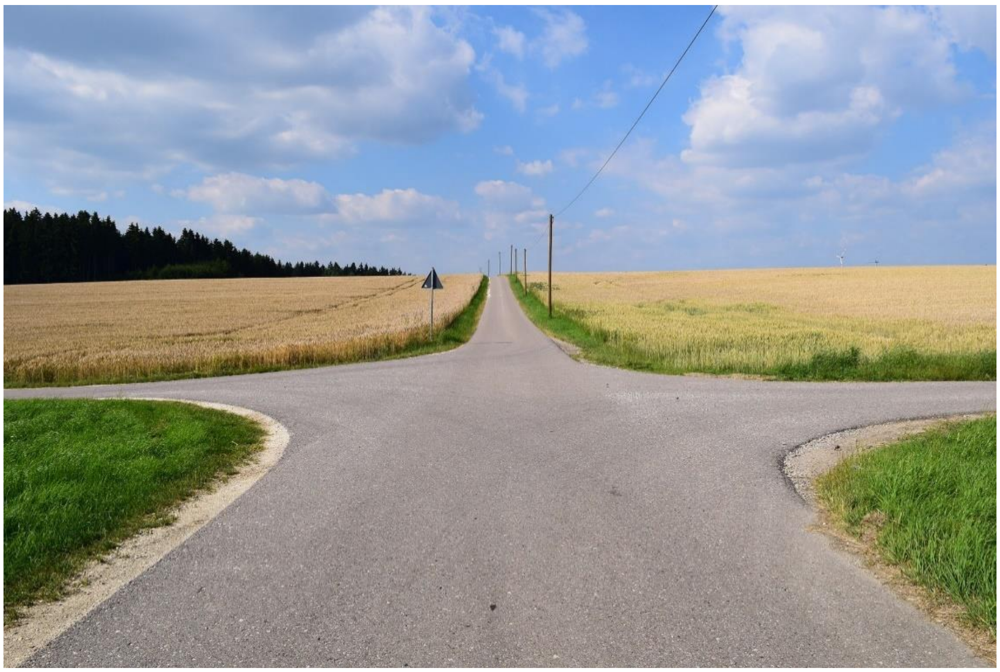
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727470

Report No ECHOES 6.5 – D6.5 Algorithm

ECHOES Report

A decision tree algorithm: To be used as an instrument to establish integrated governance and monitoring processes for the Energy Union



ECHOES Report

A decision tree algorithm: To be used as an instrument to establish integrated governance and monitoring processes for the Energy Union

KEYWORDS:
- decision tree algorithm
- instrument
- governance
- monitoring

VERSION	DATE
03	23.08.2019

AUTHOR(S)

Mehmet Efe Biresselioglu[1], Muhittin Hakan Demir[1]

[1]**IUE**

Quality insurance:
Selma Ulker **(EUAS)**, Jed Cohen **(EI)**

PROJECT NO.	NUMBER of PAGES:
727470 (H2020)	26

ABSTRACT
The developed decision tree algorithm is designed to support decision-makers at various decision-making levels to plan their processes, as well as for formulating guidelines or policies for different decision-making levels. To this end, three separate decision-tree algorithms were constructed, one pertaining to each decision-making level. Each decision tree algorithm was designed to evaluate the potential for success or utility of an energy-related process from the perspective of a particular decision-making level.

REPORT NO.	ISBN
D6.5	NA

	CLASSIFICATION	CLASSIFICATION THIS PAGE
	Public	Public

Document history

VERSION	DATE	VERSION DESCRIPTION
01	17.07.2019	First Draft
02	05.08.2019	First Version for Quality Check
03	23.08.2019	Final Version

Contents

Executive Summary

This report presents the decision tree algorithm for energy-related decisions. The decision-making levels considered in the design of the decision-tree algorithm are those defined in the scope of the ECHOES project, namely, *Formal social units* acting as policy makers and/or energy providers, *Collective decision-making units* that are more formally structured social units including TSOs, PXs, energy producers and consumer associations, and *Individual consumers engaging in joint contracts* to increase their power of negotiation, such as larger group of households, condominium management and associations of households. The project's three technological foci are those analysed in the ECHOES project: *Smart energy technologies*, *Electric mobility*, and *Buildings*.

The developed decision tree algorithm is designed to support decision-makers at various decision-making levels to plan for their processes, as well as for formulating guidelines or policies for different decision-making levels. To this end, three separate decision-tree algorithms were constructed, one pertaining to each decision-making level. Each decision tree algorithm was designed to evaluate the potential for success or utility of an energy-related process from the perspective of a particular decision-making level. The process under consideration may be a project, case, implementation, planned investment, or a similar undertaking. The decision tree algorithm combines the knowledge generated previously in WP6 with input from the decision-makers in an analytical assessment of the potential for a given undertaking. In doing so, the decision tree algorithm explicitly considers process triggers and phases, as well as key drivers, enablers, and disablers. The decision-makers are actively involved in the execution of the algorithm by estimating the relevance, magnitude of effect, and likelihood of occurrence of the key drivers, enablers, and disablers. Besides the evaluation of the relevant energy-related processes, the algorithm can also provide answers to "what if?" questions in order to reveal the effects of possible changes in the relevance, magnitude of effect, and likelihood of occurrence of the key drivers, enablers, and disablers. Such information shows the overall impact of changes in contextual factors. Hence, as well as designing the processes, it also can be utilized for policy making and identifying and planning for the critical contextual factors.

The structure of the decision trees for each decision-making level has the starting root node as the trigger affecting the energy-related process. The decision tree algorithm for each decision-making level is differentiated with respect to these triggers, which are derived from Deliverable D6.3. The two triggers for the formal social units are problem-driven processes and goal-driven processes. The three associated triggers for the collective decision-making units are: market-driven processes, sustainability-driven processes, and legislation / market-driven processes. Finally, the triggers for the individual consumers engaging in joint contracts are individual-driven process and joint benefit-driven process.

Following the triggers as root nodes, the building blocks of the decision tree algorithm are derived from work performed in WP6 of the ECHOES project. The foundation of the tree is composed of the three phases of energy-related processes derived from the best examples/successful implementations analysed in Deliverable D6.3 for formal social units, collective decision-making units, and individuals engaging in joint contracts. The five phases are problem identification, alternative selection, planning, implementation, and monitoring.

The terminal root of the decision tree refers to the assessment of the potential for success or utility of the process. The calculation of potential for success is based on the data collected and decision makers' perspectives on motivators, barriers, and other critical contextual factors. The assessment is first made for each node of the decision tree, then for the overall process.

The information regarding the key drivers, enablers, and disablers are reflected in the structure of the decision tree by superimposing nodes for each process phase representing the key drivers, enablers, and disablers that are adjacent to the node defining the process phase.

The key drivers for the processes are identified using the results of Deliverable D6.3, and the enablers and disablers corresponding to the different decision-making levels are identified using the results of the deliverables D6.1 and D6.2.

The three inputs required for each such element from the decision-makers are the weight, impact, and probability values. Within the context of the decision tree, weight corresponds to the relevance or importance of the particular key driver, enabler, or disabler, for the particular phase. The impact value refers to the magnitude of the effect of the key driver, enabler, or disabler once realized or achieved. Finally, the probability value refers to the likelihood of occurrence or realization of the particular element for the particular phase. For the importance values, the designed decision tree algorithm uses a Likert scale of 0 to 5, where 0 corresponds to none, 1 corresponds to not important, 2 corresponds to slightly important, 3 corresponds to moderately important, 4 corresponds to important, and 5 corresponds to very important. For the impact values, a similar Likert 0 to 5 scale is used. Here, 0 corresponds to none, 1 corresponds to very low, 2 corresponds to low, 3 corresponds to moderate, 4 corresponds to considerable, and 5 corresponds to severe. Regarding the probability values, the natural range of 0 to 1 is used. The decision maker can select probability values between 0 and 1

The decision maker is able to select the values from a drop-down menu in the interface of the spreadsheet for the decision tree. The values for the elements are associated with the linguistic counterparts, in order to facilitate the decision maker's choice, since linguistic terms are generally easier for the decision makers to evaluate. This part of the decision tree algorithm also helps the decision makers to translate their perceptions on the key drivers, enablers, and disablers into comparable numeric values. The excel sheet will be available upon request explaining the potential use to the authors.

1. Introduction

This report is a deliverable of WP6 of ECHOES project, with the objective of synthesizing the results of the earlier tasks and deliverables (D6.1, D6.2, and D6.3) into a decision tree algorithm, for use as a decision support tool for various decision-making units.

The decision-making units under consideration pertains to the Macro level of the ECHOES project, and are aligned with the three levels of decision-making units identified in the ECHOES perspective:

a) *Formal social units* which act as policy makers and/or energy providers, with a wider influence on energy choice decisions.
b) *Collective decision-making units*, which are more formally structured and with relatively lower information and power asymmetries.
c) *Individual consumers engaging in joint contracts* to increase their power of negotiation with the above bodies..

The report also follows the framework set by the ECHOES project in selecting the technological foci involved in the analysis for deriving the decision tree algorithm. The three technological foci are as follows:

a) *Smart energy technologies* are at the core of what the integrated roadmap for realizing the SET-plan describes as an energy revolution. These includes distributed, small-scale renewable energy production technologies, but also a range of technologies for the traditional “demand side” and energy storage.
b) *Electric mobility* as one of the core technologies to be implemented and developed further to increase road transport efficiency.
c) *Buildings* – including construction activities, insulation, energy efficiency upgrading, heating, cooling, illuminating, and energy use behaviour.

The construction of the decision tree algorithm is based on the data and analysis obtained from the earlier tasks of WP6 of the ECHOES project, which included Tasks 6.1, 6.2, 6.3, and 6.4, and the resulting deliverables D6.1 (2018), D6.2 (2018), and D6.3 (2019).

The decision tree algorithm is based on a joint analysis of the preceding deliverables D6.1, D6.2, and D6.3, with a general focus on the phases of implementations regarding the processes that pertain to energy transitions. This approach utilizes in particular the information from D6.1 and D6.2 regarding the enablers and disablers of the associated processes, and the information from D6.3 regarding the best practices/successful implementations.

Tasks 6.1, 6.2, and 6.3 resulted in the identification of the key factors, variables, and parameters pertaining to the energy choices and energy-related behaviour through the utilization of qualitative research techniques. The methodology included (1) 15 Focus groups and 67 In-depth interviews carried out in six selected countries – Austria, Bulgaria, Finland, Norway, Spain and Turkey, and (2) a-state-of-the-art and comprehensive literature review carried out in the previous tasks. The resulting deliverables, D6.1 and D6.2, revealed the components of the processes related with energy transition, in which the motivators and barriers provided a partial understanding of the conditions under which these processes were able to succeed. This information is utilized by the decision tree algorithm developed within the framework of this report. The algorithm therefore was able to develop a contextual analysis that evaluates the available information, based on the relevant decision-making levels and technological foci.

As well as the process components, key factors, motivators, and barriers, the decision tree algorithm utilizes the information from Task 6.4 and deliverable D6.3, based on the case studies regarding best practices/successful implementations in selected countries. The inquiry in D6.3 of WP6, was implemented in 7 countries, namely, Austria, Bulgaria, Finland, Germany, Italy, Norway, Spain, and Turkey. Twelve cases were selected as showcases for energy choices and energy-related behaviour. The resulting information from these 12 selected case studies was analysed to determine emerging themes, internal and external factors that derive the success of the associated implementations. Utilization of the results from Task 6.4 and deliverable D6.3 regarding the case studies enabling the decision tree algorithm to develop a more functional perspective of the energy-related processes and implementations, based on field data. The themes potentially leading to an analytical investigation of the best practices and/or successful implementations from D6.1 and D6.2 were adopted in D6.3. The foremost examples of such themes are related with the internal mechanisms. These are top-down and bottom-up mechanisms, formal structures and governance frameworks that may agree or conflict with the presumed process components as defined earlier in D6.1 and D6.2. Thus, the case studies in D6.3 resulted in process mappings including the positive and negative interrelations of the process components. These mechanisms and other relevant constructs that drive the best practices/successful implementations were analysed and the results were translated into the inputs of the decision tree algorithm, including process triggers, process mappings, and rules.

Hence, a decision tree algorithm was developed in two stages: first, taking the process components, key factors, variables, and parameters pertaining to the energy choices and energy-related behaviour as the infrastructure, and second, positioning the motivators and barriers on top of this infrastructure in order to assess their effects using the information from the case studies.

2. Background

Decision Tree algorithms are within the category of "learning algorithms". A learning algorithm is one in which available information is processed to generate patterns regarding a system, and in which these patterns are applied to make assessments and decisions for new situations (Lin and Rosasco, 2018; Faraj et al., 2018; Cha et al., 2018). That is, a learning algorithm 'learns' from examples to generate rules that can be implemented for similar future situations. Typical areas of implementation of learning algorithms are artificial intelligence and machine learning, where the artificial neural network parameters are fine-tuned so that the model behaviour resembles the expected real system behaviour. Hence, learning algorithms aim to identify the outputs of the system based on different set of inputs. The success of the learning algorithm depends on two key metrics: i) variety of the inputs involved in developing the algorithm, and the extent to which these inputs represent all possible realizations of the system scope, and ii) the extent to which the algorithm outputs matches the behaviour of the real system, considering the impact of all possible stakeholders.

To this end, the decision tree algorithm developed in this deliverable utilizes the results of the deliverables D6.1 and D6.2, as these provide a comprehensive coverage of the processes, key factors, parameter, and variables regarding the energy-related behaviour. The decision tree algorithm is also based on the 12 case studies conducted in the scope of deliverable D6.3, thus incorporating perspectives from real-life implementations into the algorithm design. That is, the developed decision tree algorithm acknowledges and accounts for the two metrics essential for the success of learning algorithms.

One further level of classification places decision tree algorithms within "supervised learning algorithms". Decision tree algorithms have the advantage over algorithms in this same category, in that they can also be used in situations where the aim is to produce regressions or establish classifications. Such algorithms use observations, as well as their outcomes, to make predictions about the system. A supervised learning algorithm constructs a model in order to arrive at conclusions under conditions of uncertainty. These predictions utilize evidence from the observations. The predictive capability of the supervised learning algorithm increases with the number of observations (Dietterich, 1998; Caruana and Niculescu-Mizil, 2006; Jordan and Mitchell, 2015; Kim, 2016)

The decision tree algorithm developed in the context of this deliverable relies on the earlier results of WP6 of ECHOES, where 15 focus group studies, 67 in-depth interviews and 12 case studies were implemented across 7 selected countries. Thus, the substantial sample base of observations utilized in the construction of the resulting algorithm considerably contributes to its predictive capability.

As discussed, the general aim of decision tree algorithms is to construct a model that will be trained by real data involving inputs and outputs to estimate the category or values of output variables. The decision tree algorithm accomplishes this aim by using the prior data as training data to generate decision rules (Song and Lu, 2015; Jadhav and Channe, 2014; Dai et al., 2016).

Decision tree algorithms use the tree representation to analyse the system under consideration and predict the outcomes. In general, the root nodes of the tree correspond to the general modes or categories of the system under consideration. The inner nodes of the tree correspond to attributes, that is, to the characteristics of the system under consideration. Finally, for classification problems, the leaf (terminal) nodes correspond to the classes of the system. For cases where the decision tree algorithm is devised to make predictions about the system outcomes, leaf nodes correspond to possible outcomes of the system (Segatori et al., 2018; Zhao and Li, 2017). The algorithm developed for this deliverable focuses on identifying the potential outcomes of energy transition processes, primarily, the degree of success. Figure 1 below depicts a generic representation of a decision tree.

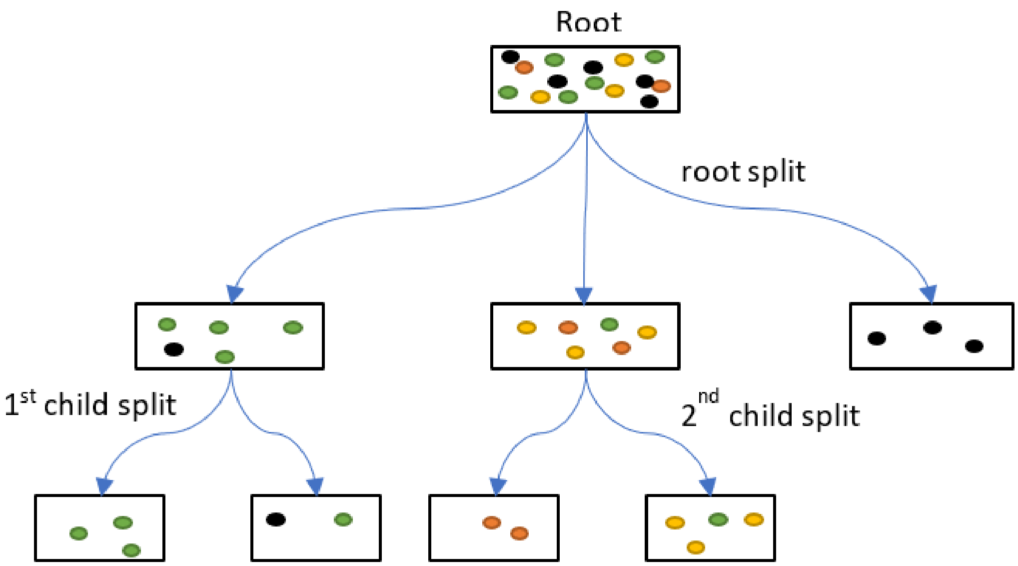


Figure 1 Sketch of a Decision Tree Algorithm

Decision trees classify information from earlier observations in order to solve regression and classification problems. The tree develops incrementally by matching of the subsets of the data with the classes, and the ordering of these classes. Finally, when the tree reaches leaf nodes, a decision, outcome, or classification is reached. The advantages of this system include ease of use, ease of explanation, a visual aspect allowing interpretation of complex processes, and the imitation of human behaviour, whereas the disadvantages are increased complexity of calculations with the number of classes, overdependence on the data characteristics, risk of bias, and risk of low accuracy (Bach and Dayan, 2017; Wu et al., 2017).

The decision tree algorithm starts from the root node and proceeds towards the leaf nodes to reach an outcome. The branches of the tree are traversed by incorporating the values associated with the internal nodes. That is, the attribute values of the internal nodes are processed to the point at which a leaf node is reached.

There are several general assumptions in constructing decision trees. To begin with, the whole system, or the categories of the system, is taken as the root. Accordingly, all data in the training set are taken as the root for the decision tree algorithm. For ease of processing, the values assigned to attributes are generally taken as the elements of a set of possible realizations that represent a category rather than continuous variables, although continuous values convey more information. In the case of category values being inherently continuous, these are discretized and input to the model (Garget al., 2016; Jaworski, 2018).

In the decision tree algorithm constructed for the purposes of this deliverable, continuous values for the attributes regarding the weights, impacts and probabilities are discretized for ease of processing, using 5 x 5 x 10 possible values to preserve as much data as possible.

The decision tree is constructed by assuming an ordering of the attributes, that is, an ordering of the nodes of the tree. If it is not implied by logical constraints, this ordering is carried out using either an algorithm or statistical analysis. Ordering of the attributes affects the efficiency of the decision tree algorithm, determining its speed and convergence, as well as the accuracy of assessments made using the decision tree algorithm. At this point, two common methods are the Information Gain and the Gini Index methods (Mathan et al., 2018; Agnihotri et al., 2017). The former is used when category values are discrete, and the latter, when continuous. Both methods consider the attributes as yet unplaced as nodes of the tree, and aim to identify how much of the output can be characterized if a particular such attribute is used as the next node. The attribute that provides the highest level of characterization is selected as the next node. In the case that the attributes inherit a natural ordering, such as the phases of a process or the steps of an operation, the ordering is implied by that of the phases or the steps.

The decision tree algorithm developed in this deliverable is cast into the process of implementations regarding energy transition on different decision making levels; therefore, ordering of the attributes are in line with the mapping of these processes. Accordingly, this approach increases the speed of the algorithm, guarantees convergence in a predetermined number of steps, and assures a more accurate estimation of the outcomes.

The generic decision tree algorithm starts from the root node and implements a top-down methodology to construct the tree. At each step, of all possible branches, the one with the maximum information gain or the minimum gini index is selected (Quinlan, 1986). This algorithm is in the class of greedy algorithms, selecting the alternative with the maximum gain at each step.

Below is a pseudocode representation of the generic decision tree algorithm:

Set all data to the root node. Set root node = current node

Branching

If there are no unprocessed nonleaf nodes, Stop!

Else for each child of the current node

For each attribute that is node assigned as a branch of the tree

Calculate the information gain (or gini index) if the attribute is selected as the next node

Identify the attribute with the highest information gain (or the attribute with the lowest gini index)

If the information highest information gain is 0 (or if the lowest gini index is 1) mark the node as leaf node

Else Branch using that attribute as the next node of the tree

Set current node = any unprocessed node with all ancestors processed and GoTo Branching

3. Decision Tree Algorithm

The foundation of the decision tree algorithm as a decision support tool is based on the below process mapping, developed in the context of D6.3 for energy-related implementations. Process mapping is augmented to include the terminal node that collects the total utility score.

The development of process mapping for the best practices/successful implementations in D6.3 included the phases of Problem Identification, Alternative Selection, Planning, Implementation, and Monitoring.

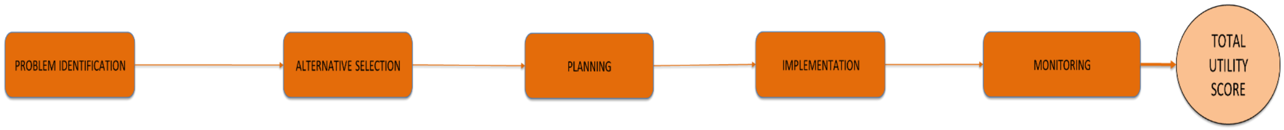


Figure 2 Augmented Process Mapping

The decision tree extends process mapping by including the process triggers. Since the triggers are not uniform over the decision-making units, three versions of the decision tree are constructed, corresponding to formal social units, collective decision making units, and individuals engaging in joint contracts. These triggers are added as root nodes to the corresponding decision trees. The below example, regarding formal social units, demonstrates the root nodes of "Problem-driven Process" and "Goal-driven Process", which represent the process triggers. These root nodes are linked to the node representing the first phase of the process, that is, 'Problem identification'.

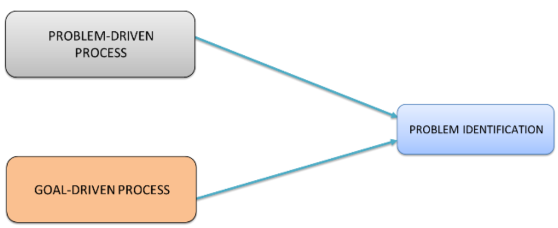


Figure 3 Triggers for the Process

Each node corresponding to a phase of the best practices/successful implementations reflects the effects of key drivers of the process pertaining to that specific phase. This is achieved by adding a node adjacent to the node that defines the process phase. Below is an example of the inclusion of the effect of mid-level management contribution for the planning phase.

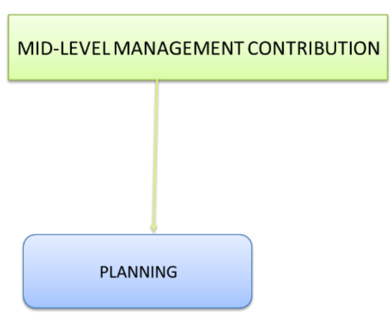


Figure 4 Key Driver and Process Phase

Deliverable D6.2 identifies the motivators (enablers) and barriers (disablers) as significant components of the process for the energy-related implementations. Hence, the construction of the decision tree is completed with the addition of nodes corresponding to the motivators and barriers for each phase of the process. This is demonstrated below for the 'Implementation' phase:

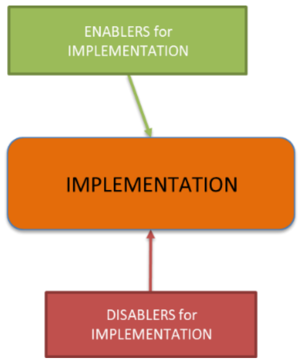


Figure 5 Inclusion of Motivators and Barriers for the process phases

The decision tree algorithm is designed as a decision-support system, and the resulting decision tree is executed by input from the decision maker(s) in order to output a total utility score for the process under consideration.

For each phase under consideration, a utility score is calculated as a joint result of the effects of key drivers, enablers, and disablers, where the disablers are considered to have negative utilities or disutilities. The utility score of a phase is the simple arithmetic sum of the utility scores of three components. That is:

$$Utility\ Score\ of\ a\ Phase = \sum_{i\ \epsilon\ \{key\ drivers, enablers, disablers\}} Utility\ by\ Component(i)$$

As an example, the planning phase with the utility score of 20 for drivers, 15 for enablers, and -12 for disablers would have a total utility score of 20+15-12 = 23.

The utility by each component is computed as follows:

For the key drivers:

$$Utility\ by\ key\ drivers = \sum_{k\ \in\ \{set\ of\ key\ drivers\ for\ the\ phase\}} \frac{Utility\ by\ key\ driver(k)}{Number\ of\ key\ drivers\ of\ the\ phase}$$

Considering a set of 2 key drivers with utility values of 16 and 8, the utility by key drivers would be calculated by: (16+8)/2 = 12.

For the enablers:

$$Utility\ by\ enablers = \sum_{e \in \{set\ of\ enablers\ for\ the\ phase\}} \frac{Utility\ by\ enabler(e)}{Number\ of\ enablers\ of\ the\ phase}$$

For a hypothetical case, considering a process phase with 4 enablers having utility values of 18, 15, 21, and 20, respectively, the utility by enablers would be equal to: (18+15+21+20)/4 = 18.5.

Finally, for the disablers:

$$Utility\ by\ disablers = \sum_{d \in \{set\ of\ disablers\ for\ the\ phase\}} (-1) * \frac{Disutility\ by\ disabler(d)}{Number\ of\ disablers\ for\ the\ phase}$$

For a phase with 3 disablers, having disutility values of 21, 10, and 8, respectively, the utility by disablers would be calculated as: ((-21)+(-10)+(-8))/3 = -13.

Calculation of the utility value for each element in the set of key drivers, enablers, or disablers requires inputs from the decision maker(s), which is necessary to obtain their contextual information. On the other hand, this phase of the decision tree algorithm directs the decision makers towards a more extensive consideration of the key drivers, enablers, and disablers, as well as developing a more analytical approach towards improving their impacts.

The input required for each such element from the decision makers are the weight, impact, and probability values. The decision makers will provide the required input based on their expertise, involvement in the particular case as well as based on the relevant data available (e.g. as the output of earlier research) on weight, impact, or probability of a key driver, enabler, or disabler. Within the context of the decision tree, weight corresponds to the relevance or importance of the particular key driver, enabler, or disabler, for the particular phase. The impact value refers to the magnitude of the effect of the key driver, enabler, or disabler, once it is realized or achieved. Finally, the probability value refers to the likelihood of occurrence or realization of the particular element for the particular phase.

For the importance values, the designed decision tree algorithm uses a Likert scale of 0 to 5, where 0 corresponds to none, 1 corresponds to not important, 2 corresponds to slightly important, 3 corresponds to moderately important, 4 corresponds to important, and 5 corresponds to very important.

For the impact values, a similar Likert 0 to 5 scale is used. Here, 0 corresponds to none, 1 corresponds to very low, 2 corresponds to low, 3 corresponds to moderate, 4 corresponds to considerable, and 5 corresponds to severe.

The decision maker is able to select the values from a drop-down menu in the interface of the spreadsheet for the decision tree. The values for the elements are associated with the linguistic counterparts, in order to facilitate the choice of the decision maker, because the linguistic choices are generally easier. This part of the decision tree algorithm also helps the decision makers to translate their perceptions on the key drivers, enablers, and disablers into comparable numeric values.

Regarding the probability values, the natural range of 0 to1 is used. The decision maker can select probability values between 0 and 1, with increments of 0.1, totalling to 11 choices. The increment values can be increased or decreased as desired; however, a smaller number of choices would hinder decision makers in reflecting precise perceptions on the probability values, while a greater number would lengthen and unnecessarily complicate the selection process.

The interface for the decision makers to select the weight, impact, and probability values is demonstrated below:

	Problem Identification			Alternative Selection		
Enabler	Weight	Impact	Probability	Weight	Impact	Probability
Globalization	3 - moderately important	4 - considerable	0,4	3 - moderately important	4 - considerable	0,3
Energy efficiency	2 - slightly important	2 - low	0,4	4 - important	1 - very low	0,5
Energy savings	3 - moderately important	3 - moderate	0,3	3 - moderately important	3 - moderate	0,4
Incentives	1 - not important	3 - moderate	0,5	2 - slightly important	2 - low	0,3
Tax benefits	4 - important	4 - considerable	0,4	2 - slightly important	3 - moderate	0,5
Climate concerns	3 - moderately important	1 - very low	0,3	3 - moderately important	3 - moderate	0,4
Environmental concerns	2 - slightly important	3 - moderate	0,5	5 - very important	1 - very low	0,3
Cost savings	2 - slightly important	2 - low	0,1	4 - important	3 - moderate	0,5
Good examples	3 - moderately important	3 - moderate	0,4	4 - important	1 - very low	0,4
Energy self-sufficiency	5 - very important	3 - moderate	0,3	4 - important	3 - moderate	0,3

Figure 6 Interface for selecting the weight, impact, and probability values

Once the weight, impact, and probability values are selected, the utility value for each element (disutility for disablers) is calculated by multiplying the corresponding selected values.

That is, for key drivers:

$$Utility\ Value\ of\ key\ driver\ (k) = \ Weight\ (k) * Impact\ (k) * Probability\ (k)$$

For a key driver with weight selected by the decision maker as 4-important, impact selected as 3-moderate, and probability selected as 0.7, the utility value would be: (4*3*0.7)=8.4

For enablers:

$$Utility\ Value\ of\ enabler\ (e) = \ Weight\ (e) * Impact\ (e) * Probability\ (e)$$

For an enabler with weight selected by the decision maker as 3-moderately important, impact selected as 2-low, and probability selected as 0.8, the utility value would be: (3*2*0.8)=4.8

For disablers:

$$Disutility\ Value\ of\ disabler\ (d) = \ Weight\ (d) * Impact\ (d) * Probability\ (d)$$

For a disabler with weight selected by the decision maker as 5-very important, impact selected as 1- very low, and probability selected as 0.9, the utility value would be: (5*1*0.9)=4.5.

Once the utility score of each process phase is computed, the overall utility score of the process is calculated by the average (or the minimum) of the phase utility scores.

That is:

$$Overall\ Utility\ Score\ of\ the\ Process = \frac{\sum_{p\ \epsilon\ \{process\ phases\}} Utility\ score\ of\ phase(p)}{Number\ of\ phases\ of\ the\ process}$$

or

$$Overall\ Utility\ Score\ of\ the\ Process = Min_{p\ \epsilon\ \{process\ phases\}}\{Utility\ score\ of\ phase(p)\}$$

For a process with utility score of the problem identification phase equal to 20, utility score of the alternative selection phase equal to 22, utility score of the planning phase equal to 17, utility score of the implementation phase equal to 19, and utility score of the monitoring phase equal to 15, the overall utility score is calculated with respect to the average method as: (20+22+17+19+16)/5 = 18.8, and with respect to the minimum method as min (20,22,17,19,16) = 16. For a more convenient perception, these scores, which have a maximum value of 25, can be converted to the percentage scale via multiplying each value by 100/25 = 4. In a such case, the percentage utility values calculated by the average method and minimum method will be 18.8*4 = 75.2% and 16*4 = 64%, respectively.

Use of the minimum utility score is justified in cases where the prevalent viewpoint is that the phases are related in a tandem manner; hence, the overall process is as successful as its least successful phase. The decision tree algorithm is flexibly designed so that other formulations of the overall utility score can be implemented, for example, by taking one of the following: the maximum of the phase utility scores, the product of the phase utility scores, or the weighted sum of the phase utility scores.

Next, the details of the decision tree algorithms for the three decision making levels are demonstrated.

3.1 Decision Tree Algorithm for Formal Social Units

There are four points of difference between the structure of the decision tree algorithm for formal social units differ and those of the other decision-making levels: the triggers, key drivers, enablers, and barriers pertaining to formal social units. The triggers and key drivers are derived from D6.3 and the enablers and disablers are derived from D6.2.

The triggers for the formal social units are problem-driven process and goal-driven process.

The key drivers for the formal social units used in the design of the decision tree algorithm are identified as iteration of top-down and bottom-up mechanisms, as well as mid-level management contribution.

The enablers for formal social units are listed in the table below:

Enabler	
Globalization	Good examples
Energy efficiency	Energy self-sufficiency
Energy savings	Prosumerism
Incentives	Local production
Tax benefits	Awareness
Climate concerns	Information
Environmental concerns	Communication
Cost savings	

Table 1 Enablers for Formal Social Units

The disablers for formal social units are listed in the table below:

Disabler	
Circumstances	
	legal
	financial
	environmental
	economic
Mismanagement	
Operational mistakes	
Lack of awareness	
Administrative	
	organizational
	capacity
	procedural
	conflicts
	trust and transparency
Percevied value of energy	
Social and individual	
	habits
	resistance to change
	status quo (inertia)
	cultural norms
Uncertainty and risk	
	technological
	regulatory
	political
	legislational

Table 2 Disablers for Formal Social Units

The resulting decision tree structure for formal social units is given below:

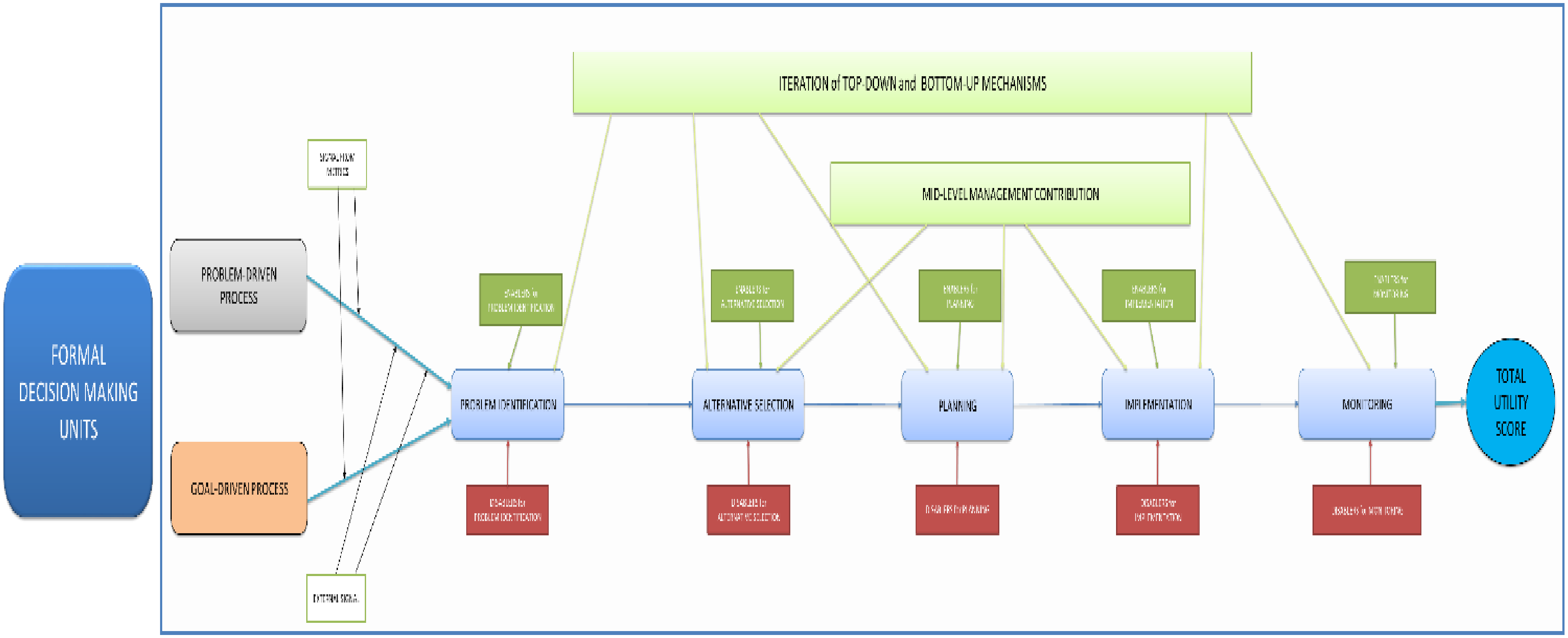


Figure 7 Decision Tree Structure for Formal Social Units

3.2 Decision Tree Algorithm for Collective Decision-Making Units

As with the formal social units, the structure of the decision tree for the collective decision-making units is differentiated by the specification of the triggers, key drivers, enablers, and disablers.

The three types of triggers for the collective decision-making units are: Market-driven process, sustainability-driven process, and legislation and market-driven process.

The key drivers for the collective decision-making units are identified as similar to those of formal social units. These are: iteration of top-down and bottom-up mechanisms and mid-level management contribution.

The enablers for collective decision-making units are as listed below:

Enabler	
Globalization	Good examples
Energy efficiency	Energy self-sufficiency
Energy savings	Prosumerism
Incentives	Local production
Tax benefits	Awareness
Climate concerns	Information
Environmental concerns	Communication
Cost savings	

Table 3 Enablers for Collective Decision-Making Units

The disablers for collective decision-making units are as listed below:

Disabler	
Circumstances	
	legal
	financial
	environmental
	economic
Mismanagement	
Operational mistakes	
Lack of awareness	
Administrative	
	organizational
	capacity
	procedural
	conflicts
	trust and transparency
Percevied value of energy	
Social and individual	
	habits
	resistance to change
	status quo (inertia)
	cultural norms
Uncertainty and risk	
	technological
	regulatory
	political
	legislational

Table 4 Disablers for Collective Decision-Making Units

The resulting decision tree structure for collective decision-making units is given below:

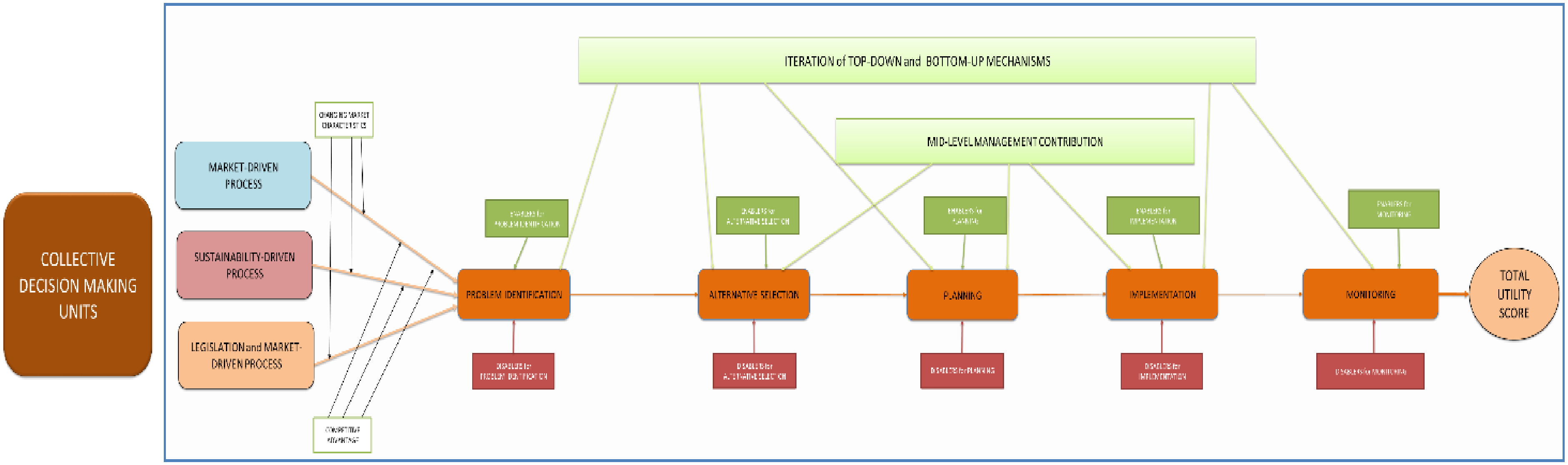


Figure 8 Decision Tree Structure for Collective Decision-Making Units

3.3 Decision Tree Algorithm for Individuals Engaging in Joint Contracts

The processes pertaining to individuals engaging in joint contracts are triggered by an individual-driven process and a joint benefit-driven process.

The key drivers for individuals engaging in joint contracts differ from those of the formal decision-making units and collective decision-making units: there is one enabling key driver, namely, lean organizational structure, and two disabling key drivers: power asymmetries and conflict of interest; and prolonged process.

The enablers for individuals engaging in joint contracts are as listed below:

Enabler	
Globalization	Good examples
Energy efficiency	Energy self-sufficiency
Energy savings	Prosumerism
Incentives	Local production
Tax benefits	Awareness
Climate concerns	Information
Environmental concerns	Communication
Cost savings	

Table 5 Enablers for Individuals Engaging in Joint Contracts

The disablers for individuals engaging in joint contracts are as listed below:

Disabler	
Circumstances	
	legal
	financial
	environmental
	economic
Mismanagement	
Operational mistakes	
Lack of awareness	
Administrative	
	organizational
	capacity
	procedural
	conflicts
	trust and transparency
Percevied value of energy	
Social and individual	
	habits
	resistance to change
	status quo (inertia)
	cultural norms
Uncertainty and risk	
	technological
	regulatory
	political
	legislational

Table 6 Disablers for Individuals Engaging in Joint Contracts

The resulting decision tree structure for individuals engaging in joint contracts is given below:

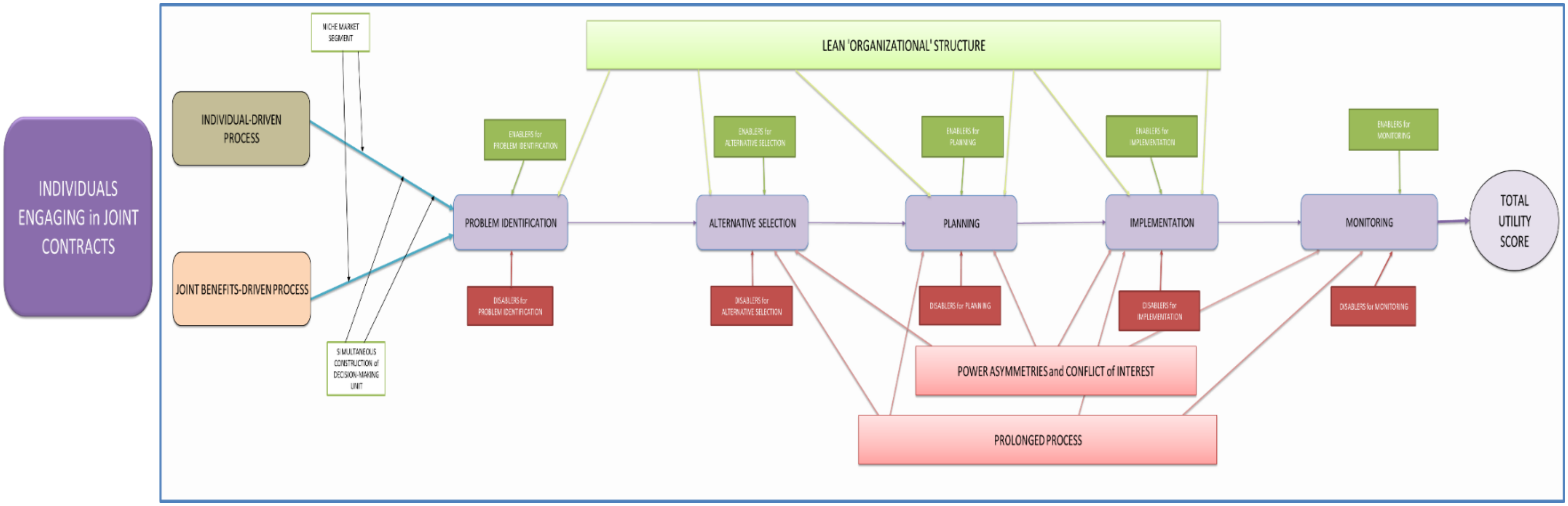


Figure 9 Decision Tree Structure for Individuals Engaging in Joint Contracts

4. Conclusion

The decision tree algorithm is constructed through integrating the results from work performed in WP6 of the ECHOES project. The resulting decision tree algorithm is designed to support planning for decision makers' processes in various decision-making units, as well as for formulating guidelines or policies for different decision-making levels.

One key advantage is its characteristic as "learning algorithm", its ability to utilize existing information and generate patterns pertaining to a system. Such patterns are then utilized to draw assessments and for decision-making in new situations. The existing data for the "learning" part of the decision tree algorithm derives from the analysis of 15 focus group studies, 67 in-depth interviews and 12 case studies in 7 selected countries within the scope of WP6 of ECHOES, providing a reliable foundation for the constructed decision tree algorithm, and increasing its predictive capability.

The structure of the constructed decision tree follows the mapping of the phases of energy-related processes. These phases are identified considering the best examples/successful implementations within the scope of Deliverable D6.3. This overall structure is common to all three decision making units, namely, formal social units, collective decision-making units, and individuals engaging in joint contracts. The consecutive phases of these processes are problem identification, alternative selection, planning, implementation, and monitoring, which form the decision tree nodes.

Additional nodes of the decision tree come from two sources: first, the enablers and disablers affecting the processes, derived from the enablers and disablers identified in Deliverables D6.1 and D6.2, and second, the information from best practices/successful implementations, as identified in Deliverable D6.3. These nodes pertain to the specific ingredients of the processes for the particular decision making unit. For formal social units and collective decision making units, these are "Iteration of top-down and bottom-up mechanisms" and "Mid-level management contribution", both of which are evaluated as enablers. For individuals engaging in joint contracts, the inputs from best practices/successful implementations are: "Lean organizational structure", considered to act as an enabler, and "Power asymmetries and conflict of interest", and "Prolonged process", considered as disablers.

The motivators and barriers at each node are explicitly evaluated by the decision maker, through a spreadsheet format, in order to assess the total utility score under a specific context. This analytical power of the decision tree algorithm makes it a valuable decision support tool for different levels of decision-makers, or from the perspectives of different decision making units.

For each process phase, the utility score is calculated as the algebraic sum of (dis)utilities implied by enablers, disablers, and additional factors. The total utility score for a particular situation is then calculated by taking the average or minimum of the utility scores at each process phase.

Decision makers can make use of the decision tree algorithm by first selecting the appropriate decision-making level, then the type of process defined by the trigger. Each assessment made by the decision maker for the motivators, barriers, and additional factors results in a total utility score. The decision maker can construct scenarios by assigning different values for weights, impacts, and probabilities of motivators, barriers, and additional factors. Scenarios can reflect different projections, base, optimistic, or pessimistic perspectives, or estimated effects of planned projects (e.g. implementation of an incentive scheme to decrease the impact of the 'cost' barrier). Comparison of the scores of different scenarios helps the decision maker in choosing the schemes to implement, planning for risk management, prioritizing projects, or even deciding on whether or not to go on with a particular implementation or project.

The decision maker provides three pieces of information. First, for each enabler, disabler, or additional factor, the utility value is computed by taking the corresponding weight, impact, and probability information from the decision maker(s). Second, the weight value refers to the relevance or importance as assessed by the decision maker, the impact value shows the decision maker's perception of the magnitude of impact of the enabler, disabler, or additional factor. The last piece of information is the probability, which refers to the likelihood of occurrence of the enabler, disabler, or additional factor.

The interface of the decision tree algorithm receives the inputs from the decision makers by associating linguistic counterparts with the numeric values of the assessments of weight, impact, and probability. This facilitates decision makers' scoring, making it easier to translate their perceptions into numeric values.

References

Agnihotri, D., Verma, K., Tripathi, P., (2017), Variable Global Feature Selection Scheme for automatic classification of text documents, Expert Systems with Applications, Vol. 81, pp. 268-281.

Bach, D.R., Dayan, P., (2017), Algorithms for survival: a comparative perspective on emotions, Nature Reviews Neuroscience, Vol. 18, pp. 311–319.

Biresselioglu, M.E., Demir, M.H., Demirbag Kaplan, M., Solak, B., (2018), An analysis of the parameters that determine the similarities and differences regarding the energy choices and energy related behaviour between different types of formal social units, Report No ECHOES 6.2 – D6.2 Parameters Analysis. Available at: https://echoes-project.eu/sites/echoes.drupal.pulsartecnalia.com/files/D6.2.pdf (Accessed 16 July 2019).

Biresselioglu, M.E., Demir, M.H., Solak, B., Nilsen, M., Sæther, S.R., Qiu, X., Manoz, L.L.,... Koljonen, T., (2019), Suggestions and Recommendations for a Better Understanding of the Factors Driving Collective Energy Choices and Energy Related Behaviour, Report No ECHOES 6.3 – D6.3 Suggestions and Recommendations.

Caruana, R., Niculescu-Mizil, A., (2006), An Empirical Comparison of Supervised Learning Algorithms, ICML '06 Proceedings of the 23rd international conference on Machine learning, pp. 161-168.

Cha, S., Jeong, S., Oh, H., (2018), A scalable learning algorithm for data-driven program analysis, Information and Software Technology, Vol.104, pp. 1-13.

Dai, Q., Zhang, C., Wu, H., (2016), Research of Decision Tree Classification Algorithm in Data Mining, International Journal of Database Theory and Application, Vol.9(5), pp. 1-8.

Dietterich, T.G., (2006), Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms, Neural Computation, Vol. 10(7), pp. 1895-1923.

Faraj, S., Pachidi, S., Sayegh, K., (2018), Working and organizing in the age of the learning algorithm, Information and Organization, Vol. 28(1), pp. 62-70.

Garg, P., Neider, D., Madhusudan, P., Roth, D., (2016), Learning invariants using decision trees and implication counterexamples, POPL '16 Proceedings of the 43rd Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, pp. 499-512.

Jadhav, S.D., Channe, , H.P., (2014), Comparative Study of K-NN, Naive Bayes and Decision Tree Classification Techniques, International Journal of Science and Research, Vol. 5(1), pp. 1842-1845.

Jaworski, M., Duda, P., Rutkowski, L., (2018), New Splitting Criteria for Decision Trees in Stationary Data Streams, IEEE Transactions on Neural Networks and Learning Systems, Vol. 29(6), pp. 2516-2529.

Jordan, M.I., Mitchell, T.M., (2015), Machine learning: Trends, perspectives, and prospects, Science, Vol. 349, pp. 255-260.

Kim, K., (2016), A hybrid classification algorithm by subspace partitioning through semi-supervised decision tree, Pattern Recognition, Vol. 60, pp. 157-163.

Lin, J., Rosasco, L., (2018), Generalization properties of doubly stochastic learning algorithms, Journal of Complexity, Vol. 47, pp. 42-61.

Mathan, K., et al., (2018), A novel Gini index decision tree data mining method with neural network classifiers for prediction of heart disease, Design Automation for Embedded Systems, Vol. 22(3), pp. 225-242.

Quinlan, J.R., (1986), Induction of Decision Trees, Machine Learning, Vol. 1, pp. 81-106.

Sæther, S.R. et al., (2018), Policy Recommendations: An analysis on collective and energy related decisionmaking processes of three formal social units, Report No ECHOES 6.1 – D6.1 Policy Recommendation papers. Available at: https://echoes-project.eu/sites/echoes.drupal.pulsartecnalia.com/files/D6.1.pdf (Accessed 17 July 2019).

Segatori, A., Marcelloni, F., Pedrycz, W., (2018), On Distributed Fuzzy Decision Trees for Big Data, IEEE Transactions on Fuzzy Systems, Vol. 26(1), pp. 174-192.

Song, Y.Y., Lu, Y., (2015), Decision tree methods: applications for classification and prediction. Shanghai Arch Psychiatry, Vol.27(2), pp. 130–135.

Wu, S. et al., (2017), Survey on Prediction Algorithms in Smart Homes, IEEE Internet of Things Journal, Vol. 4(3), pp. 636-644.

Zhao, H., Li, X., (2017), A cost sensitive decision tree algorithm based on weighted class distribution with batch deleting attribute mechanism, Information Sciences, Vol. 378, pp. 303-316.