values as follows [73]:

$$\mathrm{SI}(f) = \frac{\sum_{j=1}^{m} vjWj \times nf}{\sum_{j=1}^{m} vj}$$

where f is the reduct number, $f = 1, \ldots, n$; $v_j = 1$ if condition attribute j is selected and 0 if v_j = "x"; w_j is the weight of condition attribute j; and n_f is the number of identical reducts, f.

The same reducts are combined with a new reduct if there are any identical reducts in the changed reducts. The greatest SI reduct of each object is chosen if there are no identical reducts in the transformed reduct. The first reduction of each item is chosen, and the final rule sets are changed if there is still no such outcome. The rest of the computation process is the same as that in Case II.

Case IV: Generate new reducts and transform original reducts

Step 5.1. Reduce Tt to R_{tra}
Step 5.2. R_{new} = { R_k + R_{tra} + R_{add} }
Step 5.3. Merge the identical reducts
If { R_{tra} + R_{add} } has identical reducts in R_k
Merge the identical reducts with a new reduct
End if
Step 5.4. Go to Step 6

3.2. Illustrative Examples

The following example illustrates the proposed incremental algorithm. The original information is described in Table 3. Table 4 shows the reduct table from the original information. Tables 5 and 6 describe the results of applying the aforementioned reduct generation procedure of the multiple outcome rule extract algorithm (MORE).

Table 3. The original information.

Object No.	F_1	F_2	F_3	F_4	O_1	O_2	O_3	Object Cardinality
X_1	L	N	YA	N	I	Y	Y	6
X_2	L	N	YA	N	O	Y	N	9
X_3	L	Y	MA	N	I	N	N	7
X_4	L	N	YA	Y	O	Y	N	5
X_5	M	Y	OA	Y	I	N	N	7
X_6	M	N	OA	Y	O	N	Y	3

L: Low, M: Medium, N: No, Y: Yes, YA: Young Age, MA: Middle Age, OA: Old Age, I: In, O: Out. * refers to that it does not matter.

Table 4. The list of resulting value reducts.

Object No.	Reduct No.	F_1	F_2	F_3	F_4	O_1	O_2	O_3	Object Cardinality
X_1 & X_2	1	*	*	YA	*	*	Y	*	15
X_3	1	*	Y	*	*	I	N	N	7
	2	*	*	MA	*				
X_4	1	L	*	*	Y	O	Y	N	5
	2	*	*	YA	Y				
X_5	1	*	Y	*	*	I	N	N	7
X_6	1	M	N	*	*	O	N	Y	3
	2	*	N	OA	*				

L: Low, M: Medium, N: No, Y: Yes, YA: Young Age, MA: Middle Age, OA: Old Age, I: In, O: Out. * refers to that it does not matter.

Table 5. The final rule table.

Object No.	Rule No.	F_1	F_2	F_3	F_4	O_1	O_2	O_3	Object Cardinality
X_1 & X_2	1	*	*	YA	*	*	Y	*	15
X_3 & X_5	2	*	Y	*	*	I	N	N	14
X_4	3	*	*	YA	Y	O	Y	N	5
X_6	4	*	N	OA	*	O	N	Y	3

L: Low, M: Medium, N: No, Y: Yes, YA: Young Age, MA: Middle Age, OA: Old Age, I: In, O: Out. * refers to that it does not matter.

Table 6. The resulting concise rules.

F_3 YA $\rightarrow$ O_2Y
F_2 Y $\rightarrow$ O_1 I O_2 N O_3 N
F_3 YA F_4 Y $\rightarrow$ O_1 O O_2 Y O_3 N
F_2 N F_3 OA $\rightarrow$ O_1 O O_2 N O_3 Y

- Case I: Data change, but the original rules do not change

  The illustrative data are described in Table 7.

Table 7. The incremental data in Case I.

Object No.	F_1	F_2	F_3	F_4	O_1	O_2	O_3	Object Cardinality
X_7	M	Y	OA	Y	I	N	N	7

In this case, it is obvious that the incremental data set is identical to the original data set 5. The situation is disposed by the following steps.

Step 0. Add the new data set (object 7) into the raw data set.
Step 1. Check if the new data sets are identical to the original data set. Go to Step 2.
Step 2. The data are changed, but the original rules are not changed. Merge the identical object 5 with the new object for reduction (see Table 8).

Table 8. Merge the original data set with new data set.

Object No.	Reduct No.	F_1	F_2	F_3	F_4	O_1	O_2	O_3	Object Cardinality
X_1 & X_2	1	*	*	YA	*	*	Y	*	15
X_3	1	*	Y	*	*	I	N	N	7
	2	*	*	MA	*				
X_4	1	L	*	*	Y	O	Y	N	5
	2	*	*	YA	Y				
X_5 & X_7	1	*	Y	*	*	I	N	N	14
X_6	1	M	N	*	*	O	N	Y	3
	2	*	N	OA	*				

* refers to that it does not matter.

Step 2.3. Update the final rule (see Table 9).

Table 9. The final result rule.

Object No.	Rule No.	F_1	F_2	F_3	F_4	O_1	O_2	O_3	Object Cardinality
X_1 & X_2	1	*	*	YA	*	*	Y	*	15
X_3 & X_5 & X_7	2	*	Y	*	*	I	N	N	21
X_4	3	*	*	YA	Y	O	Y	N	5
X_6	4	*	N	OA	*	O	N	Y	3

* refers to that it does not matter.

- Case II: Data cause original rules to be changed, but no new rules are generated.

The incremental information is described in Table 10, where objects 1 to 6 are the original data, and object 8 is the incremental data.

Table 10. The incremental data in Case II.

Object No.	F_1	F_2	F_3	F_4	O_1	O_2	O_3	Object Cardinality
X_8	M	Y	OA	Y	O	Y	N	7

The features of the new data set are identical to object data set 5. The affected reduct is found, and rules are extracted.

Step 0. Add the new data set (object 8) into the raw data set.
Step 1. Check if the new data set is identical to the original data sets. Find the object with features identical to those of the new data set and then go to Step 3.
Step 3. The features of the new data set are identical to the features of the original data set, but the outcomes are not identical.
Step 3.2 Merge and re-compute the reduct identical to object 5 with the new object (see Table 11).

Step 3.3. Find the dominant reduct in the new data set and re-compute it. Then, proceed to Step 6.
Step 6. Rule extraction
Step 6.1. Compute the number of identified values from the features of each reduct.
Step 6.2. Compute the SI of all object reducts (see Table 12).

Table 11. Merging of the reducts of the original dataset with a new data set.

Object No.	Reduct No.	F_1	F_2	F_3	F_4	O_1	O_2	O_3	Object Cardinality
X_1 & X_2	1	*	*	YA	*	*	Y	*	15
X_3	1	*	Y	*	*	I	N	N	7
	2	*	*	MA	*				
X_4	1	L	*	*	Y	O	Y	N	5
	2	*	*	YA	Y				
X_5 & X_8	1	*	Y	*	*	*	*	N	14
X_6	1	M	N	*	*	O	N	Y	3
	2	*	N	OA	*				

* refers to that it does not matter.

Table 12. The final reducts from Steps 6.1 and 6.2.

Object No.	Reduct No.	F_1	F_2	F_3	F_4	O_1	O_2	O_3	SI
X_1 & X_2	1	*	*	YA	*	*	Y	*	4
X_3	1	*	Y	*	*	I	N	N	3
	2	*	*	MA	*				4
X_4	1	L	*	*	Y	O	Y	N	4
	2	*	*	YA	Y				6
X_5 & X_8	1	*	Y	*	*	*	*	N	3
X_6	1	M	N	*	*	O	N	Y	5
	2	*	N	OA	*				7
Identified value of features		2	3	4	2				

* refers to that it does not matter.

Step 6.3. If the objects have more than one reduct, then select the reduct with the maximal SI.
Step 6.4. Update the final rule (see Table 13).

Table 13. The summarized result from Steps 6.3 and 6.4.

Object No.	Rule No.	F_1	F_2	F_3	F_4	O_1	O_2	O_3	Object Cardinality
X_1 & X_2	1	*	*	YA	*	*	Y	*	15
X_3	2	*	*	MA	*	I	N	N	7
X_4	3	*	*	YA	Y	O	Y	N	5
X_5 & X_8	4	*	Y	*	*	*	*	N	14
X_6	5	*	N	OA	*	O	N	Y	3

* refers to that it does not matter.

- Case IV: New rules are generated, and original reducts are changed

The incremental information is described in Table 14, where object 1 to object 6 are the original data, and object 9 is the incremental data.

Table 14. The incremental data in Case IV.

Object No.	F_1	F_2	F_3	F_4	O_1	O_2	O_3	Object Cardinality
X_9	M	N	YA	N	O	N	Y	7

The new data set does not have identical data to those of the original data sets. The affected reduction is found, and the rules are extracted.

Step 0. Add the new data set (object 9) into the raw data set.
Step 1. Check if the new data sets are identical to the original data sets and then go to Step 4.
Step 4. Data are changed, and new rules are generated.
Step 4.2 Generate new reducts (see Table 15).

Table 15. Generation of new reducts from the original dataset with the new data set.

Object No.	Reduct No.	F_1	F_2	F_3	F_4	O_1	O_2	O_3	Object Cardinality
X_1 & X_2	1	*	*	YA	*	*	Y	*	15
X_3	1	*	Y	*	*	I	N	N	7
	2	*	*	MA	*				
X_4	1	L	*	*	Y	O	Y	N	5
	2	*	*	YA	Y				
X_5	1	*	Y	*	*	I	N	N	7
X_6	1	M	N	*	*	O	N	Y	3
	2	*	N	OA	*				
X_9	1	M	N	*	*	O	N	Y	7
	2	M	*	YA	*				
	3	M	*	*	N				

* refers to that it does not matter.

Step 4.4. Find the replaced reducts between the reducts and the new data set. Add the object of the conflicted features into temp and delete the object of the conflicted outcomes from the temp (see Table 16).

Table 16. The table to store the reduct that conflicts with features from Step 4.4.

Object No.	Reduct No.	F_1	F_2	F_3	F_4	O_1	O_2	O_3	Object Cardinality
X_1 & X_2	1	*	*	YA	*	*	Y	*	15
X_6	1	M	N	*	*	O	N	Y	3

* refers to that it does not matter.

Step 6. Rule extraction
Step 6.1. Compute the number of identified values from features of each reduct.
Step 6.2. Compute the SI of all object reducts (see Table 17).

Table 17. The 2nd final reducts from Steps 6.1 and 6.2.

Object No.	Reduct No.	F_1	F_2	F_3	F_4	O_1	O_2	O_3	No	SI
X_1 & X_2	1	L	N	*	*	X	Y	X	15	13
	2	L	*	YA	*					12
X_3	1	*	Y	*	*	I	N	N	7	6
	2	*	*	MA	*					5
X_4	1	L	*	*	Y	O	Y	N	5	10
	2	*	*	YA	Y					8

Table 17. *Cont.*

Object No.	Reduct No.	F_1	F_2	F_3	F_4	O_1	O_2	O_3	No	SI
X_5	1	*	Y	*	*	I	N	N	14	6
X_6	1	M	N	*	*	O	N	Y	3	13
	2	*	N	OA	*					11
X_9	1	M	N	*	*	O	N	Y	7	13
	2	M	*	YA	*					12
	3	M	*	*	N					10
Identify value of features		7	6	5	3					

* refers to that it does not matter.

Step 6.3. If the objects have more than one reduct, then select the reduct with the maximal SI.
Step 6.4. Update the final rule (see Table 18).

Table 18. The 2nd summarized result from Steps 6.3 and 6.4.

Object No.	Rule No.	F_1	F_2	F_3	F_4	O_1	O_2	O_3	Object Cardinality
X_1 & X_2	1	L	N	*	*	*	Y	*	15
X_3 & X_5	2	*	Y	*	*	I	N	N	14
X_4	3	L	*	*	Y	O	Y	N	5
X_6 & X_9	4	M	N	*	*	O	N	Y	3

* refers to that it does not matter.

From the case study of the IMORE incremental algorithm, the results of Tables 9, 13 and 18 are identified with the MORE algorithm.

3.3. The Time Complexity

The time complexity of this proposed approach is composed of four cases. The algorithm adds the new data accordingly via a loop and then obtains the reducts when a set of data is added to the original data set. According to this method, the technique presented in this paper reduces the time complexity while simultaneously solving incremental problems. The algorithm contains pre-processing time. The time complexity of Case II is the worst case in the whole IMORE algorithm (Table 19).

Table 19. Time complexity.

Case Number	Description	Time Complexity in the Worst Case
I	Data are changed, but original rules do not change.	O (mpq(n))
II	That Data is changed causes original rules to be changed. No new rules are generated.	O (mpq(n) + rp(Nncor))
III	Data is changed and new reducts are generated.	O (mp + r + t(Nr))
IV	New rules are generated and original reduct rules are changed	O (mp + r + t(Nr + Ndr))

m: number of original data sets; p: number of feature s; q: number of outcomes; r: number of reduct data sets; t: number of R_{tmp}; n:denote the number of new data sets; Nncor: denotes the total number of new reducts from the object that generated the reduct that conflict with the new data set; Nr: denotes the total number of new generated reducts; Ndr: denotes the total number of reducts that conflict with the new data set.

When new objects are added to the information system without the IMORE algorithm, the time complexity in the worst case is $O((mqr(mpq)) + pqp(r + Nncor))$. The worst case of the application of the proposed algorithm is in Case II. The time complexity of the MORE algorithm is $O((mqr(mq)) + p^2q(r + Nr + Nnco))$. In a comparison of these two complexities, it is obvious that the proposed one is efficient. Since "*m*" denotes the total number of objects, in the worst case, the number must be less than or equal to *m*. "*p*" denotes the features, "*q*" denotes the outcomes of the features, and "mpg" indicates that the number is bigger than "m". Moreover, "*pqp*" in MORE is bigger than "*rp*" in Case II. Therefore, the proposed algorithm surpasses the one without the IMORE algorithm.

The proposed solution approach induces decision rules incrementally. However, some limitations are observed:

- The IMORG algorithm focuses on particular data being added-in. The data move-out situation is not considered in this study and could be the focus of future work.
- The number of data sets added in cannot be greater than the number of original data sets. Otherwise, the computation time will be not be less than that of the MORE algorithm. To solve this situation, new data sets added in can be divided into two sets and implemented twice.
- The solution approach does not consider data which are typos or missing or otherwise in need of pre-processing. To implement this IMORE algorithm, data should be cleaned first.

4. Case Study of Green Investment in Energy

To observe the relationship of investors' characteristics and their preferences in their approach to investing in green energy companies, the solution approach in Section 3 can be used to collect and processes qualitative data and induce rules. It can incrementally generate rules as new objects are added. The approach not only has the ability to deal with multiple outcomes and features but can also specifically identify the rules in a more effective way, where massive times are required as new data are added.

This study aims to simplify the complexity of data and features to uncover the relationship between investor characteristics and decision-making behaviors. Through these simplifications, we aim to categorize investors based on their distinct characteristics while understanding their genuine sentiments toward green investment decisions. Such a study aids green energy companies in designing strategies and making decisions that cater to investors' needs. This study aids green energy companies in designing strategies, which deduce the risk during the process, and in making decisions that cater to investors' needs. To investors, this approach offers decision rules induced from data that help them decide upon their investment and enhance their efforts in decision-making and their ability to make earnings.

Account investor characteristics such as business income, share of renewables in the investment portfolio, experience in the renewable energy investing, age, education or training related to renewable energy, and exposure to the renewable energy investing domain, may influence the decision-making process regarding investment in renewable energy [30].

The raw data from the social media captured by the agent is truly massive, more than 1 TG for ten years, and 14–15% of raw data are updated monthly. This case study aims to reduce the complexity of the data and feature space and to extract the rules for the relationship between investors' features and decision-making behavior. The resulting description uses rules to classify the different features of investors and determines how investors really feel about investment decision-making tasks to help green investment service providers make policy decisions that satisfy investor requirements of an investment package.

In this study, the task list for green investment decisions is categorized into four main groups. The first one involves the initial energy investment task. Green energy investment tasks entail selecting companies leading in environmental conservation within their respective industries to build a green investment portfolio and enforce corporate social responsibility [12]. The second category is financing. Large-scale green energy projects, including solar energy and hydroelectric power, are long-term initiatives that call for prolonged funding, such as with bank loans. However, green energy projects face two main obstacles that pose significant challenges to the development of green energy projects: (a) lower return on investment in comparison with fossil fuel projects and (b) higher investment risks in comparison with fossil fuel projects [72]. Therefore, it is crucial to formulate suitable financial plans and strategies to balance investment returns and risks. The third aspect involves pre-investment tasks, such as preliminary selection of investment targets; aligning with investors' preferences for green energy investment; considering various methods and types of renewable energy investment; and valuation of green energy. The final category is the objective tasks. Once investors decide to invest, they typically make numerous decisions. These may include selecting green energy investment portfolios, gathering market data on different renewable energies, and identifying investment objectives.

A summary of the green investment decision-making task list is provided, comprising 12 characteristics categorized into 4 tasks. In this study, through expert selection, 6 features were chosen as conditional characteristics, and 6 were selected as decision outcomes, as depicted in Tables 20 and 21.

Table 20. A list of green investment decision-making tasks.

1. Initial Green Investment Tasks	I. Enforcing Corporate Social Responsibility II. Impacting Corporate Financial Performance III. Entrepreneur and Investor Interests
2. Financing Tasks	I. Arranging Financing II. Arranging Financial Consultation
3. Pre-investment Tasks	I. Preliminary Selection of Investment Targets II. Aligning with Investors' Preferences for Green Energy Investment III. Considering Various Methods and Types of Renewable Energy Investment IV. Valuation of Green Energy
4. Destination Tasks	I. Selecting Green Energy Investment Portfolios II. Gathering Market Data on Various Renewable Energies III. Green Energy Investment Objectives

Table 21. Categorization of the features.

Feature Name and Domain					
F_1	F_2	F_3	F_4	F_5	F_6
Education/Training Related to RE	**Exposure to the RE Investing Domain**	**Business Income**	**Share of Renewables in the Investment Portfolio**	**Experience in the RE Investing**	**Age**
Yes (Y) No (N)	Yes (Y) No (N)	<100,000 (L) 100,000∼500,000 (M) >500,000 (H)	<5% (N) 5∼9% (L) 10∼49% (M) 50%∼100% (H)	No experience (N) <5 years (L) 5∼10 years (M) >10 years (H)	<25 (YO) 25∼50 (MA) >50 (OA)

The characteristics of investors are categorized in Table 21 and include business income, share of renewables in the investment portfolio, experience in the renewable energy investing, age, education/training related to renewable energy, and exposure to the renewable energy investing domain.

Table 22 categorizes the outcomes selected from the green investment decision-making task list through an expert system. It includes enforcing corporate social responsibility, positively impacting corporate financial performance, investor interests, arranging financing, aligning with investors' preferences, and investment objectives.

Table 22. Categorization of the outcomes.

Outcome Name and Domain					
O_1	O_2	O_3	O_4	O_5	O_6
Arranging Financing	**Aligning with Investors' Preference**	**Investment Terms**	**Enforcing CSR**	**Impacting Corporate Financial Performance**	**Investor Interests**
Yes (Y) No (N)	Yes (Y) No (N)	Short (S) Medium (M) Long (L)	Yes (Y) No (N)	>10% Profit (VP) Very Positive 10%∼2% (P) Positive −2%∼2% (NI) No Impact −2%∼−10% (N) <−10% (N) Very Negative	Yes (Y) No (N)

Initially, the MORE algorithm was applied, and Table 23 provides a list of the resulting reducts.

Table 23. The list of the resulting value reducts (Rk).

Object No.	Reduct No.	Rule No.	F_1	F_2	F_3	F_4	F_5	F_6	O_1	O_2	O_3	O_4	O_5	O_6	SI
X_1 & X_2 & X_3	1	1	*	Y	*	*	*	*	*	Y	*	*	N	Y	10
X_4 & X_5	1	2	*	*	M	*	M	*	*	Y	S	N	N	Y	34
	2	3	*	Y	*	M	*	*	*	Y	S	N	N	Y	23
	3	4	*	*	*	*	M	MA	*	Y	S	N	N	Y	33
	4	5	*	Y	*	*	M	*	*	Y	S	N	N	Y	33
X_{15}	1	33	*	*	H	*	N	*	Y	Y	L	N	VN	N	34
	2	34	*	*	*	H	N	*	Y	Y	L	N	VN	N	36
	3	35	*	*	*	*	N	YO	Y	Y	L	N	VN	N	33
X_{17} & X_{18} & X_{19} & X_{20}	1	36	Y	*	*	*	*	*	*	Y	*	*	*	Y	2
Identify value of features			2	10	11	13	23	10							

* refers to that it does not matter.

Due to changes in government regulation, a new data set is then added. Table 24 illustrates the data when a financial advisor has additional green investment tasks. Add-ins improve the customer's experience with green investment services and update the decision rules.

Table 24. The new data set.

Object No.	F_1	F_2	F_3	F_4	F_5	F_6	O_1	O_2	O_3	O_4	O_5	O_6	Object Cardinality
X_{21}	N	N	M	L	L	YO	N	Y	S	N	VP	N	12
X_{22}	N	N	H	H	L	OA	N	Y	M	Y	VP	Y	9
X_{23}	N	N	L	L	L	MA	Y	N	M	Y	P	N	6

First, it is important to determine which case the newly formed data sets belong to. Object X_{21} belongs to Case II (the condition features are distinguishable to object X_{12}, but the outcomes are not), and objects X_{22} and X_{23} belong to Case IV (the condition features are different to those in the original data set, and they do not conflict with the original reducts). The IMORE is then applied, and the reducts of objects X_{21} $[N,N,M,L,L,YO]_{f1, f2, f3, f4, f5, f6}[N,Y,S,N,VP,N,]_{o1, o2, o3, o4, o5, o6}$ and $X_{12}[H,H,L,OA,N.N]_{f1, f2, f3, f4, f5, f6}[M,Y,B,N,N,Y]_{o1, o2, o3, o4, o5, o6}$ are re-computed.

To utilize the suggested IMORE, the extracted rules are displayed in Table 25. For example, rule 5 indicates that if business income > 500,000 and share of renewables in the investment portfolio = no, then arranging financing = yes and enforcing CSR = no. That is, if an investor has more than US$ 500,000 per/year in income and has a share of renewables in the investment portfolio, then he/she usually arranges financing well and enforces CSR from the investment perspectives.

Table 25. The list of rule extraction by IMORE.

Object No.	Rule No.	F_1	F_2	F_3	F_4	F_5	F_6	O_1	O_2	O_3	O_4	O_5	O_6	Object Cardinality
X_1 & X_2 & X_3	1	*	Y	*	*	*	*	*	Y	*	*	N	Y	23
X_4 & X_5	2	*	Y	*	*	M	*	*	Y	S	N	N	Y	12
X_6	3	*	Y	*	*	H	*	Y	Y	M	Y	N	Y	3
X_7 & X_8	4	*	N	*	*	H	*	Y	*	S	*	NI	Y	15
X_9 & X_{16}	5	*	*	L	N	*	*	Y	*	*	N	*	*	7
X_{10} & X_{11} & X_{14}	6	*	*	*	*	N	*	*	*	L	*	VN	*	25
X_{12} & X_{23}	7	*	*	H	*	L	*	*	Y	M	Y	P	*	13
X_{13}	8	*	*	*	N	M	*	N	N	S	N	NI	N	5
X_{15}	9	*	*	H	*	N	*	Y	Y	L	N	O	N	2
X_{17} & X_{18} & X_{19} & X_{20}	10	Y	*	*	*	*	*	*	Y	*	*	*	Y	23
X_{21}	11	*	*	L	L	L	*	N	Y	*	M	Y	B	5
X_{22}	12	YO	*	M	M	*	*	N	N	*	S	N	K	7

* refers to that it does not matter.

In the rule extraction procedure, it is preferable to use a reduct with a higher strength index value rather than one with a lower value. At least one reduct from each object is extracted into the decision rules. Obviously, decision rules do not omit special cases. Therefore, for the case study, when there are multiple customers that have different decision outcomes and new objects are added to an information system, the IMORE can deal with the problem flexibly, without re-computing the entire MORE and MORG.

To show the drawbacks of the traditional RS reduct generation that cannot deal with the multiple outcome problem, using the data from the case study, Table 26 compares the coverage of the IMORE method with that of the conventional rough set reduct generation process. Here is how the coverage index is defined [73,74]:

$$\text{Coverage } \psi_R(D) = \frac{|[x]_R \cap D|}{|D|}, \text{ and } 0 \ < \ \psi_R(D) \leq 1$$

where ∣A∣ denotes the cardinality of set A, and $\psi_R(D)$ denotes the coverage of R as to categorization of D. That is, $e_{ij} = v_{ij}$, in a decision table. The element (e_{ij}) denotes the value of attribute (A_j) that an object (tuple) (X_j) contains.

Table 26. The coverage of the IMORE and traditional RS approaches.

Algorithm	Number of Reducts	Number of Objects Covered by Reducts	Number of Data Covered by the Reducts	Coverage
IMORE	36	20	148	100%
Traditional rough set reduct generation	18	5	28	18.92%

The result shows that only five objects (X_6, X_{13}, X_{15}, X_{21}, and X_{22}) can generate reducts. That is, unused features are redundant and therefore can be ignored. Although each reduct generated by the traditional rough set and MORE has a 100% accuracy rate, it is particularly noteworthy that the coverage by the traditional rough set is only 18.92%.

When new data sets are added, since the MORE and the IMORE use the same computation procedure, the accuracy and final rules are the same. However, the reductive percentage for the MORE is 0%, and that for the IMORE is 42.15%. This study created a prototype system in R for testing and validation. The system execution environment was an Intel i7 13700 5.2GHz CPU with 32 GB RAM. The server was a Tomcat 9.0.85. To test the proposed approach that is less time-consuming in large data sets since it does not re-compute entire data sets for newly added data, four experiments were run.

To validate the proposed algorithm, we conducted four experiments with a large random data set (1.2 TG), with 50 features being assigned. The new data, of which 20% were from original data set, were updated. The combination of the varied percentages of the four cases was predicted. The first experiments combined 75% in Case I and 25% in Case II, with an average reduction time of 37%. In this experiment, the added data did not require the computation of any reducts, and its desired segments were investment groups. The second experiment combined 100% in Case II, with an average reduction time of 35%. In these experiments, the added data only required the computation of the affected reducts, and its desired segment was the particular investor. The third experiment combined 75% in Case III and 25% in Case II, with the average reduction time being 38%. In this experiment, the added data only required the computation of the reducts of new objects, and the desired segment was the individual investor. The fourth experiment combined 75% in Case IV and 25% in Case II, with the average reduction time being 38%. In this experiment, the added data required the computation of the reducts of new objects and the objects affected, and its desired segment was the individual investor.

In this case study, the provided data set were limited into particular features in a certain time frame, as were the outcomes. Additional features and types of outcomes may be considered. Evaluating its performance over longer time frames may be of interest. A wider research scope is suggested before this approach is applied to fields other than green energy investment. Before taking this approach into other investment targets and sectors, we hope that different kinds of features will be added and several case studies conducted.

In this case study, the proposed approach not only worked to induce rules from qualitative data with multiple features and outcomes but also generated rules in a more effective way when new data were added, saving more times compare with the traditional RS method. The outcomes of the case also showed positive relations between characteristics of investors and their investment behaviors. Additional cases in numerous sectors could be studied to enhance, validate, and improve this approach.

5. Conclusions

The shortcomings of previous studies show that (1) only one outcome feature is considered, and it is unrealistically assumed that all investment data are numeric, and that (2) rules are induced without the consideration of this dynamic structure in the data set added. This paper proposes an incremental rough set rule induction approach for multiple outcomes to solve dynamic database problems. This approach can efficiently handle updated data and provide rules to make green investment decisions agile. The impact and contribution of the study are summarized as follows:

- The proposed approach can observe and identify differences between the original reducts/rules and the updated reducts/rules in green investment after new (or upcoming) objects are added incrementally, where previous approaches implemented as black boxes cannot.
- With the aforementioned differences, it is not required to recompute the reducts for rules that are unaffected by the incremental data set while extracting reduct rules from big data. The affected rule sets are updated by changing the original rule sets slightly, saving a great deal of processing time via the proposed approach. The key idea satisfies the agility requirement in decision-making.
- The case study shows that the green investment requires decision-making to be agile in response to environmental changes of government regulation (reductive percentage 42.15%) and that the management level should be focused on updated rules, particularly corresponding to objects, e.g., X_6, X_{13}, X_{15}, X_{21}, and X_{22}. Feature 5, experience in the RE investing, is the most important factor in green investment decision-making since these objects have Feature 5 in common.
- In future study, (1) additional and diverse energy cases using different factors/feature are required to recognize the nature and categories of green investment, specifically under regulatory uncertainty. The uncertainty regarding subsidies, tax incentives, or

carbon-pricing mechanisms may be the factors affecting the financial viability of green energy projects. (2) More complex, dynamic of data, not only data added in but also data moved out, may be of interest in regard to understanding the efficiency of the proposed approach, specifically under the technological risks. Updated technology could cause objects to become unavailable in the market. Consequently, the data should be excluded. (3) The use of hybrid qualitative and quantitative methods is encouraged to explore and address the challenges in green investment, specifically under market volatility, especially in sectors heavily subject to government policies or public sentiment. Fluctuations in commodity prices, currency exchange rates, or geopolitical tensions can qualitatively and quantitatively affect the financial performance of green projects and portfolios.

Author Contributions: Conceptualization, C.-C.H.; Methodology, H.-F.C. and T.-L.T.; Formal analysis, W.-Y.L.; Writing—review & editing, Y.-C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science and Technology Council of Taiwan, grant number (NSTC 112-2410-H-260-010-MY3, MOST 111-2410-H-018-017-MY2, MOST 110-2410-H-260-016-MY2).

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. Long, C.; Jiang, Z.; Shangguan, J.; Qing, T.; Zhang, P.; Feng, B. Applications of carbon dots in environmental pollution control: A review. *Chem. Eng. J.* **2021**, *406*, 126848. [CrossRef]
2. Shen, Y.; Su, Z.-W.; Malik, M.Y.; Umar, M.; Khan, Z.; Khan, M. Does green investment, financial development and natural resources rent limit carbon emissions? A provincial panel analysis of China. *Sci. Total Environ.* **2021**, *755*, 142538. [CrossRef] [PubMed]
3. Inderst, G.; Kaminker, C.; Stewart, F. *Defining and Measuring Green Investments*; OECD Publishing: Paris, France, 2012.
4. Li, Z.; Kuo, T.-H.; Siao-Yun, W.; Vinh, L.T. Role of green finance, volatility and risk in promoting the investments in Renewable Energy Resources in the post-COVID-19. *Resour. Policy* **2022**, *76*, 102563. [CrossRef]
5. Sachs, J.D.; Woo, W.T.; Yoshino, N.; Taghizadeh-Hesary, F. Importance of green finance for achieving sustainable development goals and energy security. *Handb. Green Financ. Energy Secur. Sustain. Dev.* **2019**, *10*, 1–10.
6. Wang, C.; Wang, Y.; Shao, M.; Qian, Y.; Chen, D. Fuzzy Rough Attribute Reduction for Categorical Data. *IEEE Trans.* **2020**, *28*, 818–830. [CrossRef]
7. Pawlak, Z. Rough Sets. *Int. J. Comput. Inf. Sci.* **1982**, *11*, 341–356. [CrossRef]
8. Xu, W.; Yang, Y. Matrix-Based Feature Selection Approach Using Conditional Entropy for Ordered Data Set with Time-Evolving Features. *Knowl.-Based Syst.* **2023**, *279*, 110947. [CrossRef]
9. Kuczera, K. Application of rough set theory to identify priorities in project management. *Procedia Comput. Sci.* **2023**, *225*, 4410–4415. [CrossRef]
10. Campagner, A.; Ciucci, D.; Dorigatti, V. Uncertainty representation in dynamical systems using rough set theory. *Theor. Comput. Sci.* **2022**, *908*, 28–42. [CrossRef]
11. Wąsikowska, B.; Linowska, A.A. Application of the rough set theory to the analysis of food safety in fish processing. *Procedia Comput. Sci.* **2021**, *192*, 3342–3350. [CrossRef]
12. Patra, S.; Barman, B. A novel dependency definition exploiting boundary samples in rough set theory for hyperspectral band selection. *Appl. Soft Comput.* **2021**, *99*, 106944. [CrossRef]
13. Rasyidah; Efendi, R.; Nawi, N.M.; Deris, M.M.; Burney, S.M.A. Cleansing of inconsistent sample in linear regression model based on rough sets theory. *Syst. Soft Comput.* **2023**, *5*, 200046. [CrossRef]
14. Li, H.; Sun, H.; Fan, F.; Liu, H.; Li, L.; Yin, F. Research on an intelligent disassembling method for multi-type mobile phones based on rough set theory. *Procedia CIRP* **2022**, *105*, 547–552. [CrossRef]
15. Shasteen, K.; Kacira, M. Predictive Modeling and Computer Vision-Based Decision Support to Optimize Resource Use in Vertical Farms. *Sustainability* **2023**, *15*, 7812. [CrossRef]
16. Song, Y.; Lin, H.; Li, Z. Outlier detection in a multiset-valued information system based on rough set theory and granular computing. *Inf. Sci.* **2024**, *657*, 119950. [CrossRef]
17. Sarkar, S.; Pramanik, A.; Maiti, J. An integrated approach using rough set theory, ANFIS, and Z-number in occupational risk prediction. *Eng. Appl. Artif. Intell.* **2023**, *117*, 105515. [CrossRef]

18. Wu, D.; Yang, Z.; Wang, N.; Li, C.; Yang, Y. An Integrated Multi-Criteria Decision Making Model and AHP Weighting Uncertainty Analysis for Sustainability Assessment of Coal-Fired Power Units. *Sustainability* **2018**, *10*, 1700. [CrossRef]
19. Leng, Y.-J.; Zhang, H. A novel evaluation method of renewable energy development based on improved rough set theory and grey cloud model. *J. Clean. Prod.* **2023**, *428*, 139299. [CrossRef]
20. Huang, C.-C.; Liang, W.-Y.; Gung, R.R.; Wang, P.-A. Rough-Set-Based Rule Induction with the Elimination of Outdated Big Data: Case of Renewable Energy Equipment Promotion. *Sustainability* **2023**, *15*, 14984. [CrossRef]
21. Kang, X. Combining rough set theory and support vector regression to the sustainable form design of hybrid electric vehicle. *J. Clean. Prod.* **2021**, *304*, 127137. [CrossRef]
22. Lu, T.; Xiao, X.; Yu, H.; Ren, D. The integration of technological advancements in fossil fuels towards environmental sustainability: Insights from big data analytics. *Resour. Policy* **2023**, *86*, 104196. [CrossRef]
23. Chen, K.-Y. Combining linear and nonlinear model in forecasting tourismdemand. *Expert Syst. Appl.* **2011**, *38*, 10368–10376. [CrossRef]
24. Cakranegara, P.A.; Sidjabat, F.M. Green Investment: Incorporate Environment, Social, and Government Factors in Investment Decision. *Eksis J. Ris. Ekon. Dan Bisnis* **2021**, *16*, 17–28. [CrossRef]
25. Ringel, M.; Mjekic, S. Analyzing the Role of Banks in Providing Green Finance for Retail Customers: The Case of Germany. *Sustainability* **2023**, *15*, 8745. [CrossRef]
26. Chițimiea, A.; Minciu, M.; Manta, A.-M.; Ciocoiu, C.N.; Veith, C. The Drivers of green investment: A bibliometric and systematic review. *Sustainability* **2021**, *13*, 3507. [CrossRef]
27. Lingyun, T.; Liping, A. Incremental learning of decision rules based on rough set theory. In Intelligent Control and Automation, 2002. In Proceedings of the 4th World Congress, Calgary, AB, Canada, 4–9 August 2002; IEEE: New York, NY, USA, 2002; Volume 1, pp. 420–425.
28. Bian, J.; Zhang, G.; Zhou, G. Manufacturer vs. consumer subsidy with green technology investment and environmental concern. *Eur. J. Oper. Res.* **2020**, *287*, 832–843. [CrossRef]
29. Li, Z.-Z.; Li, R.Y.M.; Malik, M.Y.; Murshed, M.; Khan, Z.; Umar, M. Determinants of carbon emission in China: How good is green investment? *Sustain. Prod. Consum.* **2021**, *27*, 392–401. [CrossRef]
30. Chen, Y.; Ma, Y. Does green investment improve energy firm performance? *Energy Policy* **2021**, *153*, 112252. [CrossRef]
31. Stucki, T. Which firms benefit from investments in green energy technologies?—The effect of energy costs. *Res. Policy* **2019**, *48*, 546–555. [CrossRef]
32. Albuquerque, L.G.; de Oliveira Roque, F.; Valente-Neto, F.; Koroiva, R.; Buss, D.F.; Baptista, D.F.; Hepp, L.U.; Kuhlmann, M.L.; Sundar, S.; Covich, A.P.; et al. Large-scale prediction of tropical stream water quality using Rough Sets Theory. *Ecol. Inform.* **2021**, *61*, 101226. [CrossRef]
33. Meiyou, D.; Ye, Y. Establishment of big data evaluation model for green and sustainable development of enterprises. *J. King Saud Univ.-Sci.* **2022**, *34*, 102041. [CrossRef]
34. Ikram, M.; Ferasso, M.; Sroufe, R.; Zhang, Q. Assessing green technology indicators for cleaner production and sustainable investments in a developing country context. *J. Clean. Prod.* **2021**, *322*, 129090. [CrossRef]
35. Yang, C.-C.; Ou, S.-L.; Hsu, L.-C. A Hybrid Multi-Criteria Decision-Making Model for Evaluating Companies' Green Credit Rating. *Sustainability* **2019**, *11*, 1506. [CrossRef]
36. Tsai, J.-F.; Shen, S.-P.; Lin, M.-H. Applying a Hybrid MCDM Model to Evaluate Green Supply Chain Management Practices. *Sustainability* **2023**, *15*, 2148. [CrossRef]
37. Taherdoost, H.; Madanchian, M. Multi-Criteria Decision Making (MCDM) Methods and Concepts. *Encyclopedia* **2023**, *3*, 77–87. [CrossRef]
38. Wang, M.; Li, X.; Wang, S. Discovering research trends and opportunities of green finance and energy policy: A data-driven scientometric analysis. *Energy Policy* **2021**, *154*, 112295. [CrossRef]
39. McAfee, A.; Brynjolfsson, E.; Davenport, T.H.; Patil, D.J.; Barton, D. Big data. *Manag. Revolut. Harv. Bus. Rev.* **2012**, *90*, 61–67.
40. Lam, C.; McKercher, B. The tourism data gap: The utility of official tourism information for the hospitality and tourism industry. *Tour. Manag. Perspect.* **2013**, *6*, 82–94. [CrossRef]
41. Goldman, S.L.; Nagel, R.N. Management, technology and agility: The emergence of a new era in manufacturing. *Int. J. Technol. Manag.* **1993**, *8*, 18.
42. Kidd, P.T. *Agile Manufacturing: Forging New Frontiers*; Addison-Wesley: Wokingham, UK; Reading, MA, USA, 1994.
43. Le, T.T.; Ferasso, M. How green investment drives sustainable business performance for food manufacturing small-and medium-sized enterprises? Evidence from an emerging economy. *Corp. Soc. Responsib. Environ. Manag.* **2022**, *29*, 1034–1049. [CrossRef]
44. Torres, F.C.; Almenar, J.B.; Rugani, B. Photovoltaic-green roof energy communities can uphold the European Green Deal: Probabilistic cost-benefit analyses help discern economically convenient scenarios. *J. Clean. Prod.* **2023**, *414*, 137428. [CrossRef]
45. Anagnostopoulos, T.; Kyriakopoulos, G.L.; Ntanos, S.; Gkika, E.; Asonitou, S. Intelligent predictive analytics for sustainable business investment in renewable energy sources. *Sustainability* **2020**, *12*, 2817. [CrossRef]
46. Du, H.S.; Zhan, B.; Xu, J.; Yang, X. The influencing mechanism of multi-factors on green investments: A hybrid analysis. *J. Clean. Prod.* **2019**, *239*, 117977. [CrossRef]
47. Wu, H. Evaluating the role of renewable energy investment resources and green finance on the economic performance: Evidence from OECD economies. *Resour. Policy* **2023**, *80*, 103149. [CrossRef]

48. Murwaningsari, E.; Rachmawati, S. The Effect of Green Banking and Green Investment on Firm Value with Eco-Efficiency as Moderation. *Int. J. Soc. Manag. Stud.* **2023**, *4*, 141–150.
49. Lempp, H.; Kingsley, G. Qualitative assessments. *Best Pract. Res. Clin. Rheumatol.* **2007**, *21*, 857–869. [CrossRef]
50. Osman, I.; Maâ, M.; Muda, R.; Husni, N.S.A.; Alwi, S.F.S.; Hassan, F. Determinants of behavioural intention towards green investments: The perspectives of muslims. *Int. J. Islam. Bus.* **2019**, *4*, 16–38. [CrossRef]
51. Rana, M.W.; Zhang, S.; Ali, S.; Hamid, I. Investigating green financing factors to entice private sector investment in renewables via digital media: Energy efficiency and sustainable development in the post-COVID-19 era. *Sustainability* **2022**, *14*, 13119. [CrossRef]
52. Weihong, J.; Kuo, T.-H.; Wei, S.-Y.; Hossain, M.S.; Tongkachok, K.; Imran, A. Relationship between trade enhancement, firm characteristics and CSR: Key mediating role of green investment. *Econ. Res.-Ekon. Istraživanja* **2022**, *35*, 3900–3916. [CrossRef]
53. Yang, M. Green investment and e-commerce sales mode selection strategies with cap-and-trade regulation. *Comput. Ind. Eng.* **2023**, *177*, 109036. [CrossRef]
54. Shankar, R. *Potential of Blockchain Based Tokenized Securities for Green Real Estate Bonds. Infrastructure Development–Theory, Practice and Policy*; Routledge: Oxfordshire, UK, 2022; pp. 88–99.
55. Heckerman, D.; Mannila, H.; Pregibon, D.; Uthurusamy, R. Applying data mining and machine learning techniques to submarine intelligence analysis. In Proceedings of the Third International Conference on Knowledge Discovery and Data Mining, Newport Beach, CA, USA, 14–17 August 1997; AAAI Press: Menlo Park, CA, USA, 1997.
56. Rebolledo, M.R. Integrating rough sets and situation-based qualitative models for processes monitoring considering vagueness and uncertainty. *Eng. Appl. Artif. Intell.* **2005**, *18*, 617–632. [CrossRef]
57. Pawlak, Z. Rough classification. *Int. J. Man-Mach. Stud.* **1984**, *20*, 469–483. [CrossRef]
58. Wong, S.K.M.; Ziarko, W.; Ye, R.L. Comparison of rough-set and statistical methods in inductive learning. *Int. J. Man-Mach. Stud.* **1986**, *25*, 53–72. [CrossRef]
59. Boulatoff, C.; Boyer, C.M. Green recovery: How are environmental stocks doing? *J. Wealth Manag.* **2009**, *12*, 9–20. [CrossRef]
60. Tran, T.; Do, H.; Vu, T.; Do, N. The factors affecting green investment for sustainable development. *Decis. Sci. Lett.* **2020**, *9*, 365–386. [CrossRef]
61. Tan, A.; Wu, W.; Qian, Y.; Liang, J.; Chen, J.; Li, J. Intuitionistic Fuzzy Rough Set-Based Granular Structures and Attribute Subset Selection. *IEEE Trans.* **2019**, *27*, 527–539. [CrossRef]
62. Shu, W.; Qian, W.; Xie, Y. Incremental Approaches for Feature Selection from Dynamic Data with the Variation of Multiple Objects. *Knowl.-Based Syst.* **2019**, *163*, 320–331. [CrossRef]
63. Wang, G. Valid Incremental Attribute Reduction Algorithm Based on Attribute Generalization for an Incomplete Information System. *Chin. J. Electron.* **2019**, *28*, 725–736. [CrossRef]
64. Zhang, X.; Mei, C.; Chen, D.; Yang, Y.; Li, J. Active Incremental Feature Selection Using a Fuzzy-Rough-Set-Based Information Entropy. *IEEE Trans. Fuzzy Syst.* **2020**, *28*, 901–915. [CrossRef]
65. Pan, Y.; Xu, W.; Ran, Q. An incremental approach to feature selection using the weighted dominance-based neighborhood rough sets. *Int. J. Mach. Learn. Cybern.* **2023**, *14*, 1217–1233. [CrossRef]
66. Ni, P.; Zhao, S.; Wang, X.; Chen, H.; Li, C.; Tsang, E.C.C. Incremental feature selection based on fuzzy rough sets. *Inf. Sci.* **2020**, *536*, 185–204. [CrossRef]
67. Raza, M.S.; Qamar, U. A parallel approach to calculate lower and upper approximations in dominance based rough set theory. *Appl. Soft Comput.* **2019**, *84*, 105699. [CrossRef]
68. Li, W.; Pedrycz, W.; Xue, X.; Xu, W.; Fan, B. Fuzziness and incremental information of disjoint regions in double-quantitative decision-theoretic rough set model. *Int. J. Mach. Learn. Cybern.* **2019**, *10*, 2669–2690. [CrossRef]
69. Zhong, N.; Dong, J.Z.; Ohsuga, S.; Lin, T.Y. An incremental, probabilistic rough set approach to rule discovery. In Proceedings of the 1998 IEEE International Conference on Fuzzy Systems Proceedings. IEEE World Congress on Computational Intelligence, Anchorage, AK, USA, 4–9 May 1998; IEEE: New York, NY, USA, 1998; Volume 2, pp. 933–938.
70. Huang, C.C.; Tseng, T.L.B. Rough set approach to case-based reasoning application. *Expert Syst. Appl.* **2004**, *26*, 369–385. [CrossRef]
71. Masini, A.; Menichetti, E. The impact of behavioural factors in the renewable energy investment decision making process: Conceptual framework and empirical findings. *Energy Policy* **2012**, *40*, 28–38. [CrossRef]
72. Taghizadeh-Hesary, F.; Yoshino, N. Sustainable solutions for green financing and investment in renewable energy projects. *Energies* **2020**, *13*, 788. [CrossRef]
73. Tseng, T.L.B. Quantitative Approaches for Information Modeling. Ph.D. Thesis, University of Iowa, Iowa, IA, USA, 1999.
74. Tsumoto, Automated extraction of medical expert system rules from clinical databases based on rough set theory. *Inf. Sci.* **1998**, *112*, 67–84. [CrossRef]