Problem Chosen	2023	Team Control Number
D	MCM/ICM Summary Sheet	2321591

Multi-Criteria Decision Making for SDGs using K-Means and AHP

Summary

The United Nations has established **17 Sustainable Development Goals (SDGs)**, which represent humanity's vision for a better future world. These 17 goals may have **positive or negative impacts** on each other, and therefore, it is crucial to establish a prioritization model for handling these goals. To accomplish this, we require a model that involves both **network and graph theory** and a model that can determine the most prioritized items.

In Task 1, we constructed a **network graph** of the 17 SDGs, where we assigned the same color to points representing goals belonging to the same category and used lines to indicate their relationships. To avoid the intersection of points and lines, we let the 17 targets form a circle, so that our diagram will look clearer.

In Task 2, we used the traditional **Analytic Hierarchy Process (AHP)** method to establish a model for selecting the optimal goal from the 17 SDGs. First, we provided 5 criteria for evaluating the goals as the decision layer, and then the 17 SDGs as the alternative layer. By comparing the judgment matrices through the **AHP algorithm**, we obtained the evaluation weights of the 5 criteria and the priority levels (represented by numerical values) of each goal. The goal with the highest priority value was considered the most prioritized item.

In Task 3, we aimed to determine the new most prioritized item after a particular goal had been accomplished. Considering that the traditional AHP method would result in complicated operations and substantial human input, and that removing a goal would require changing the parameters of the original matrix, we used the **K-Means clustering algorithm** to classify the 17 goals. After classifying the goals, we placed goals from the same class in the same **AHP alternative layer**, and each class's local most prioritized item was determined. Finally, we included these local most prioritized items in the **AHP model** to determine the overall most prioritized item.

In Task 4, due to unexpected events, we realized that the original decision layer might no longer be appropriate, so we replaced it with the United Nations' official **5P classification** for the 17 SDGs to adapt to the new situation. The other procedures were similar to those in Task 3. Here we discuss what happens to the 5Ps, which are indicators at the decision-making level, when economic, technological, political, and environmental factors change; we also plot pie charts comparing these changes to the original state.

In Task 5, we discussed in detail what kind of problems other companies and organizations can solve using our model and how to solve them using our model. We also considered other situations that might arise in practical problems, and therefore, we provided optimization suggestions for these situations.

Finally, we conducted sensitivity analysis, analyzed the strengths and weaknesses of the model, and provided prospects for the model's future development.

Keywords: 17 SDGs   Graph and Network   Analytic Hierarchy Process(AHP)   K-Means cluster algorithm   5P classification

Contents

1 Introduction

1.1 Problem Background

The 17 Sustainable Development Goals (SDGs) are not standalone objectives, but rather a network of interdependent goals that interact with each other in various ways. The success of achieving one goal often depends on the success of achieving other goals. For instance, ensuring access to clean water and sanitation (SDG 6) is critical to reducing poverty (SDG 1) and improving health and well-being (SDG 3). Similarly, promoting affordable and clean energy (SDG 7) is essential to combatting climate change (SDG 13) and promoting sustainable cities and communities (SDG 11).

Conversely, the failure to achieve one goal can also hinder progress towards other goals. For example, if education (SDG 4) is not accessible to all, then achieving gender equality (SDG 5) and reducing poverty (SDG 1) become more difficult. Moreover, the failure to address climate change (SDG 13) may exacerbate hunger (SDG 2) and limit the availability of clean water and sanitation (SDG 6).

Therefore, it is essential to recognize the interconnectedness of the SDGs and address them holistically, rather than separately. To achieve sustainable development, we must prioritize integrated approaches that consider the impacts of achieving one goal on other goals. This requires collaboration and partnerships among governments, international organizations, civil society, and the private sector to develop strategies and implement solutions that address the interrelated challenges faced by the world. By working together and taking a systemic approach, we can achieve the SDGs and create a better world for all.

Figure 1: The 17 SDGs of UN

1.2 Restatement of the Problem

Considering the background information and restricted conditions identified in the problem statement, we need to establish a model that is universal in its applicability to different athletes and complete the following tasks using the model:

- **In Task 1** We need to create a **network** of relationships among the 17 SDGs. Since each of these 17 goal relationships has its own uniqueness, we treat each of the 17 goals as 17 nodes of the network and construct the undirected network. In a certain two nodes, i.e., between certain two goals, there is a unique relationship function. We indicate that they have certain connections by connecting lines in the relational network. The details are described in Task 1.

- In **Task 2**, We need to use the individual SDGs as well as the network structure to set priorities, and we need to prioritize each event. This task requires a generalization of the 17 SDGs, as well as a decision model that can effectively calculate priorities and make stage-by-stage projections of future development trends.
- In **Task 3**, We need to test the scalability of the model. We need to consider how the structure of our network will change when one of the SDGs is achieved. We need to compare the structural changes in the network model before and after the hypothesis, as well as think about related issues that may occur in the future, and suggest possible sustainability topics that may arise in the future.
- **In Task 4**, We need to think about the evolution of models under special events. When global progress or major disasters occur, we need to explore the changes in the model and how these changes affect the results of our model calculations, i.e., the priorities of the Sustainable Development Goals. We also need to think about how these changes will affect the future of the UN.
- **In Task 5**, We need to extend the model and explore how it can be extended from the 17 UN Sustainable Development Goals to other companies or organizations for goal prioritization.

1.3 Overview of Our Work

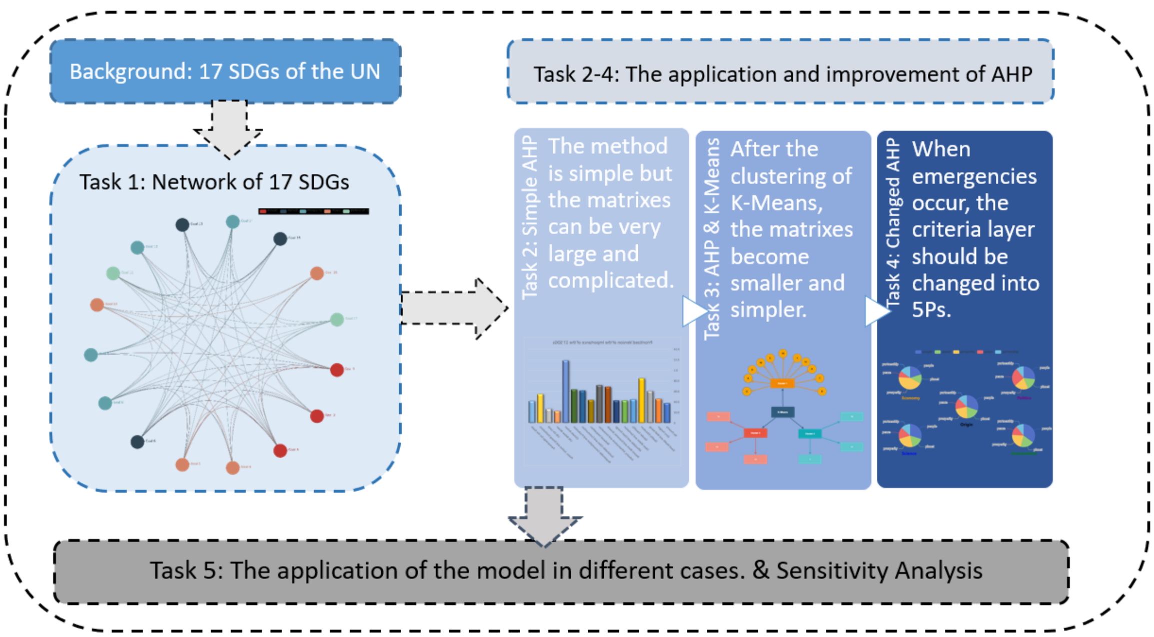


Figure 2: Our Work Structure

2 Assumptions and Notations

2.1 Assumptions

- **Task 1:** The goals can be represented by points and lines between the points can show their relationship.
- **Task 2:** The goals can be measured by several dimensions, and their priority can be measured by number.
- **Task 3:** The goals can be clustered by K-Means with several metrics, and when a goal is achieved, it will be removed from our model.
- **Task 4:** The emergencies makes the metrics change in a large scale, so we need to change the metrics.
- **Task 5:** The organizations or companies' problems will be similar to this problem.

2.2 Notations

Symbol	Definition
$\mathbf{w}$	Eigenvectors
$\mathbf{v}$	Sum of column vectors of the discriminant matrix
λ_{max}	Maximum eigenvalue of the judgement matrix
$\mathbf{CR}$	Consistency ratio
$\mathbf{A}$	Judgment matrix
$\mathbf{W}$	Weight matrix
$\mathbf{S}$	Solution score matrix
$\mathbf{R}$	Decision matrix
$\mathbf{N}$	scheme normalized score matrix
$\mathbf{CI}$	Consistency index
$\mathbf{RI}$	Random consistency index

3 Task1: Network Model based on Network Flows

Task one requires us to complete a 17-node network. According to the information given in the article, the 17 nodes are the 17 sustainable development goals officially proposed by the United Nations. They have already been described in section 1.1 of the article, so we follow Goal 1,2 here... 17 names to replace them.

According to the official United Nations paper, we can classify the 17 SDGs into 5 broad categories, namely the 5 Ps: **People, Prosperity, Peace, Planet and Partnership**, which together encompass the

17 SDGs. Of course, this does not mean that each goal is independent of the other, rather, each goal is more or less related to the 5 Ps, but is classified according to its main content.

The 17 Sustainable Development Goals (**SDGs**) outlined in the 2030 Agenda for Sustainable Development can be classified into the following **5 Ps** based on the official report on un.org:

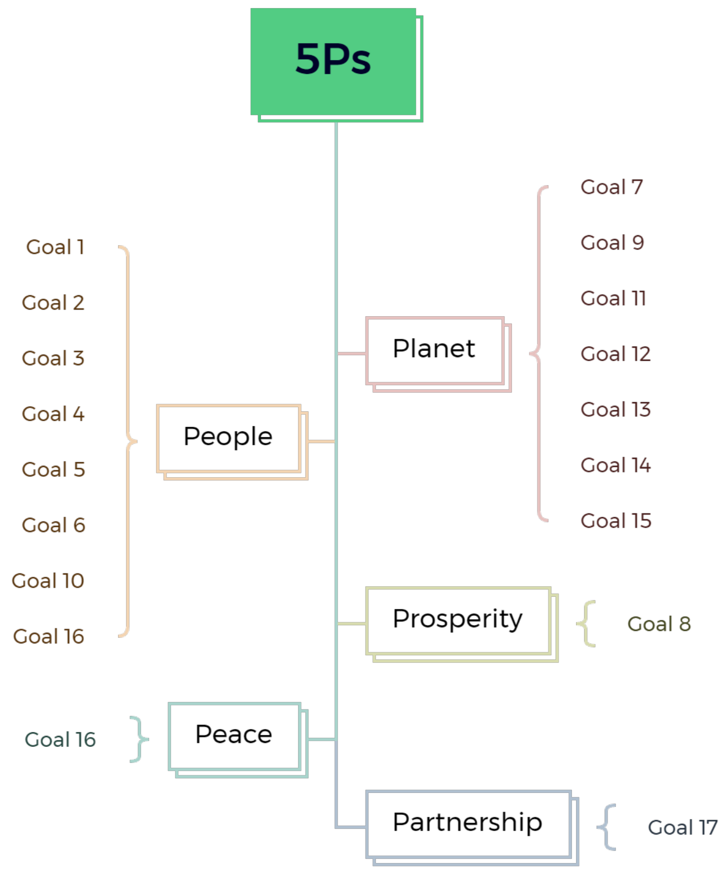


Figure 3: SDGs by Categories

Subsequently, we rationally completed the network structure as shown in the figure, based on the nature of the association of each SDG and the reality of the situation.

This network of relationships is a very visual representation of highly relevant issues, while the five different colors show the main directions and elements to which the SDG belongs. The color of the line between two SDGs indicates the type of relationship between the two goals, and the goal with the same color as the connected curve has a strong relevant influence on the connected goal. If two goals have the same color as the connected curve, it means that they have a mutual influence.

In summary, this circular undirected relational network simply and clearly depicts the inductive and action relations between different objectives.

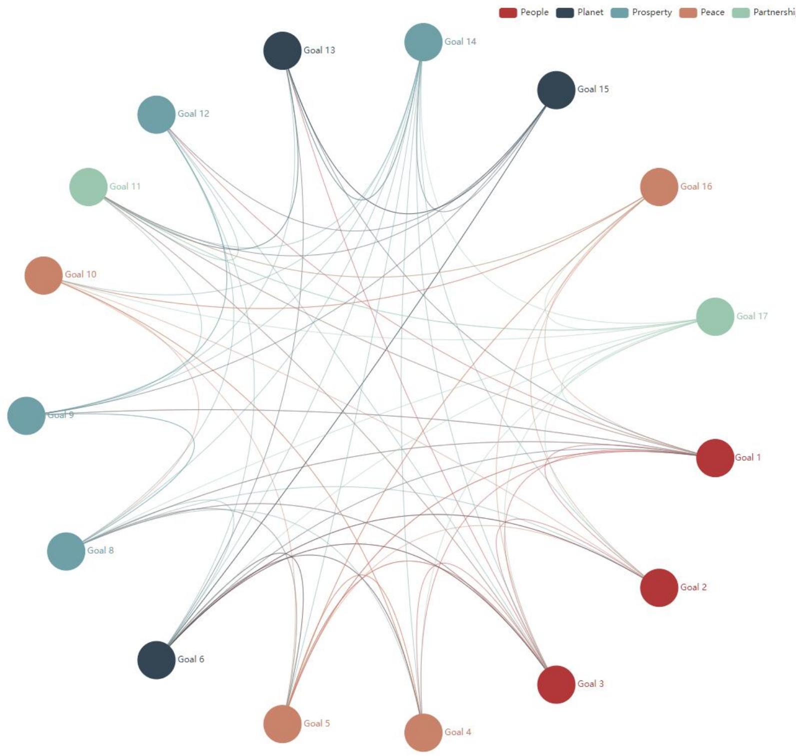


Figure 4: The Network of 17 SDGs

4 Task2: Hierarchical Analytic Decision Model Based on AHP

4.1 Preparation for Model I

According to the topic, we need to analyze 17 sustainable development goals for decision making. We will model the analysis based on the **Analytic Hierarchy Process** (AHP), which is an easy way to quantify qualitative problems and is useful for abstracting policy decisions.

Here is how to apply the Analytic Hierarchy Process (AHP) to the 17 Sustainable Development Goals (SDGs) that were put forward by the United Nations:

- **Step 1: Define the problem and objectives**

  The problem is to prioritize the 17 SDGs to determine which ones are most important to focus on. The objective is to identify the most pressing SDGs that require the most attention and resources.

- **Step 2: Create a hierarchy**

  The hierarchical structure for the SDGs could look like this:

  - **Goal:** Sustainable Development
  - **Criteria:** Economic, Social, and Environmental Sustainability
  - **Sub-criteria:** Targets of each SDG

- **Step 3: Pairwise comparisons**

For each level of the **hierarchy**, pairwise comparisons are made using a scale of 1 to 9. Decision-makers compare each element to every other element to determine their relative importance. For example, decision-makers might compare Economic Sustainability to Social Sustainability and Environmental Sustainability to determine their relative importance.

- **Step 4: Create a weight matrix**

  Using the results of the pairwise comparisons, a weight matrix is created for each level of the hierarchy. The weight matrix reflects the relative importance of each element. For example, the weight matrix for the criteria level might show that Economic Sustainability is more important than Social Sustainability and Environmental Sustainability.

- **Step 5: Evaluate the alternatives**

  In this case, the alternatives are the 17 SDGs. For each criterion and sub-criterion, decision-makers evaluate each SDG using a scale of 1 to 9. For example, decision-makers might evaluate SDG 1 (No Poverty) in terms of its contribution to Economic Sustainability.

- **Step 6: Create a score matrix**

  Using the results of the pairwise comparisons and the evaluations of the alternatives, a score matrix is created for each alternative. The score matrix reflects the overall desirability of each SDG. For example, the score matrix might show that SDG 13 (Climate Action) has a high overall desirability score because it contributes strongly to Environmental Sustainability.

- **Step 7: Check for consistency**

  A consistency check is performed by calculating the consistency index and consistency ratio. If the consistency ratio exceeds a pre-defined threshold, the decision-makers are required to revise their pairwise comparisons until a consistent set is obtained.

Once the consistency check is complete, decision-makers can use the results to prioritize the SDGs and focus their attention and resources on the most important ones. For example, the results might show that SDG 3 (Good Health and Well-Being) is the most important because it contributes strongly to all four criteria of **Economic, Social, Technological and Environmental** Sustainability.

In order to meet the requirements of Task 1, we set up Model 1, the details of which have been described in detail in Chapter 2. The ideas and methods of model setup are described here.

4.2 Establishment of Model I

4.2.1 Parameter Settings and Calculation about AHP

Here we set these parameters: the discriminant matrix for comparison is **X**=$(x_{ij})_{n\times n}$, which has dimensions.

The eigenvalues of this matrix are λ, **CI** is the consistency index, **RI** is the random consistency index, **CR** is the consistency ratio rate.

Firstly, we need to perform the calculation of relevant parameters based on Analytic Hierarchy Process (AHP). The calculation process is shown below.

- **Step1: Calculate the eigenvectors of the judgment matrix**

$$\mathbf{w} = \frac{\mathbf{v}}{\sum_{i=1}^{n} v_i} \tag{1}$$

Here, **w** stands for the eigenvectors of the discriminant matrix , **v** stands for the sum of column vectors of the discriminant matrix.

- **Step2: Calculate the Consistency Ratio (CR) of the Judgment Matrix**

$$\text{CR} = \frac{\lambda_{\text{max}} - n}{n - 1} \tag{2}$$

Here, λ_{max} represents the maximum eigenvalue of the judgment matrix, and n represents the order of the matrix.

- **Step3: Calculate the Maximum Eigenvalue of the Judgment Matrix**

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0 \tag{3}$$

Here, **A** represents the judgment matrix, λ represents the maximum eigenvalue, and **I** represents the identity matrix.

- **Step4: Calculate the Weight Matrix**

$$\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_m] \tag{4}$$

Here, **W** represents the weight matrix, and $\mathbf{w}_i$ represents the weight vector of the ith layer.

- **Step5: Calculate the Solution Scores**

$$\mathbf{S} = \mathbf{WR} \tag{5}$$

Here, **S** represents the solution score matrix, **W** represents the weight matrix, and **R** represents the decision matrix.

- **Step6: Calculate the Normalized Solution Scores**

$$\mathbf{N} = \frac{\sum_{i=1}^{n} s_i}{S} \tag{6}$$

where **N** denotes the scheme normalized score matrix, and here s_i denotes the score or weight of scheme i under some hierarchy. Specifically, it can be the score or weight of a criterion or sub-criterion under some upper-level criterion, or the score or weight of a scheme under some criterion or sub-criterion.

4.2.2 Comparison of Criteria Level

Once the criterion and solution are determined, we need to perform two levels of comparison, namely the criterion level comparison and the solution level comparison. Both of these two levels of comparison need to be represented using matrices, which are described separately below.

Suppose there are n criteria and we need to determine the weights among them. We can construct a matrix A of $n \times n$, where A_{ij} denotes the importance of criterion i with respect to criterion j. Since the importance of criterion i with respect to criterion j and the importance of criterion j with respect to the ratio of the importance of the criterion i should be equal, so the matrix A is a symmetric matrix, i.e., $A_{ij} = 1/A_{ji}$.

We need to measure the importance between criteria using a **comparison scale**, which is generally a number from 1 to 9, where 1 means that both criteria are equally important, 3 means that one criterion is slightly more important than the other, 5 means that one criterion is significantly more important than the other, 7 means that one criterion is very important than the other, and 9 means that one criterion is absolutely more important than the other.

In order to determine the matrix A, we need to let the user make a two-by-two comparison for each pair of criteria to obtain the comparison scale a_{ij} and then calculate A_{ij} according to the following formula.

$$A_{ij} = \begin{cases} 1 & i = j \\ \alpha_{ij} & i \neq j \end{cases} \tag{7}$$

For example, suppose we have five criteria: **Importance, Achievability, Urgency, Time Horizon and Alignment**.We need to ask users to compare them two by two. If the user thinks that urgency is slightly more important than importance and significantly more important than social benefits, then the comparison scales are 3 and 5. Thus, we can calculate the matrix A as follows.

4.2.3 Comparison of Alternative Level and Development

Suppose there are m scenarios and we need to determine the weights among them. We can construct a matrix B of $n \times m$, where B_{ij} denotes the importance of criterion i to scheme j.

Similarly, we need to use a comparison scale to measure the effect of the criterion on the program

In the following, I will describe the hierarchical analysis method and model it in the context of the problem.

Hierarchical analysis is divided into three parts, which are the upper middle level and the lower level. The middle level is the decision level and the lower level is the solution level. The decision level is divided into different core decision points and each solution is related to the decision level, so that the upper and lower levels have a "comparative discriminant matrix" with the decision level respectively.

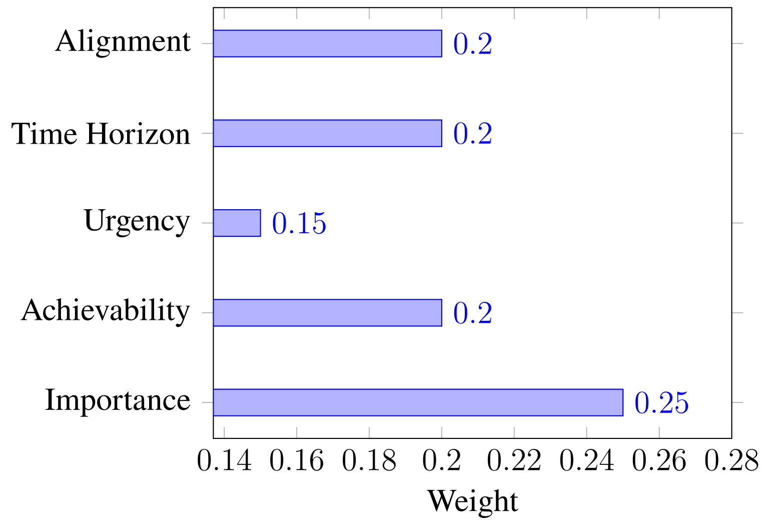
$$\mathbf{X} = (x_{ij})_{n \times n} \tag{8}$$

where n represents the number of various decision indicators or solutions in the middle or lower level, and x_{ij} denotes the relative importance of each i element with respect to the quantification of the jth element with respect to some indicator in the upper level, where the element $x_{ii} = 1, i = 1, ..., n$ located on the diagonal; for the elements in the matrix, the following conditions need to be satisfied:

$$x_{ij} > 0, x_{ij} = \frac{1}{x_{ji}}, x_{ii} = 1, i, j \in \{1, ..., n\}$$

We need to decide the weight proportion of the decision level through discussion and analysis, and we refer to relevant papers in this regard.

The results are shown below:



For this weight icon, we conclude, based on the consistency test, that the icon is consistent with the relevant logic and has some scientific reference value.

Next, we provide an in-depth analysis of the relevant content.

According to the theory of AHP, we can get the final weight ratios of each decision by solving the characteristic roots of the weight matrix. Subsequently, we need to discuss the corresponding **weight ratios** for each solution, in a certain characteristic, so that we can calculate the weight analysis relationship for each solution, and thus get the final weight relationship for each objective. Due to the large content of the correlation matrix, we list it in the final code section. Below are the final values of the weight relationships between the different sustainable development goals that we obtained.

Looking at the data, we can see that SDG 13, Climate Action, has the highest priority value for this item, so Goal 13 has the highest priority.

Climate Action is a broad topic that typically refers to a variety of measures taken to address climate change. These measures can be divided into the following categories.

- **Reduce greenhouse gas emissions:** Reduce greenhouse gas emissions: This is one of the most basic measures to combat climate change. By using clean energy, improving energy efficiency, and promoting low-carbon travel patterns, we can reduce greenhouse gas emissions such as carbon dioxide and methane.

- **Adaptation to Climate Change:** Adaptation to climate change has become an important task as climate change is already causing a range of impacts, such as extreme weather events and sea level rise. Measures to adapt to climate change include building healthier and safer cities, protecting ecosystems, and reducing risks.

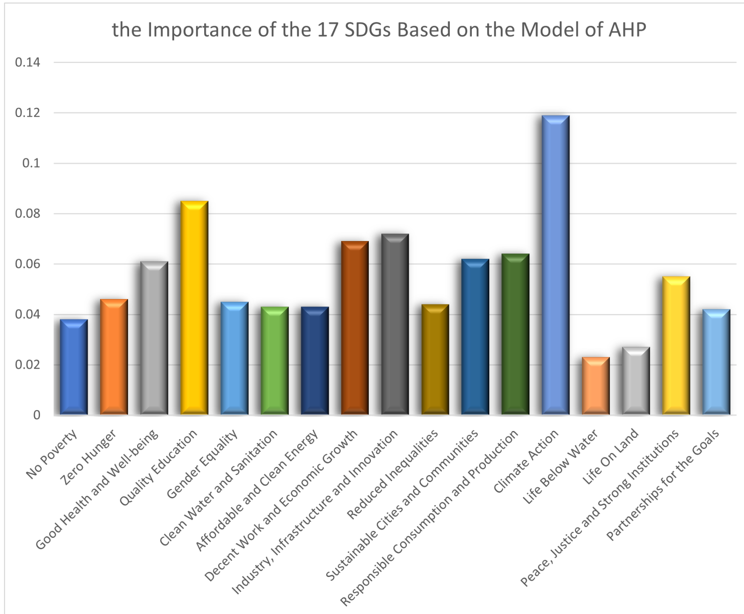


Figure 5: the Importance of the 17 SDGs Based on the Model of AHP

- **Transition to a sustainable development model:** Since existing development models often come at the expense of the environment, a transition to a sustainable development model becomes a must. This includes an economic model that values environmental protection, efficient resource use, and encourages innovation and investment in sustainability.
- **Climate Awareness:** While more and more people are recognizing the seriousness of climate change, there is still a need to increase global public awareness and understanding of climate change in order to better contribute to it.

Climate Action covers many areas, including energy, urban planning, transportation, agriculture, water management, etc., and the measures taken should be diverse.

If the **implementation of climate action** is initiated as a priority, the following objectives can be achieved.

- **Reducing greenhouse gas emissions:** Carbon emissions can be reduced and global warming slowed by promoting renewable energy and clean energy technologies, and by enhancing energy efficiency.
- **Protecting ecosystems:** By promoting sustainable development and conservation measures, we can mitigate the effects of climate change and promote the restoration and protection of ecosystems.
- **Promoting sustainable economic development:** By investing in renewable energy and other green technologies, new jobs and economic growth can be created, driving sustainable economic development.

- **Raising public awareness:** Through education and publicity, public awareness of climate change and environmental protection can be raised, action and cooperation can be promoted, and the awareness and action of the whole society to tackle climate change together can be formed.

In conclusion, implementing climate action can promote sustainable social and economic development, mitigate the effects of climate change, and increase public awareness and attention to environmental protection, leading to a cleaner, greener, healthier and more sustainable lifestyle for the next decade.

5 Task 3: Hierarchical Clustering Decision Model based on AHP and K-Means

5.1 Cluster Optimization based on Model I

The AHP model we have just built is a good representation of the most important target decisions we need to make, but it has the obvious drawback that our model requires a lot of data and a lot of work to process the data. This model is not perfect in terms of ease of use and simplicity. Therefore, we have improved the AHP model, especially to enhance its ease of use and computational simplicity.

The refinement of the model has two parts:

- **On the one hand,** we have changed the overall data processing process and added a data pre-processing process to facilitate better and more complete classification of the data in the end and to perform the final calculations.

- **On the other hand,** we optimized the pre-processing content of the model for better simplification of the data.

Our optimizations are as follows.

- **Step1:** We use the 5Ps as the evaluation criteria for the 17 goals, and each goal will be scored from 0-100 based on these 5 evaluation indicators, which in turn leads to a scoring matrix composed of 17 goals $\mathbf{X} = [x_1, x_2, ..., x_n]$, Then each target will have a score vector for these 5 criteria
$$x_i = (x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5}), i = 1, ..., 17, x_{ij} \in 0, ..., 100 \tag{9}$$
A vector representing the score assigned to the ith goal under the People, Planet, Prosperity, Peace, Partnership evaluation metric.

- **Step2:** Then the 17 targets are classified into N classes by K-means clustering algorithm with 5 evaluation metrics as features. Also, because the K-means algorithm classification has uncertainty, we need to use this algorithm several times in this dataset and choose the one with the smallest loss function among multiple results as our final classification.

- **Step3:** The objectives in each class are put into the solution layer of AHP separately by class, so that the relatively more important one in these classes can be solved by AHP, and then the most important one in each class is taken out and put into AHP again for solving, so that the second optimal solution after a certain objective is completed can be calculated.

5.2 Preparing and Parameter Setting for Model II

First, the 17 scenarios are represented as 17 points in the feature space, each point representing a scenario, and the dimension of the feature space is chosen as needed. For AHP, each scheme corresponds to multiple criteria/factors, and each criterion/factor can be used as a dimension of the feature space. While for **K-Means**, each scheme may have only one or several numerical indicators as features, and each indicator can be used as one dimension of the feature space.

These points are then clustered into several groups using the K-Means algorithm. the basic idea of the K-Means algorithm is to randomly select k initial clustering centers and then attribute each point to the cluster of its nearest clustering center. Then the centers of each cluster are recalculated and each point is reattributed to the cluster of its nearest cluster center, and so on iteratively until the cluster centers no longer change or the preset number of iterations is reached

For each cluster, AHP is used to further analyze it for decision making. For each cluster, it can be considered as a whole for AHP's hierarchy building and judgment matrix construction. In the judgment matrix, for the solutions in the same cluster, the relative importance between them is determined by the AHP method. Eventually, the weighted score of each cluster can be obtained, indicating the relative merits of the solutions represented by that cluster.

Finally, the weighted scores of all clusters are considered together, and the cluster with the highest score and the solution in it are selected as the final decision solution.

$$\begin{cases} \lambda_{max} = \frac{\mathbf{w}^T \mathbf{A} \mathbf{w}}{\mathbf{w}^T \mathbf{w}} \\ CI = \frac{\lambda_{max} - n}{n-1} \\ CR = \frac{CI}{RI} \\ w_i = \frac{\prod_{j=1}^{n} a_{ij}^{1/n}}{\sum_{i=1}^{n} \prod_{j=1}^{n} a_{ij}^{1/n}} \\ \mathbf{w} = (w_1, w_2, ..., w_n) \end{cases} \qquad RI = \begin{cases} 0, & n = 1 \\ 1, & n = 2 \\ \frac{0.58(n-2)}{n-1} + 1, & n \geq 3 \end{cases}$$

Figure 6: Analytic Hierarchy Process (AHP) Parameters Setting

The parameters of this model are similar to those of model 1 and will not be repeated here.

5.3 Establishment of Model II:

Suppose there are n schemes and each scheme i is represented by m indicators, we consider each scheme as a point $\boldsymbol{x_i} = (x_{i1}, x_{i2}, \cdots, x_{im})$ in m-dimensional space, where x_{ij} denotes the value of scheme i on the jth indicator.

First, these points are clustered into k groups using the k-means algorithm to obtain the clustering center of each group $\boldsymbol{c_j} = (c_{j1}, c_{j2}, \cdots, c_{jm})$, where $j \in 1, 2, \cdots, k$. The clustering centers are calculated as follows.

We use the following equation to calculate the centroid c_j of the jth cluster:

$$c_j = \frac{1}{n_j} \sum_{i=1}^{n} x_i \cdot I(y_i = j) \tag{10}$$

Here, n_j represents the number of samples in the jth cluster, and $\mathbb{I}(\cdot)$ is an indicator function that takes the value of 1 when $y_i = j$ and 0 otherwise. Initially, we can randomly select k points as the initial clustering centers.

Next, for each cluster j, we use AHP for further decision analysis. Firstly, we construct the hierarchy of AHP based on the schemes in the cluster and obtain the judgment matrix $\boldsymbol{A_j}$. In the judgment matrix, the relative importance among the schemes in the same cluster is determined by the AHP method.

The calculation steps are as follows:

1 **Construct** the judgment matrix $\boldsymbol{B}$, where b_{ij} represents the importance of scheme i relative to scheme j, which can be determined by experts or subjective evaluation. The judgment matrix should satisfy the positive reciprocal property, i.e.,

$$b_{ij} = 1/b_{ji} \tag{11}$$

2 **Normalize** the judgment matrix to obtain the weight vector $\boldsymbol{w} = (w_1, w_2, \cdots, w_n)$, where w_i represents the weight of scheme i.

$$w_i = \sum_{k=1}^{n} \left(\prod_{j=1}^{n} b_{kj}^{1/n} \right) \left(\prod_{j=1}^{n} b_{ij}^{1/n} \right) \tag{12}$$

3 **Calculate** the weighted score s_{ij} of each scheme in the cluster, which represents the relative importance of scheme i in cluster j.

$$s_{ij} = \sum_{k=1}^{n_j} w_k \cdot w_i \tag{13}$$

We have illustrated the **Framework Diagram** below.

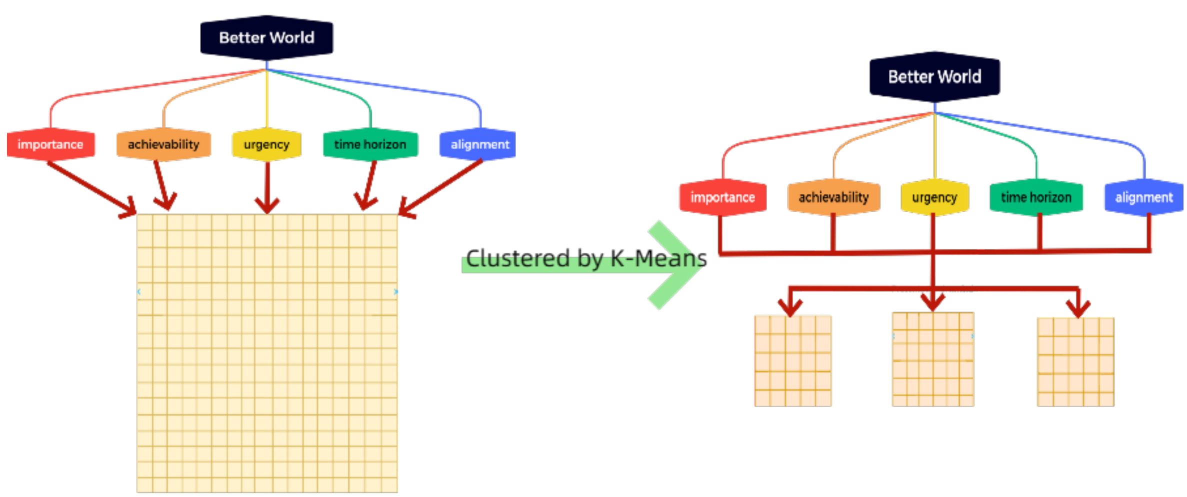


Figure 7: The Change of AHP After the Goals Clustered by K-Means

5.4 Application for Model II

According to the flowchart, we used K-means algorithm to preprocess the data and then used AHP to calculate the final weight proportions. After clustering the 17 sustainable development goals with K-means, we obtained K=3, and the clustering results are as follows:

Table 1: Clusters of Sustainable Development Goals

Cluster 1	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 15
Cluster 2	1, 16, 17
Cluster 3	12, 13, 14

Based on the K-means clustering and subsequent AHP calculation, we can divide the 17 sustainable development goals into four categories, each containing the goals from the respective clusters. Then, we can calculate the weights for each category using AHP, resulting in the following:

- **Cluster 1:** Sustainability goal
- **Cluster 2:** Infrastructure and inequality goals
- **Cluster 3:** Education and health goals

Table 2: Weights for Goal Clusters

Goal Cluster	Weight
Social sustainability goals	0.479
Economic sustainability goals	0.319
Environmental sustainability goals	0.202
Infrastructure goals	0.556
Inequality goals	0.444
Education goals	0.560
Health goals	0.440

Therefore, the weights for the 17 sustainable development goals are as follows.

Assuming there is no hunger, we have conducted K-means clustering on the 17 sustainable development goals based on the evaluation matrix

$$\mathbf{X} = [\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}] \tag{14}$$

where $\mathbf{x_i} = (x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5}), i = 1, \ldots, 17, x_{ij} \in 0, \ldots, 100$ represents the score vector of the ith goal in the People, Planet, Prosperity, Peace, Partnership dimensions.

Next, we use the AHP method to calculate the weights of each sustainable development goal in the five dimensions. Let $\mathbf{w} = (w_1, w_2, w_3, w_4, w_5)$ be the weight vector, and the specific values are as follows:

$$\mathbf{w} = (0.176, 0.192, 0.281, 0.099, 0.253) \tag{15}$$

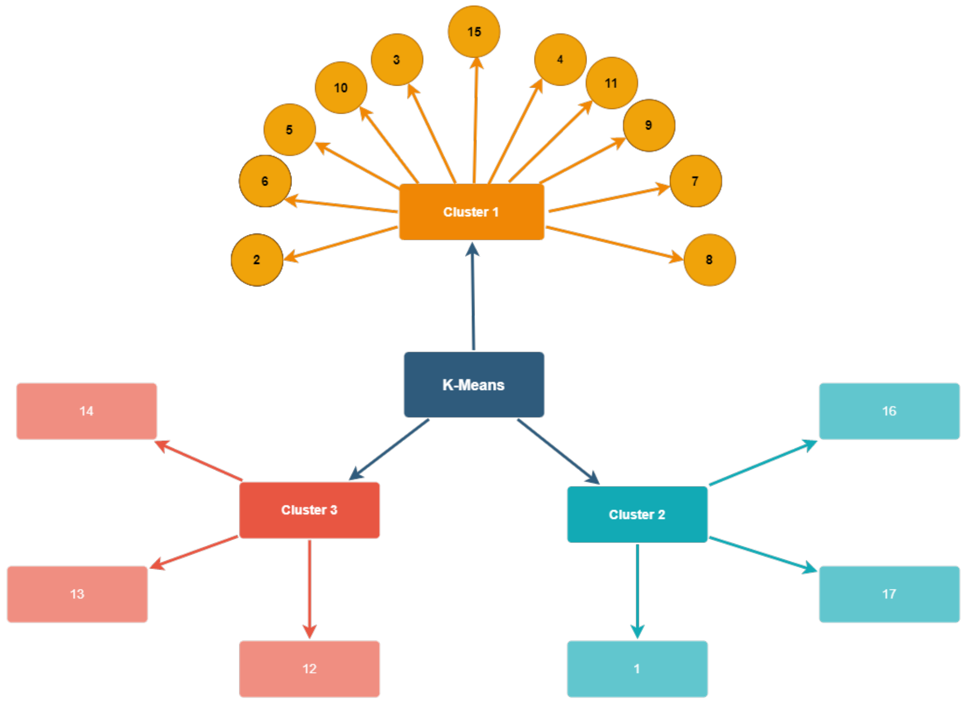


Figure 8: Clusters Divided by K-Means

and finally, the weights of Importance,Achievability,Urgency,Time Horizon,Alignment are 0.176, 0.192, 0.281, 0.099, and 0.253, respectively.

After calculation, in the absence of hunger, the weights of the 17 sustainable development goals are:

According to the results, quality education and climate action should be the first priorities to promote sustainable development in a hunger-free world.

Similarly, when there is no poverty, the weight values for Importance, Achievability, Urgency, Time Horizon, and Alignment are 0.180, 0.197, 0.289, 0.101, and 0.233 respectively.

Some weights of the 17 sustainable development goals are:

- **No Hunger**: 0.065
- **Good Health and Well-being**: 0.064
- **Quality Education**: 0.066

When poverty is eradicated, it becomes crucial to prioritize quality education in order to ensure technological progress and civilized development, which in turn can promote sustainable development. Quality education provides individuals with the knowledge and skills needed to actively participate in the workforce, innovate new technologies, and make informed decisions about their lives and communities. This can lead to a more sustainable and equitable society, with improved health, reduced inequality, and increased economic growth.

More detailed information can be found in the bar chart below. The chart provides a clear view of the weights assigned to different goals, as well as their comparisons and trends over time.

Figure 9: Weights for No Poverty and No Hunger

6 Task 4: Clustering Optimization Prediction Decision Model based on AHP and K-Means,

6.1 Overview of Model III

Task 4 requires us to conduct predictive analysis based on major events. Our previous model lacked this predictive capability and focused more on analyzing the current situation and the evolution of the process under specific hypothetical scenarios, lacking the ability to predict future developments under certain special circumstances. Therefore, we optimized Model 2 by changing the calculation method of the weight layer and optimizing the AHP model based on K-means, which enables us to achieve predictive capabilities under certain special circumstances.

6.2 Establishment of Model III

Model 3 differs from Model 12 mainly in that we changed the decision layer of AHP in Model 3. The decision layer of AHP was changed to 5Ps (People, Planet, Prosperity, Peace, Partnership) for model decision-making. There are several reasons why the use of 5Ps for AHP decision-making is reasonable:

- **Scientificity:** 5Ps is a framework used to comprehensively consider the five main aspects of sustainable development: People, Planet, Prosperity, Peace, and Partnership. It was developed based on the United Nations Sustainable Development Goals (SDGs) and is widely recognized and used.
- **Comprehensiveness:** 5Ps covers 17 sustainable development goals and has sufficient predictive ability for future major events.
- **Applicability:** It is compatible with the calculation method of AHP and is well suited for the model's adaptability.

After changing the decision layer and using 5Ps, we can conduct a highly summarized analysis of relevant content. Any major event, such as technological advancement, global pandemics, etc., can change the importance of the 5Ps, i.e., change the weights of the 5Ps, thereby optimizing our final calculations and satisfying our prioritization estimation method.

For example, when a global pandemic occurs, the weight of People will continuously increase, while the weights of other relevant content will decrease. Based on the estimated weights, we recalculate the 17 SDGs through our Model 2 K-means clustering analysis and AHP method calculations. We can clearly obtain the priority of each time.

To facilitate unified management and prediction, we categorized potential global issues into the following:

- **Environmental factors:** including climate change and global pandemics, which are usually global in nature and have profound impacts on the global ecosystem and human health.
- **Technological factors:** including global technological advancements and major breakthroughs, which are usually related to significant progress in STEM fields and industrial breakthroughs, and have important impacts on human health, ecological environment, and social life.
- **Political factors:** including regional wars and refugee flows, which are usually related to political situations and geopolitics, and have serious impacts on the economy, society, and population of specific regions.
- **Economic factors:** including technological advancements and international crises, which are usually related to the global economy and have far-reaching impacts on global trade, industry, and employment.

Next, we will conduct weight analysis and data processing for each of the four categories of issues in order to obtain priority changes.

6.3 Application and Development for Model III:

The main change in Model III is in the processing logic and decision-making module, while the overall formula calculation and parameter processing are similar, so we will not repeat them here.

First, we consider how the 5Ps weights are in today's society. Here is our preset weight matrix.

Category	Weight
People	0.30284
Planet	0.19048
Prosperity	0.22393
Peace	0.15436
Partnership	0.12839

Table 3: Weights for the 5 Ps with updated values

We will consider how the weights of the 5 Ps will change under the influence of the four major factors: environment, technology, politics, and economy.

6.3.1 Political Factors

We will now take the political factor as an example to demonstrate the process of content analysis.

First, we need to analyze the influence matrix based on relevant materials. The following is a possible matrix of the impact of political factors on the 5 Ps:

$$\begin{bmatrix} 1 & 0.8 & 0.9 & 0.7 & 0.6 \\ 1.2 & 1 & 1.1 & 0.8 & 0.7 \\ 1.1 & 0.9 & 1 & 0.7 & 0.6 \\ 1.3 & 1.2 & 1.3 & 1 & 0.8 \\ 1.4 & 1.3 & 1.4 & 1.2 & 1 \end{bmatrix} \tag{16}$$

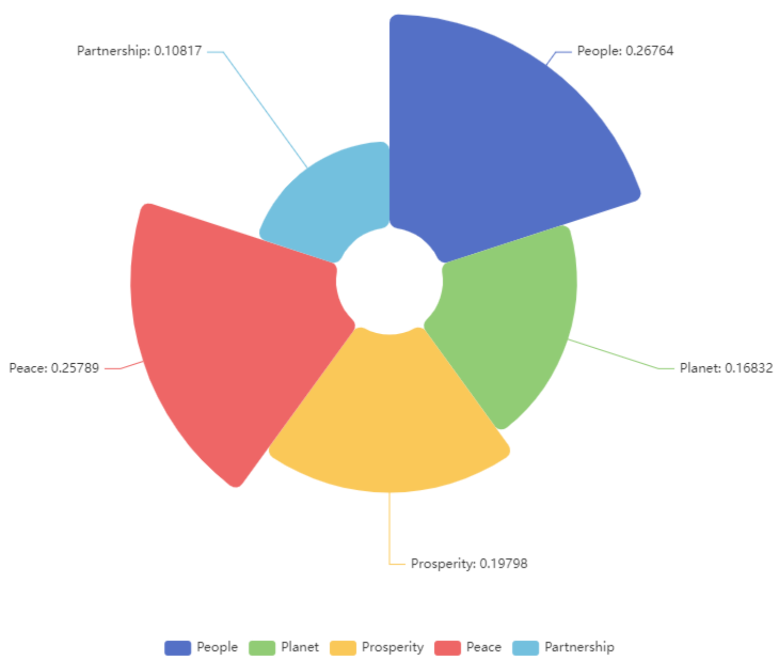


Figure 10: Weights under the influence of political factors

Based on this matrix, we can obtain the final numerical values:

Category	Weight
People	0.26764
Planet	0.16832
Prosperity	0.19798
Peace	0.25789
Partnership	0.10817

Table 4: Weights for the 5 Ps with updated values

6.3.2 Others

Category	Environment	Technology	Economy
People	0.35	0.2969	0.20
Planet	0.19	0.1693	0.13
Prosperity	0.22	0.2382	0.43
Peace	0.13	0.1608	0.09
Partnership	0.11	0.1348	0.15

Table 5: Combined weights for environment, technology economy

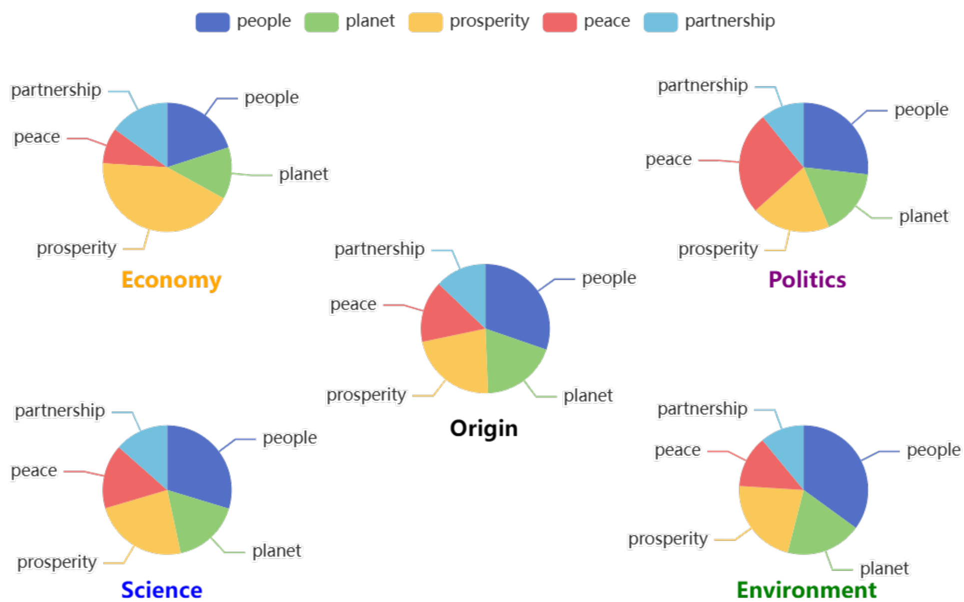


Figure 11: Pie Charts: Combined Weights of Origin, Economy, Science, Politics, Environment

Environmental issues can lead to climate change, economic recession or depression, social unrest or conflict, and so on. Therefore, such weight ratios were derived.

When significant scientific and technological progress emerges, it can be seen that compared to the original weights, the weight of People decreases slightly, while the weight of Prosperity increases slightly, with relatively little relative change in the other three Ps.

When encountering economic depression and other issues, it will have a profound impact on global trade, industry, and employment.

7 Task5: Application of Predictive Modeling Methods in Organizational Decision Making

First, other companies and organizations can elaborate on the interrelationship of the goals through a network similar to the one in Task 1; they can also color the dots representing the goals based on an initial categorization of the goals. This network diagram template helps other companies and organizations to clarify the relationship between the goals they need to accomplish, and to mark similar goals with the same color.

Second, the **K-Means algorithm** and **AHP** used by our group can be useful for other organizations and enterprises. For example, when the decision level of these organizations and enterprises involves more issues, we can use the K-Means algorithm to classify each issue in the decision level, and then use the classification to put the split objectives into the decision level of AHP, which can reduce the workload and make the results more scientific.

At the same time, we believe that these companies and organizations may also encounter another situation: for example, the variance of the weights assigned to the indicators at the decision-making level in these companies may be large, some may be large, and some may be much smaller than others. Although we did not encounter this situation in these tasks, none of the evaluation indicators had a "much smaller importance" than the others. However, it is worth considering this situation.

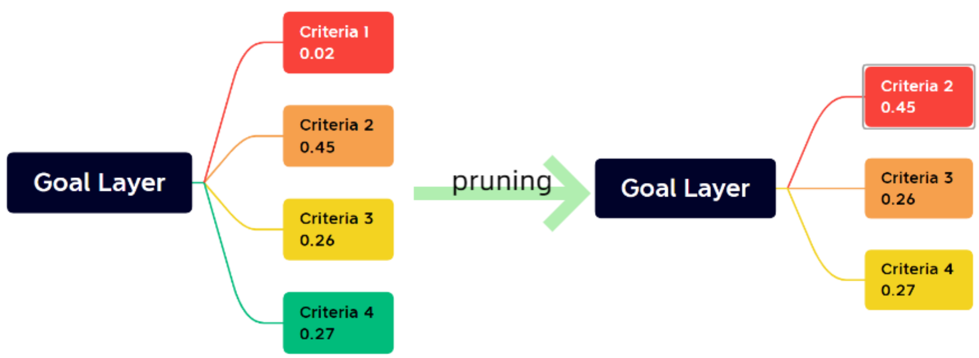


Figure 12: Pruning of Unimportant Criteria Elements

We think that if this situation occurs, we can remove all the indicators that have much less weight than other indicators at the decision level by a method similar to "pruning", so that on the one hand, the weight of these indicators is relatively small and has little impact on the final decision, and on the other hand, the number of matrices between the decision and solution levels can be reduced. The number of matrices between decision and solution layers is reduced, which further improves the speed of computation.

Here we think that if the full time of a certain indicator is 1/10 of the average of other indicators, then we can consider it as a **"small impact indicator"** and do not consider it; when calculating the average, only those indicators that seem to have a relatively large weight are selected, and the size of small indicators are not considered in the calculation.

8 Sensitivity Analysis

8.1 Model 1

Model 1 is an AHP-based method network, which has two parameters:

- **Parameter 1:** Weight matrix
- **Parameter 2:** Plan weight matrix

Changes in both parameters can affect the final result, but the weight matrix has a greater impact. For each plan, its requirement for the weight matrix is usually fixed. For example, the requirement for Sustainable Development Goal 1: No Poverty is fixed, and we have calculated that the Importance weight for Goal 1 is the largest, at 0.302. These data are related to the nature of each SDG and are generally fixed. They may vary slightly in different countries or regions, but this is not within the scope

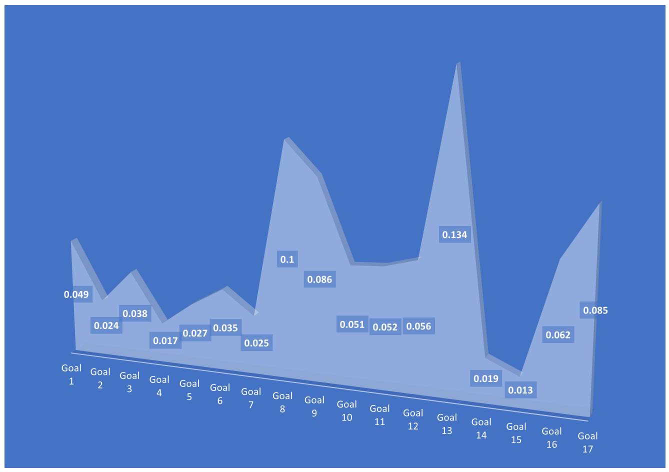


Figure 13: Weight of changes after minor perturbations

of our consideration. The impact of the weight matrix is significant, especially for Model 1, which requires 17 plans to be changed through the assignment of weight matrix values.

If we change the parameters of the weight matrix.then the 17 plans would change as above. Therefore, slight changes in the weight matrix have high sensitivity in Model 1 and can lead to many changes in the final results.

8.2 Model 2 and Model 3

Model 2 is based on K-means and AHP, and has three parameters:

- **Parameter 1:** Weight Matrix
- **Parameter 2:** Solution Weight Matrix
- **Parameter 3:** Number of Clusters in K-Means

The result of K-means clustering directly affects the calculation of AHP weights, so the size of K will have a significant impact on the final result. We can try changing the value of K and observe the change in SDG weights.

In the AHP method, the composition of the decision layer and the criterion layer will also affect the final weights. We can try changing the composition of the decision layer or criterion layer and observe the change in SDG weights. This is similar to the sensitivity analysis in Model 1.

Task 2 requires us to consider the **scalability and development** of our model when a certain indicator is achieved. This actually involves changing the types and weight matrices of solutions. K-means clustering will introduce some changes, and AHP weights will also change. Among them, the weight matrix is still the most important. Similar to Model 1, when hunger is completely eliminated, the individual properties of different sustainable development goals will differ, thereby affecting the allocation of weights.

Model 3 is an **optimized version** of Model 2, based on K-means and AHP. The parameter settings are the same as in Model 2, but the weight module has been changed to use the 5Ps for weight analysis. The most important parameter in this part is undoubtedly the changes in the weight of the 5Ps, as Model 3 is aimed at solving the requirements of task 4, which involves addressing the demand situation in the event of global climate change. We made tentative changes to the weight of the 5Ps, and the overall data changes are shown in the figure below:

Figure 14: Weight of changes after minor perturbations

Our changes to the weight of the 5Ps demonstrate the impact of changes in global events on our model or the changes in weight that may occur due to the evolving global situation. By adjusting the weights of the 5Ps, our model can effectively adapt to changes in global priorities and shifting focus on specific goals, making it a robust tool for sustainable development planning and decision-making. The flexibility and adaptability of our model make it a valuable asset for policymakers and stakeholders to assess and evaluate the progress of the SDGs and make informed decisions accordingly.

In addition, the model provides valuable insights for decision-makers to design effective policies and strategies to achieve sustainable development goals in different countries and regions. It also serves as a useful reference for future research on sustainable development and decision-making, providing a systematic and quantitative approach to address complex global issues. Overall, our model has the potential to make a significant contribution to promoting sustainable development and achieving a better future for all.

9 Discussion

9.1 Strength

- **Our Model** has a high degree of scalability and modularity, and is highly practical for handling special hypothetical scenarios and future predictions. It quantifies abstract concepts and provides a demonstration and guidance for decision-making in problem-solving.
- **Model I** fully utilizes the AHP method for decision analysis, and simultaneously carries out innovative and summarizing comprehensive planning for the 17 sustainable development goals, providing reference solutions for future decision-making regarding sustainable development goals.
- **Model II** integrates the use of K-Means model for comprehensive decision analysis. K-Means can not only perform integrated processing but also preprocess the complex data to achieve a more convenient decision-making model. This is a highly innovative combination and an excellent solution for multi-option decision-making.
- **Model III** has made a secondary innovation based on Model 2, optimizing the decision direction of the decision-making layer, ensuring the predictability and overall fluidity of the model, and possessing high practicality for long-term forecasting of the problem.

9.2 Weakness

- It is preferable to use the Delphi method to obtain the weights of the decision layer in the AHP model. However, due to the tight schedule of the competition, we had to refer to relevant materials for analysis, which may inevitably lead to errors.
- The K-Means method is best to undergo multiple training attempts to obtain the final clustering solution. Due to the tight schedule of the competition and the insufficient relevant data, we were unable to perform multiple iterative training for comparison. Instead, we adopted the clustering solution that performed well in the test data.

9.3 Future Work

Considering the shortcomings of the model, we also hope that some improvements can be made in the future:

- **Data preprocessing:** In K-Means, data preprocessing has an important impact on the accuracy and stability of clustering results. Data cleaning, normalization, and dimensionality reduction can be used to process data and improve the effectiveness of the model.

- **Parameter selection:** In K-Means and AHP methods, the selection of model parameters also has a significant impact on the accuracy of the results. For example, in K-Means, the optimal number of clusters can be determined by methods such as the elbow method and silhouette coefficient. In AHP, the subjectivity of decision-makers and the completeness of information collection also affect the results and require careful consideration.
- **Model integration:** In K-Means and AHP methods, more advanced model integration methods can be used to improve the effectiveness of the model. For example, ensemble learning and deep learning can be used to integrate the results of multiple models and improve prediction accuracy and stability.
- **Model evaluation:** Model evaluation is a critical step in K-Means and AHP methods, and various indicators can be used to evaluate the effectiveness of the model to determine if it meets expected performance. For example, in K-Means, indicators such as SSE and silhouette coefficient can be used to evaluate the accuracy of clustering results. In AHP, consistency indicators and consistency ratios can be used to evaluate the reliability of decision-making results.

References

[1] Shiping Cheng, Ming Chen and Yue Ying, The combination of AHP and K-means clustering Method for determining the weight of evaluation indicators for doctoral dissertations,Changsha Hunan: Graduate School of Central South University, 2021:2-3

[2] SDGs.Wikipedia: https://en.wikipedia.org/wiki/

[3] Our World in Data: https://ourworldindata.org/

[4] Statista: https://www.statista.com/

[5] The UN:https://sdgs.un.org/goals

[6] Bojie Fu, Shuai Wang, Junze Zhang, Zengqian Hou, Jinghai Li, Unravelling the complexity in achieving the 17 sustainable-development goals, National Science Review, Volume 6, Issue 3, May 2019, Pages 386–388, https://doi.org/10.1093/nsr/nwz038

[7] George Halkos, Eleni-Christina Gkampoura, Where do we stand on the 17 Sustainable Development Goals? An overview on progress, Economic Analysis and Policy, Volume 70, 2021, Pages 94-122, ISSN 0313-5926, https://doi.org/10.1016/j.eap.2021.02.001.

[8] Luis Ospina-Forero, Gonzalo Castañeda, Omar A. Guerrero, Estimating networks of sustainable development goals, Information and Management, Volume 59, Issue 5, 2022, 103342, ISSN 0378-7206, https://doi.org/10.1016/j.im.2020.103342.

[9] Bellantuono, L., Monaco, A., Amoroso, N. et al. Sustainable development goals: conceptualization, communication and achievement synergies in a complex network framework. Appl Netw Sci 7, 14 (2022). https://doi.org/10.1007/s41109-022-00455-1

[10] Our **ANONYMOUS** Code Repository:https://anonymous.4open.science/r/MCM-2023-2321591-Codess-740F