People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research

Abdelhamid Ibn Badis University of Mostaganem
Faculty of Exact Sciences and Computer Science
Department of Mathematics and Computer Science

DOCTORATE (L.M.D) THESIS IN COMPUTER SCIENCE

Specialty: Artificial Intelligence and its applications

Identification And Detection of Outliers in Multicriteria Decision Aid.

Presented by: ACHIR Toufik

On 10/14/2025 in front of the committee composed of

SEHABA Karim	**President**	**Prof**	**University of MOSTAGANEM**
ROUBA Baroudi	**Supervisor**	**MCA**	**University of MOSTAGANEM**
MEROUFEL Bakhta	**Reviewer**	**MCA**	**University of SIDI BELABES**
MECHAOUI Moulay Driss	**Reviewer**	**MCA**	**University of MOSTAGANEM**
LAREDJ Mohamed Adnane	**Reviewer**	**MCA**	**University of MOSTAGANEM**

Academic Year: 2024 – 2025

Acknowledgement

I would like to begin by expressing my gratitude to Almighty God, whose infinite mercy and strength enabled me to overcome challenges and complete this thesis.

My deepest appreciation goes to my thesis supervisor, **Dr. Baroudi ROUBA**, Senior Lecturer at the University of Mostaganem. His thoughtful supervision, insightful feedback, and methodological guidance have been essential to this research. Throughout the years, he remained a dedicated and patient mentor, always ready to provide support and direction. I am truly thankful for his availability, encouragement, and intellectual generosity. Thank you, Sir, for your commitment, academic wisdom, and the respect and care you showed during this journey.

I would like to express my sincere gratitude to the members of the jury who kindly agreed to evaluate my doctoral work:

To Professor **SEHABA Karim**, from the University of MOSTAGANEM, for accepting to serve as the president of the jury.

To Senior Lecturer **MEROUFEL Bakhta**, from University of DJILLALI LIABES SIDI BELABES, whom I warmly thank for her kind acceptance to be part of the jury and for evaluating this work.

To Senior Lecturer **MECHAOUI Moulay Driss**, also from University of MOSTAGANEM, to whom I am deeply grateful for participating in the jury despite his demanding administrative responsibilities.

To Senior Lecturer **LAREDJ Mohamed Adnane**, from the University of MOSTAGANEM, who honored me by agreeing to examine this thesis.

Figures List

Tables List

Content Table

Publications and Communications

International Publication

- Toufik Achir , Baroudi Rouba .“A Novel Weighted Preference Relation Approach to Detect Outliers in Multi-Criteria Decision Aid Context”. Foundations of Computing and Decision Sciences (FCDS) Vol 50, N°2, pp.117-156, 2025.
  https://sciendo.com/article/10.2478/fcds-2025-0005?tab=abstract

National Communication

- Toufik Achir, Baroudi Rouba. “Towards identifying multicriteria outliers: An approach based on PROMETHEE γ and DBSCAN algorithm”. CSNT 2024 Computer Science and New Technologies Workshop, Mostaganem, Algérie, 8 Décembre 2024.

International Communication

- Toufik Achir, Baroudi ROUBA and Mohamed Midoun. « A novel approach to detect outliers in MCDA problems ». Proceedings of Scientific Workshop On Preliminary Research Work in Computer Science In 7th International Symposium on Modelling and Implementation of Complex Systems MISC 2022. Mostaganem, Algérie, 30-31 Octobre 2022.
  https://misc2022.misc-lab.org/workshop-proceedings.php

Abstract

Outlier detection, also known as anomaly detection, is a crucial process in data analysis used to identify data points that significantly deviate from the general pattern of a dataset. The problem of outlier detection in the Multicriteria Decision Aid (MCDA) field has not been extensively explored in the current literature. As far as we know, only three papers have addressed this topic.

This study proposes a novel approach for outlier detection in MCDA. The contribution is twofold. The first part introduces the importance degree, a measure that captures the strength of a preference relation between alternatives in PROMETHEE. Each alternative is represented as a vector of its preference values with respect to all other alternatives. The importance degree is calculated using the Euclidean distance between these vectors, providing a clear metric for how strongly one alternative dominates another.

The second part of the contribution applies this concept to outlier detection in MCDA. The proposed approach evaluates how similar an alternative is to others by aggregating importance degrees, thus identifying how typical or atypical it is in the decision space. Alternatives with significantly low similarity are flagged as outliers using statistical tools such as the interquartile range or the standard deviation method.

Together, these contributions offer a novel framework for both modeling preference intensity and improving the detection of anomalous alternatives in MCDA. This enhances the reliability and interpretability of decision support systems that rely on outranking methods like PROMETHEE.

Keywords: degree of importance, preference relation, multicriteria decision aid, PROMETHEE, outlier detection, statistical methods

Résumé

La détection des outliers, également appelée détection d'anomalies, est un processus essentiel en analyse de données visant à identifier les points qui s'écartent significativement du schéma général d'un ensemble de données. Le problème de la détection des outliers dans le domaine de l'Aide Multicritère à la Décision (AMCD) n'a pas été largement exploré dans la littérature actuelle. À notre connaissance, seulement trois travaux ont abordé cette thématique.

Cette étude propose une nouvelle approche pour la détection des outliers en AMCD. La contribution se décline en deux volets. Le premier introduit la notion de degré d'importance, une mesure qui capte l'intensité d'une relation de préférence entre alternatives dans la méthode PROMETHEE. Chaque alternative est représentée sous forme d'un vecteur de ses valeurs de préférence par rapport à toutes les autres alternatives. Le degré d'importance est alors calculé à l'aide de la distance euclidienne entre ces vecteurs, fournissant ainsi une mesure claire de la domination d'une alternative sur une autre.

Le second volet de la contribution applique ce concept à la détection des outliers en AMCD. L'approche proposée évalue la similarité d'une alternative par rapport aux autres en agrégeant les degrés d'importance, permettant ainsi d'identifier dans quelle mesure elle est typique ou atypique dans l'espace de décision. Les alternatives présentant une similarité très faible sont considérées comme aberrantes à l'aide d'outils statistiques tels que l'intervalle interquartile ou la méthode de l'écart-type.

Ensemble, ces contributions offrent un cadre innovant pour modéliser l'intensité des préférences et améliorer la détection des alternatives outliers en AMCD. Cela renforce la fiabilité et l'interprétabilité des systèmes d'aide à la décision reposant sur des méthodes d'agrégation par surclassement comme PROMETHEE.

Mots clés : Degré d'importance, relation de préférence, Aid multicritère à la décision, PROMTHEE, détection des outliers, méthodes statistiques

ملخص

تُعدّ كشف القيم الشاذة، المعروفة أيضًا باكتشاف الحالات غير الطبيعية، عملية أساسية في تحليل البيانات تُستخدم لتحديد النقاط التي تنحرف بشكل كبير عن النمط العام لمجموعة البيانات. لم يتم تناول مشكلة اكتشاف القيم الشاذة في مجال المساعدة على اتخاذ القرار متعدد المعايير (MCDA) بشكل واسع في الأدبيات الحالية. وحسب علمنا، لم يُناقش هذا الموضوع إلا في ثلاث دراسات فقط.

تقترح هذه الدراسة نهجًا جديدًا لكشف القيم الشاذة في .MCDA وتنقسم المساهمة إلى جزئين رئيسيين. يُقدّم الجزء الأول درجة الأهمية، وهي مقياس يُعبّر عن قوة علاقة التفضيل بين البدائل ضمن طريقة .PROMETHEE يتم تمثيل كل بديل على شكل متجه يحتوي على قيم تفضيله مقارنة بجميع البدائل الأخرى. تُحسب درجة الأهمية باستخدام المسافة الإقليدية بين هذه المتجهات، مما يوفر مؤشرًا واضحًا لمدى تفوق بديل ما على الآخرين.

أما الجزء الثاني من المساهمة، فيُطبّق هذا المفهوم على كشف القيم الشاذة في .MCDA يُقيّم النهج المقترح مدى تشابه كل بديل مع بقية البدائل من خلال تجميع درجات الأهمية، مما يُمكّن من تحديد ما إذا كان البديل نموذجيًا أو شاذًا في فضاء القرار. وتُحدّد البدائل التي تُظهر تشابهًا منخفضًا بشكل ملحوظ على أنها شاذة باستخدام أدوات إحصائية مثل نطاق الربيعيات أو طريقة الانحراف المعياري.

تُقدّم هذه المساهمات معًا إطارًا جديدًا لنمذجة شدة التفضيلات وتحسين عملية كشف البدائل الشاذة فيMCDA ، مما يُعزّز من موثوقية وقابلية تفسير نظم دعم القرار التي تعتمد على طرق التجاوز مثل.PROMETHEE

الكلمات المفتاحية: درجة الأهمية، علاقة التفضيل، اتخاذ القرار متعدد المعايير، PROMETHEE، كشف القيم الشاذة، أدوات إحصائية

GENERAL INTRODUCTION

General context

Outlier detection, often referred to as anomaly detection, plays a pivotal role in data analysis by identifying data points that deviate significantly from the overall pattern of a dataset. These unusual observations may indicate errors, rare events, or novel phenomena, and are critically important in fields such as statistics, machine learning, and data mining. Their presence can distort results, reduce model accuracy, and obscure underlying patterns.

A wide range of techniques has been developed to address this issue, which can be broadly classified based on methodological foundations [1]. These include statistical methods [2], which rely on distributional assumptions; distance-based approaches [3], which measure deviation based on proximity; density-based techniques [4] that assess local data density; clustering-based methods [5] that analyze group consistency; graph-based models [6]; ensemble strategies [7] that combine multiple detectors; and learning-based methods [8], including supervised and unsupervised machine learning algorithms.

Outlier detection has found widespread application across numerous domains, such as fraud detection [9], network intrusion detection [10], and wireless sensor networks [11], among many others. These applications benefit from the ability to identify rare or abnormal events that may indicate errors or threats. Despite its broad utility, outlier detection remains a relatively underexplored topic in the field of Multicriteria Decision Aid (MCDA).

In the field of MCDA, methods are specifically designed to support decision-makers (DMs) in evaluating and comparing a set of alternatives based on multiple, often conflicting, criteria. MCDA typically addresses three fundamental types of decision problems[12]. The first involves choice problems, where the objective is to select the most suitable alternative. The second concerns ranking problems, which aim to order the alternatives from the most to the least preferable. The third relates to sorting problems, where alternatives are assigned to predefined categories based on their characteristics.

According to Vincke [13], MCDA approaches can be grouped into three main families: aggregation methods, outranking methods, and interactive methods. Interactive methods [14] involve ongoing interaction between the decision-maker and the system. A preliminary solution is proposed, and if it is deemed unsatisfactory, the decision-maker may provide additional preference information, such as prioritizing certain criteria over others. The model is subsequently updated, and this iterative process continues until a satisfactory solution is achieved. Aggregation methods [15] consolidate the performance scores of alternatives into a single utility value, typically through mathematical functions or weighted averages, which allows for a unified evaluation. In contrast, outranking methods establish pairwise comparisons between alternatives based on their performance across all criteria, aiming to define dominance relationships without collapsing the information into a single numerical score.

The present research focuses specifically on PROMETHEE, a widely used and well-established family of outranking methods [16]. A detailed discussion of PROMETHEE and its relevance to this study will be provided in the subsequent sections of this thesis.

Outlier detection within the field of MCDA has received limited attention in the literature, with only three notable studies addressing the issue. The first study [17] introduces a distance-based approach grounded in the profiles of alternatives by comparing their preference relations (preference, indifference, incomparability) across the set. To reduce the impact of outliers, the method samples subsets of alternatives and analyzes the resulting distance distributions (bimodal patterns) signal potential outliers. The second paper [18] employs a multicriteria method such as PROMETHEE to derive preference relations, and then represents each alternative using a distribution vector that summarizes these relations. Outliers are detected using the LOF algorithm, which identifies local density anomalies. The third approach [19] involves computing the net flow for each alternative and testing for normality. If the values are normally distributed, the standard deviation (SD) method is used; if not, the interquartile range (IQR) method is applied to identify outliers.

Identified problems

Many MCDA methods rely on subjective parameters (such as preference, indifference, weight, and veto thresholds) that reflect the decision maker's preferences. Variations in these parameters can significantly influence results, sometimes producing alternatives that diverge notably from the rest. Such alternatives are known as multicriteria outliers [17]. Detecting these outliers is crucial for improving method robustness and helping decision makers adjust parameters effectively. However, this task poses several challenges:

- MCDA's reliance on preference relations,
- The context-specific definition of outliers,
- The need to adapt existing detection methods for the MCDA framework.

This research is structured around three core questions that guide our contribution. In the context of outranking methods, it is common for multiple alternatives to exhibit similar relationships with a given alternative. For instance, when two alternatives, a_i and a_j are both preferred over a third alternative, a_k , a central question arises: which of the two (a_i or a_j) demonstrates a stronger or more significant preference over a_k? Put differently, is the preference of a_i over a_k equally important as that of a_j over a_k , or does one represent a more substantial dominance? This leads to our first research question: how can we compare the relative strength or importance of such preference relations?

The second question addresses the methodological aspect: how can we quantify this degree of importance in a consistent and meaningful way within the outranking framework? Developing a metric or a function capable of capturing this nuance is crucial for advancing preference modeling in MCDA.

The third and final question explores the practical application of this measure: how can the computed importance degree be used to detect outliers in MCDA evaluations? Furthermore, how does this approach compare in effectiveness to existing outlier detection methods?

By systematically exploring these questions, this study proposes and formalizes the concept of the importance degree of a preference relation. It introduces a straightforward method to compute this degree and evaluates its utility in enhancing outlier detection within MCDA. The proposed approach offers a novel perspective for improving the interpretability and reliability of preference-based decision-making.

Contribution

The contribution of this thesis is structured into two main components, each addressing a distinct but interconnected aspect of multicriteria decision analysis (MCDA), particularly within the PROMETHEE outranking framework.

Part 1: Defining and Quantifying the Importance Degree of Preference Relations

In the first part, we introduce the concept of the importance degree, which captures the relative significance of a preference relation between two alternatives as produced by the PROMETHEE method. The underlying idea is that not all preference relations carry equal weight; some preferences may be more influential or meaningful than others in the decision-making process. To formalize this concept, we represent each alternative a_i in a high-dimensional space using a vector that encodes all its multicriteria preference interactions with the other alternatives. This vectorized representation encapsulates both how an alternative is preferred to others and how it is viewed by others. The importance degree of the preference relation between two alternatives is then computed as the Euclidean distance between their respective vectors. This metric reflects the overall dissimilarity in preference patterns and allows us to quantify how pronounced a preference relation truly is within the decision space.

Part 2: A Novel Framework for Outlier Detection in MCDA

The second part of our contribution presents a new perspective on outlier detection in the MCDA context, along with a refined definition of what constitutes an outlier in such settings. Traditional outlier detection methods may not align well with the unique structure of MCDA problems, where preferences and relations between alternatives play a central role. To address this, we propose a similarity-based approach grounded in the previously defined importance degree. Specifically, we assess how similar each alternative is to all other alternatives by aggregating their pairwise importance degrees. This similarity measure serves as a proxy for how typical or atypical an alternative is with respect to the overall set.

To detect outliers, we analyze the distribution of similarity values across the set of alternatives. Alternatives that exhibit significantly lower similarity (meaning they differ substantially from the majority in terms of their preference relations) are flagged as potential outliers. We apply appropriate statistical techniques, such as interquartile range method or standard deviation method, to formally identify these anomalies.

Together, these two contributions advance the state of MCDA by introducing a nuanced understanding of preference strength and offering a robust, data-driven method for identifying outlier alternatives that may otherwise distort or obscure the decision-making process.

Structure of the thesis

Broadly speaking, this thesis is structured into two main parts: the first provides a review of the state of the art, while the second presents the proposed contributions. The organization of the thesis is as follows:

The first part: State of the art.

Chapter 1:

This chapter presents a comprehensive overview of outlier detection methods, outlining the various types of outliers typically encountered in data analysis. It also examines the diverse applications of outlier detection across several domains and identifies key challenges in effectively applying these techniques.

Chapter 2:

Here, we introduce the foundational concepts of MCDA, followed by an exploration of its practical applications and a classification of the main families of MCDA methods.

Chapter 3:

In this chapter, we investigate existing outlier detection approaches within the MCDA framework and conduct a comparative analysis to evaluate their strengths and limitations.

The second part: Contribution.

Chapter 4:

In this chapter, we outline our principal contribution. We begin by proposing the importance degree metric, designed to assess the significance of preference relations between alternatives. An illustrative example is provided to clarify its computation. Subsequently, we introduce an innovative method for identifying outliers within the MCDA framework.

Chapter 5:

We conduct empirical tests to validate our contribution in this chapter. The first part examines the performance of the importance degree in distinguishing preferences. The second part evaluates the robustness of the overall outlier detection approach by applying it to a different problem instance.

This thesis concludes with a general conclusion and some perspectives.

Part I:
STATE OF THE ART

CHAPTER 1:
OUTLIER DETECTION

1.1 Introduction

Outlier detection is an important concept in statistics, data analysis, and machine learning. It is the process of identifying data points in a dataset that deviate significantly from the majority of the data. Such points can be of interest because they signify some important pattern, possible errors, or an unusual occurrence that needs follow-up investigation. Outlier detection finds application in various fields, including finance, cybersecurity, health, and so forth.

Detecting anomalies is deemed important because, in the related fields of application, often those anomalies represent precious pieces of information. For instance, an outlier in network intrusion may be characterized by unpermitted access to computer networks, resulting in irregular traffic patterns [20]. Outliers in credit card transactions can indicate potential credit card or identity theft [21]. Similarly, the spread of fake news that adversely influences individuals and society as a whole represents an outlier in the domain of misinformation on social networks [20].

The study of outlier detection in data began in the 19th century within the statistics community [22]. Over time, various anomaly detection techniques have been created across multiple research fields. Some of these techniques are specifically designed for certain applications, while others are more general-purpose.

1.2 What are Outliers

The term "outlier" is defined in various ways depending on the field of study and the perspective of the author. Ayadi et al. [23]presented twelve distinct interpretations of outliers as viewed by various authors. For example,[24] defined an outlier as "*one that appears to deviate markedly from other members of the sample in which it occurs*". Additional definitions describe outliers as "*an observation (or subset of observations) which appears to be inconsistent with the remainder of that set of data*," or "*a data point which is significantly different from other data points, does not conform to expected normal behavior, or conforms well to a defined abnormal behavior*"[25]. As shown in Figure 1.1, outliers are commonly characterized based on the following assumptions[25]:

- Outliers differ significantly in their features or values compared to the majority of the data.
- Outliers appear rarely compared to the regular instances in a dataset.

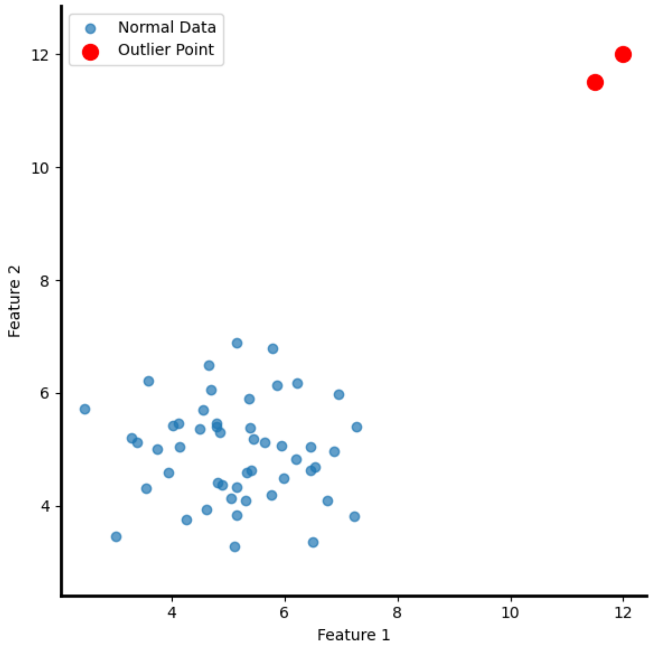


Figure 1.1 A simple example of outliers in a two-dimensional data set.

1.3 Outlier types

Outliers are categorized into three main types: global outliers, contextual outliers, and collective outliers [26]

- *Global Outliers*

In global outliers, an extremely distant data point from others is set to become an outlier in the overall dataset. It is probably the simplest type of outlier and therefore to be at the heart of most studies on outlier detection. For instance, given a human height dataset with most values between 5 and 6 feet, then 9 feet will be regarded as a global outlier: it is an extremely rare value compared to the rest of the data.

- *Contextual Outliers*

Contextual outliers are data points that become outliers within some context or under specific conditions. For example, for a temperature dataset throughout a year, a record of 100°F in winter could be a contextual outlier. Though 100°F might be normal in summer, it is rather unusual in winter.

- *Collective Outliers*

The groups of data points comprising collective outliers cause significant non-conformance to the whole dataset, even if no point within the group itself might be outlying in one dimension or another. In other words, the sale numbers for a certain month are abnormally high, and sales for the subsequent months drop below normal; this collective pattern could be a collective outlier. Each month's data may not present suspiciousness individually, but their combined behavior over time surely is.

1.4 Outlier detection

Outlier detection is the process of finding unusual data points, usually during the step of preparing data. This process is important because:

- Outliers can significantly affect the results of analysis
- Looking closely at outliers can uncover valuable new information.

Finding outliers is very important in many areas because it helps to spot unusual cases that could show mistakes, fraud, rare occurrences, or new information. The process helps people make better choices, makes data more trustworthy, and boosts the security of the system.

1.5 Applications Of Outlier Detection

In this section, we explore how anomaly detection is employed in different sectors, the notion of outliers and the challenges present.

1.5.1 Intrusion detection

Outliers in intrusion detection refer to abnormal activities or patterns that can hint at possible threats such as cyberattacks or the compromising of systems. The major problem with intrusion detection systems is their data dimensionality being too high for efficiently differentiating normal from abnormal patterns. In addition, as both user and system behaviors undergo transformations, complemented systems must be put in place to enable themselves in the learning and detection phases of new types of intrusions.

Intrusion detection systems are classified into two broad categories, namely Host-Based Intrusion Detection Systems (HIDS) [27] and Network-Based Intrusion Detection Systems (NIDS)[28].

1.5.2 Fraud Detection

Outliers in fraud detection are often transactions that carry exorbitantly high amounts deviating from a user's typical spending patterns. They would also include any transactions that do not show regularity in behavioral patterns, such as multiple purchases in very short time intervals or activities that occur at odd hours. The transactions that happen in atypical places for the user or sudden changes in the user behavior like starting a type of transaction they have never done before may also be signs of fraud.

Outlier detection has been found to be paramount in detecting fraudulent activities across different domains ranging from credit card transactions [29], insurance claims [30], telecommunications [31], and financial statements [32]. It searches for anomalies such as unusual spending patterns, high claim frequencies, irregular calling behaviors, and inconsistencies in financial data. This particular area, however, finds itself in a struggle as the number of fraudulent cases pales in comparison to legitimate ones, thus making the training of models quite difficult. What is discouraging, also, is the fact that fraudsters are constantly changing their modus operandi; hence, the models have to be changed often.

1.5.3 Sensor Networks

A curious data, or a set of data, is considered an outlier if it departs from the usual sensor reading patterns. Outliers may arise due to sensor failures, environmental changes, communication channel issues, or even malicious interference [33]. Challenges facing anomaly detection in sensor networks include limited sensor resources, rapidly changing environmental conditions, and noisy or variable data. Furthermore, the large scale of sensor networks requires detection algorithms that are both efficient and scalable.

1.5.4 Medical health care analysis and medical diagnosis

For health care analysis and medical diagnosis, anomalies correspond to serious diseases, errors in data, or even cases of fraud. That is to say, in this domain, outliers are data points that deviate from customary medical patterns, and their detection may pave the way to early diagnosis of diseases, a new set of possibilities for radically changing patient care, and medical decision-making [34]. Challenges faced by outlier detection in healthcare may include data imbalance wherein rare diseases find limited representation in the data and the other being the multi-dimensional complexities of health data. Other major concerns hindering anomaly detection techniques include the interpretability of AI models and the utmost need for the privacy and security of the data.

1.5.5 Falsified News or Misinformation

In the modern digital era, the main issues related to the discovery of fictional news and misinformation arise. Fake news is deliberately ideologically inclined for or misinformed in nature with the aim to sway public opinion to gain clout, kill political outcomes, or earn revenue through clickbait options. Outlier detection techniques are very good at spotting irregularities, inconsistencies, or a strange type of pattern within news articles, social media posts, and other kinds of digital content [35]. Further, detecting either misinformation or truth may remain an issue because of adaptive methods of fake news perpetrators, limited datasets for training, and deep contextual understanding requirements. Ethical issues concerning these automated models and certain implicit biases inherent in them could mislabel genuine content.

1.5.6 Data Integrity and Cleansing

Introduced by various authors, data integrity ensures that data accuracy, consistency, and reliability are maintained through its life. Such outliers in datasets may arise from different reasons: human mistakes, system failures, or malafide data manipulations [36][37]. Detecting and treating these anomalies improves data accuracy by removing incorrect or inconsistent records, strengthens decision-making by ensuring business insights emanate out of clean and trustworthy data, prevents data corruption by identifying forged or manipulated information, and serves machine learning models well by excluding anomalies to enable models to train over top-quality relevant data points.

1.6 The Challenges in Detecting Outliers

Conceptually, an outlier is a pattern that varies from said expected normal behavior. Yet, the detection of outliers is much challenging due to many factors, namely:

- It is difficult to define a normal region that enshrines every possible form of normal behavior. Even so, the boundary between normal and anomalous behavior is often indefinite, so an observation placed slightly on the boundary edge might be erroneously classified.
- Noise usually creeps into the data, be it random errors or sudden fluctuations, sometimes being mistaken for true outliers. Well-robust algorithms should be in place to differentiate out real outliers from mere random variations.
- The very definition of outliers can differ, depending on context and the peculiarities of the data. In one setting, an outlier might be deemed normal in another. This kind of variability means that methods of detecting outliers need to be attuned to the specific data characteristics and the specific problem.
- Availability of labeled data for training and testing models is an anomaly detection problem. These data are often limited-the rarer the anomaly, the more limited the data. This data limitation tends to lead to poorly trained models that perform poorly and produce unreliable detection results in the anomaly detection domain. Sometimes, putting together or annotating enough labeled data is a time-intensive and costly process, worsening the situation even further.

1.7 Outlier detection approaches

Outlier detection approaches can be categorized using various criteria; in our study, we classify them based on the techniques they employ. According to H. Wang et al [20], outlier detection methods can be categorized into several approaches. Statistical-based techniques rely on data distribution models, using parametric or non-parametric methods. Distance-based methods identify outliers by measuring the distance between data points, while density-based methods focus on regions of low data density. Clustering-based approaches classify outliers as points excluded from major clusters, and graph-based methods detect deviations in interconnected data structures. For improved accuracy, ensemble-based techniques combine multiple models, whereas learning-based methods (including deep learning) train on data to adaptively detect complex anomalies (see Figure 1.2). Each approach has distinct strengths depending on data characteristics and application needs.

In the following, we present a concise review each approach, along with an analysis of their benefits, limitations, and challenges.

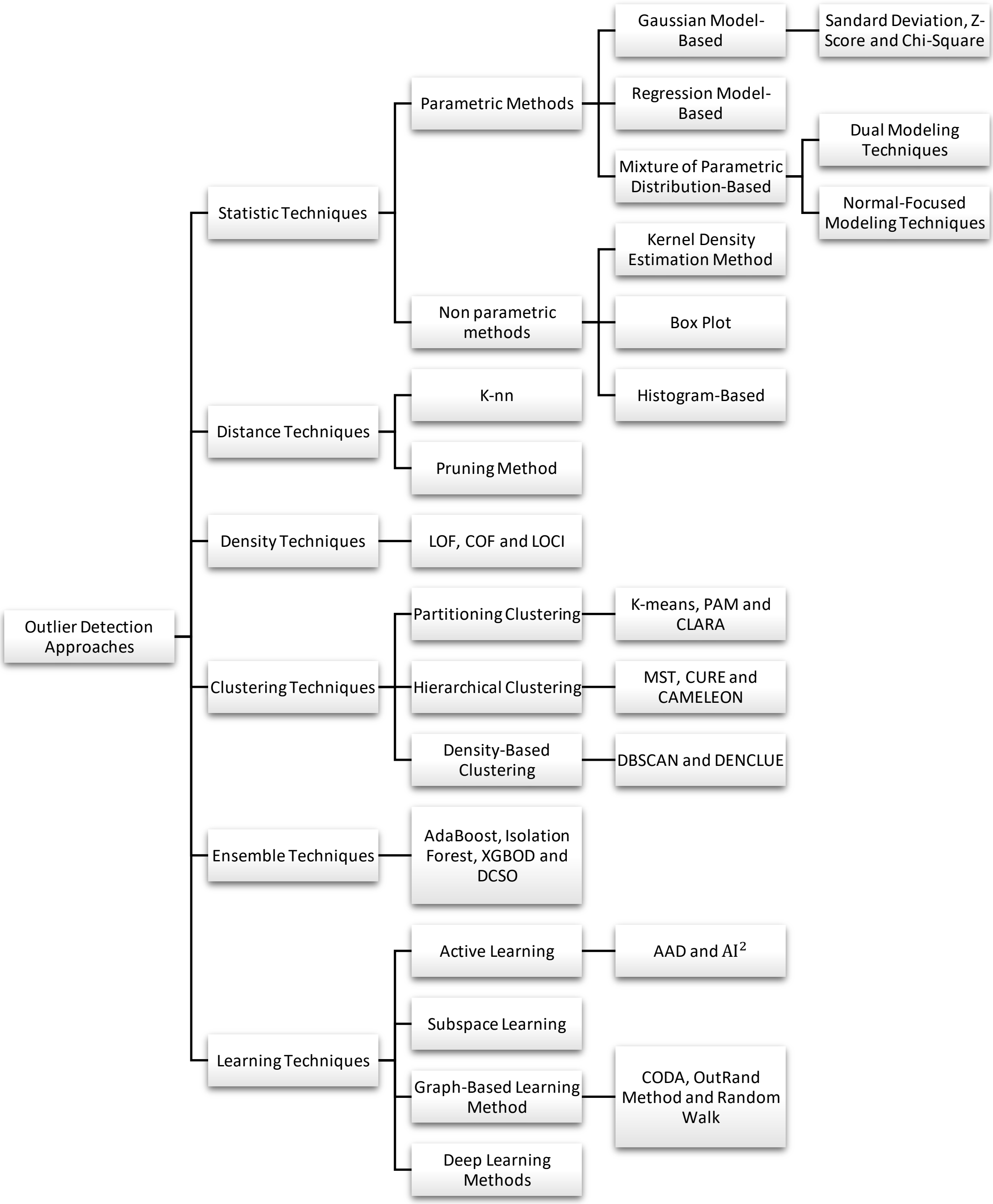


Figure 1.2 Outlier detection approaches

1.7.1 Statistical Anomaly Detection Techniques

When applying statistical techniques for detecting outliers, they may be applied as supervised, semi-supervised, and unsupervised approaches. In a statistical-based method of outlier detection, data points are often modeled by some stochastic distribution, and points for which the given model does not apply well are called outliers. The identification of outliers and inliers depends on the specific distribution model used. These methods assume that the normal data follows some distribution, and if some particular data point does not conform to this distribution model, it will be classified as an outlier. Such approaches can broadly be divided into parametric and non-parametric methods.

1.7.1.1 Parametric methods

These methods assume that the data follows some form of a known distribution like the normal (Gaussian), Poisson, or exponential distributions. They estimate the parameters of the distribution (mean, variance, etc.) and consider as outliers those data points that deviate significantly from the modeled distribution.

- *Gaussian Model-Based.*

These types of anomaly detection methods presuppose that the dataset follows a Gaussian distribution. Parameters of this distribution, such as mean and variance, are generally estimated using maximum likelihood estimation (MLE). A data instance is given an anomaly value based on its distance from the estimated mean: farther distances mean potential anomalies. Then, a particular value, above or below which an object considered to be largely normal or anomalous, is imposed on that anomaly score. The variation of this method differs in how the distance is computed and how to set this threshold for identifying anomalies.

- **Standard Deviation Method :** Outliers are detected in this method by checking to see how far data points deviate from the mean of the dataset. This method assumes that the data is normally distributed, with the bulk of values being close to the mean. The standard deviation measures the average deviation from the mean for the points in the distribution. Usually within this method, a threshold set to 2 or 3 standard deviations from the mean is considered. In this respect, any point that falls outside this threshold is said to be an outlier.
- **Z-Score Method**: is also called the standard score technique. This method evaluates the relative position of a data point by measuring how far its value deviates from the mean in terms of the number of standard deviations. Any data point with an absolute value of its z-score greater than 3 is generally considered an outlier. This technique works best in detecting outliers when data are normally distributed but will be less reliable if data are skewed in any way or have heterogeneous variance. The generic formula to compute z-score for the data-point X is given by (1).

$$Z = \frac{X - \mu}{\sigma} \tag{1}$$

- **The Chi-Square Test** [38] : is a statistical method used for outlier detection by evaluating whether an observed dataset significantly deviates from expected values. It is commonly applied in categorical data analysis but can also be adapted for numerical data outliers. The test works by calculating the Chi-Square statistic:

$$\chi^2 = \sum \frac{(O - E)^2}{E} \tag{2}$$

where *O* represents observed values and *E* represents expected values.

A high Chi-Square value usually translates into great differences between observed and expected distributions; thus, possible outliers may be indicated. In an outlier detection process, data points exhibiting extreme Chi-Square values (beyond the critical value obtained from the Chi-Square distribution table) can be labeled as anomalies. This method works best when the underlying data is known to follow a certain expected distribution so that one can pinpoint deviations signaling possibilities of errors, rare events, or fraudulent activities.

- *Regression Model-Based*

In such techniques, regression methods are utilized to identify deviations in the data. The underlying philosophy is that any regression model well fitted to the training instances should be able to characterize reasonably well the relationship between independent and dependent variables. In Figure 1.3, titled "Regression-Based Outlier Detection," any point with a residual that largely differs from zero is considered a potential anomaly.

The steps in a regression model-based anomaly-detection procedure are:

1. Regression-based anomaly detection first consists of training a regression model on data exhibiting normal behavior. The model to use (linear, polynomial, time-series regression, etc.) is chosen considering the type of data available. The model is then fitted to minimize prediction errors so as to represent well the expected patterns in the data.
2. After the model is trained, it makes predictions on new data. The residual is the difference between the actual value and the predicted one; it indicates how much a given instance deviates from the expected pattern:

   Residual = Actual Value – Predicted Value

   Small absolute value of residual indicates normal behavior, but large values probably indicate the unusual behavior.

3. The residuals are further translated into anomaly scores and added to the detection process. The simplest form of detection is setting a threshold of anomaly score values above which instances would be interpreted as outliers-common are multiples of standard deviation. While the most straightforward method is thresholding, inferential statistics tests such as Z-score, Grubbs test, or even an interquartile range (IQR) could also be used to find outlying residuals. How one chooses to find anomalies depends on its relevance to the application at hand and the acceptable level of false positives.
4. After anomaly detection, statistical validation methods are used to check the reliability of the results, with hypothesis tests like the Chi-square test comparing whether anomalies significantly differ from normal data. Furthermore, criteria for model selection such as the Akaike Information Criterion (AIC) [39] and Bayesian Information Criterion (BIC) [40] work to check the goodness of fit of the regression model, helping it to not over-fit.

Traditional regression methods, for example, Ordinary Least Squares (OLS), are highly susceptible to outliers considering the underlying mechanism of minimizing squared errors, and squared errors exaggerate the influences of extreme values. Robust regressions like Least Trimmed Squares (LTS) [41] and Huber Regression [42] are employed to diminish the influence of outliers, raising the trustworthiness of the model whilst detecting such deviations. This is especially important in medical diagnostics and sensor monitoring where outliers may be very significant.

In the field of time-series data, manifestation of anomalies is abrupt changes or deviation from the anticipated trend. Special regression models—like Autoregressive (AR), Autoregressive Moving Average (ARMA), and Autoregressive Integrated Moving Average (ARIMA) [43] —are used for capturing temporal patterns and detecting anomalies when the observed values deviate from trend-predicted ones. There are more complex methodologies, such as multivariate time-series models (by Tsay et al.) [44], allowing anomalies to be detected over interconnected variables in finance, industrial monitoring, etc.

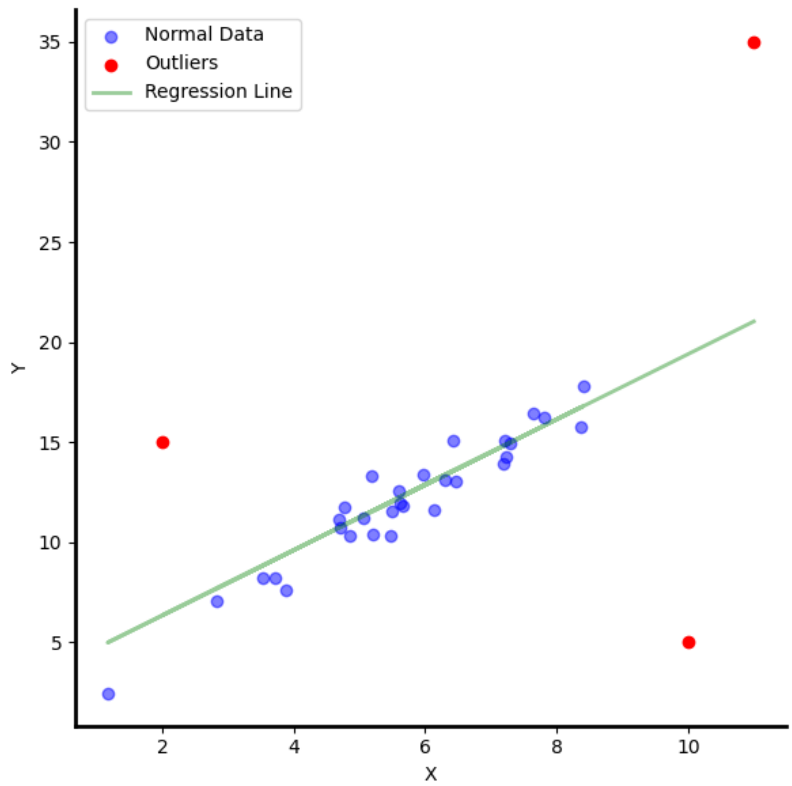


Figure 1.3 Regression-based outlier detection.

- *Mixture of Parametric Distributions-Based*

Mixture of Parametric Distributions Models represent a statistical modeling approach where complex data patterns are captured by combining multiple predefined probability distributions. These methods assume that the observed data is generated from multiple underlying distributions, each characterized by specific parameters. This modeling approach captures heterogeneity within the data and is commonly divided into two main categories:

- **Dual Modeling Techniques**: These methods explicitly represent both normal data and anomalies as distinct parametric distributions. This dual representation enables direct detection of anomalies as deviations from the established normal patterns (Eskin [45], Abraham and Box [46]).
- **Normal-Focused Modeling Techniques**: These approaches concentrate exclusively on modeling normal data using a mixture of parametric distributions. Anomalies are then identified as data points that show poor fit with the learned normal model [47].

1.7.1.2 Nonparametric Techniques

Anomaly detection via nonparametric statistics is a mighty technique that does not assume any particular distribution for data. While parametric methods work according to fixed probability distributions, the nonparametric ones allow the data to form the model. Such methods are most flexible and robust and thus fit perfectly when data-sets showcase complicated, multimodal, skewed, or irregular distributions.

- *Kernel Density Estimation Method*

Kernel Density Estimation (KDE) [48] is a well-known nonparametric technique used to estimate the probability density function (PDF) of a random variable based on a finite sample of data points.

KDE identifies anomalies as points in low-probability regions by estimating the data's underlying density distribution. It uses smooth kernel functions (like Gaussian) to model the probability density without assuming a specific data shape. Points with significantly lower density than their neighbors are flagged as outliers.

- *Box plot*

Considered to be one of the basic graphical tools, the box plot [49], or a whisker plot, is used to check the distribution and variability of a data set. It consists of five important summary statistics: minimum, first quartile (Q1), median (Q2), third quartile (Q3), and maximum. The middle 50% of the data is represented by the interquartile range (IQR = Q3 – Q1). Data points that lie outside 1.5 times the IQR from Q1 or Q3 are usually regarded as outliers (see Figure 1.4). It is a non-parametric method that is extremely good at revealing skewness, spread, and outliers for a sample of data.

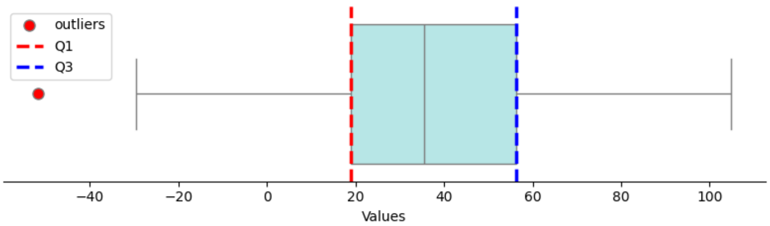


Figure 1.4 Boxplot used to visualize outlying points.

- *Histogram-Based*

By exploiting the frequency distribution of data values (see Figure 1.5), Histogram-Based Outlier Detection (HBOD) is a technique employed to find outliers in a dataset. It is especially useful when working with univariate data, but it may also be applied to multivariate situations. The approach is straightforward, understandable, and computationally efficient; creating it appropriate for big datasets. The basic theory behind HBOD is that outlier appear seldom while normal data points occur often inside a dataset. These are the procedures in HBOD:

1. Histogram Construction: The dataset is split into bins (intervals) such that the frequency of each bin indicates how many data points fall within its range, therefore reconciling interpretability and granularity.
2. Calculation of density: Highlighting sparsely inhabited areas, the density of each bin is determined by the fraction of points it holds in connection to the total dataset.
3. Outlier Identification: Abnormal Bins having densities under a pre-set threshold (e.g., bottom 5% or statistically obtained cutoff) are marked as containing outliers.

For multivariate data [50][51], one popular method is to create individual histograms for every feature. The height of its matching bin in the histogram is used to give each attribute in a test instance an anomaly score during testing. The entire anomaly score for the test instance is calculated by aggregating these separate scores.

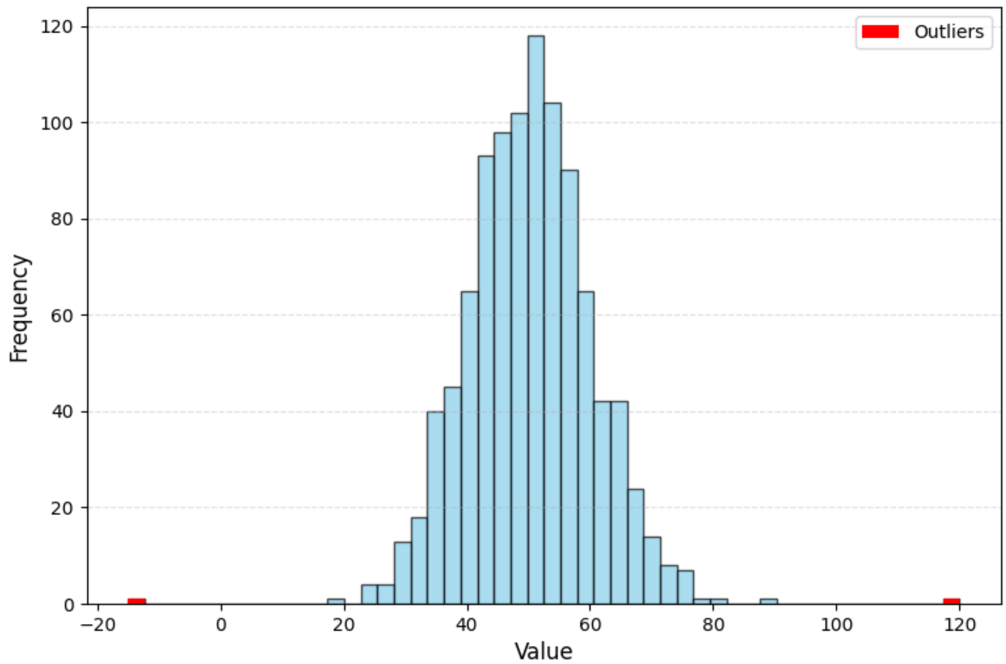


Figure 1.5 Histogram to visualize outliers.

1.7.1.3 Statistical-Based Approaches: Strengths, Limitations, and Challenges

Using standard probability distributions and hypothesis testing, statistical-based outlier detection techniques provide a mathematically correct framework for finding anomalies, therefore ensuring consistency and clarity. Since they rely on measures like standard deviation and interquartile range, their primary advantages are applicability to smaller datasets, computational speed, and straightforward decision-making. without the need of intensive instruction. Still, these methods have great drawbacks, including fixed distributional requirements (such as normality) that lower robustness in practical, non-parametric, or high-dimensional settings. Further limiting their usefulness are difficulties such the curse of dimensionality, sensitivity to concept drift in dynamic data, and dependence on manually defined thresholds. Constant challenges include the demand for adaptive thresholding techniques and scalable solutions to preserve accuracy in changing data contexts.

1.7.2 Distance Anomaly Detection Techniques

Distance-based methods for outlier detection identify outliers by measuring the distances between data points using various metrics, such as Euclidean, Manhattan, or Mahalanobis distances, depending on the dataset's nature and distribution. These methods assume that normal data points are located in dense regions, while outliers appear in sparse regions, far from most other points. Distance-based methods for outlier detection can be roughly divided into two sorts: k-nearest neighbor computation and pruning techniques.

1.7.2.1 K-Nearest Neighbor Method (k-NN method)

The k-nearest neighbor (k-NN) method is a frequently employed distance-based technique for outlier detection, relying on spatial relationships within a point's neighborhood. It assumes that normal data points reside in dense regions, while outliers are relatively isolated (see Figure 1.6). A key parameter is the number of neighbors (k), which significantly influences performance: small k values may lead to sensitivity to noise, whereas large k values may overlook local anomalies.

Outlier scores are typically based on distance, either to the k-th nearest neighbor (Ramaswamy et al. [52]) or the sum of distances to the k nearest neighbors [53][54][55], with applications in domains such as fraud detection [9]. Alternatively, some methods estimate density by counting the number of neighbors within a fixed radius, identifying low-density regions as anomalous [56].

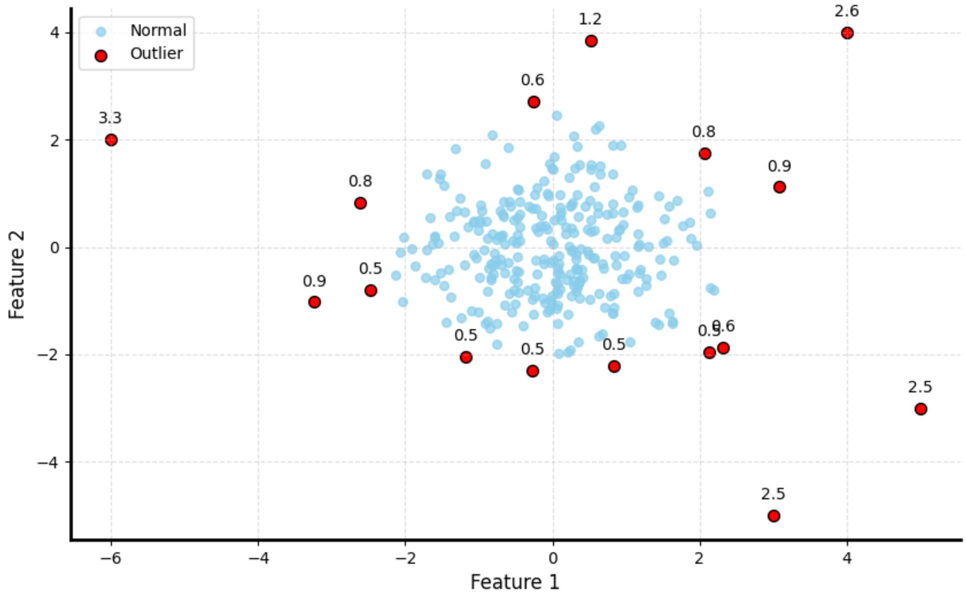


Figure 1.6 Outlier Detection Using the k-NN Method

1.7.2.2 Pruning Method

The pruning approach in distance-based outlier detection is designed to reduce computational cost by eliminating data points or computations that are guaranteed not to influence the final outlier result. It is especially valuable for large-scale datasets, where brute-force distance calculations are too expensive. Several pruning-based methods have been developed to improve the scalability of distance-based outlier detection in large datasets. Ren et al. [57] enhance classical definitions by introducing a neighborhood-based model combined with vertical data representation (P-Trees) and a by-neighbor labeling strategy to reduce redundant computations. The DOLPHIN algorithm [58] addresses disk-resident data challenges by limiting processing to two scans and leveraging indexing with early inlier pruning to minimize I/O and computational costs. Vu and Gopalkrishnan [59] propose MIRO, a two-phase method that applies multiple pruning rules to effectively reduce the search space and execution time.

1.7.2.3 Distance-Based Approaches: Strengths, Limitations, and Challenges

Distance-based methods are good at detecting anomalies by measuring how far a data point lies from its neighbors and are fairly intuitive and easy to explain. These methods work best in higher-dimensional spaces since they do not assume a particular distribution for the data, giving them flexibility for different datasets. They work better in global outlier detection and can treat both numerical and categorical data, provided that a suitable distance measure is used.k-NN-type methods are fast enough to run effectively and have vast applications in fraud detection, network security, and medical diagnostics for small to moderate-sized datasets.

The downside of distance-based techniques is that they become computationally expensive on large datasets. The choice of the distance metric is critical, since a poor choice can give poor results. These methods do not perform well when the dimensionality of the data increases, and local outliers go undetected. Furthermore, they are heavily dependent on parameters that require a lot of domain knowledge to be properly set. Distance-based approaches also assume a uniform data distribution, so this assumption can be terribly limiting when it comes to datasets with varying densities. Lastly, many methods rely on labeled data for maximum deliverability, thereby indicating an occasion to hybridize methods in order to achieve higher accuracy and adaptability.

1.7.3 Density Anomaly Detection Techniques

Using the hypothesis that normal data points reside inside a defined area, density-based anomaly detection methods find outliers by examining the distribution of data points. Exist in high-density areas; abnormalities show up in low-density zones. Using approaches like k-nearest neighbors (k-NN) and kernel density estimation (KDE), these approaches estimate local density by calculating the number of nearby points inside a predetermined distance. If its local density is substantially less than those of neighboring points, a data point is regarded to be an outlier. We will thoroughly cover density-based anomaly detection methods in this area.

1.7.3.1 Local Outlier Factor (LOF)

The Local Outlier Factor (LOF) [4] is one of the most well-known density-based approaches that measures how much a data point deviates from its local neighborhood. Unlike traditional distance-based methods, LOF considers the relative density of a point compared to its neighbors, making it particularly effective in datasets with varying densities.

Introduced by Breunig et al. [4], LOF assigns an outlier score to each data point, with higher values indicating a greater likelihood of being an anomaly (see Figure 1.7). The method is based on the concept of reachability distance, local reachability density (LRD), and LOF score, which are detailed below along with the algorithm's key steps.

1. For a given point p, the k-distance, denoted as $d_k(\mathrm{p})$, is the distance to its k-th nearest neighbor. It is defined as:

$$d_k(\mathrm{p}) = distance(p, p_k) \tag{3}$$

   where p_k is the k-th nearest neighbor of p.

2. To make the LOF more stable, the authors define the reachability distance between two points p and q:

$$reach - dist_k = max\{d_k(q), distance(p,q)\} \tag{4}$$

This ensures that the reachability distance is never smaller than the k-distance of q ($d_k(q)$), preventing unstable calculations when dealing with points that are very close.

3. The local reachability density (Lrd) of a point p, denoted as Lrd (p), is the inverse of the average reachability distance of p to its k-nearest neighbors:

$$Lrd\ (\mathrm{p}) = 1/\left(\frac{\sum_{q \in N_k(p)} reach-distance(p,q)}{|N_k(p)|}\right) \tag{5}$$

where $N_k(p)$ is the set of k-nearest neighbors of p.

This density measure captures how closely packed a point is relative to its neighbors. A low Lrd indicates a sparsely populated region, while a high Lrd suggests a dense region.

4. Finally, The LOF of a point p quantifies how much it deviates from its neighbors by comparing their local densities. It is calculated as the ratio of the average local reachability density of p to the local reachability densities of its surrounding points. A high LOF value suggests that p resides in a sparser region relative to its neighbors, indicating it may be an outlier:

$$LOF(P) = \frac{\sum_{q \in N_k(p)} \frac{Lrd(q)}{Lrd(p)}}{|N_k(p)|} \tag{6}$$

LOF is a powerful method for detecting outliers, especially in datasets with varying densities. However, it requires careful parameter tuning. The choice of k is crucial; too small a k might make LOF sensitive to noise, while too large a k might smooth out local anomalies.

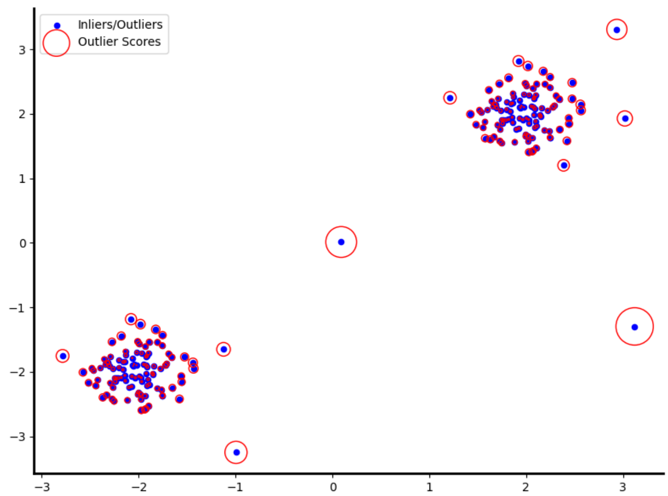


Figure 1.7 Outlier Detection using Local Outlier Factor (LOF)

1.7.3.2 Connectivity-Based Outlier Factor (COF)

The COF algorithm [60] works in enhancing the LOF algorithm with a connectivity-based distance measurement scheme. While LOF considers mainly the density of neighboring points, COF uses the chaining distance that indicates the strength of connectivity of a point to its immediate surroundings.

Given this, the method would be useful in situation where variations in local density exist, outliers being more sensitively detected in distributions that are non-uniform. The COF computes the outlier scores based on the average chaining distance of a point to its k-nearest neighbors, so that points with somewhat irregular connectivity patterns will get higher outlier scores. The COF score of p within its k-neighborhood is given by:

$$COF_k(p) = \frac{|N_k(p)|\ ac{-}dist_{N_k(p)}(p)}{\sum_{o \in N_k(p)} ac{-}dist_{N_k(o)}(o)} \tag{7}$$

where $ac - dist_{N_k(p)}(p)$ is the average chaining distance from p to $N_k(p)$.

The chaining distance is defined recursively, meaning that points forming a continuous structure have lower COF scores, while isolated points or those breaking the connectivity pattern are assigned higher scores. By focusing on connectivity rather than just density, COF can identify outliers in complex spatial distributions where LOF might fail. The algorithm is made of three main steps: one, building the minimum spanning tree over the k-nearest neighbors; two, computing the chaining distance for each point; and three, overriding the COF score by normalizing this chaining distance.

The COF score helps to discern a real outlier from natural density variations within the dataset. A core advantage of COF is that one can identify outliers in regions of irregular density, thus making COF useful for high-dimensional data analysis. Yet COF is computationally more expensive than LOF, and that MST computation potentially puts a limit on its scalability for very large datasets.

1.7.3.3 Local Correlation Integral (LOCI)

The LOCI method [61] addresses a limitation in LOF and COF, as these methods fail to properly handle the issue of multi-granularity and rely on global density estimates (see Figure 1.8). In contrast, LOCI focuses on local neighborhoods, making it particularly effective for datasets with varying densities. The key idea behind LOCI is to compare the density of a point with the densities of its neighboring points using a multi-granularity approach. This method introduces the concept of *m-neighborhoods*, where *m* represents a radius-dependent sampling of neighboring points. LOCI calculates the correlation integral, which measures the number of points within a given radius, and then determines the average and standard deviation of these values. A critical metric in LOCI is the *multi-granularity deviation factor (MDEF)*, which quantifies how much a point's density deviates from the local average and is defined as:

$$MDEF(p_i, m, \alpha) = 1 - \frac{n(p_i, \alpha m)}{\hat{n}(p_i, m, \alpha)} \tag{8}$$

where $n(p_i, \alpha m)$ and $\hat{n}(p_i, m, \alpha)$ are the numbers of αm neighborhoom objects and the average of all the objects p in the m-neighborhood of p_i.

If the MDEF value is significantly high, the point is considered an outlier. One major advantage of LOCI is its ability to detect not only individual outliers but also micro-clusters of anomalies, which many traditional methods fail to identify. Furthermore, LOCI is parameter-free regarding the choice of neighborhood radius since it automatically determines optimal radius values based on data distribution. This makes it more robust compared to methods like LOF, which require careful parameter tuning. However, LOCI can be computationally expensive due to its extensive neighborhood density calculations, limiting its scalability for very large datasets.

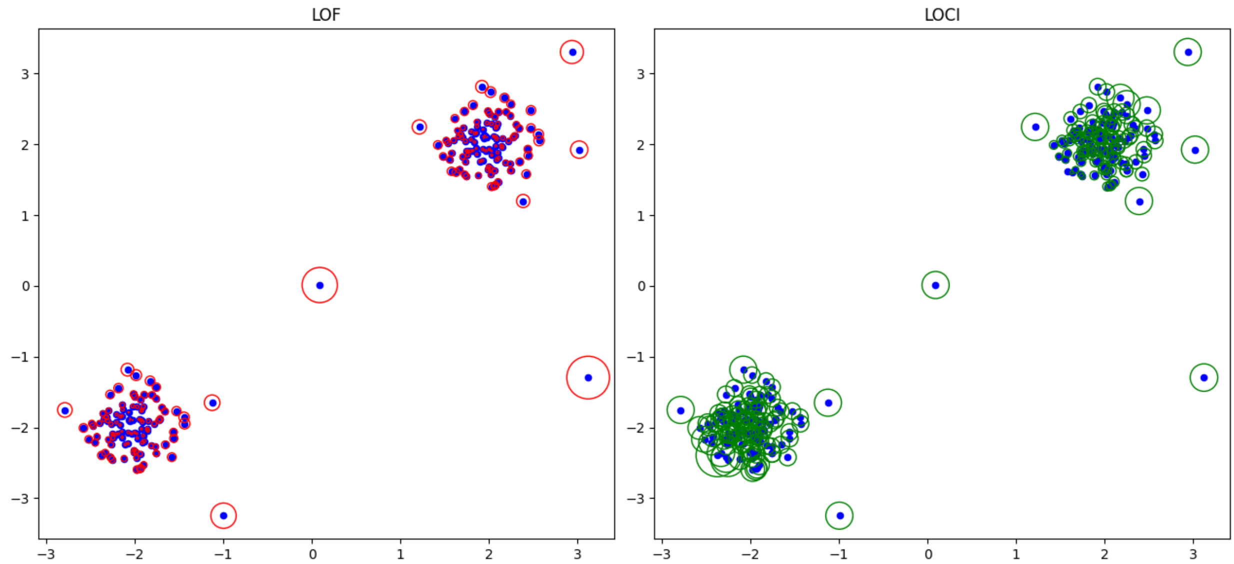


Figure 1.8 Comparison of LOF and LOCI in outlier detection with circle size representing outlier scores

1.7.3.4 Density-Based Approaches: Strengths, Limitations, and Challenges

Density-based approaches for outlier detection, though many advantages accrue to them with regard to dealing with local anomalies. If global approaches are used, then the density of data points is computed with respect to one global density region. Those with higher density may be termed potential outliers when the density of that region is lower and vice versa. This is a plus when it comes to complex real-life data where outliers can be seen only in terms of specific subgroups. Also, density-based methods are forgiving of noise and irregularities in data distribution because they do not consider simple distance measures but rather the relative density of neighboring data points.

Second, their capacity to handle clusters of arbitrary shapes only adds to their strength because most conventional clustering algorithms such as k-means rely on the assumption of spherical groupings. This ability turns out to be quite useful in application domains like image processing and spatial data analysis. Third, they are robust against outlier detection: they can differentiate well between high-density clusters of normal data and low-density regions of anomalous data, hence avoiding false positives in cases of market segmentation and medical diagnostics.

Density-based techniques have certain drawbacks even with these advantages. Their great computational complexity, especially in large or high-dimensional datasets, which results from the laborious computations needed for local density, is one of the main issues. estimation (e.g. in Local Outlier Factor). These techniques are also vulnerable to parameter selection, including neighborhood size and density thresholds, therefore needing manual tuning and maybe affecting performance if erroneously adjusted.

Performance deterioration in high-dimensional spaces, caused by the "curse of dimensionality," moreover limits the dependability of density-based techniques since differences in local densities become less apparent. Their capacity for scaling and flexibility to real-time or dynamic data streams also remain underutilized. Real-time updating of density predictions adds a computational load that limits their application in time-critical cases, therefore stressing the necessity of more scalable and adaptive algorithm approaches.

1.7.4 Clustering Anomaly Detection Techniques

Techniques based on clustering help to find aberrant patterns by using clustering algorithms to examine data behavior. These methods suppose that anomalies manifest as tiny, sparse clusters or solitary dots whereas normal data points form thick clusters. The basic premise is that a cluster with far fewer data points than others is more likely to be outliers than a significant subgroup. Although clustering is mostly known as an unsupervised method, there have been semi-supervised forms advancements [62]. Since their goals vary, it is imperative to distinguish clustering from outlier identification. While outlier detection seeks to find anomalies from the norm, clustering seeks to group comparable data points based on shared characteristics.

The literature provides several clustering methods, which can be divided into the following groups:

- Partitioning Clustering
- Hierarchical Clustering

- Density-based Clustering

1.7.4.1 Partitioning clustering

Partitioning clustering techniques split a dataset into a specified number of clusters such that every data point belongs exactly one such cluster. These techniques try to maximize inter-cluster distance while minimizing intra-cluster distance, hence improving a given criterion. Among popular approaches are PAM, CLARA, and k-means among the partitioning clustering algorithms.

- *K-Means*

One well-known unsupervised learning technique, K-Means clustering [5] can be adjusted for outlier detection by examining how well data points fit into clusters. Minimizing the sum of squared distances between data points and their respective cluster centroids (see Figure 1.9), the algorithm divides data into k clusters. Anomalies in outlier detection are seen as points that don't fit well into any cluster, usually with significant distances from their assigned cluster centroid. The approach starts with randomly choosing k first centroids, then every data point is assigned to the closest centroid according on Euclidean distance. The centroids are next recomputed as the average of all points in a cluster, and the assignment process is repeatedly done until convergence.

The K-Means algorithm is fairly scalable and efficient for handling large datasets since its computational complexity is $O(nkt)$, where n represents the total number of data points, k is the number of clusters, and t denotes the number of clustering iterations.

Using K-Means presents a difficulty in that it is sensitive to the selection of k; choosing too few or too many clusters might affect its performance. Furthermore, K-Means presumes spherical clusters with equal variance, so reducing its usefulness for complicated or non-linearly separable data. Outliers can also skew cluster centroids, therefore lowering the algorithm's precision.

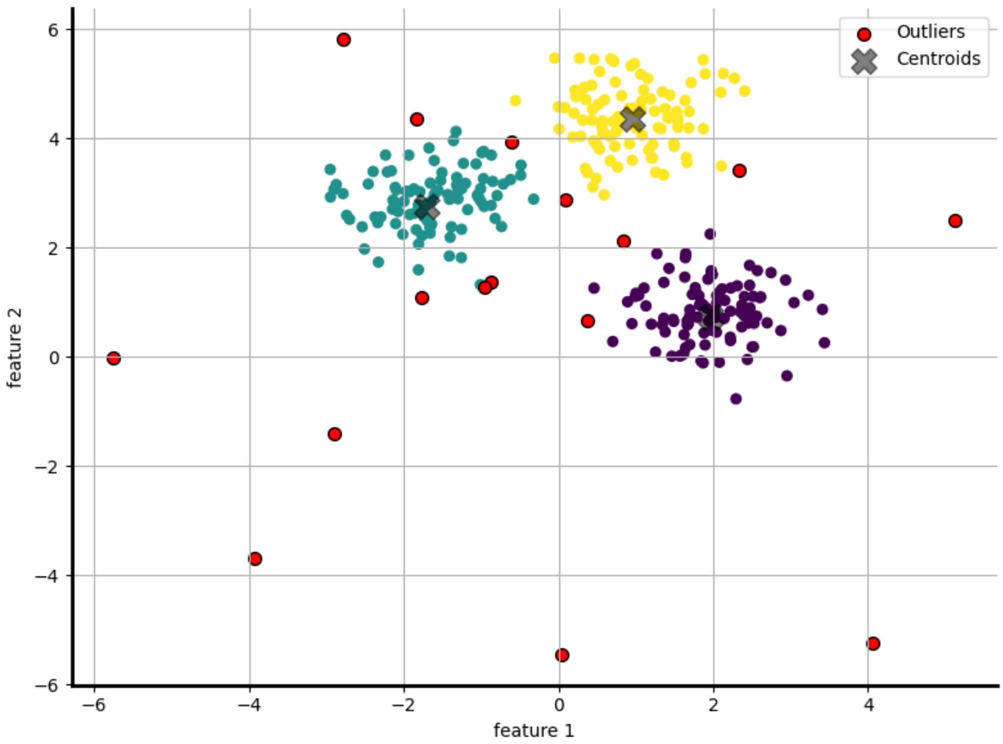


Figure 1.9 Detecting Outliers Using KMeans Clustering

○ *Partitioning Around Medoids (PAM)*

Particularly in the presence of noise and outliers, PAM algorithm or K-Medoids [63] offers a more resilient alternative to k-means. PAM picks actual data points as representative cluster centers—medoids—instead of calculating cluster centroids (centroids) from the mean of data points. Minimizing the total of distances between itself and all other points in the cluster, a medoid is the most centered point inside of one. The PAM algorithm first chooses an initial set of medoids then repeatedly exchanges them with non-medoid points to lower the general clustering. Measured by the entire distance between data points and their appointed medoids, cost is With a complexity of $O(k(n-k)^2)$**,** where n is the number of data points and k is the number of clusters, PAM becomes computationally more expensive than k-means through this iterative optimization. Its opposition to outliers, nevertheless, makes it especially helpful in uses including medical diagnosis, customer segmentation, and anomaly detection where strength is crucial.

○ *Clustering Large Applications (CLARA)*

Extending the k-medoids method, CLARA [63] is an effective clustering method meant to manage big datasets. Unlike k-means, which is sensitive to outliers and relies on centroids, CLARA uses medoids to be more resistant to noise and outliers. Introduced to overcome the scalability problems of PAM, CLARA becomes computationally costly for big datasets.

CLARA aims to cooperate with samples rather than the whole dataset. It picks several little subsets (samples) from the dataset, uses the k-medoids technique to each sample, then assesses the quality of clustering across the whole dataset. Based on a predetermined measure like the average dissimilarity, the best clustering result—chosen as the ultimate output. This increases the computational practicality of CLARA relative to PAM while yet preserving acceptable clustering quality.

Because CLARA does not directly process the entire dataset, computational complexity is greatly decreased. Its efficacy, however, depends on the nature of the chosen subsets. The clustering outcomes might not be ideal if the samples do not properly represent the whole collection of data. Though it has this constraint, CLARA is still a rather often used option for clustering huge datasets when conventional k-medoids techniques are computationally impractical.

1.7.4.2 Hierarchical clustering

Hierarchical clustering produces a tree-like structure (dendrogram) that shows nested clusters of data points. It permits for a hierarchical, flexible perspective on data connections. This approach is split into two basic strands:

- Starting with every data point as a separate cluster, agglomerative hierarchical clustering (Bottom-Up) merges them recursively according to likeness until only one cluster is left. Single-linkage, complete-linkage, and average-linkage are among the most used linkage methods.
- Starting with the entire dataset as one cluster, top-down divisive hierarchical clustering recursively divides it into smaller clusters. Although this method is computationally demanding, it gives a thorough analysis of data structure.

Agglomerative techniques are more often used in actual use. Among well-known hierarchical clustering methods are CURE, MST clustering, and CHAMELEON.

o *Minimum Spanning Tree (MST)*

MST clustering [64] is a a type of algorithm that uses graphs to divide data points. It does this by taking advantage of the layout of a minimum spanning tree. It starts by creating a Minimum Spanning Tree (MST) from a set of data. Each data point acts like a node, and the connections between these nodes have weights based on how far apart they are from each other. The Minimum Spanning Tree (MST) is created using methods such as Prim’s or Kruskal’s algorithms. These methods make sure that all points are linked together with the smallest total weight of connections, while also preventing any loops. Instead of depending on fixed shapes for groups like other techniques, MST clustering discovers natural groups by carefully removing the longest connections in the tree. This process works on the idea that the longest edges usually link different clusters. By cutting these edges, we can split the dataset into important groups. One big benefit of MST clustering is that it can find groups of different shapes and sizes without needing to know how many groups there are in advance. It is also very good at finding outliers because they usually show up as nodes that are not closely connected and have long edges. MST clustering can be affected by noise, so it often needs careful adjustment of the edge-cutting threshold to get the best results. Additionally, it can take a lot of computer power to handle large datasets, especially when creating the Minimum Spanning Tree (MST).

o *Clustering Using Representatives (CURE)*

CURE [65] is another clustering method that organizes data into groups. It is made to work better with large datasets that have clusters of different shapes and also deal with unusual data points. To solve problems with scalability and sensitivity to outliers, CURE picks several representative points from each cluster instead of depending on just one center point or medoid. These representative points are pulled closer to the center of the cluster by a set amount. This keeps the clusters looking good and organized, even when there is some noise around them. The algorithm starts by looking at each data point as its own separate group. Then, it gradually combines these groups by looking at how close they are to one another, using chosen representative points to measure the distance. CURE simplifies calculations by using only a small set of key points instead of every single data point. This makes it a good choice for working with large datasets. The shrinking factor also helps to balance tightness and spacing, which allows CURE to group together clusters that have different densities and shapes that are not round. To improve efficiency, CURE uses random sampling and partitioning methods. This helps to lower the number of distance calculations needed for clustering. This makes it much quicker than other hierarchical methods, while still keeping accuracy intact. CURE works really well when dealing with large sets of complex data. Its strength in managing unusual data points and odd-shaped groups makes it a strong choice compared to k-means, DBSCAN, and other clustering methods.

o *CHAMELEON*

CHAMELEON [66] is a smart clustering algorithm that changes based on the natural patterns in the data to make clustering more accurate. CHAMELEON concentrates on two main points: how similar different groups are to each other and how well the members of the same group stick together. It starts by creating a k-NN graph, which connects each data point to its closest neighbors, making a simple version of the dataset. The graph is divided into small groups called clusters using a method known as a graph partitioning algorithm, like METIS. This process makes sure that each cluster has strong connections within itself. CHAMELEON uses a flexible model to combine clusters. Instead of just looking at how far apart they are or how dense they are, it checks how close and tight the clusters are before bringing them together. This flexible method helps it recognize complicated cluster shapes, making it a good choice for datasets that have different densities and unusual structures. It provides a strong and adaptable system by keeping a balance between local and global clustering features, which helps it fit different types of data. Additionally, CHAMELEON works very well with complex data and large sets of information. However, it can be more complicated to compute because it requires building and dividing the graph.

1.7.4.3 Density-based Clustering

Density-based clustering algorithms find groups of data by looking for areas where many points are close together. This makes them good at finding clusters that have different shapes and dealing with noisy data. These methods are especially helpful for datasets that have different cluster densities. Important clustering algorithms that focus on density are:

o *Density-Based Spatial Clustering of Applications with Noise (DBSCAN)*

DBSCAN [67] is a popular density-based clustering algorithm. It is particularly effective for discovering clusters of arbitrary shapes and handling datasets with noise and outliers. DBSCAN does not require specifying the number of clusters beforehand and can identify clusters based on the density of data points. DBSCAN groups data points based on density. It defines clusters as dense regions of points separated by areas of lower density. The algorithm classifies data points into three categories:

- *Core Points* – Points that have at least *MinPts* (a user-defined threshold) neighbors within a given radius (*Eps*).
- *Border Points* – Points that do not meet the *MinPts* requirement but are within *Eps* of a core point.
- *Noise (Outliers)* – Points that are neither core nor border points and do not belong to any cluster.

The DBSCAN algorithm begins by selecting an unvisited data point from the dataset. It then checks the point's density by determining if it has at least *MinPts* neighbors within a given radius (*Eps*). If so, the point is classified as a core point, and a new cluster is initiated. The algorithm then expands the cluster by adding all directly reachable points. If any of these points are also core points, the expansion process continues until no more points can be added. Points that lie within Eps of a core point but do not meet the *MinPts* requirement are labeled as border points. Any remaining points that do not belong to any cluster are identified as noise. This process repeats until all points in the dataset have been processed (see Figure 1.10).

The main factors of the DBSCAN algorithm are *Eps* (ε) and *MinPts*. *Eps* sets the distance that nearby points are looked at. If Eps is small, it can lead to a lot of tiny groups. If *Eps* is large, it might cause separate groups to combine into one. *MinPts* is the smallest number of points needed within *Eps* for a point to be considered a core point. A higher *MinPts* value makes DBSCAN stricter, needing clusters to be denser. Selecting the best values for *Eps* and *MinPts* is important for good clustering. This choice is usually made by using a k-distance plot or by relying on expert knowledge in the field.

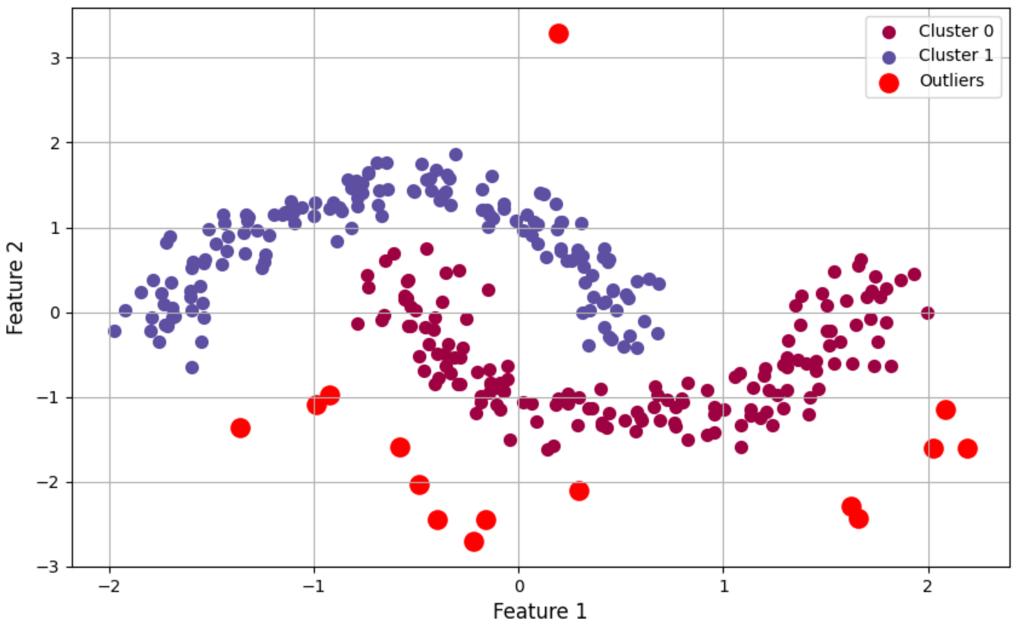


Figure 1.10 Application of DBSCAN for anomaly detection.

○ *DENsity-based CLUstEring (DENCLUE)*

DENCLUE [68] is a clustering method where the cluster building process is based on the density distribution of the data points. It models the distribution of data through a kernel density estimation, where each data point contributes to the density based on a mathematical function (e.g., Gaussian function). The algorithm works based on gradient ascent on the density function, wherein points jump iteratively toward local maxima of densities that serve as cluster centers.

For outlier detection, DENCLUE treats points that occur in regions of very low density and do not lie in any region of high density as outliers. These points do not get swept along with the gradient paths that lead to clusters and remain stranded on some lonely hill of sparse regions in the data space. Thus, the approach helps in separating the outliers from real structures by imposing a density threshold below which points are considered as noise. Instead, DENCLUE is robust to noise and can trace clusters of arbitrary shapes that effectively differentiate anomalies from genuine data patterns in contradistinction with some of the other clustering algorithms. The density-based nature of the algorithm ensures that outliers remain excluded naturally from the main clusters so that it can be reliably used for anomaly detection in complex datasets.

1.7.4.4 Clustering-Based Approaches: Strengths, Limitations, and Challenges

In general, clustering-based approaches to outlier detection offer several advantages. First, they can detect arbitrary-shaped anomalies without making assumptions about their distributions, which makes them very useful for complex real-world datasets. Being unsupervised also means that they can be used in domains where even labeled anomalous instances are scarce, giving them adaptability. Methods that service clustering well, for example, DBSCAN and DENCLUE, are highly noise-resistant and perform well in high-dimensional spaces where distance-based approaches may fail. Moreover, since many clustering algorithms work efficiently by clustering large datasets into groups of similar points, they reduce the computational complexity of any operation performed on those clusters.

Parameter setting is another challenge, as it weighs heavily towards detection performance. Overlapping clusters make true outlier distinction difficult in certain cases, especially in high-dimensional spaces. Some require more computational time than is feasible for big data, for instance, hierarchical. Most clustering methods are geared toward static data and do poorly when faced with evolving data streams, highlighting a gap in their adaptability toward real-time clustering.

1.7.5 Ensemble Anomaly Detection Techniques

Outlier detection can be improved using ensemble techniques, whereby several models are combined for added robustness and accuracy. Using methods like bagging (e.g., Isolation Forest); boosting (e.g., AdaBoost); stacking; or voting, one aggregates different model outputs to enhance performance of the system. They all collectively try to fix the weaknesses of the individual models but run into problems with unsupervised anomaly detection and lack of labeled data. Also, the different notions of outlier according to some models may become hurdles in further aggregation of results. Hybrid ensemble methods have been suggested to alleviate such problems and to yield further improvements in detection. Several types of methods have been proposed, such as:

1.7.5.1 Adaptive Boosting (AdaBoost)

It is one of the most notorious ensemble learning algorithms that undergoes training sample reweighting methods iteratively based on the classification accuracy of a weak classifier [69]. This formulation of weak to strong classification was first introduced by Freund and Schapire in 1996. Generally, the weak learner is thought of as being a decision stump so AdaBoost ensures that instances previously misclassified receive a heavier consideration. Predictions are finally accomplished by utilizing a weighted majority vote, whereby greater consideration is given to models that showed more testing accuracy during training. It reduces bias and variance ; however, it is sensitive to noise and outliers which degrade performance. Its effectiveness is dependent on the choice of base learners, on the number of iterations, and on the quality of the training data.

1.7.5.2 Isolation Forest

Isolation Forest [70](iForest) is a tree-based anomaly detection technique that identifies outliers by isolating them through random recursive partitioning of the data (see Figure 1.11). The principle leveraged here is that anomalies require fewer random splits to be isolated, resulting in shorter average path lengths within isolation trees (iTrees). The algorithm works efficiently in linear time while requiring very little memory, making it appropriate for big data. It is noise tolerant and makes no particular assumption about data distribution. However, the effectiveness of the model depends on the correct setting of contamination parameters, and the performance may vary with the algorithm's randomness.

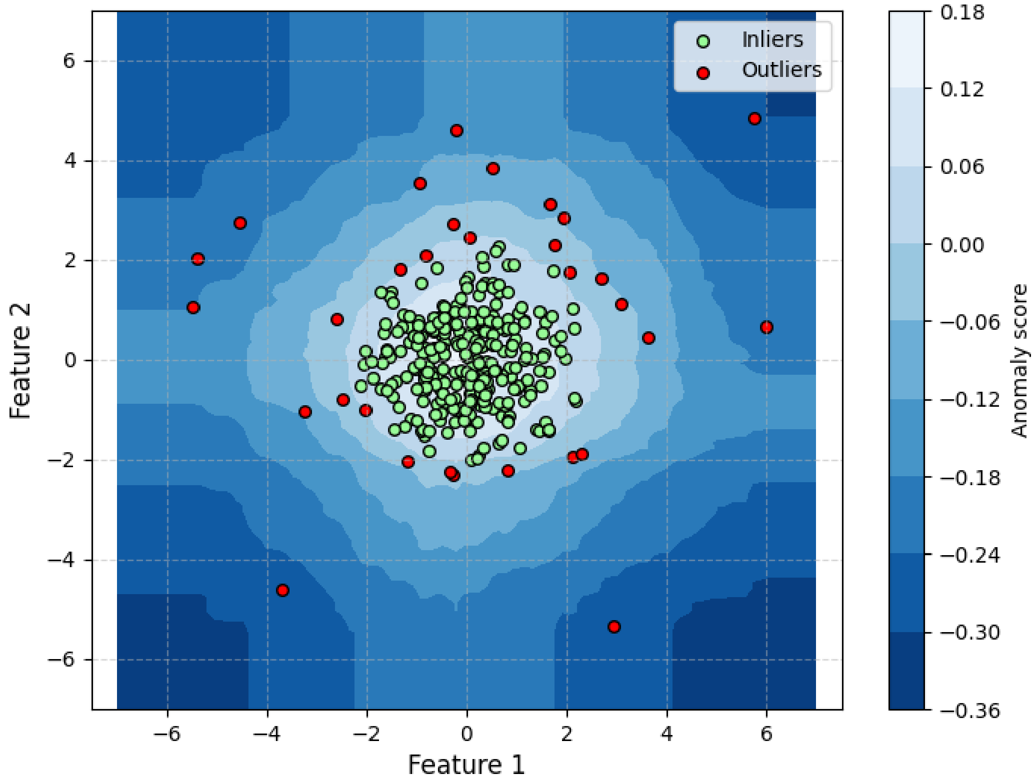


Figure 1.11 Visualization of Isolation Forest-Based Anomaly Detection

1.7.5.3 Extreme Gradient Boosting Outlier Detection (XGBOD)

Being a hybrid anomaly detection framework, XGBOD [71] synergizes unsupervised outlier scores with supervised learning for improved anomaly detection performance. It contains three phases: generating outlier scores with numerous unsupervised methods; selecting, transforming, and selecting the most informative scores; and finally training an XGBoost classifier on the augmented feature space. With weak detectors such as KNN and One-Class SVM, different anomaly characteristics are learned and make XGBOD more robust. By learning the outputs given by multiple models, XGBOD manages to give stronger anomaly detection ability. Thus, the method, on the other hand, can be quite demanding on resources and is easily prone to overfitting if not well regularized especially when trying to learn on small datasets.

1.7.5.4 Dynamic Combination of Detector Scores for Outlier Detection (DCSO)

DCSO [72] is an unsupervised outlier detection framework that maximizes detection accuracy by dynamically selecting the most skillful base detectors with respect to the local data characteristics. A set of detectors for whom a large diversity in parameter choice is sought is developed along two phases. The top detectors are chosen in the local region of every test instance using k-NN. This dynamic, localized strategy enables DCSO to adapt locally to the behavior of data regions without requiring any labeled data, hence making it applicable to all kinds of real-world anomaly detection scenarios. However, DCSO suffers from certain shortcomings, such as giving rise to higher computational complexity and being sensitive to parameter tuning. Aggarwal et al. [73] provide a detailed theoretical and practical overview that serves as a compilation of the main testing procedures.

1.7.5.5 Ensemble-Based Approaches: Strengths, Limitations, and Challenges

An ensemble outlier detection provides a greater level of reliability by employing a variety of models in an attempt to capture different types of anomalies, thereby lessening the chance of missing outliers. They provide better generalization and avoid overfitting by combining several models, hence providing better performances on various datasets. They also allow mixing supervised with unsupervised approaches, thereby making them flexible enough to fit different data situations. Their ensemble nature also helps bagging or boosting methods to be scalable for conducting real-time analyses on bigger data sizes. However, several drawbacks are posed by ensemble methods. The prominent ones among them are computational complexity arising from training and aggregating multiple models.

With heightened redundancy between base classifiers, the effectiveness of an ensemble goes down; hence, one has to choose the classifiers carefully so as to keep the redundancy low. They may even have limitations on imbalanced datasets and usually end up failing at spotting rare anomalies. Another downside of an ensemble model is that usually, a set-based approach would provide hardly any straightforward explanations on how two patterns were deemed similar, or why certain points were so fiercely declared to be anomalies.

1.7.6 Learning Anomaly Detection Techniques

Outlier detection is a fundamental task in machine learning, aimed at identifying data points that significantly deviate from normal patterns. learning-based approaches have gained traction due to their ability to model complex structures and adapt to various data distributions. These methods have been applied across different sub-disciplines, including active learning, graph-based learning, and deep learning. In the following section, we will present research on outlier detection that utilizes these learning approaches.

1.7.6.1 Active Learning

Active Learning is a semi-supervised learning technique where an algorithm selectively queries the labeling of the most informative data points to obtain better models with fewer labeled data points. With respect to outlier detection, the objective of active learning is to pinpoint instances of anomalies with as little human input as possible. This becomes extremely helpful when labeled anomalies are rare, expensive, or time-consuming to acquire. Active learning functions like a system where the algorithm can query the user for instance labels to enhance prediction accuracy. Active learning strategies have been widely explored for outlier detection in various techniques, In the following section, we will discuss common approaches such as AAD and AI^2.

o *The Active Anomaly Discovery (AAD)*

The Active Anomaly Detection (AAD) algorithm [74] is designed to assist analysts in efficiently identifying true outliers within a dataset under a limited query budget. It ranks data instances based on anomaly scores generated by the LODA ensemble-based method [75] and iteratively presents the top-ranked, unlabeled instances for analyst feedback. As labels are provided, AAD dynamically updates its ranking model, adjusting the weights of ensemble members to promote true outliers and demote false positives. A notable strength of AAD is its ability to leverage even limited feedback, such as nominal labels, to refine its outlier ranking. This feedback-driven adaptation makes AAD particularly effective in interactive anomaly detection settings.

o $\boldsymbol{AI^2}$ *(Artificial Intelligence + Analyst Intuition)*

The AI^2method [76] is a human analyst-in-the-loop security framework that incorporates human expertise with machine learning toward improving the capacity to detect cyber threats. The system follows a closed-loop scheme, with three phases: first, one builds initial models by training on historical and behavioral data; then, in real-time, those models are used to detect anomalies; finally, after feedback, the models are updated. The bi-directional feedback from the analyst's evaluation of the flagged anomalies serves to gradually improve the system's accuracy and reduce the false-positive rate. In this way, AI^2 works with a mixture of supervised, unsupervised, and semi-supervised learning, and thus remains adaptive to currently unknown threats. AI^2 has been tested on a dataset encompassing 3.6 billion log entries and shows good capabilities for spotting emerging attacks. As suggested, however, its performance depends so much on the quality of analyst feedback that bad influence can arise from the analyst feedback labeled incorrectly.

1.7.6.2 Subspace Learning

Subspace learning addresses the challenges of high-dimensional data by identifying lower-dimensional feature subsets where anomalies are more detectable and patterns more meaningful. It enhances both computational efficiency and detection accuracy by filtering out irrelevant features and focusing on informative subspaces. Several methods have been proposed based on this strategy. CLOM [77] uses a concept lattice to find sparse subspaces and identify outliers based on sparsity and density thresholds. RODS [78] employs sparse representations via dictionary learning and is effective in both batch and online settings using probabilistic scoring. RBDA [79] replaces traditional density measures with rank-based comparisons of data point proximity to neighbors, improving performance in non-uniform data. OUTRES [80] adaptively selects statistically significant subspaces and calculates outlierness using deviation and density metrics. Methods like SOD [81] and HiCS [82] further refine subspace outlier detection by analyzing weighted deviations from subspace projections and identifying high-contrast subspaces via statistical tests, respectively.

1.7.6.3 Graph-Based Learning Methods

In graph anomaly detection, we look for abnormal nodes, edges, or behavior patterns in graph-structured data that do not conform to normal behavior. The matter is crucial because many data types are inherently relational, and any anomaly may arise from the interaction between connected elements. Graphs can naturally view these relationships, making them tremendously capable of detecting such irregularities. Akoglu et al. [83] reviewed a number of graph-based methods such as CODA and gOutRank, illustrating their promise while further charting the issues in scalability and dynamic graph analysis.

- *Community Outlier Detection Algorithm (CODA)*

CODA [84] is a probabilistic model for detecting outliers in information graphs by combining data attributes and graph structure. It assumes nodes belong to normal communities and models these using distributions like Gaussian or multinomial, treating outliers as deviations from them. Using a Hidden Markov Random Field, CODA captures local dependencies, enabling context-aware detection. It is optimized through EM and ICM algorithms. CODA has shown high precision on real datasets, such as identifying atypical researchers in co-authorship networks, making it effective for detecting anomalies relative to local community behavior.

- *OutRank method*

The OutRank method [6] is a graph-based outlier detection method that identifies anomalies using random walks on a similarity graph. It constructs a weighted, undirected graph where nodes are data points and edges reflect similarities. By applying a Markov chain to the graph, it computes outlier scores using the dominant eigenvector of the transition matrix—lower scores indicate potential outliers. OutRank captures both isolated and clustered anomalies and does not require manual parameter tuning thanks to automatic thresholding. It consistently outperforms traditional methods in accuracy and robustness across various datasets.

o *Outlier Detection Model Using Random Walk on Local Information Graph*

This approach [83] introduces a local information graph-based outlier detection model that improves on traditional graph methods by incorporating asymmetric local dependencies. It constructs a graph capturing both local and global data relationships and performs random walks to assign outlier scores. To handle dangling nodes and improve accuracy, two algorithms with different restart vector strategies are used. These restart vectors bias the walk toward likely outliers. The method iteratively refines probabilities until convergence, effectively identifying both local and global anomalies. Experiments show strong, consistent performance across diverse datasets with lower false positive rates.

1.7.6.4 Deep Learning Methods

Deep learning has gained prominence in outlier detection due to its ability to handle complex, high-dimensional data, surpassing traditional methods (Chalapathy & Chawla,[85]). Its effectiveness stems from learning intricate patterns adaptively, supporting supervised, semi-supervised, and unsupervised paradigms. Supervised methods excel with labeled data but face challenges like class imbalance and annotation costs, while semi-supervised approaches leverage limited labels and autoencoders to model normal data distributions. Unsupervised techniques, such as autoencoders ([86][87]), identify anomalies via reconstruction errors but risk performance degradation if trained on contaminated data. Ensemble methods [3] mitigate this by combining diverse autoencoders, enhancing robustness through median-based scoring and adaptive sampling. Robust Deep Autoencoders (RDA, [88]) further improve resilience using anomaly-regularizing penalties and optimization techniques like ADMM [89].

Generative Adversarial Networks (GANs, [90]) employ adversarial training for anomaly detection, exemplified by AnoGAN [91], which maps medical images to latent spaces for anomaly scoring. Its successor, f-AnoGAN [92], accelerates detection via learned mappings. RAMODO [93] addresses representation limitations in traditional methods by integrating outlier-aware objectives, with REPEN enhancing distance-based detection via triplet sampling. While deep learning automates feature extraction and improves accuracy, it demands large datasets to prevent overfitting. Despite computational costs, these advancements underscore deep learning's superiority in outlier detection, particularly for high-dimensional and unstructured data.

1.7.6.5 Learning-Based Approaches: Strengths, Limitations, and Challenges

The major advantages of learning-based approaches include the automation of feature extraction with deep learning, which means less dependency on manual feature engineering and thus greater scale. These approaches can detect complex, non-linear relationships in the data, which classical statistical approaches cannot; meanwhile, incremental learning adaptation to changes in data distribution allows for robustness in evolving environments. Also, using such methods along with explanatory AI and NLP would be quite complementary to extend applicability. Its problems lie in the need for large labeled datasets-interpreting and explaining black-box models is essential for trust in critical systems, but outliers being rather rare creates serious data scarcity-perils of overfitting, the risk of ethical infringements on using sensitive data, all require appropriate regularization and an appropriate data protection framework, so performance can be measured with accountability.

1.8 Conclusion

This chapter deals with the methods of outlier detection and stresses the need to detect data points that significantly deviate from normal patterns, as these outliers may present interesting patterns, or errors, or events requiring further investigation. There have been various options offered: statistical, distance-based, density-based, graph-based, and ensemble-based methods, each having its own pros and cons. The chapter emphasized the importance of outlier detection in various fields such as finance, cybersecurity, and healthcare, where it contributes to decision making, improves data integrity, and strengthens system security.

The next chapter then introduces the basic theoretical and methodological tools of Multiple Criteria Decision Analysis (MCDA), on which a new approach to outlier detection in the MCDA context is founded.

CHAPTER 2:
MULTICRITERIA DECISION AID

2.1 Introduction

Multicriteria Decision Aid (MCDA), also known as Multicriteria Decision Making (MCDM), is a vital area within operations research and management science designed to address complex decisions involving multiple, often conflicting criteria. MCDA provides structured methods for evaluating alternatives (whether to select the best option, rank them, or classify them into categories). It effectively integrates both quantitative and qualitative information under varying levels of certainty, uncertainty, or risk. The methodology typically involves clearly defining objectives, identifying relevant criteria, evaluating alternatives, and assigning importance weights (often based on expert input) [1]. These inputs are then organized into a performance matrix that shows how each alternative performs against each criterion. Mathematical models and preference aggregation rules are applied to this matrix to derive recommendations. When objectives conflict, MCDA often seeks non-dominated (efficient) solutions (options where no criterion can be improved without worsening another). This helps narrow the decision space to a manageable set of alternatives. In this chapter, we aim to present the essential concepts of decision aid and MCDA, along with other key concepts, examine the problematics in MCDA, and provide an overview of its applications and method families.

2.2 Decision Aid

With the aim of mitigating difficult decisions that entail multiple parameters, criteria, and stakeholders, decision aid approaches have been formulated to assist individuals or groups in making such decisions. Decision-making proves difficult when, for instance, an antenna is being chosen for investments, a team member is being hired, a landfill site is being located, or even a medical diagnosis is being prescribed. The decision cases involve reasoning, possible pros and cons, and dimensions of alternatives. Decision-making is mostly not a matter of snapping one's fingers but rather goes through a long route laden with partial steps-one large step of early consideration, one step of progressive learning, one step of discussions, and finally one step of the intermittent search for more new information. During the introduction, new alternatives could be generated, whereas some initially conceptualized alternatives may lose their relevance or feasibility with time due to changing circumstances.

Often it will happen that decision-making requires consideration of many conflicting criteria-globally considered, maybe economic, technical, social, and environmental with some elements of objective data and some subjective judgment. There may also be incomplete, uncertain, or imprecise data, making it difficult to assess the viability of different alternatives. Limitations from economic, political, or social constraints could also tip feasible solutions in one way or the other. Further complexity is added by the presence of various trade-offs between the parties that have a bearing on the priorities and objectives under consideration. To deal with this complex situation, decision-makers generally use models as tools for analysis, exploration, and discussion.

Such models assist in creating an orderly presentation of information that brings about clarity but are, by their nature, limited, imprecise, and unable to include all real-world nuances [1]. Therefore, it is of utmost importance to treat any model vis-a-vis decision support as being purely a tool and, certainly, not a perfect reality. Acknowledging this, Bernard Roy [1], becomes shy about using a model. He gave an unambiguous interpretation of decision aid by claiming that "it is not a means of arriving at the 'best' solution, but a process to help the decision-maker better understand the problem, better examine the options available, and clarify personal preferences." Decision aid does not intend to replace human judgment but intends to become its partner by laying down difficult choices in a way that is clear, systematic, and easy to communicate [1]. It acts as a lubricant for discussions among the participants, fosters the inclusion of different viewpoints into the consideration, and gives support to balanced choices. Hence, the decision aid becomes the sine qua non of today's problem-solving, where complexity, uncertainty, and diversity of opinions are commonplace.

2.3 Multicriteria Decision Aid

Multicriteria Decision Aid (MCDA) is a concept that was born out of the observed limitations of the classical OR methods when faced with complex, real-world problems. In 1950, OR founders believed that mathematics could solve decision problems naturally and optimally. While there were successes in some areas (e.g., scheduling production, logistics, and inventory management), in many other areas they could not. According to Schärlig [94], classical OR methods do well in the case of problems that can be abstracted away from their real-world contexts, such as traveling in route optimization or production in product mixes, but the decisions that most real-world scenarios are much more subtle and involve various often conflicting criteria. The principal concern was that most decision problems are embedded in dynamic and uncertain environments, and the classical models ran with rigid assumptions that did not reflect reality. These assumptions included the very idea of a fixed and known set of alternatives, an alternative fully representing the decision problem itself, and, most importantly, the very assumption of transitivity in preferences-that is, if A is preferred to B and B to C, then A must be preferred to C. But real life hardly conforms to this idealized scenario. In dynamic situations, new options are created, others become outdated, and preferences vacillate. One of the major criticisms is that human preferences are not always transitive or consistent, especially when decisions have to weigh cost against quality, time against satisfaction. Hence comes the MCDA: the methods do not seek "the absolute best solution" but rather "compromise solutions" that consider multiple dimensions of value. For instance, going for a holiday is not about finding a single optimal destination but rather judging according to cost, cultural interest, climate, and options for entertainment. If price were the sole criterion, possibly hiking in the mountains would win; if partying were up there, Ibiza would undoubtedly be the horse to back. Whereas in practice, most individuals consider several criteria, much of which is done subconsciously, leading to the necessity of making trade-offs. In this situation, MCDA helps to sort, choose, or categorize choices based on how closely they match the decision maker's main goals and values. The method does not pretend there is a universal truth or single best choice ; instead, it embraces complexity and subjectivity. One solution might not dominate others in every aspect but could still be the most acceptable given the criteria at hand. For example, visiting the pyramids might be cheaper and richer in cultural value than a cruise, making it a more favorable option overall, even if it's not the best for relaxation. This explains the main idea of MCDA: understanding that decisions rely on many different factors, and no single factor can fully capture the complexity of human choices on its own.

2.4 Principal Concepts and Terminology

In full-fledged decision-making scenarios, the DM tries to put structure and clarity, a necessary step that helps convert vague or overwhelming contexts into a coherent and tractable problem space. The immediate step in the rationalization process is to identify and define the core components that would influence the decision. Generally, these components would include the object or goal of the decision, the full set of all possible alternatives or options available, and a consistent way of evaluating or comparing those alternatives. More questions may emerge: what criteria or factors can influence the decision?

Indeed, the initial phase in setting up a decision problem proves to be crucial to the efficacy of any MCDA approach. With an ill-defined structure, the application of formal methods or the drawing of sound conclusions becomes difficult. The structuring process allows the DM to translate the complexity of the real world into a systematic model amenable to logical analysis. It also aids in comprehending the interrelationships among various aspects of the problem and in tailoring the decision process toward the DM's objectives and values.

Under the MCDA methodology, the structuring phase is not a mere adjunctive step; rather, it constitutes the vital foundation to the entire decision support process. It dictates the selection of solution techniques, construction of evaluation models, and subsequent analysis of the results. It ensures transparency and consistency in the course of final decision, especially when declarations require simultaneous consideration of various criteria. The present section is devoted to presenting the fundamental vocabulary in common use within the MCDA community. This will acquaint the reader with the very basic ingredients that constitute the formulation of an MCDA problem. This foundational understanding is necessary for engaging with the more technical aspects of MCDA methods and for appreciating the rationale behind various decision support tools.

2.4.1 The Alternatives

As a starting point, it is important to define the core concept of an alternative (an alternative) in the context of MCDA. More broadly, an alternative denotes any object, option, or alternative that is subject to evaluation within the decision-making framework. These alternatives may represent tangible or abstract entities such as products, candidates, strategies, investment projects, or any other elements upon which a choice, ranking, or evaluation is to be performed. In other words, alternatives are the potential options the DM must consider when aiming to resolve a particular decision problem.

More formally, according to Roy, B. [95], *an alternative* is a generic term that identifies the object toward which the decision aiding effort is directed. It encapsulates the element that the decision maker must choose, evaluate, prioritize, or assess using a structured methodology. In the notation that follows, we denote the complete set of alternatives as $A = \{a_1, a_2, \ldots, a_n\}$, where each a_i represents a distinct alternative under consideration.

As reported by Vincke [13], the very nature of the set A can change depending on the context and the dynamics of the decision-making process. The alternative set may be considered stable if its definition is given beforehand and it is expected not to undergo any changes during the course of the whole decision-making process. This kind of stability generally arises when a problem is very well defined-one that is static-where options are known and fixed. In contrast, A is evolutive if it is subject to change during the process. This is typical of many real-world decisions wherein the environment is inevitably dynamic and reactive; new alternatives may arise as new pieces of information come into light while some considered alternatives may lose their relevance or feasibility.

In addition, A can be categorized by its inner internal structure. Should each of the alternatives inside the set be considered mutually exclusive (thus selecting one alternative excludes all others), then the set is said to be globalized. Usually, this would be the case in decision problems where an exclusive choice must be made. Conversely, if the decision outcome can be composed of a combination of several alternatives, then the set is said to be fragmented. In fragmented scenarios, the final solution may consist of multiple alternatives selected together, making the structure of the decision space more complex.

Another major distinction regarding the set A is how the set A is defined. If A is relatively small and all elements can be explicitly listed, it is defined by extension. This is conventionally the case when the decision involves a finite number of specific options. In the other case, when it is very large or infinite, it has to be defined by comprehension: its elements are described by giving them a set of constraints or properties which they have to fulfill. This can be the case in more abstract or generalized decision situations: e.g., free selection on a continuous spectrum of alternatives or the search for solutions in an optimization scenario.

Having an understanding of the characteristics of the alternative set and of its definition hence becomes crucial in the structuring of the decision problem, in particular because it determines the modelling approach and thereby influences the application of MCDA methods throughout the whole decision process.

2.4.2 The Criteria

After identifying the set of possible alternatives, denoted as *A*, the next essential step in a multicriteria decision-making process involves describing these alternatives from various perspectives. This is achieved through the concept of **criteria**. A **criterion** is formally defined [13] as a function f, which maps elements of *A* (i.e., the alternatives) to a totally ordered set *E*. This function reflects the decision maker's preferences with respect to a specific point of view. Mathematically, this is represented as

$$f : A \rightarrow E \tag{9}$$

where *E* is structured in such a way that any two elements within it can always be compared: for any e_i, e_j ∈ *E*, it must be possible to determine whether $e_i < e_j$, $e_i = e_j$, or $e_i < e_j$. This ordering allows for meaningful preference expressions between alternatives based on their evaluations.

All criteria are usually assumed to be minimized for the sake of simplicity and consistency. Thus, for each criterion f_j, the value $f_j(a_i)$ denotes how well alternative a_i performs on the criterion. If the decision problem involves q different criteria, they can be gathered into a set F= $\{f_1, f_2, \ldots, f_q\}$, consisting of all relevant evaluative views for the problem.

2.4.3 Preference modelling

In many real-world situations involving MCDA, relying solely on the concept of dominance is not enough to determine an optimal solution. The dominance relation, which compares alternatives based on whether one is at least as good as another in all criteria and better in at least one, often leads to an excessively large set of efficient alternatives. This makes the selection process impractical. As a result, additional information must be elicited from the decision maker to reduce this set and guide the decision-making process more effectively.

The true value of a decision is usually unmanageable, hence a mathematical representation must be formulated in compliance with the decision maker's preferences. The model must be parameterized by the decision maker's input, so that it really expresses what the decision maker values and really puts into practice such values in its prioritization. A preference model is one of the most important elements of, or is used in, most multicriteria decision analysis frameworks. It is used to describe decision-making in a formal way, rather than relying on general assumptions or mere technical background logic.

Preference modelling involves capturing and translating the decision maker's judgments about various alternatives into a structured and logical form. This structured representation allows analysts to process and analyze preferences in a consistent manner. The goal is to understand not only which options are favored over others but also which options are considered equivalent and which cannot be compared at all due to lack of information or subjective uncertainty.

To structure preferences formally, three fundamental binary relations are commonly used [13] : Preference (P), Indifference (I), and Incomparability (R). These relations help categorize the decision maker’s stance when comparing any two alternatives, say a_i and a_j, within the available set A. Specifically:

- a_i P a_j implies that a_i is preferred over a_j
- a_i I a_j indicates the decision maker sees no significant difference between them
- a_i R $a_j$suggests that the decision maker is unable or unwilling to compare the two.

These three relations (preference, indifference, and incomparability) are not arbitrary. They follow a set of logical rules that ensure consistency and coherence within the model, as follows:

- P is asymmetric, meaning if a_i is preferred to a_j, then it cannot simultaneously be true that a_j is preferred to a_i. This ensures that a clear direction is established in preference comparisons.
- I is both reflexive and symmetric. Reflexivity means that each alternative is considered indifferent to itself, i.e., a_i I a_i always holds true. Symmetry implies that if the decision

maker considers a_i and a_j to be equally good, then the reverse must also be true: a_j I a_i.

- R is distinct in nature. It is irreflexive, which means an alternative cannot be incomparable with itself (i.e., a_i R a_i is never valid). Additionally, this relation is symmetric, indicating that if a_i R a_j, then a_j R a_i.

Formalizing such relations permits multicriteria decision analysis to develop a sound approach to human judgments and complex choices. Hence, preference modelling evolved to fill the gap between mathematical decision theory and the preference evaluations by the decision maker, facilitating more applicable, workable, and customized decision support systems.

2.4.4 Performance table

The fundamental component that organizes and presents the set of data needed for analysis and comparison of alternatives with reference to multiple criteria turns out to be the performance table (Table 2.1), also known as the decision matrix or evaluation table. Each alternative, say A_i occupies a row in the table (e.g., a project, product, or policy option), and each column represents a particular criterion, say C_j (e.g., cost, efficiency, environmental impact). The score contained in a cell of the table at the intersection of row i and column j, denoted by a_{ij}, is the performance score of the respective alternative versus the respective criterion, possibly numeric or qualitative (ratings, marks). Thus, the performance table gives the Editors a well-structured perspective on all the alternatives in terms of different criteria so that it can be further used for preference modelling and eventual decision making. Weights can also be attached to these tables when the decision maker gives different importance to the various criteria and later uses these weights to calculate aggregated scores or to produce rankings.

Criteria / Alternative	Criterion 1	Criterion 2	…	Criterion n
Alternative A1	a_{11}	a_{12}	…	a_{1n}
Alternative A2	a_{21}	a_{22}	…	a_{2n}
…	…	…	…	…
Alternative Am	a_{m1}	a_{m2}	…	a_{mn}

Table 2.1 Performance table

2.5 Problematics in Multicriteria Decision Aid

The decision-aiding procedure occurs in stages, the first of which is that one must identify the decision-maker's actual orientation of intent sometimes known as prescription. At this stage, it is important to figure out what the decision-maker wants to achieve. Hence, the stage is to establish the type of decision problem, or "problematic," in shorthand. There are four classes of problematics considered basic in structuring the process [1]: choice, sorting, ranking, and description. The choice problematic is about choosing the best alternative(s) from a given set. The sorting problematic assigns the alternatives into predefined categories or classes. The ranking problematic strives for placing the alternatives from the best to the worst according to the decision-maker's preferences. The description problematic aims at understanding and characterizing the alternatives without putting forth a definitive judgment. In other words, this first step frames the whole process of decision-making. The method, tools, and data necessary for the analysis are then shaped around this framework. Proper identification of the problematic lets the decision support speak to the actual goals of the decision-maker.

2.5.1 Choice problematic P.α

This classic formulation of decision problems, often referred to as the most traditional approach, is centered around identifying the "best choice" among a set of possible alternatives (see Figure 2.1). This problematic, labeled Pα as defined in [1], refers to the process of guiding a decision-maker toward selecting a subset A' of the global set of alternatives A, ideally as small as possible, which can inform the next critical stage of the decision process. The aim is to assist the decision-maker by focusing attention on the most promising alternatives, understanding that the overall set A might be dynamic or subject to evolution. The outcome of this analysis could be a firm recommendation on which alternative to take or could involve the development of an automated or semi-automated method capable of repeatedly identifying optimal alternatives as contexts change.

In the choice problematic, each alternative in A is evaluated in relation to the others, with the goal of eliminating less favorable alternatives. The ideal scenario would involve narrowing down the possibilities to a single alternative that is at least as desirable as all others

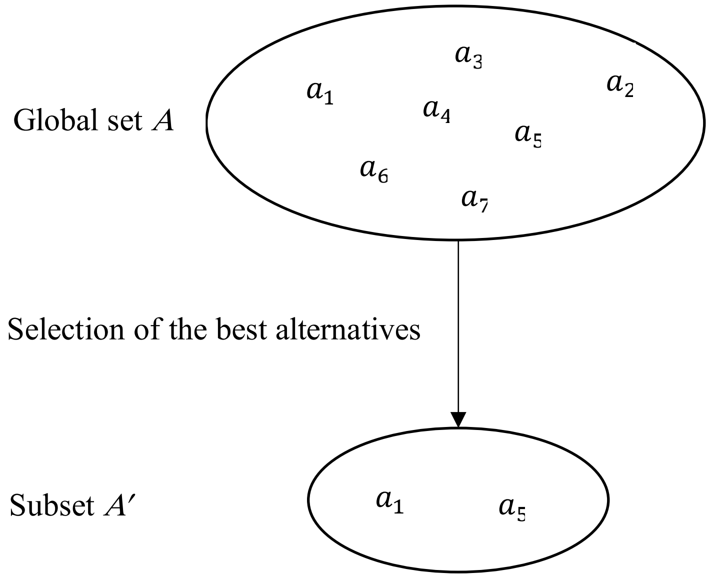


Figure 2.1 The choice problematic

2.5.2 Sorting problematic P.β

Sorting problems concern assigning a set of alternatives *A* to certain distinctive categories considered as alternatives for decision-making or subsequent handling thereof (see Figure 2.2). These categories are not arbitrarily considered. Instead, they are defined with respect to a normative or intrinsic value being set upon each alternative, acknowledging that the set *A* could very well change over time. The principal intention is to further assist the decision-making process so that the alternatives may then be accepted or rejected or farther examined depending on their classification.

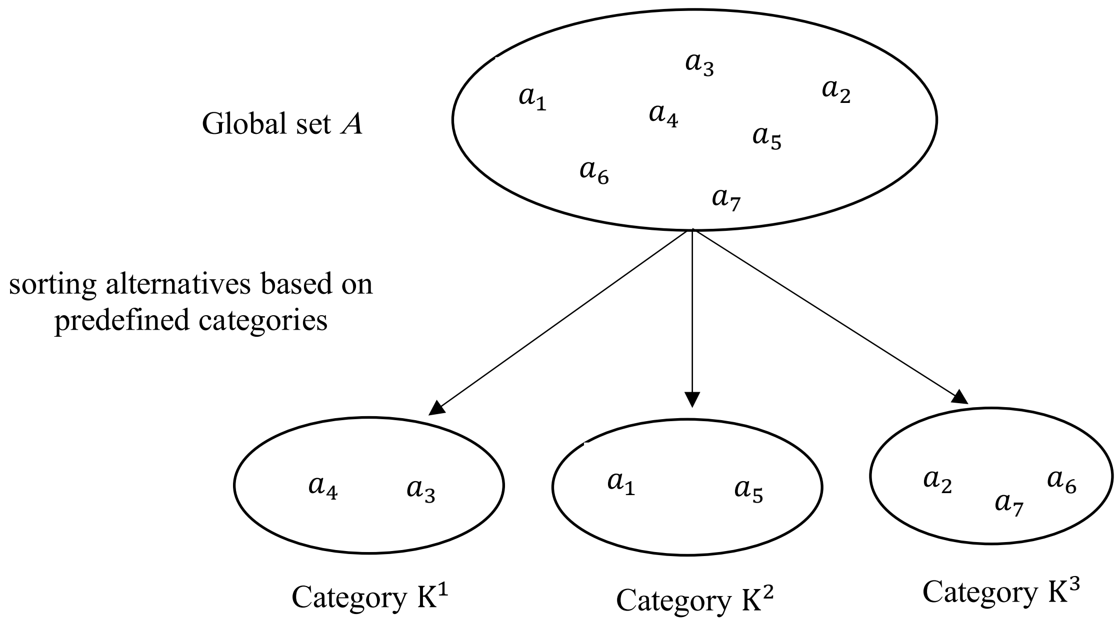


Figure 2.2 the sorting problematic

2.5.3 Ranking problematic P.γ

The ranking problematic, denoted as P.$\boldsymbol{\gamma}$, addresses the task of arranging the alternatives in a set *A* (or a subset thereof) in a manner that reflects a decreasing order of preference (see Figure 2.3). Rather than categorizing alternatives based on intrinsic values, as in sorting, this approach emphasizes comparison among alternatives to reveal those that are "sufficiently satisfactory" according to a preference model. This comparative evaluation leads to the construction of an ordered structure, either partial or complete, where alternatives are grouped into equivalence classes reflecting similar levels of preference. These classes are then ranked, typically from most to least preferred. The objective is not to assign rigid categories, but to establish a meaningful hierarchy that highlights the most promising alternatives. Notably, the lowest-ranked alternatives (especially if they do not form true equivalence classes in terms of preference) may be excluded from detailed consideration, as refining distinctions among less satisfactory options often yields little practical value. This type of problematic supports the use of automated or iterative ranking methodologies, facilitating repeated application and adaptation as new alternatives emerge or existing ones change. The dynamic nature of *A* is thus accounted for, ensuring the ranking remains relevant over time. In contrast to sorting, which orients decisions around predefined acceptance or rejection categories, ranking supports competitive assessment, offering a more nuanced prioritization.

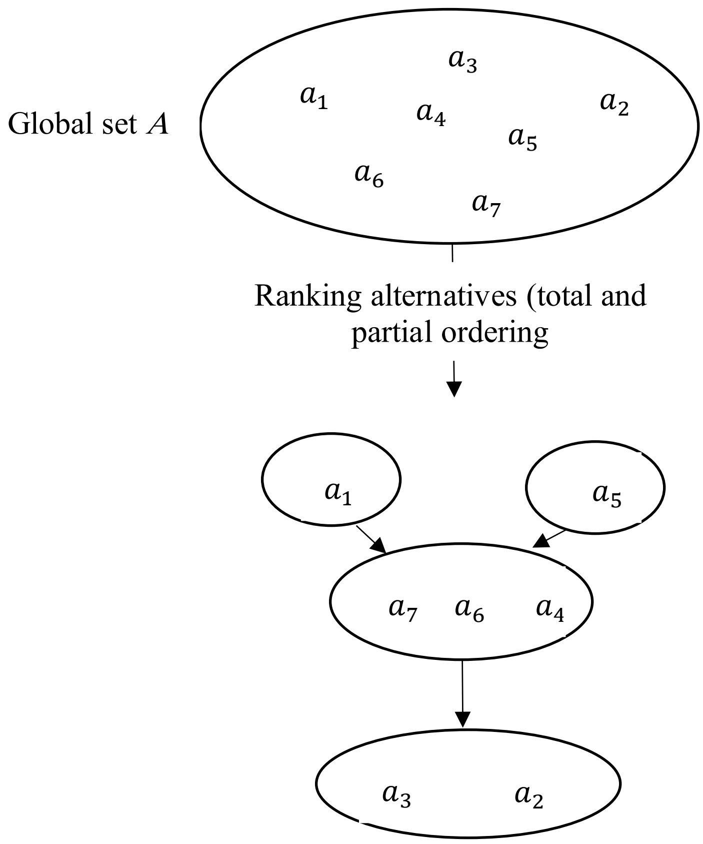


Figure 2.3 The ranking problematic

2.5.4 Description problematic P.δ

The descriptive problematic involves helping clients clarify alternatives and their possible impacts in a systematic, often formalized way. In many situations, clients may not need a final decision but rather a structured way to frame their thinking. The analyst thus provides a tool for reflection, not a direct answer. This approach applies broadly—whether for individual decisions like applying for a loan, or institutional planning such as infrastructure projects. The task is to make relevant information about choices and outcomes explicit. This support can come as either a formal description or a repeatable, often automated cognitive tool. While related to other decision-making frameworks, this problematic stands apart due to its focus on aiding understanding over choosing. It is a valuable contribution in its own right. Tools from systems theory offer useful concepts and models to guide such efforts.

2.6 Application of MCDA methods

Notwithstanding the growth of MCDA methods across various domains with their ability to address complex decision-making situations involving conflicting criteria, some of the key contexts where MCDA methods have found wide applications are highlighted, with references to foundational works and recent studies in each field.

In environmental management, MCDA helps evaluate environmental policies, energy alternatives, waste management, and conservation efforts. Hajkowicz, S. A. et al [96]. addressed water management challenges by reviewing 113 cases from 34 countries. Their analysis shows that MCDA methods are vital tools for evaluating water policies, strategic planning, and project selection. Huang, I. B. et al. [97] conducted a review of over 300 academic papers published between 2000 and 2009 to map the application of MCDA in environmental domains. These papers span various environmental challenges and decision types. The results demonstrate a growing interest in using MCDA for addressing environmental issues across all sectors.

The last years have shown a way toward MCDA emerging as a promising field of healthcare. The International Society for Pharmacoeconomics and Outcomes Research gave two major task force reports [98][99], on the topic, with numerous academic reviews also exploring this methodology. MCDA introduces a structured approach to the evaluation of options involving conflicting criteria, which makes it exceedingly relevant to healthcare settings. At the hospital level, it has become crucial in aiding decisions that necessitate a balanced appraisal of diverse considerations.

The energy domain presents several fields where MCDA approaches are widely applied: renewable energy, energy resource allocation, building energy management, and electric utility operations. In [100], MCDA is used for sustainable energy planning on the island of Crete, Greece. Several alternatives are considered for setting up renewable energy on the island according to five criteria. The analysis is considered being an exploratory support for regional energy planners in ranking and comparing sustainable energy options.

Green supply chains represent another domain where MCDA methods are highly applicable. As green supply chains introduce additional environmental and social considerations, the decision-making process becomes increasingly complex—particularly with the rising importance of intangible factors such as corporate image, business continuity, and social responsibility. The authors in [101] present a comprehensive review of MCDA approaches applied to supplier evaluation and selection, covering literature from 1997 to 2011. A wide range of individual and integrated methods have been proposed within this context.

Additionally, MCDA methods are commonly employed in business and finance due to their capacity to address a variety of solution types. Numerous studies confirm that financial decision-making is inherently multifactorial, prompting experts to utilize MCDA techniques for resolving complex financial issues. These methods have been applied to a range of financial domains, including credit scoring and failure prediction [102] and [103], portfolio selection and management [104] and [105], corporate performance assessment [106], investment appraisal [107], and fund selection for asset investment [108].

2.7 Families of Multicriteria Methods

MCDA methods are generally categorized into three principal families: interactive methods, multiple attribute utility theory (MAUT)-Based methods, and outranking methods [13]. Each family represents a different approach to handling decision-making problems (see Figure 2.4).

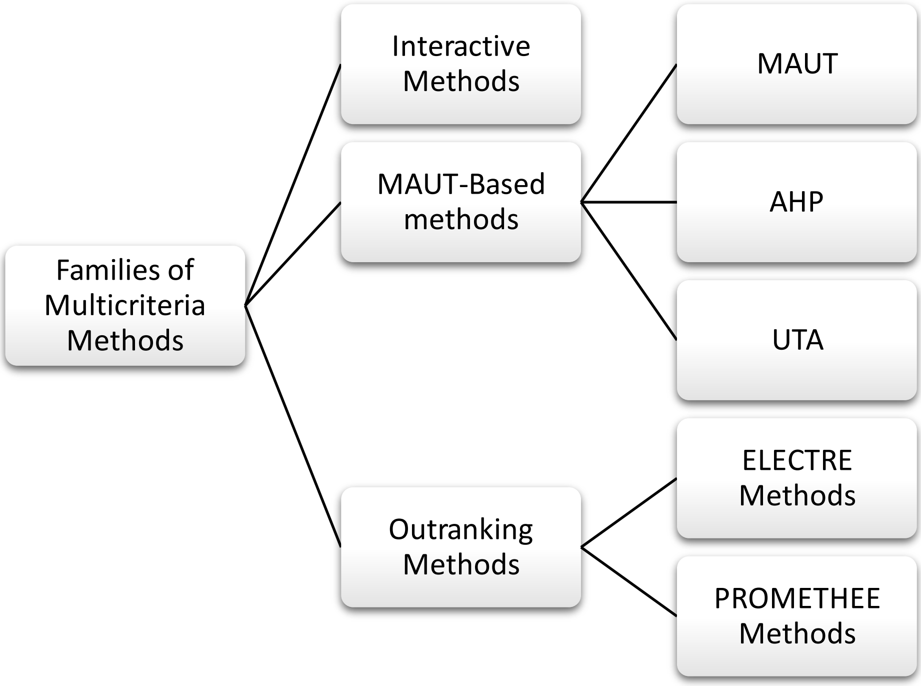


Figure 2.4 Families of Multicriteria Methods

2.7.1 Interactive Methods

Interactive methods are a group of decision-making techniques that focus on keeping the decision-maker involved all the time while solving a problem. These methods recognize that preferences may not be fully known at the outset; instead, they evolve during exploration. Initially, the system proposes a tentative solution based on partial information or default parameters. The DM then evaluates this solution and provides feedback, either in the form of additional preferences, constraints, or adjustments to the importance of criteria. This new input is integrated into the decision model, and the solution is recomputed. This iterative loop continues, enhancing mutual understanding between the model and the DM until a satisfactory solution is achieved. Such methods are particularly valuable in complex or poorly structured decision problems where quantifying preferences upfront is difficult. Key contributions include those by Geoffrion et al. [109] and Zionts and Wallenius [110], with an introductory overview of interactive methods is available in [14]. Interactive methods are especially suited to strategic or long-term decision-making scenarios where stakeholder involvement and consensus are crucial.

2.7.2 Multiple Attribute Utility Theory (MAUT)-Based Methods

MAUT (Multiple Attribute Utility Theory) is a well-structured technique for solving multicriteria optimization problems. It assumes that all criteria of relevance may be quantified and initially brought together in a single utility function. This function expresses the preferences of the Decision Maker and is therefore used to find the best alternative. One finds utility values for the levels of each criterion, weighs their importance relative to one another, and then sums up the weighted utility scores of each alternative. Under this system, the alternative with the highest utility score wins. In the subsequent section, the three major MAUT-based methods will be discussed, namely: Analytic Hierarchy Process (AHP), Multi-Attribute Utility Theory (MAUT), and UTA (Utilités Additives).

2.7.2.1 MAUT Method

MAUT [111] is an approach that converts qualitative preferences into a single numerical utility value, typically scaled between 0 and 1. This transformation enables standardized comparisons among alternatives.

The overall utility value of an alternative *x*, denoted as *v*(*x*), is represented as a weighted sum of individual evaluations across multiple attributes:

$$v(x) = \sum_{i=1}^{n} w_i . v_i(x) \tag{10}$$

- $v_i(x)$: Evaluation score of alternative *x* for the i^{th} attribute.
- w_i: Weight assigned to the i^{th} attribute, reflecting its relative importance.
- n: Total number of attributes considered.

All weights w_i are non-negative and sum up to 1:

$$\sum_{i=1}^{n} w_i = 1 \tag{11}$$

To evaluate each dimension more precisely, the utility value $v_i(x)$ can itself be defined as a weighted sum over several relevant attributes associated with that dimension:

$$v_i(x) = \sum_{a \in A_i}^{n} w_{ai} . v_i(I(a)) \tag{12}$$

Where:

- A_i: Set of attributes associated with the i^{th} dimension.
- w_{ai}: Weight of attribute *a* within dimension *i*.
- $v_i(I(a)$: Value function that quantifies the preference level of *x* with respect to attribute a.

2.7.2.2 AHP Method

AHP [112] being a potent MCDA method is based on an eigenvalue analysis of pairwise comparisons, thus allowing decision-makers to hierarchically structure problems and assess elements with respect to their relative importance. The unique characteristics of AHP lie in calibrating a numerical scale for performance evaluation of both objective/quantitative and subjective/qualitative measurements. The scale runs between 1 (equal importance) and 9 (extreme importance).

Some of the major procedural steps in the methodology are:

1. ***Define the Problem Clearly***

   Explicitly articulate the decision problem to be addressed, including its scope, objectives, etc.

2. ***Expand the Decision Context***

   Take into consideration all relevant stakeholders, objectives, and possible outcomes to gain a holistic understanding of the problem.

3. ***Determine Evaluation Criteria***

   Identify the set of criteria affecting the decision, either tangible or intangible.

4. ***Structure the Problem Hierarchically***

   The decision problem is organized into a hierarchical structure containing multiple levels, with the overall goal at the top, followed by criteria and sub-criteria, and finally the set of alternatives at the bottom.

5. ***Conduct Pairwise Comparisons***

   Assess the elements at each level in the light of their influence on the element above using the standardized 1 to 9 scale. For n elements, $n(n-1)/2$ comparisons must be made. Every element must be compared against each other, with diagonal elements set at 1 (representing equal importance) and reciprocal values for inverse comparisons.

6. ***Calculate Priorities and Consistency Measures***

   Apply matrix operations to calculate the overall λ_max and then obtain the priority vector (weights). Calculate the Consistency Index CI and Consistency Ratio CR measures to determine whether the judgments are logically consistent. If CR is less than 0.10, consistency is deemed acceptable.

7. ***Finalize the Decision***

   Once the measures of consistency are acceptable, proceed to determine the solution on the basis of the computed priority weights. If the judgments are not acceptable, go back and modify the judgments until they satisfy consistency criteria.

2.7.2.3 UTA Method

The UTA is in-between [113] utility theory and tries to build an additive utility function best fitting the preference structure of the decision-maker, expressed through a series of reference alternatives endowed with a given ranking. Its generalized idea is to infer for each criterion a piecewise linear marginal utility function such that the global utility of each alternative calculated as the weighted sum of these marginal utilities reflects the decision-maker’s judgments.

$$U(a_i) = \sum_{j=1}^{n} w_j \, u_j(g_j(a_i)) \quad (13)$$

Where:

- $U(a_i)$: overall utility of alternative a_i
- w_j : weight of criterion j (usually normalized: $\sum w_j$=1)
- $u_j(.)$: marginal utility function for criterion j
- $g_j(a_i)$: performance of alternative a_i on criterion j

The process starts with the decision-maker providing a ranking or pairwise comparisons of several reference alternatives. These are then used to determine the marginal value functions through a mathematical optimization process, typically involving linear programming. The objective is to find functions that are monotonic and piecewise linear, satisfying the additive model assumption while minimizing inconsistencies with the expressed preferences. Once the utility functions are established, they can be applied to assess and rank new alternatives not included in the original set.

However, it can become computationally very expensive if the number of criteria or alternatives increases. On the other hand, gaining an accurate representation of utility functions may be difficult, especially when DMs have unstable preferences or have trouble expressing them. Nonetheless, MAUT is still one of the cornerstones of MCDA when rational and justifiable outcomes are necessary.

2.7.3 Outranking Methods

Outranking methods offer an alternative philosophy to MCDA by avoiding full aggregation of criteria into a single score and instead constructing a pairwise preference relation between alternatives. The key idea is to determine whether one alternative "outranks" another based on evidence from the criteria, without necessarily requiring complete or transitive ordering. An outranking relation indicates that, for most criteria and given thresholds of indifference and preference, one alternative is at least as good as another. These methods tolerate imprecision, conflicting criteria, and incomplete information, making them highly adaptable to real-world decision problems. Outranking models involve the definition of concordance and discordance indices to assess the degree of support and opposition to an outranking relation. The result is often a partial or complete ranking, or a classification of alternatives. These methods are particularly useful when criteria are qualitative or ordinal, and when trade-offs are complex or controversial. Prominent examples include ELECTRE (ELimination Et Choix Traduisant la REalité) and PROMETHEE (Preference Ranking Organization Method for Enrichment Evaluations).

2.7.3.1 ELECTRE methods

ELECTRE is a family of outranking methods developed in the mid-1960s by Bernard Roy and his colleagues as part of the European school of MCDA. It was designed to support complex decision-making where the decision-maker may not be willing or able to express all preferences in a precise, quantitative manner. ELECTRE [114] methods emphasize the concept of outranking—whether one alternative is at least as good as another according to a majority of criteria, without being significantly worse on any of them.

The fundamental idea behind ELECTRE is to compare alternatives pairwise based on their performances across multiple criteria and determine if one can be said to "outrank" the other. The term "outranking" implies a relation of credibility, not absolute dominance. The principal steps underlying the ELECTRE method are as follows:

A. Normalization of the Weighted Matrix

To begin the analysis, it is essential to normalize the decision matrix based on the nature of the evaluation criteria—either benefit-type (where higher values are preferable) or cost-type (where lower values are preferable). In cases involving mixed criteria, cost-type values should be transformed to benefit-type by appropriate scaling or inversion techniques.

Let x_{ij}denote the performance value of the i^{th}alternative with respect to the j^{th} criterion. The normalized value R_{ij} is computed using the following formula:

$$R_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}}, \; i = 1, 2, \ldots, n; \; j = 1, 2, \ldots, m \tag{14}$$

This operation produces a dimensionless matrix where all criteria are scaled appropriately for comparison. Once the normalization is completed, the next step involves weighting each normalized value by its corresponding criterion weight w_j. This results in a weighted normalized matrix V_{ij}, calculated as:

$$V_{ij} = R_{ij} \cdot w_j \tag{15}$$

In matrix form:

$$V = R \cdot W = \begin{bmatrix} r_{11}w_1 & r_{12}w_2 & \cdots & r_{1m}w_m \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1}w_1 & r_{n2}w_2 & \cdots & r_{nm}w_m \end{bmatrix} \tag{16}$$

All entries in matrix *V* are constrained to the range [0, 1].

B. Formation of Concordance and Discordance Sets

For each pair of alternatives $a_p$and a_q, determine the concordance set *C (p, q),* which includes all criteria where a_p is preferred over or equal to a_q. This can be expressed mathematically as:

$$C(p,q) = \{j \mid v_{pj} \geq v_{qj}\} \tag{17}$$

Conversely, the discordance set $D(p, q)$ is the complement of the concordance set and consists of all criteria where a_p performs worse than a_q:

$$D(p,q) = \{j | \, v_{pj} < v_{qj}\} \tag{18}$$

These sets are foundational for further calculations in the ELECTRE method.

C. Computation of Concordance and Discordance Indices

The concordance index C_{pq} quantifies the collective strength of preference of alternative a_p over a_q, based on the total weight of criteria in the concordance set:

$$C_{pq} = \sum_{j \epsilon C(p,q)} w_j \tag{19}$$

On the other hand, the discordance index D_{pq} captures the degree of opposition or disagreement with the assertion that a_p is preferred over a_q. It is defined as the maximum relative deviation among the discordant criteria:

$$D_{pq} = \frac{max_{j \epsilon D(p,q)} |v_{qj} - v_{pj}|}{max_j |v_{qj} - v_{pj}|} \tag{20}$$

This ratio reflects the extent to which a_p underperforms compared to a_q.

D. Establishment of Outranking Relationships

An outranking relation is said to exist from $a_p$to a_q if both of the following conditions are satisfied:

$$C_{pq} \geq \bar{C} \; and \;\; D_{pq} < \bar{D} \tag{21}$$

Where:

- $\bar{C}$ is the average concordance index across all alternative pairs,
- $\bar{D}$ is the average discordance index across all pairs.

Based on these conditions, the nature of the relationship between any two alternatives may be classified as:

- *Outranking (p outranks q):* both conditions hold true.
- *Indifference*: neither condition holds.
- *Incomparability*: only one of the two conditions is met.

This step forms the core of the ELECTRE I methodology in identifying dominance relations among alternatives.

E. Derivation of Overall Rankings

To determine the global preference order, compute the net concordance index C_p and the net discordance index D_p for each alternative a_p :

$$C_p = \sum_{k=1, k \neq p}^{n} \left(C_{pk} - C_{kp}\right) \tag{22}$$

$$D_p = \sum_{k=1, k \neq p}^{n} \left(D_{pk} - D_{kp}\right) \tag{23}$$

A higher value of C_p suggests that the alternative a_p exhibits greater cumulative dominance over others, whereas a lower D_p indicates lesser opposition. The final ranking of performance indicators is derived by prioritizing alternatives with higher net concordance and lower net discordance scores.

The ELECTRE family includes several variants, such as ELECTRE I , II, III, IV, IS, and TRI [115][13], each suited for different decision problems like choice, ranking, or sorting. For example, ELECTRE I is typically used for choice problems, while ELECTRE III is used for ranking, and ELECTRE TRI is applied to sorting alternatives into predefined categories.

2.7.3.2 PROMETHEE methods

A wide range of methods are included in the PROMETHEE family to consider different types of MCDA problems [116][117]. The very first two techniques were PROMETHEE I and PROMETHEE II. PROMETHEE I, for example, provides a partial ranking of alternatives, alternatives not necessarily have to be ordered all the way down in situations where some alternatives cannot be directly compared with each other. PROMETHEE II, instead, will give a complete ranking that scores all the options in full view.

Some versions had to be created to make the method fit for more complex situations. PROMETHEE III, for example, deals with decisions involving interval data, in which evaluations are communicated as intervals rather than as exact values. PROMETHEE IV extends the methodology to problems involving a continuous set of alternatives, and is capable of generating both partial and complete rankings in such settings, making it suitable for optimization problems with infinite solution spaces.

In decision problems concerning segmentation constraints, e.g., grouping or clustering requirements, PROMETHEE V offers a framework specifically adapted for such conditions. This version constitutes a solution derivation methodology that contains the additional constraints view in order to establish feasible and practical solutions under complicated situations.

Then come PROMETHEE VI, modeling an aspect of human cognitive behavior, mainly how individuals approach decision-making under uncertainty and preference conflict. The Group Decision Support System (GDSS) can also be applied in a classic group decision-making setting, which allows multiple stakeholders to compose preference statements and reach consensus. The GAIA (Geometrical Analysis for Interactive Aid) module further complements the PROMETHEE techniques by providing a visual and interactive graphical representation of the decision problem. GAIA maps the alternatives and criteria onto a multi-dimensional plane, helping users visually interpret the influence of criteria and the positioning of alternatives, especially in complex and multidimensional decision scenarios.

More recently, PROMETHEE TRI and PROMETHEE CLUSTER [118] were introduced to handle classification and categorization tasks. PROMETHEE TRI is designed for sorting problems, where alternatives are assigned to predefined categories (e.g., good, average, poor), typically based on profiles. In contrast, PROMETHEE CLUSTER is intended for nominal classification, grouping alternatives without any predefined order, often used for market segmentation or organizational clustering.

Every variation of the PROMETHEE method has proven to be effective and robust with respect to several decision environments. In the following, a brief review of procedural and algorithmic steps in the implementation of PROMETHEE I is provided.

To begin, each pair of alternatives (a_i, a_j) from the set of alternatives A is compared across each of the k decision criteria. The difference in evaluation between alternatives a_i and a_j on criterion c is calculated as:

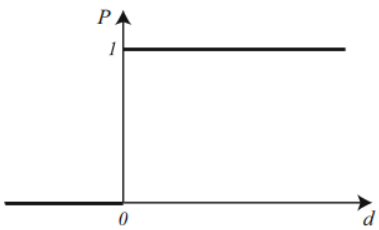
$$d_c(a_i, a_j) = f_c(a_i) - f_c(a_j) \;\; \forall\, c \in 1, \dots, k \qquad (24)$$

Here, $f_c(a_i)$ is the score of performance of the alternative a_i concerning criterion c. This difference indicates how much more a_i is preferred or considered less than a_j under criterion c.

The raw difference is translated into degrees of preference by means of a preference function P_c defined for each criterion. The value $P_c[d_c(a_i, a_j)]$ indicates the strength of preference for a_i over a_j on criterion c, where:

- P_c is a monotonically non-decreasing function, bounded between 0 and 1.
- A value of 0 denotes no preference, while a value of 1 signifies a strong or complete preference.

PROMETHEE stands for six types of preference functions represented in Table 2.2 that enable decision makers to choose from among the models presented, selecting one that best fits their judgment behavior and perception of difference in performances.

Type	**Preference function**	**Definition**	**Parameters**
Usual Criterion		$P(d) = \begin{cases} 0 & d \leq 0 \\ 1 & d > 0 \end{cases}$	-

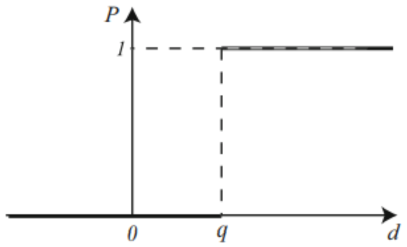
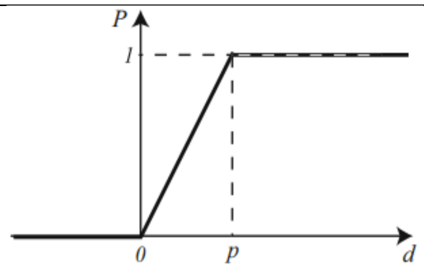
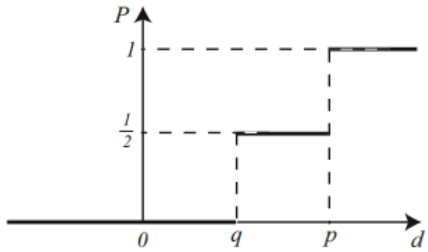
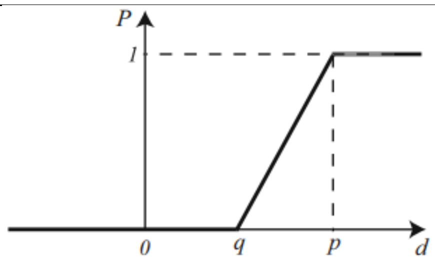
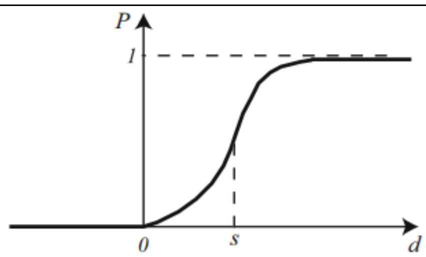
U-shape Criterion		$P(d) = \begin{cases} 0 & d \leq q \\ 1 & d > q \end{cases}$	q
V-shape Criterion		$P(d) = \begin{cases} 0 & d \leq 0 \\ \dfrac{d}{p} & 0 \leq d \leq p \\ 1 & d > 0 \end{cases}$	p
Level Criterion		$P(d) = \begin{cases} 0 & d \leq q \\ \dfrac{1}{2} & q < d \leq p \\ 1 & d > p \end{cases}$	p, q
V-shape with indifference Criterion		$P(d) = \begin{cases} 0 & d \leq 0 \\ \dfrac{d-q}{p-q} & q < d \leq p \\ 1 & d > p \end{cases}$	p, q
Gaussian Criterion		$P(d) = \begin{cases} 0 & d \leq 0 \\ 1 - e^{-\frac{d^2}{2s^2}} & d > 0 \end{cases}$	s

Table 2.2 Types of preference functions.

After computing the individual preference degrees for all criteria, they are combined into a global preference index $\pi(a_i, a_j)$, which expresses the overall preference of a_i over a_j across all criteria:

$$\pi(a_i, a_j) = \sum_{c=1}^{k} w_c \, P_c(a_i, a_j), \quad \left(\sum_{c=1}^{k} w_c = 1\right) \tag{25}$$

In this equation, w_c denotes the normalized weight associated with criterion c, reflecting its relative importance in the decision-making process. The aggregated index $\pi(a_i, a_j)$ is a real number between 0 and 1, representing the composite preference of one alternative over another.

Using the global preference indices, we then calculate the positive and negative preference flows for each alternative:

- The positive flow $\phi^+(a_i)$ indicates the degree to which alternative a_i is preferred over all other alternatives:

$$\phi^+(a_i) = \sum_{a_j \in A} \pi(a_i, a_j) \tag{26}$$

- The negative flow $\phi^-(a_i)$ indicates how much the other alternative are preferred to alternative a_i:

$$\phi^-(a_i) = \sum_{a_j \in A} \pi(a_j, a_i) \tag{27}$$

These flow scores represent the strength of preference relationships and form the basis of the ranking.

Based on the computed flows, PROMETHEE I establishes a partial preorder among alternatives. This means that some alternatives may be incomparable if the preference flows do not indicate a clear dominance. The rules for the partial preorder are defined as follows:

$a_i \; P \; a_j$ (preference)

$$\text{if} \begin{cases} \phi^+(a_i) \geq \phi^+(a_j) \wedge \phi^-(a_i) < \phi^-(a_j) \\ \vee \\ \phi^+(a_i) > \phi^+(a_j) \wedge \phi^-(a_i) \leq \phi^-(a_j) \end{cases}$$

$a_i \; I \; a_j$ (indefference)

$$\text{if} \;\; \phi^+(a_i) = \phi^+(a_j) \wedge \phi^-(a_i) = \phi^-(a_j)$$

$a_i \; R \; a_j$ (incomparability)

otherwise.

2.8 Conclusion

This chapter has examined the fundamental principles and methodological diversity that characterize the field of MCDA, emphasizing its role in facilitating structured decision-making in complex and multidimensional contexts. Among the various families of MCDA methods, outranking techniques are particularly notable for their flexibility and their ability to handle imprecise, conflicting, and qualitative information. These methods operate through pairwise comparisons of alternatives, establishing preference relations based on concordance and discordance indices. Within this category, the PROMETHEE method distinguishes itself as both a powerful and intuitive outranking approach, enabling decision-makers to derive partial or complete rankings through the use of preference flows. In this chapter, we have provided a detailed presentation of PROMETHEE I in view of its application in the present research.

In the next chapter, we present a literature review of existing outlier detection approaches within the MCDA context, highlighting their main contributions, limitations, and areas for further development.

CHAPTER 3: OUTLIERS DETECTION IN MCDA FIELD

3.1 Introduction

A standard method of MCDA is one that is dependent on a parameter set that is often entirely subjective and reflects the preferences and judgments of decision makers. Usually, the parameters are related to preference, indifference, weights, and veto thresholds, each of which models some aspect of the decision maker's valuation system. An important candidate parameter is indeed important to choose and correctly set because it greatly influences the results that are produced by the method. When some changes occur in these values, there will hardly be any minor changes caused by the others in the final ranking or classification of alternatives. Specific instances of parameter settings may lead to situations where some alternatives behave in ways that are completely different from the norm, thus becoming anomalies with respect to the rest of the data set. In the context of MCDA, these unusual alternatives are called multicriteria outliers [17]. These outliers may either represent potentially overlooked opportunities or reflect inconsistencies in the parameter settings.

Multicriteria outliers, when detected, could therefore greatly contribute to the quality and robustness of the MCDA process. Outlying alternatives offer the decision makers a chance to question and modify their parameter assignments so that the final rankings truly represent their deep preferences. This reviewing process makes validation of the robustness of the results possible and also improves the transparency and interpretability of the procedure.

The notion of detecting outliers in an MCDA context holds great importance but remains underexplored. As far as we know, very few papers have addressed this issue directly [17][18][19]. Hence, the domain of multicriteria outliers may be a fertile and relatively unexplored research arena with the possibility to enhance existing MCDA frameworks and methodologies.

Depending upon the definition of outliers in the MCDA context, there are multiple ways of looking at the problem of detecting outliers, with some attempting to resolve these definitions. We will analyze the different definitions further in the next section.

3.2 Outlier definitions in MCDA

Defining outliers rigorously and in context is crucial in MCDA due to its unique characteristics. As far as we know, the only definition of an outlier proposed in the MCDA field is that of Rouba et al. [18]. In general, the farther an object deviates from the rest of the dataset, the more likely it is to be perceived as an outlier. However, the authors in [18] note that comparisons between alternatives are typically not based on distance metrics, but rather on preference relations. They propose incorporating these relations directly into the definition of an outlier.

To this end, they introduce the notion of a relation-based outlier, which considers how an alternative is situated with respect to others in terms of preference and incomparability relations. Let $A = \{a_1, a_2, ..., a_n\}$ be a set of *n* alternatives, and let *P* and *R* denote the preference and incomparability relations, respectively. An alternative $a_i \in A$ is considered a relation-based outlier if it satisfies at least one of the following conditions:

1. a_i is *preferred to* at least a certain percentage (*pct)*% of the other alternatives in *A*. in other words, the number of alternatives $a_j \in A \setminus \{a_i\}$ for which a_i P a_j holds must be greater than or equal to *(pct)*% of (n−1).
2. Alternative a_i is preferred by at least a percentage *pct* of the alternatives in *A*. That is, the number of alternatives $a_j \in A \setminus \{a_i\}$ such that a_j P a_i is at least *(pct)*% of (n−1).
3. Alternative a_i is incomparable with at least a percentage *pct* of the alternatives in *A*. Formally, the number of alternatives $a_j \in A \setminus \{a_i\}$ such that a_j *R* a_i holds is at least *(pct)*% of (n−1).

This definition aims to capture various ways in which an alternative can be considered anomalous in the context of MCDA, by leveraging the fundamental structure of preference relations.

3.3 Key approaches for detecting outliers in MCDA

As noted, the topic of outlier detection in MCDA has been scarcely explored. To date, only three papers have addressed it: De Smet et al. [17], Rouba Baroudi and Nait Bahloul [18], and Rouba Baroudi [19]. The following section examines these works in detail.

3.3.1 Sampling based approach

In De Smet et al [17], the authors propose a novel approach to identify outliers by leveraging a distance-based model combined with sampling techniques. Their method relies on two main concepts: the profile of an alternative (an idea introduced by De Smet and Montano [119]) and a distance measure between alternatives.

The profile $P(a_i)$ of an alternative a_i is defined as a 4-tuple that captures its relations with all other alternatives:

- $R(a_i)$: Set of alternatives incomparable to a_i
- $P^-(a_i)$: Set of alternatives preferred over a_i
- $I(a_i)$: Set of alternatives indifferent to a_i
- $P^+(a_i)$: Set of alternatives that a_i is preferred to

The distance d_A (a_i, a_j) between two alternatives a_i and a_j is given by:

$$d_A\left(a_i, a_j\right) = 1 - \tfrac{1}{2}\textstyle\sum_{k=1}^{4}\left|P_k(a_i) \cap P_k\left(a_j\right)\right| \quad (28)$$

where n is the total number of alternatives. This measure quantifies how dissimilar the profiles of a_i and a_j are.

The proposed method involves the following steps:

1. ***Sampling Subsets:***Randomly generate k subsets $B_l \subset A$, each containing mm alternatives. Since outliers are assumed to be rare, they are unlikely to appear in most subsets, which reduces their influence.
2. ***Distance Calculation:***For each subset B_l, compute $d_{B_l}\left(a_i, a_j\right)$ for all pairs of alternatives. Then, for each alternative a_i, calculate the maximum distance to any other alternative within the subset:
$$\delta(a_i) = \max_{a_i \in A} d_{B_l}\left(a_i, a_j\right) \tag{29}$$
3. ***Aggregate Measure:***For each subset B_l, compute the aggregate dissimilarity as:
$$\Delta(B_l) = \sum_{i=1}^{m} \delta(a_i) \tag{30}$$
4. ***Distribution Analysis:***Analyze the distribution of $\Delta(B_l)$ values across all k subsets. If the dataset contains no outliers, the distribution should be unimodal. A bimodal distribution suggests the presence of one or more outliers, with the second mode corresponding to subsets containing them.
5. ***Outlier Identification:***Subsets with the highest $\Delta(B_l)$ values are examined. The alternative that appears most frequently in these subsets is identified as the outlier.

The approach increases the reliability of parameter estimation by diminutive the influence of outliers through repeated sampling of the subsets. The detection of bi-modal distributions in the distance values is a vivid and clear indicator of the presence of outliers. However, the method has some drawbacks, too. The detection process is still qualitative and does not provide anyone with formalized thresholds to identify bi-modal distributions, which could result in different interpretations. Also, the performance is dependent on the selection of parameters (e.g., subset size and number of repetitions), which are hard to set up correctly. The method may face challenges in the case of multiple outliers or when a significant part of the dataset is made up of outliers, thus, its effectiveness might be reduced.

3.3.2 Relation-based approach

In this study Rouba and Nait Bahloul [18], the authors propose a novel method for outlier detection within the MCDA framework, introducing the concept of a relation-based outlier. The approach integrates preference relations derived from multicriteria outranking methods such as PROMETHEE or ELECTRE with the LOF algorithm to identify anomalous alternatives.

The process begins with the construction of a preference matrix, where each pair of alternatives is evaluated using a multicriteria outranking method. This evaluation classifies their relationship as one of preference (P), indifference (I), or incomparability I.

Next, for each alternative a_i, a distribution vector $V(a_i)$ is computed. This vector is defined by the following four components:

- $V_1(a_i)$: Number of alternatives indifferent to a_i.
- $V_2(a_i)$: Number of alternatives preferred by a_i.
- $V_3(a_i)$: Number of alternatives that prefer a_i.
- $V_4(a_i)$: Number of alternatives incomparable to a_i.

These vectors provide a compact representation of each alternative's relational profile.

In this method, the LOF algorithm measures how isolated each alternative is within its local context, based on a distribution vector. It does so by identifying k-nearest neighbors using Euclidean distance, calculating the LRD, and finally determining the LOF score. Alternative with notably high LOF value is flagged as potential outlier.

The use of the LOF algorithm strengthens the approach by enabling the detection of density-based anomalies, which is particularly useful for identifying local deviations in the data. Additionally, the method is flexible, as it can be adapted to various multicriteria outranking techniques, such as PROMETHEE or ELECTRE. However, the approach has some limitations. Its performance and stability are highly sensitive to parameter choices, such as the number of nearest neighbors (k) in the LOF algorithm and the selected multicriteria method. This sensitivity may necessitate extensive tuning to ensure reliable results. Furthermore, the method always identifies the alternative with the highest LOF score as an outlier. This can be problematic in situations where multiple outliers are present or when no true outlier exists, potentially leading to misclassification.

3.3.3 Net-flow based approach

In [19], the author presents another approach for detecting outliers in MCDA problems by leveraging the net-flow values derived from the PROMETHEE II method, combined with statistical techniques. The central theme is to change the multicriteria decision-making process into a one-dimensional scale by means of the net-flow values of PROMETHEE, which express the total preference structure of the alternatives. To be more precise, the net-flow for each alternative a_i is derived from the difference between its outgoing flow $\phi^+(a_i)$ (the degree of its preference over the rest of the alternatives) and its incoming flow $\phi^-(a_i)$ (the degree of others' preference over it). This net-flow acts as an all-inclusive measure that indicates the relative performance of each alternative in terms of all criteria, thus maintaining the multipled criteria problem characteristics.

After the net-flow values have been calculated, the method applies statistical tests to reveal their distribution. Firstly, a Kolmogorov-Smirnov test is conducted to check if the net-flow values are distributed normally. The normal distribution of the net-flow values leads to the use of the standard deviation (SD) method to flag outliers. Thus, the values beyond the mean ± 2 SD are accepted as outliers. On the other side, in case the data is non-normally distributed, the interquartile range (IQR) method is put into action. In this case, outliers are those values that are lower than Q1 - 1.5 IQR or those that are greater than Q3 + 1.5 IQR, where Q1 and Q3 denote the first and third quartiles, respectively.

The proposed method offers several advantages, including its parameter-free nature—which eliminates the need for subjective tuning—and its ability to detect multiple outliers simultaneously. However, it is sensitive to tied net-flow values, which can compromise the normality assumption. Furthermore, the method is specifically based on PROMETHEE II for net-flow computation, which may limit its applicability to other MCDA methods.

3.4 Comparison between the approaches

The comparative analysis of the three outlier detection approaches in MCDA (namely, De Smet et al. (2017) [17], Rouba and Nait Bahloul (2018) [18], and Rouba (2021) [19]) highlights distinct methodologies, each with specific strengths and limitations, as summarized in Table 3.1. The method proposed by De Smet et al. is based on a distance-driven model utilizing relational profiles and repeated random sampling. Outliers are identified through the detection of bimodal patterns in dissimilarity metrics. While this approach is robust to outlier influence and offers a qualitative signal of their presence, it lacks formal decision thresholds and requires careful tuning of parameters such as subset size and the number of repetitions. In contrast, the method introduced by Rouba and Nait Bahloul combines multicriteria preference relations with the LOF algorithm to capture local deviations in the relational structure of alternatives. This method is adaptable and well-suited for identifying density-based anomalies. However, it depends heavily on the choice of parameters, particularly the number of neighbors (k), and its tendency to always classify the highest LOF scorer as an outlier can lead to instability or false positives, especially when true anomalies are absent or multiple are present. Rouba's 2021 approach offers a statistically grounded and parameter-free alternative. It transforms multicriteria data into a one-dimensional net flow space using PROMETHEE II, and applies standard deviation or interquartile range-based thresholds depending on the data distribution.

This method is notable for its simplicity and scalability, particularly in detecting multiple outliers. Nevertheless, its performance may be affected by the presence of tied net flow values and its reliance on PROMETHEE II, which may limit its applicability across other MCDA frameworks.

Criteria	De Smet et al. (2017)	Rouba & Nait Bahloul (2018)	Rouba (2021)
Core Idea	Distance-based outlier detection using subset sampling and profile dissimilarities.	Outlier detection using preference relation vectors and Local Outlier Factor (LOF) algorithm.	Statistical analysis of PROMETHEE II net-flow values to detect outliers.
Data Representation	Profile 4-tuples based on preference, indifference, incomparability.	Distribution vector with 4 components (P, I, R-based counts).	Scalar net-flow value summarizing all preferences.
MCDA Integration	PROMETHEE or ELECTRE to build preference matrix.	PROMETHEE or ELECTRE to build preference matrix.	Uses PROMETHEE II to compute net-flows directly.
Detection Technique	Analyze bimodal distribution of aggregated distances from multiple random subsets.	Apply LOF algorithm to distribution vectors.	Apply SD or IQR outlier detection.
Sensitivity to Parameters	High — needs careful tuning of subset size (m) and number of repetitions (k).	High — depends on LOF parameters like k and the outranking method used.	Low — parameter-free (except for normality test threshold).
Robustness to Multiple Outliers	Weak — may fail when multiple outliers are present or not rare.	Weak — identifies only the most isolated point (highest LOF score).	Strong — can detect multiple outliers simultaneously.
Scalability	Moderate — relies on repeated sampling and pairwise distance computation.	Moderate — requires pairwise comparison and LOF computation.	High — statistical calculations on a 1D vector are computationally efficient.

Limitation	No formal threshold for outlier identification; not ideal for multiple outliers.	Always returns the top LOF-scoring alternative as outlier; limited with multiple outliers.	Dependent on PROMETHEE II; sensitive to tied net-flow values.

Table 3.1 A comparison between the three approaches

3.5 Conclusion

The chapter focused on the significant but frequently overlooked problem of outlier detection in MCDA, with a particular focus on hw subjective parameter choices inherent to many MCDA methods can result in alternatives that diverge significantly from the majority. The main prominent approaches developed in this context were analyzed, illustrating a methodological diversity that spans distance-based models, density-based detection, and statistical analysis of net flows. Every method has its pros and cons in terms of insights and capabilities, such as being sensitive to parameters, not being able to scale up, or not being able to adapt to other MCDA frameworks. Based on these drawbacks, the following chapter introduces a new method that is intended to overcome the problems mentioned.

Part II:
CONTRIBUTION

CHAPTER 4: Weighted preference relation-based approach for outlier detection in MCDA context

4.1 Introduction

This chapter presents our contribution, structured into two principal components. The first introduces the concept of importance degree, designed to quantify the significance of preference relations between pairs of alternatives within the PROMETHEE framework. In this approach, each alternative is represented as a vector in a multidimensional space defined by its multicriteria preference indices. The importance degree is subsequently determined by computing the Euclidean distance between these alternatives.

In the second part, we develop a novel methodology for outlier detection in the context of multicriteria decision analysis (MCDA). We first introduce the notion of similarity, derived from the importance degree, to evaluate the degree of resemblance between each alternative and the rest of the set. Based on this measure, we then propose a formal definition of an outlier within this framework. The identification of outliers is conducted using statistical techniques that analyze the distribution of similarities across the set of alternatives.

4.2 Problem Statement

Multicriteria decision analysis (MCDA) methods often rely on subjective parameters (such as preference, indifference, weight, and veto thresholds) which serve as formal representations of the decision maker's (DM's) preferences. These parameters play a critical role in shaping the outcomes of the decision process. Variations in their values can substantially alter the results produced by MCDA methods, potentially leading to the emergence of alternatives that deviate significantly from the majority of the other elements. Within the MCDA context, such alternatives are referred to as multicriteria outliers [17], Consequently, detecting outliers in MCDA not only enhances the quality and robustness of the results but also assists the DM in refining the selection of optimal parameter values. Addressing these outliers contributes to a more reliable and transparent decision-making process.

Despite its importance, this area of research faces several challenges. First, the specific characteristics of MCDA (particularly its reliance on preference relations) necessitate the development of context-specific definitions and methodologies for outlier detection. Second, existing outlier detection techniques often require significant adaptation to be effectively applied in multicriteria settings. Finally, while outlier detection has been extensively studied in other fields, its application within MCDA remains relatively unexplored. To the best of our knowledge, only a limited number of studies have addressed this topic [17], [18], [19]. This paper represents a novel contribution to this field. This work seeks to advance the field by offering a novel contribution in this direction.

Our contribution is guided by the following research questions:

Our contribution is guided by the following research questions:

1. In an outranking context, multiple alternatives may exhibit similar comparative behavior with respect to a particular alternative. For instance, if two alternatives, a_i and a_j, are both preferred to a third alternative, a_k, a key question arises: which of the two, a_i or a_j, exhibits a stronger preference over a_k? More precisely, does the preference of a_i over a_k hold the same level of importance as that of a_j over a_k?
2. How can this degree of importance be formally computed?
3. How can this concept of importance be leveraged to identify outliers within the MCDA context, and how does its performance compare to existing approaches?

By addressing these questions, this work introduces the concept of the importance degree of a preference relation and proposes a straightforward approach for its computation. Furthermore, it examines the benefits of incorporating this measure into the outlier detection process in MCDA, thereby contributing to a more robust and insightful decision-making framework.

4.3 Positioning of Our Work

In this research, we aim to address the limitations and challenges identified in previous studies. A key element of our proposed approach is the integration of statistical methods that provide two major advantages:

- The capacity to detect multiple outliers within a dataset, and
- The ability to avoid incorrectly identifying outliers in cases where none are present.

These features distinguish our method from those proposed in the first and second referenced studies. Additionally, the reliance on net flow values in the third study introduces certain limitations, particularly in cases where multiple alternatives exhibit identical net flow values. Such scenarios can compromise the accuracy of the analysis and may violate the assumption of normality. To overcome this, our approach employs similarity measures instead of net flow values, allowing for a more refined differentiation between alternatives, even when they share identical net flow scores.

4.4 Importance Degree of Preference Relations

This section introduces and formalizes the concept of the *importance degree* associated with a preference relation between two alternatives, as derived from the PROMETHEE method. Furthermore, it details the methodology employed to quantify this importance.

In the PROMETHEE framework, the classification of a preference relation between two alternatives is determined by their respective positive flow score (ϕ^-) and positive flow score

(ϕ^{+}). To quantify the importance degree of a given preference relation, each alternative is represented as a point within an orthogonal plane, where the coordinates correspond to its positive and negative flow scores. This graphical representation facilitates the analysis of preference relations between alternatives.

For illustrative purposes, we designate the alternative a_i as the reference point for comparison, as shown in Figure 4.1. To streamline notation, the points (0,0), (1,0), (1,1), and (0,1) are denoted as *x, y, z,* and *w*, respectively.

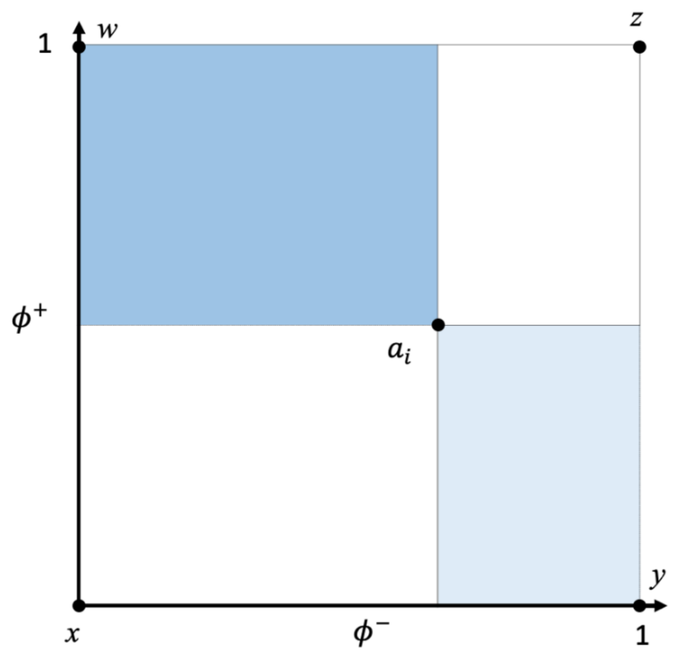


Figure 4.1 Preference relation regions for each alternative in the graph of ϕ^+ and ϕ^-

The preference relation between the reference alternative a_i and any other alternative is determined by their relative positions within the graph. To facilitate interpretation, the graph is divided into distinct colored regions, each representing a specific type of preference relation with respect to a_i:

- The dark blue region corresponds to all alternatives that are preferred to a_i.
- The light blue region represents all alternatives to which a_i is preferred.
- The white region includes all alternatives that are incomparable to a_i.
- Alternatives that are indifferent to a_i occupy the same coordinates as a_i

In our approach, we define the importance degree of a preference relation between two alternatives as the normalized distance between their respective positions in the plane. As this distance increases, the importance of the associated preference relation becomes correspondingly more significant.

To ensure comparability, we standardize these distances by identifying the maximum possible distance for each type of preference relation. For incomparability, it is evident that the greatest possible distance between two incomparable alternatives corresponds to the distance between points x and z Any alternative located within the white region will necessarily maintain a distance to $a_i$that does not exceed this value, regardless of its exact position within the region. Similarly, for cases of preference, the maximum distance between two alternatives, where one is preferred to the other, corresponds to the distance between points y and y. In contrast, for indifference, any two indifferent alternatives coincide at the same point, resulting in a distance of zero; consequently, the importance of such a relation is considered undefined.

While positive and negative flow scores offer a useful representation of alternatives for analyzing preference relations, this approach presents limitations in capturing all possible preference scenarios. Therefore, alternative representations will be explored in the following sections to address these shortcomings.

4.4.1 Review of Existing Alternative Representations

In this section, we examine the various representations of alternatives proposed in the existing literature. For example, in the PROMETHEE-GAIA framework [120], the authors employ unicriterion net flow scores as coordinates to represent alternatives within a multidimensional space.

$$\phi^c(a_i) = \frac{1}{n-1}\sum_{j\in A}(\pi^c(a_i,a_j) - \pi^c(a_j,a_i)) \quad (31)$$

where $\phi^c(a_i)$represent the unicriterion net flows of a_i on criterion *c*.

Similarly, in [121], the authors employed unicriterion net flows as coordinates to represent alternatives. By calculating and analyzing the distances between alternatives, they assessed the effectiveness of their proposed approach (PROMETHEE γ) in comparison to PROMETHEE I. However, utilizing unicriterion net flows as coordinates introduces certain limitations. Specifically, different alternatives may be mapped to identical coordinates, potentially leading to inaccurate evaluations of the importance of preference relations in specific scenarios. For instance, consider the alternatives x (ϕ^+= 0, $\phi^- = 0$) and z (ϕ^+= 1, $\phi^- = 1$). According to formulas (25), (26), (27), and (31), all their unicriterion net flows are equal to zero ($\forall\ c \in k$, $\phi^c(x)$= 0 and $\phi^c(z)$= 0). As a result, in cases of incomparability, the maximum distance between these alternatives is zero—specifically, the distance between *x* and *z*. Consequently, the importance of preference relations cannot be defined in such cases, highlighting a critical limitation of this representation.

4.4.2 Proposed Representation of Alternatives

In this section, we propose a novel representation of alternatives specifically designed to facilitate the evaluation of the importance of their preference relations. The core idea involves representing each alternative, denoted as a_i, in a multidimensional space, where its coordinates correspond to its multicriteria preference indices.

$$a_i = (\pi(a_i, a_1), \dots, \pi(a_i, a_n), \pi(a_1, a_i), \dots, \pi(a_n, a_i)) \quad (32)$$

where $\pi(a_i, a_j)$ denotes the preference value of a_i over a_j, and the following properties obviously hold for this preference values, $\forall\, a_i, a_j \in A$:

(1) $\pi(a_i, a_i) = 0$

(2) $0 \leq \pi(a_i, a_j) \leq 1$

For illustration, consider alternative y ($\phi^+ = 1$, $\phi^- = 0$) (Figure 4.1), meaning that it holds a multicriteria preference index of 1 over every other alternative, while all other alternatives have a multicriteria preference index of 0 when compared to y. Consequently, alternative y is represented by the point: $(1_{y,1}, 1_{y,2} \dots, 1_{y,n}, 0_{1,y}, 0_{2,y} \dots, 0_{n,y})$

Where:

$1_{i,j}$: denotes the multicriteria preference index of alternative *i* over alternative *j* being equal to 1

$0_{i,j}$: represents the multicriteria preference index of alternative *i* over alternative *j* being equal to 0.

By utilizing multicriteria preference indices, each alternative (including *x, y, z,* or *w*) is mapped to a unique point in the multidimensional space. This level of distinction is not achievable using unicriterion net flows. As a result, the maximum distances in cases of both incomparability and preference can be determined. However, in cases of indifference, the maximum distance remains undefined, since two indifferent alternatives correspond to identical points. It is noteworthy that, for indifferent alternatives, their separation distance does not reach that of *[x, z] or [y, w]*. Therefore, we establish the maximum distance across all scenarios as:

$$[x, z] = [y, w] = \sqrt{2n - 2}$$

Where *n* represents the number of alternatives.

Definition 1

The importance of a preference relation of two alternatives a_i and a_j, is denoted as imp_{ij}, where $imp_{ij} \in]0; 1]$, defined by [122]:

In case of indifference

$$imp_{ij} = 1 - \frac{1}{\sqrt{2n-2}} \sqrt{\sum_{k \in A} ((\pi_{ik} - \pi_{jk})^2 + (\pi_{ki} - \pi_{kj})^2)} \tag{33}$$

in other cases:

$$imp_{ij} = \frac{1}{\sqrt{2n-2}} \sqrt{\sum_{k \in A} ((\pi_{ik} - \pi_{jk})^2 + (\pi_{ki} - \pi_{kj})^2)} \tag{34}$$

In cases of indifference, the importance of the preference relation is assigned a value of 1 when the two alternatives exhibit identical multicriteria preference indices, resulting in a distance of zero between them. Conversely, for incomparability or preference relations, the importance is considered equal to 1 when the distance between the two alternatives reaches its maximum possible value. Accordingly, the importance of a preference relation between two alternatives a_i and a_j is defined as follows:

$$a_i \; imp_{i,j} \; S \; a_j \tag{35}$$

where *S* is a preference relation $\in \{ P^+, P^-, R, I \}$, and $imp_{i,j}$ is the importance of preference relation between a_i and a_j.

The degree of importance adheres to the following fundamental properties:

- **Monotonicity**:
  The importance degree varies monotonically with the distance between alternatives, as defined by their multicriteria preference indices.
  - *In the case of indifference:*
    When alternatives are considered indifferent, the importance degree increases as the distance between them decreases. It reaches a maximum value of 1 when the distance is zero. Conversely, as the distance increases, the importance degree decreases, tending toward 0.
  - *In other cases:*
    When alternatives are not indifferent, the importance degree increases with greater distance between them, reaching 1 at the maximum distance. Conversely, as the distance shrinks, the importance degree decreases and approaches 0.
- **Reflexivity**:
  The importance degree is reflexive in the case of self-comparisons. That is, when an

alternative a_i is compared to itself, the importance degree is always 1:

$$imp_{i,j} = 1$$

This is because there is no distance between an alternative and itself.

- **Symmetry**:

  The importance degree is symmetric by definition:

  $$imp_{i,j} = imp_{j,i}$$

  This property holds because the preference relation between alternatives a_i and a_j is unaffected by their order, given that the distance metric used is symmetric.

- **Non-Compensation**:

  As defined by equations (33) and (34), the importance degree is computed using a sum of independent squared differences. Since squared terms are always non-negative, an increase in one term cannot be offset by a decrease in another. As a result, the importance degree reflects a non-compensatory nature.

4.4.3 Case Illustration

In this section, we will present an illustrative example to elucidate the computation of preference relation importance. To this end, we consider the decision problem summarized in Table 4.1.

Criteria	a_1	a_2	a_3	a_4	a_5	a_6	Type of Criteria	Parameters	Max Min
f_1	80	65	83	40	52	94	Quasi-criterion	q = 10	Min
f_2	90	58	60	80	72	96	Criterion with linear preference	p =30	Max
f_3	6	2	4	10	6	7	Criterion with - linear preference and indifference area	q =0.5 p =5	Min
f_4	5.4	9.7	7.2	7.5	2.0	3.6	Level criterion	q = 1 p = 6	Min
f_5	8	1	4	7	3	5	Usual criterion	-	Min
f_6	5	1	7	10	8	6	Gaussian criterion	σ = 5	Max

Table 4.1 A decision problem.

Using PROMETHEE, the multicriteria preference indices for each alternative are provided in Table 4.2, while the corresponding positive and negative flow scores are presented in Table 4.3.

	a_1	a_2	a_3	a_4	a_5	a_6
a_1	-	0.296	0.250	0.269	0.100	0.185
a_2	0.463	-	0.389	0.333	0.296	0.500
a_3	0.235	0.180	-	0.333	0.056	0.429
a_4	0.399	0.506	0.305	-	0.224	0.212
a_5	0.444	0.515	0.487	0.380	-	0.448
a_6	0.287	0.399	0.250	0.431	0.133	-

Table 4.2 The multicriteria preference indexes.

	a_1	a_2	a_3	a_4	a_5	a_6
ϕ^+	0.220	0.396	0.247	0.329	0.455	0.300
ϕ^-	0.366	0.379	0.336	0.349	0.162	0.355

Table 4.3 Positive and negative outranking flows.

The first step in computing the degree of importance consists of representing each alternative by its corresponding vector, as defined in Equation (25). Table 4.4 provides a summary of the vectors associated with all alternatives.

Alternative	*Vector*
a_1	(0.296, 0.250, 0.269, 0.100, 0.185, 0.463, 0.235, 0.399, 0.444, 0.287)
a_2	(0.463, 0.389, 0.333, 0.296, 0.500, 0.296, 0.180, 0.506, 0.515, 0.399)
a_3	(0.235, 0.180, 0.333, 0.056, 0.429, 0.250, 0.389, 0.305, 0.487, 0.250)
a_4	(0.399, 0.506, 0.305, 0.224, 0.212, 0.269, 0.333, 0.333, 0.380, 0.431)
a_5	(0.444, 0.515, 0.487, 0.380, 0.448, 0.100, 0.296, 0.056, 0.224, 0.133)
a_6	(0.287, 0.399, 0.250, 0.431, 0.133, 0.185, 0.500, 0.429, 0.212, 0.448)

Table 4.4 Alternative vectors.

Using Equations (33) and (34), the degree of importance for each preference relation is computed, with the results presented in Table 4.5.

	a_1	a_2	a_3	a_4	a_5	a_6
a_1	I	0.158 R	0.125 P^-	0.130 P^-	0.247 P^-	0.191 P^-
a_2	0.158 R	I	0.164 R	0.134 R	0.214 P^-	0.201 R
a_3	0.125 P^+	0.164 R	I	0.161 R	0.215 P^-	0.208 R
a_4	0.130 P^+	0.134 R	0.161 R	I	0.183 P^-	0.122 P^+
a_5	0.247 P^+	0.214 P^+	0.215 P^+	0.183 P^+	I	0.220 P^+
a_6	0.191 P^+	0.201 R	0.208 R	0.122 P^-	0.220 P^-	I

Table 4.5 Importance degrees of preference relations.

As an illustrative example, a_1 is incomparable to a_2, and the importance of this relation, quantified using the proposed approach, is given by: $imp_{1,2} = \frac{1}{\sqrt{2n-2}} * 0.500$ = 0.158 (with n = 6). According to definition 1, their relation is expressed by a_1 0.158 R a_2.

Table 4.5 illustrates that alternative a_5 is preferred over all other alternatives. Although the preference relations between $a_5$the other alternatives are of the same type, their corresponding levels of importance vary. This variation allows for distinguishing between seemingly similar preference relations.

In the subsequent section, this concept of importance serves as the foundation for introducing a novel approach to outlier detection based on preference relations.

4.5 Outlier detection

In this section, we present a method for detecting outliers within the MCDA domain. The proposed approach is based on measuring the similarity between alternatives by utilizing the concept of the importance degree of preference relations. Relying on the distribution of these similarity values, an appropriate statistical technique is applied for outlier detection, specifically either the interquartile range (IQR) method or the standard deviation (SD) method.

4.5.1 Outlier in MCDA field

In general, an outlier is defined as an observation that deviates significantly from the other observations in a dataset [123]. However, the definition of an outlier may vary depending on the field of study or the specific context in which it is applied. In the context of MCDA, the evaluation of an alternative is determined by its preference relations with other alternatives, making these relations the key element for identifying outliers in this domain. For instance, in [120], the authors define an outlier as any alternative that possesses a number of specific preference relations (P+, P−, or R) with other alternatives that is equal to or exceeds a predetermined percentage threshold.

In this study, we propose a novel conceptualization of outliers in the MCDA context, founded on the notion of similarity. This concept measures the extent to which an alternative resembles the remainder of the dataset, providing a new perspective for outlier detection in multicriteria decision analysis.

Definition 2

Let's denote three alternatives as a_i, a_j, and a_k. We consider that alternatives a_i and a_j are similar with respect to a_k. If a_i and a_j have the same preference relation with a_k, we term this similarity ($similarity(a_i, a_j)_{a_k}$), and it is computed as follows [122]:

$$similarity(a_i, a_j)_{a_k} = 1 - \frac{|imp_{i,k} - imp_{j,k}|}{\max(imp_{i,k}, imp_{j,k})} \quad (36)$$

where $imp_{i,k}$ is the importance degree of the preference relation between alternatives a_i and a_k

According to Definition 2, we can deduce the similarity between two alternatives a_i and a_j by averaging their similarities with respect to all other alternatives:

$$similarity(a_i, a_j) = \frac{\sum_{a_k \in A} similarity(a_i, a_j)_{a_k}}{n - 2} \quad (37)$$

where *n* represents the number of alternatives of the set *A*

Likewise, the similarity of an alternative a_i and all the remaining alternatives within A is given by:

$$similarity(a_i)_A = \frac{\sum_{a_j \in A} similarity(a_i, a_j)}{n-1} \quad (38)$$

As previously established, the similarity between two alternatives relative to a third serves as the basis for both pairwise similarity and the overall similarity of an individual alternative. Consequently, the following analysis will focus exclusively on the properties of this foundational similarity measure:

- **Monotonicity**:
  The similarity increases as the absolute difference $|imp_{i,k} - imp_{j,k}|$ decreases. Since this absolute difference is subtracted from 1 in the similarity formula, a smaller difference between $imp_{i,k}$ and $imp_{j,k}$ results in a higher similarity value. This behavior reflects the principle of monotonicity.
- **Reflexivity**:
  When $a_\mathrm{i} = a_j$, it follows that $imp_{i,k} = imp_{j,k}$. Therefore:

$$similarity(a_i, a_i)_{a_k} = 1 - \frac{0}{\max(imp_{i,k}, imp_{i,k})} = 1$$

  This confirms the reflexivity property, as the similarity of an alternative with itself is always equal to 1.

- **Symmetry**:
  The similarity measure is symmetric due to the presence of the absolute difference $|imp_{i,k} - imp_{j,k}|$, which is invariant under the interchange of a_i and a_j. Thus:

$$similarity(a_i, a_j)_{a_k} = similarity(a_j, a_i)_{a_k}$$

- **Non-Compensation**:
  The similarity between a_i and a_j with respect to a_k is determined exclusively by the magnitude of the difference $|imp_{i,k} - imp_{j,k}|$. As a result, the measure does not allow for compensation (no additional factors can offset the effect of a large difference, which inherently reduces similarity).
- **Normalization**:
  When $|imp_{i,k} - imp_{j,k}| = 0$, the similarity reaches its maximum value:

$$similarity(a_i, a_j)_{a_k} = 1 - \frac{0}{\max(imp_{i,k}, imp_{j,k})} = 1.$$

  Conversely, as the difference $|imp_{i,k} - imp_{j,k}|$ approaches its theoretical maximum (i.e., $\max(imp_{i,k}, imp_{j,k})$), the similarity tends toward 0. However, it never reaches zero because the importance degrees are strictly positive (as defined in Definition 1). Therefore:

$$similarity(a_i, a_j)_{a_k} \in]0; 1]$$

Having established the theoretical properties of the proposed similarity measure, the next step involves identifying outliers within the MCDA framework. This is achieved by evaluating the similarity of each alternative relative to all others, as this reflects the degree to which an alternative conforms to the overall set.

Definition 3

A multicriteria outlier is an alternative whose similarity significantly deviates from the general pattern of similarities within the set of alternatives. It is characterized by a near-zero similarity score relative to other alternatives, indicating a substantial lack of resemblance to the rest [122].

Table 4.6 presents the similarity scores of alternatives of our example, calculated through the application of formula (38) to the data in table 4.5.

Alternative	a_1	a_2	a_3	a_4	a_5	a_6
Similarity(.)	0.432	0.454	0.638	0.631	0.148	0.659

Table 4.6 The similarity of alternatives

According to Definition 3, an outlier is characterized by possessing a similarity value that is significantly closer to zero compared to the remainder of the alternatives. To assess the degree to which an alternative can be considered an outlier, we apply statistical techniques based on the following procedure:

1. Testing the normality of the similarity distribution**.**
2. Selecting the appropriate outlier detection method based on the result:
    - If the similarity values follow a normal distribution, the Standard Deviation (SD) method is applied.
    - Otherwise, the Interquartile Range (IQR) method is employed.

To verify the normality of the data, we use the *Shapiro-Wilk test* [124]. This test evaluates the null hypothesis that the data are normally distributed. A p-value below a chosen significance level (typically 0.05) leads to the rejection of the null hypothesis, indicating that the data are not normally distributed. Conversely, a p-value exceeding the significance threshold suggests insufficient evidence to reject the null hypothesis, implying that the data may be normally distributed.

In our illustrative example, the Shapiro–Wilk test applied to the similarities presented in Table 6 produces a p-value of 0.1556, indicating that the similarities follow a normal distribution. Therefore, the SD method is employed, yielding a mean (μ) of 0.493 and a standard deviation (σ) of 0.178. As a result, alternatives with similarity values outside the range [0.226, 0.760] are identified as outliers. Under this criterion, alternative a_5, with a similarity of 0.148, is classified as an outlier.

4.6 Conclusion

This chapter presented a novel approach for outlier detection within the domain of multicriteria decision analysis (MCDA). We introduced a new metric for quantifying the importance of preference relations, providing a means not only to distinguish whether alternatives are preferred, indifferent, or incomparable, but also to evaluate the intensity with which these relations are expressed.

Building upon this concept, we proposed an outlier detection method tailored to the MCDA context. This method defines similarity between alternatives based on the importance degrees of their preference relations with other alternatives. Through statistical analysis of these similarity measures, the approach enables the identification of alternatives that exhibit significant deviations from the general distribution, thereby classifying them as potential multicriteria outliers.

In the following chapter, we will validate our proposed approach through its application to two real-world decision problems. The first part of the evaluation will focus on assessing the contribution of the importance degree concept in isolation, while the second part will evaluate the overall performance of the complete approach across varying decision-making scenarios.

CHAPTER 5:
APPLICATION AND RESULTS

5.1 Introduction

This chapter evaluates the proposed approach through two distinct phases, each designed to highlight specific aspects of its effectiveness.

In the first phase, the objective is to assess the contribution of the importance degree concept independently. To this end, we incorporate the importance degree into an existing outlier detection framework, specifically the Local Outlier Factor (LOF)-based method described in [18] The evaluation is conducted across three distinct scenarios, each formulated to examine the behavior and added value of the importance degree under varying conditions. For this purpose, we employ a real-world dataset from the World Happiness Report, which ranks countries based on multiple well-being indicators. By comparing the results of outlier detection with and without the integration of the importance degree, we aim to empirically demonstrate the contribution of the proposed metric to enhancing outlier identification.

The second phase involves a comprehensive evaluation of the complete proposed framework. This phase is carried out using a different real-world dataset—the Human Development Index (HDI)—which consolidates metrics related to life expectancy, education, and income to assess national development levels. Here, two scenarios are explored to evaluate the performance of the full approach under varying parameter settings and structural assumptions. The primary objective of this phase is to validate the utility, robustness, and adaptability of the proposed outlier detection framework in practical multicriteria decision-making contexts.

5.2 Effectiveness Testing of Importance Degree Metric

This section is dedicated to assessing the effectiveness of the *importance degree* concept through an empirical case study based on the *World Happiness Ranking*. The objective of this case study is to rank countries according to multiple criteria that collectively reflect happiness and overall quality of life. Specifically, six key indicators are employed in the evaluation: Gross Domestic Product (GDP) per capita, Social Support (Family), Healthy Life Expectancy (Life), Freedom to make life choices (Freedom), Trust (in government and institutions), and Generosity. The experimental analysis is conducted using data from the 2015 edition of the *World Happiness Report* [125], which comprises information on 158 countries.

The experimentation was conducted through three distinct scenarios, each designed to evaluate the impact of integrating the *degree of importance* into the LOF-based outlier detection method. In each scenario, we applied the LOF approach twice: once in its original form and once with the incorporation of the importance degree concept.

The scenarios were defined as follows:

- **Scenario 1**: Included the top 20 ranked alternatives along with the lowest-ranked alternative.
- **Scenario 2**: Included 20 alternatives from the middle of the ranking, in addition to both the highest-ranked and lowest-ranked alternatives.
- **Scenario 3**: Included the 20 lowest-ranked alternatives, along with the highest-ranked alternative.

This sampling strategy was deliberately chosen to emphasize particular alternatives as outliers depending on the scenario:

- In **Scenario 1**, the expectation was that the lowest-ranked alternative would be identified as the outlier.
- In **Scenario 2**, both the highest and lowest-ranked alternatives were anticipated to stand out as outliers.
- In **Scenario 3**, the highest-ranked alternative was expected to emerge as the outlier.

To perform the rankings, the PROMETHEE method was employed in the initial stage, with the results presented in *Appendix A*. The PROMETHEE method was further utilized within each scenario to:

1. Establish the preference relations among the selected alternatives (see *Appendices B, C, and D*).
2. Compute the multicriteria preference indices of each alternative, which are essential for determining the *degree of importance* of the established preference relations.

For the PROMETHEE configuration, the *level preference function* was adopted, with the first and third quartiles of the evaluation differences used as the *indifference* and *preference* thresholds, respectively. The detailed parameter settings for the PROMETHEE method are summarized in Table 5.1.

To conduct a sensitivity analysis, the LOF-based approach was implemented using various values of the parameter *k* to assess its influence on the detection of outliers. For each experimental scenario, the distribution of alternatives was analyzed both with and without incorporating the *degree of importance* concept. The corresponding distributions are presented in Tables 5.2, 5.4, and 5.6 for Scenarios 1, 2, and 3, respectively.

Additionally, the respective results of the LOF algorithm for each scenario—highlighting the three alternatives with the highest LOF scores—are provided in Tables 5.3, 5.5, and 5.6. These results allow for a comparative analysis of the LOF method's performance with and without the integration of the *degree of importance*, thereby illustrating its contribution to the effectiveness of outlier detection.

	GDP	Family	Life	Freedom	Trust	Generosity
Weight	0.1666667	0.1666667	0.1666667	0.1666667	0.1666667	0.1666667
Indifferent threshold (q)	0.105	0.072	0.042	0.044	0.034	0.052
Preference threshold (p)	0.412	0.295	0.216	0.206	0.125	0.205
Type of Criteria	Level criterion	Level criterion	Level criterion	Level criterion	Level criterion	Level criterion
Max/Min	max	max	max	max	max	max

Table 5.1 Parameters used for the execution of the PROMETHEE method

5.2.1 Integrating the importance degree

As previously discussed, this outlier detection technique employs the LOF algorithm, which is applied to the *distribution vectors* of alternatives. These vectors are constructed to capture the preference relations of each alternative with respect to the entire set of alternatives, denoted as *A*.

For each alternative *ai* ∈ *A*, its corresponding distribution vector encodes the degrees of importance associated with its preference relations relative to every other alternative in *A*. This representation enables the LOF algorithm to assess the degree to which each alternative deviates from the general distribution pattern, thereby facilitating the identification of potential outliers within the multicriteria decision-making context.

$V_1(a_i) = |\{a_j \in \mathrm{A}, a_i \neq a_j \wedge a_i \mathrm{I} a_j\}|$

$V_2(a_i) = |\{a_j \in \mathrm{A}, a_i \neq a_j \wedge a_i \mathrm{P} a_j\}|$

$V_3(a_i) = |\{a_j \in \mathrm{A}, a_i \neq a_j \wedge a_j \mathrm{P} a_i\}|$

$V_4(a_i) = |\{a_j \in \mathrm{A}, a_i \neq a_j \wedge a_i \mathrm{R} a_j\}|$

By integrating the importance degree, instead of calculating a preference relation as 1 in the distribution vectors, we compute its degree of importance. Consequently, the distribution vector will be as follows:

$V_1(a_i) = \sum_{a_j \in \mathrm{A}, a_i \neq a_j \wedge a_i \mathrm{I} a_j} imp_{i,j}$

$V_2(a_i) = \sum_{a_j \in \mathrm{A}, a_i \neq a_j \wedge a_i \mathrm{P} a_j} imp_{i,j}$

$V_3(a_i) = \sum_{a_j \in \mathrm{A}, a_i \neq a_j \wedge a_j \mathrm{P} a_i} imp_{i,j}$

$V_4(a_i) = \sum_{a_j \in \mathrm{A}, a_i \neq a_j \wedge a_i \mathrm{R} a_j} imp_{i,j}$

5.2.2 Scenario 1

In this first scenario, we evaluated the LOF-based outlier detection approach by applying it both with and without the integration of the importance degree concept. The analysis was restricted to a subset of the dataset consisting of the 20 highest-ranked alternatives, along with the lowest-ranked alternative, as presented in Table 5.2 and Table 5.3.

This scenario was specifically designed to highlight the distinctiveness of the lowest-ranked alternative, aiming to determine whether incorporating the importance degree enhances the ability of the LOF algorithm to identify it as an outlier.

Alternatives	Without the degree of importance				With the degree of importance			
	$V_1(.)$	$V_2(.)$	$V_3(.)$	$V_4(.)$	$V_1(.)$	$V_2(.)$	$V_3(.)$	$V_4(.)$
New Zealand	1	14	0	6	0	2.272	0	0.716
Australia	1	12	1	7	0	1.960	0.072	0.755
Canada	1	7	5	8	0	1.488	0.388	0.897
Ireland	1	6	5	9	0	1.333	0.465	0.886
Norway	1	12	1	7	0	2.077	0.082	0.705
Netherlands	1	6	7	7	0	1.296	0.664	0.766
Sweden	1	8	2	10	0	1.600	0.144	1.060
United Kingdom	1	7	5	8	0	1.433	0.538	0.856
Denmark	1	13	1	6	0	2.185	0.063	0.595
Switzerland	1	15	0	5	0	2.422	0	0.505
Luxembourg	1	5	7	8	0	1.111	0.707	0.741
Hong Kong	1	5	3	12	0	1.120	0.313	1.209
Iceland	1	6	4	10	0	1.188	0.435	0.961
Qatar	1	5	0	15	0	1.243	0	1.887
Finland	1	5	10	5	0	1.081	0.998	0.516
Singapore	1	3	0	17	0	0.905	0	2.573
Austria	1	1	15	4	0	0.530	2.239	0.453
Malta	1	2	16	2	0	0.626	2.445	0.183
Germany	1	1	15	4	0	0.523	2.370	0.456
United States	1	1	17	2	0	0.462	3.238	0.208
Burundi	**1**	**0**	**20**	**0**	**0**	**0**	**11.694**	**0**

Table 5.2 The distribution matrix for two cases, with and without the degree of importance, in Scenario 1 [122]

	Without the degree of importance			With the degree of importance		
	1^{st}	2^{nd}	3^{rd}	1^{st}	2^{nd}	3^{rd}
2	Singapore	Qatar	Finland	**Burundi**	Singapore	United States
3	Singapore	Qatar	**Burundi**	**Burundi**	Singapore	United States
4	Singapore	**Burundi**	United States	**Burundi**	United States	Singapore
5	**Burundi**	United States	Malta	**Burundi**	United States	Singapore
6	**Burundi**	United States	Malta	**Burundi**	United States	Malta
7	**Burundi**	United States	Malta	**Burundi**	United States	Malta

Table 5.3 The outcomes of the approach proposed in [18] with and without importance degree in Scenario 1 [122]

5.2.3 Scenario 2

In the second scenario, we modified the selection of alternatives by including the 20 middle-ranked alternatives, along with both the highest- and lowest-ranked alternatives. This configuration was intended to evaluate the ability of the LOF-based approach, with and without the integration of the importance degree, to detect outliers positioned at both extremes of the ranking. The results obtained from this scenario are presented in Table 5.4 and Table 5.5.

Alternatives	Without the degree of importance				With the degree of importance			
	$V_1(.)$	$V_2(.)$	$V_3(.)$	$V_4(.)$	$V_1(.)$	$V_2(.)$	$V_3(.)$	$V_4(.)$
New Zealand	**1**	**21**	**0**	**0**	**0**	**9.595**	**0**	**0**
Slovakia	1	17	1	3	0	3.023	0.384	0.535
Venezuela	1	17	1	3	0	2.553	0.382	0.525
Philippines	1	13	3	5	0	2.004	0.692	0.745
Jamaica	1	11	4	6	0	1.712	0.754	0.854
Vietnam	1	10	4	7	0	1.527	0.792	0.984
Hungary	1	6	7	8	0	1.142	1.141	1.362
Jordan	1	9	4	8	0	1.317	0.768	1.171
Latvia	1	10	3	8	0	1.452	0.665	1.254
Rwanda	1	1	1	19	0	0.300	0.453	3.458
Cambodia	1	1	3	17	0	0.313	0.849	2.776
China	1	5	8	8	0	0.778	1.201	1.046
Myanmar	1	1	1	19	0	0.304	0.446	3.447
Kyrgyzstan	1	2	7	12	0	0.446	1.212	1.460
Bolivia	1	1	12	8	0	0.350	1.938	1.077
Lebanon	1	3	11	7	0	0.545	1.751	0.872
Turkey	1	6	6	9	0	0.860	1.186	1.210
Azerbaijan	1	1	12	8	0	0.348	1.810	1.026
Macedonia	1	2	9	10	0	0.435	1.678	1.308

Russia	1	1	11	9	0	0.333	1.902	1.253
Peru	1	1	10	10	0	0.364	1.743	1.339
Burundi	**1**	**0**	**21**	**0**	**0**	**0**	**7.955**	**0**

Table 5.4 The distribution matrix for two cases, with and without the importance degree, in Scenario 2 [122]

k	Without the degree of importance			With the degree of importance		
	1^{st}	2^{nd}	3^{rd}	1^{st}	2^{nd}	3^{rd}
2	**Burundi**	Rwanda	Myanmar	**Burundi**	**New Zealand**	Rwanda
3	**Burundi**	Rwanda	Myanmar	**Burundi**	**New Zealand**	Rwanda
4	**Burundi**	Rwanda	Myanmar	**Burundi**	**New Zealand**	Rwanda
5	**Burundi**	Rwanda	Myanmar	**Burundi**	**New Zealand**	Rwanda
6	**Burundi**	Rwanda	Myanmar	**Burundi**	**New Zealand**	Rwanda
7	**Burundi**	**New Zealand**	Rwanda	**New Zealand**	**Burundi**	Rwanda

Table 5.5 The outcomes of the approach proposed in [18] with and without importance degree in Scenario 2 [122]

5.2.4 Scenario 3

In the third and final scenario, we selected a different set of alternatives consisting of the 20 lowest-ranked alternatives, along with the highest-ranked alternative. This configuration was designed to assess the effectiveness of the LOF-based approach in identifying outliers when the majority of alternatives are concentrated at the lower end of the ranking. The corresponding results for this scenario are presented in Table 5.6 and Table 5.7.

Alternatives	Without the degree of importance				With the degree of importance			
	$V_1(.)$	$V_2(.)$	$V_3(.)$	$V_4(.)$	$V_1(.)$	$V_2(.)$	$V_3(.)$	$V_4(.)$
New Zealand	**1**	**20**	**1**	**1**	**0**	**11.497**	**0**	**0**
Egypt	1	10	1	9	0	1.857	0.485	1.186
Lesotho	1	13	4	3	0	1.579	0.870	0.323
Sierra Leone	1	7	8	5	0	0.878	1.336	0.500
Ghana	1	17	1	2	0	2.459	0.499	0.208
Pakistan	1	10	2	8	0	1.493	0.601	0.866
Comoros	1	15	2	3	0	2.172	0.605	0.276
Zimbabwe	1	12	1	7	0	1.583	0.553	0.805
Guinea	1	7	8	5	0	0.801	1.367	0.499
Malawi	1	4	14	2	0	0.480	2.228	0.260
Benin	1	9	6	5	0	1.083	1.142	0.596
Afghanistan	1	6	13	1	0	0.741	1.930	0.132
Yemen	1	14	3	3	0	1.875	0.714	0.319

Congo (Kinshasa)	1	3	15	2	0	0.349	2.456	0.188
Central African Republic	1	0	19	1	0	0	3.678	0.108
Liberia	1	8	7	5	0	0.909	1.225	0.562
Angola	1	2	7	11	0	0.299	1.413	1.332
Madagascar	1	7	8	5	0	0.940	1.349	0.582
Togo	1	2	16	2	0	0.236	2.626	0.208
Chad	1	2	14	2	0	0.275	2.302	0.396
Burundi	1	0	19	1	0	0	4.127	0.108

Table 5.6 The distribution matrix for two cases, with and without the importance degree, in Scenario 3 [122].

k	Without the degree of importance			With the degree of importance		
	1^{st}	2^{nd}	3^{rd}	1^{st}	2^{nd}	3^{rd}
2	Liberia	**New Zealand**	Egypt	**New Zealand**	Burundi	Liberia
3	**New Zealand**	Liberia	Ghana	**New Zealand**	Burundi	Central African Republic
4	**New Zealand**	Central African Republic	Burundi	**New Zealand**	Burundi	Central African Republic
5	**New Zealand**	Central African Republic	Burundi	**New Zealand**	Burundi	Central African Republic
6	**New Zealand**	Central African Republic	Burundi	**New Zealand**	Burundi	Central African Republic
7	**New Zealand**	Central African Republic	Burundi	**New Zealand**	Burundi	Central African Republic

Table 5.7 The outcomes of the approach proposed in [18] with and without importance degree in Scenario 3 [122].

5.3 Validation of the entire approach

After evaluating the effectiveness of the concept of the degree of importance independently, this section demonstrates the application of our complete outlier detection approach on a real-world problem: the Human Development Index (HDI), as presented in [126]. The HDI, introduced by the United Nations Development Programme (UNDP), assesses 179 countries based on three main criteria: life expectancy, education, and income index. In this study, we do not aim to reproduce or analyze the exact rankings of countries. Instead, our focus is on calculating similarities between alternatives in order to detect potential outliers.

To implement our approach, we begin by applying the PROMETHEE method to establish preference relations among the alternatives and compute the corresponding degrees of importance. Table 5.8 provides a summary of the generalized criteria types used for each evaluation criterion, along with their respective parameter values, as specified by the decision maker (DM). For further details regarding these parameters, please refer to De Smet et al.[126].

Parameters	**Life expectancy**	**Adult literacy index**	**GDP**
preference threshold	0.704	0.719	0.828
Indifference threshold	0	0	0
Weight of criterion	0.333	0.333	0.333
preference function	V-shape	V-shape	V-shape

Table 5.8 The indifference and preference thresholds, as well as the weight and preference function of each criterion [122].

We aim to evaluate our proposed technique using two distinct scenarios: one dataset without outliers and another containing an artificially introduced outlier. In the first scenario, we apply our method to the original list of countries, which does not include any outliers (the evaluations for these countries are provided in [126]) In the second scenario, we introduce an artificial outlier, labeled **ARTIF**, by assigning it the minimum possible value for each criterion, as illustrated in Table 5.9. By conducting experiments on both datasets (with and without outliers) we can thoroughly assess the sensitivity and robustness of our approach in different conditions.

	Life expectancy	**Adult literacy index**	**GDP**
ARTIF	0.253	0.274	0.172

Table 5.9 The evaluations of the artificial outlier (ARTIF) [122].

5.3.1 Scenario 1

This scenario involves the calculation of country similarities without any outliers, as illustrated in Table 5.10.

Country	**similarity**	**Country**	**similarity**	**Country**	**similarity**
Iceland	0.296	Venezuela	0.404	Guatemala	0.37
Norway	0.298	Romania	0.406	Kyrgyzstan	0.373
Canada	0.299	Malaysia	0.406	Vanuatu	0.366
Australia	0.3	Montenegro	0.406	Tajikistan	0.366

Ireland	0.309	Serbia	0.405	South Africa	0.353
Netherlands	0.309	Saint_Lucia	0.406	Botswana	0.348
Sweden	0.309	Belarus	0.406	Morocco	0.346
Japan	0.312	Macedonia TFYR	0.411	Sao Tome and Principe	0.346
Luxembourg	0.314	Albania	0.411	Namibia	0.341
Switzerland	0.314	Brazil	0.411	Congo	0.335
France	0.313	Kazakhstan	0.406	Bhutan	0.332
Finland	0.315	Ecuador	0.411	India	0.331
Denmark	0.318	Russian Federation	0.409	Lao People	0.331
Austria	0.317	Mauritius	0.413	Solomon Islands	0.325
United States	0.32	Bosnia and Herzegovina	0.413	Myanmar	0.324
Spain	0.319	Turkey	0.413	Cambodia	0.319
Belgium	0.319	Dominica	0.412	Comoros	0.319
Greece	0.321	Lebanon	0.412	Yemen	0.315
Italy	0.322	Peru	0.411	Pakistan	0.314
New Zealand	0.322	Colombia	0.413	Mauritania	0.313
United Kingdom	0.324	Thailand	0.411	Swaziland	0.301
Hong Kong, China(SAR)	0.327	Ukraine	0.413	Ghana	0.304
Germany	0.325	Armenia	0.414	Madagascar	0.307
Israel	0.331	Iran	0.414	Kenya	0.304
Korea(Republic of)	0.333	Tonga	0.414	Nepal	0.306
Slovenia	0.336	Grenada	0.414	Sudan	0.3
Brunei Darussalam	0.343	Jamaica	0.414	Bangladesh	0.302
Singapore	0.342	Belize	0.409	Haiti	0.3
Kuwait	0.347	Suriname	0.412	Papua New Guinea	0.297
Cyprus	0.344	Jordan	0.414	Cameroon	0.295
United Arab Emirates	0.35	Dominican Republic	0.412	Djibouti	0.295
Bahrain	0.355	Saint Vincent and the Grenadines	0.413	Tanzania	0.291
Portugal	0.353	Georgia	0.412	Senegal	0.292
Qatar	0.357	China	0.411	Nigeria	0.289
Czech Republic	0.355	Tunisia	0.41	Lesotho	0.287
Malta	0.357	Samoa	0.411	Uganda	0.289
Barbados	0.359	Azerbaijan	0.409	Angola	0.274
Hungary	0.372	Paraguay	0.408	Timor-Leste	0.284
Poland	0.371	Maldives	0.406	Togo	0.281
Chile	0.371	Algeria	0.405	Gambia	0.276
Slovakia	0.374	El Salvador	0.404	Benin	0.267
Estonia	0.377	Philippines	0.406	Malawi	0.27
Lithuania	0.378	Fiji	0.401	Zambia	0.266
Latvia	0.381	Sri Lanka	0.402	Eritrea	0.262
Croatia	0.38	Syria	0.399	Rwanda	0.257
Argentina	0.382	Occupied Palestinian Territories	0.398	Cote d’Ivoire	0.251
Uruguay	0.383	Gabon	0.376	Guinea	0.249
Cuba	0.384	Turkmenistan	0.394	Mali	0.234
Bahamas	0.389	Indonesia	0.393	Ethiopia	0.234
Costa Rica	0.39	Guyana	0.39	Chad	0.232
Mexico	0.394	Bolivia	0.392	Guinea Bissau	0.231
Libyan Arab Jamahiriya	0.396	Mongolia	0.393	Burundi	0.233
Oman	0.396	Moldova	0.393	Burkina Faso	0.225
Seychelles	0.401	Viet_Nam	0.393	Niger	0.226
Saudi Arabia	0.399	Equatorial	0.339	Mozambique	0.221

Bulgaria	0.402	Guinea Egypt	0.387	Liberia	0.224
Trinidad and Tobago	0.401	Honduras	0.385	Congo Democratic	0.223
Panama	0.402	Cape Verde	0.379	Central African	0.213
Antigua and Barbuda	0.404	Uzbekistan	0.379	Sierra Leone	0.202
Saint Kitts and Nevis	0.404	Nicaragua	0.376	**ARTIF**	-

Table 5.10 The similarities of the alternatives in the scenario 1[122].

5.3.2 Scenario 2

In this scenario, we compute the similarities between countries after introducing the artificial outlier ARTIF, which was assigned the lowest possible value for each criterion, as shown in Table 5.11.

Country	similarity	Country	similarity	Country	similarity
Iceland	0.297	Venezuela	0.405	Guatemala	0.371
Norway	0.299	Romania	0.406	Kyrgyzstan	0.374
Canada	0.299	Malaysia	0.406	Vanuatu	0.367
Australia	0.301	Montenegro	0.406	Tajikistan	0.367
Ireland	0.31	Serbia	0.406	South Africa	0.354
Netherlands	0.31	Saint_Lucia	0.406	Botswana	0.349
Sweden	0.309	Belarus	0.406	Morocco	0.347
Japan	0.312	Macedonia TFYR	0.411	Sao Tome and Principe	0.347
Luxembourg	0.314	Albania	0.411	Namibia	0.342
Switzerland	0.315	Brazil	0.411	Congo	0.336
France	0.313	Kazakhstan	0.406	Bhutan	0.333
Finland	0.315	Ecuador	0.411	India	0.333
Denmark	0.318	Russian Federation	0.409	Lao People	0.334
Austria	0.318	Mauritius	0.413	Solomon Islands	0.326
United States	0.321	Bosnia and Herzegovina	0.413	Myanmar	0.327
Spain	0.319	Turkey	0.413	Cambodia	0.32
Belgium	0.319	Dominica	0.413	Comoros	0.321
Greece	0.321	Lebanon	0.412	Yemen	0.317
Italy	0.322	Peru	0.412	Pakistan	0.316
New Zealand	0.323	Colombia	0.412	Mauritania	0.315
United Kingdom	0.324	Thailand	0.411	Swaziland	0.305
Hong Kong, China(SAR)	0.327	Ukraine	0.414	Ghana	0.306
Germany	0.326	Armenia	0.414	Madagascar	0.309
Israel	0.331	Iran	0.414	Kenya	0.306
Korea(Republic of)	0.333	Tonga	0.413	Nepal	0.308
Slovenia	0.336	Grenada	0.414	Sudan	0.302
Brunei Darussalam	0.344	Jamaica	0.414	Bangladesh	0.305
Singapore	0.343	Belize	0.409	Haiti	0.302
Kuwait	0.347	Suriname	0.414	Papua New Guinea	0.299
Cyprus	0.345	Jordan	0.414	Cameroon	0.297
United Arab Emirates	0.35	Dominican Republic	0.413	Djibouti	0.296
Bahrain	0.355	Saint Vincent and the Grenadines	0.413	Tanzania	0.294
Portugal	0.353	Georgia	0.412	Senegal	0.294

Qatar	0.358	China	0.411	Nigeria	0.291
Czech Republic	0.356	Tunisia	0.412	Lesotho	0.289
Malta	0.357	Samoa	0.412	Uganda	0.292
Barbados	0.359	Azerbaijan	0.41	Angola	0.276
Hungary	0.371	Paraguay	0.408	Timor-Leste	0.287
Poland	0.371	Maldives	0.405	Togo	0.284
Chile	0.371	Algeria	0.405	Gambia	0.279
Slovakia	0.374	El Salvador	0.404	Benin	0.27
Estonia	0.377	Philippines	0.406	Malawi	0.272
Lithuania	0.378	Fiji	0.401	Zambia	0.268
Latvia	0.381	Sri Lanka	0.402	Eritrea	0.265
Croatia	0.38	Syria	0.4	Rwanda	0.259
Argentina	0.382	Occupied Palestinian Territories	0.398	Cote d'Ivoire	0.254
Uruguay	0.382	Gabon	0.377	Guinea	0.252
Cuba	0.384	Turkmenistan	0.395	Mali	0.237
Bahamas	0.389	Indonesia	0.394	Ethiopia	0.237
Costa Rica	0.39	Guyana	0.391	Chad	0.235
Mexico	0.394	Bolivia	0.393	Guinea Bissau	0.235
Libyan Arab Jamahiriya	0.396	Mongolia	0.392	Burundi	0.236
Oman	0.396	Moldova	0.394	Burkina Faso	0.229
Seychelles	0.402	Viet_Nam	0.393	Niger	0.23
Saudi Arabia	0.399	Equatorial Guinea	0.341	Mozambique	0.224
Bulgaria	0.402	Egypt	0.388	Liberia	0.227
Trinidad and Tobago	0.401	Honduras	0.385	Congo Democratic	0.226
Panama	0.402	Cape Verde	0.38	Central African	0.217
Antigua and Barbuda	0.404	Uzbekistan	0.379	Sierra Leone	0.205
Saint Kitts and Nevis	0.404	Nicaragua	0.376	**ARTIF**	**0.17**

Table 5.11 The similarities of the alternatives in the scenario 2 [122].

In the first experiment, the similarities between countries did not follow a normal distribution, as indicated by a p-value of 5.589e-9. Using the IQR method, we calculated the upper bound ($Q3 + 1.5 \times IQR$) as 0.5415 and the lower bound ($Q1 - 1.5 \times IQR$) as 0.1695. Since no country's similarity value fell outside this range, we concluded that no outliers were present in the dataset. In the second experiment, we introduced the artificial outlier 'ARTIF' into the dataset and again tested for normality. The updated sample continued to exhibit a non-normal distribution, with a p-value of 6.882e-9. After recalculating the IQR bounds, we found the interval to be [0.171, 0.540]. ARTIF, with a similarity score of 0.170, fell outside this range, confirming its status as an outlier.

5.4 Discussion

In the initial phase of the experiment, we assessed the results by defining accuracy as the proportion of cases in which the LOF algorithm correctly identified the primary outlier (or both the first and second outliers, depending on the scenario) relative to the total number of tested values for the parameter *k*, which was fixed at seven across all scenarios.

The experimental results demonstrated that integrating the degree of importance into the preference relations significantly enhanced performance, yielding substantially higher accuracy

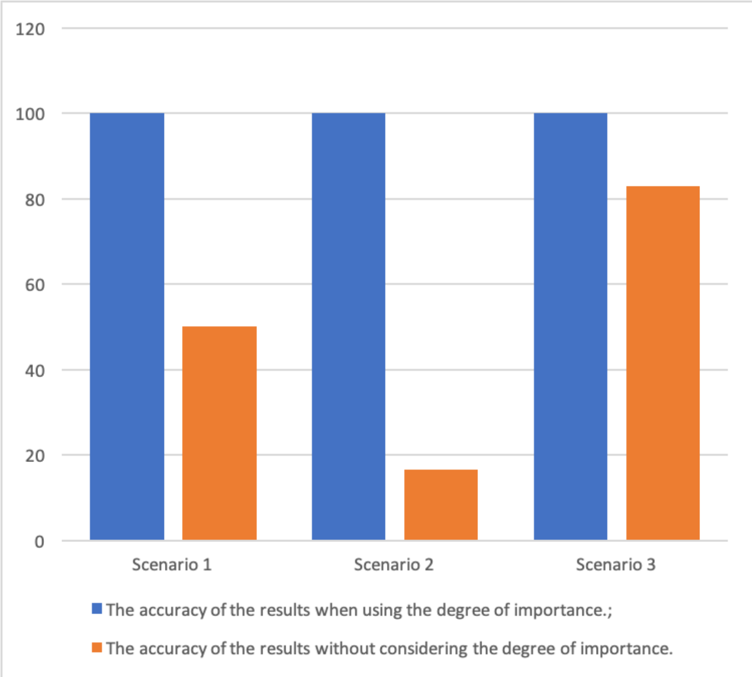


Figure 5.1 Comparing the accuracy of results in all scenarios [122].

rates in all three scenarios, as depicted in Figure 5.1. Specifically, when the importance degree was incorporated, the approach consistently achieved 100% accuracy across all scenarios. In contrast, considerably lower accuracy rates were observed when the degree of importance was excluded. These findings clearly underscore the critical role of the importance degree in improving the reliability and precision of outlier detection in MCDA contexts.

Notably, the three scenarios exhibited a substantial decline in accuracy when the degree of importance was excluded. Specifically, Scenario 1 dropped to 42%, Scenario 2 to 14%, and Scenario 3 to 85%.

Furthermore, when representing alternatives in three-dimensional space using their distribution matrices (excluding the first coordinate, as it holds the same value for all alternatives), the effect of incorporating the degree of importance becomes more visually evident. As illustrated in Figure 5.2, the spatial configuration of alternatives reveals that integrating the importance degree leads to clearer separation of outliers from the main clusters, thereby facilitating their detection. Conversely, omitting the degree of importance results in alternatives being positioned more closely together, making the identification of outliers more challenging.

This visual evidence reinforces the earlier findings, demonstrating that considering the degree of importance significantly enhances both the precision and interpretability of outlier detection in multicriteria decision analysis.

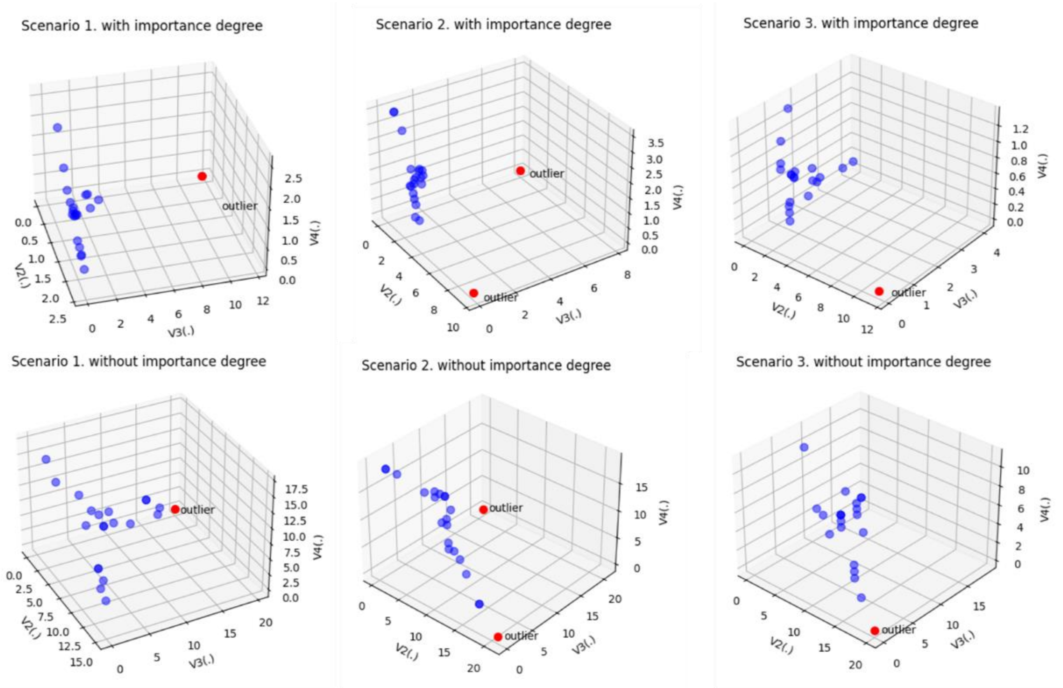


Figure 5.2 Representing alternatives using their distribution matrix in all scenarios [122].

In the second phase of the experiment, the results from both scenarios further confirmed the effectiveness of the proposed approach for detecting outliers within the MCDA framework.

In the first scenario, the method was applied to a dataset without any outliers to assess its capacity for avoiding false positives. The dataset, characterized by consistent and homogeneous similarity measures between countries, was thoroughly analyzed. The results confirmed that the data did not follow a normal distribution, and importantly, no outliers were detected within the defined statistical bounds, demonstrating the method's reliability in stable contexts.

In the second scenario, an artificial outlier ('ARTIF') was introduced to rigorously test the robustness of the approach. The successful detection of ARTIF as an outlier, after it was found outside the calculated bounds, highlights the method's strong discriminative capability. As illustrated in Figure 5.3, the outlier is clearly distinguished from the remaining alternatives, providing visual evidence of the method's effectiveness in separating anomalies from regular observations.

These findings underscore the robustness and adaptability of the proposed method for outlier detection in multicriteria decision analysis.

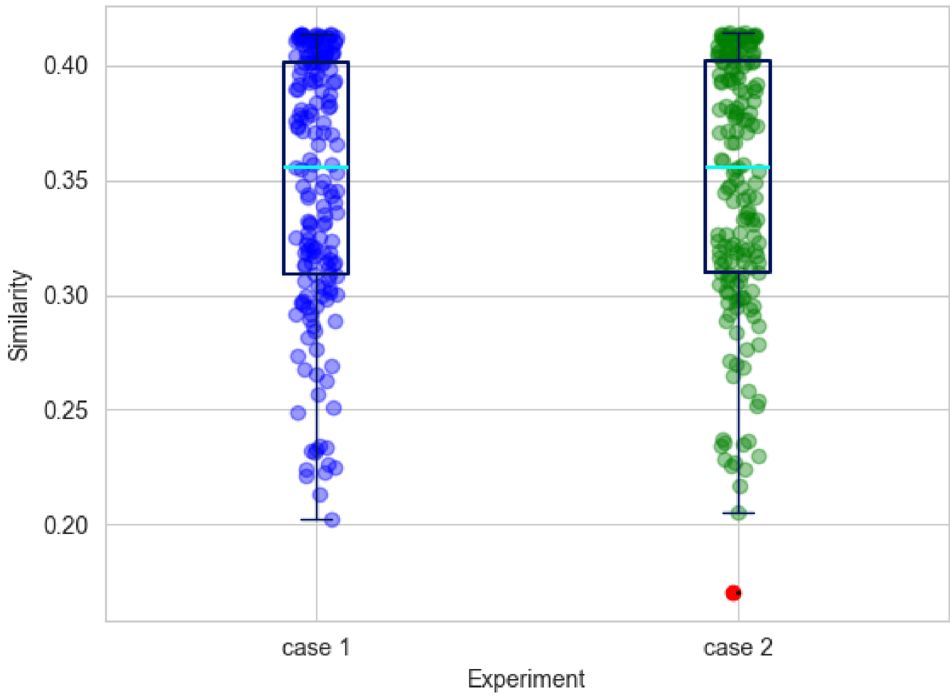


Figure 5.3 Data representation of the first and second experimentations, with outliers marked in red [122].

5.5 Conclusion

This chapter presented a series of experimental studies aimed at evaluating the effectiveness of the proposed approach, applied to two real-world problems. The experimental design was structured in two main parts.

The first part focused on assessing the contribution of the degree of importance concept, which has the potential to enhance the performance of existing outlier detection methods. This was demonstrated by integrating the importance degree into the LOF-based outlier detection method. To thoroughly evaluate its contribution, we constructed three distinct scenarios that intentionally positioned certain alternatives as outliers. These scenarios allowed us to observe whether incorporating the degree of importance improves the ability to accurately detect these outliers. The experimental results clearly demonstrated that the integration of this concept significantly enhances the accuracy of outlier detection.

In the second part, the complete outlier detection approach was applied to another real-world problem. Here, two evaluation scenarios were considered. The first involved a dataset without any real outliers, aimed at verifying that the method does not produce false positives. The second scenario introduced an artificial outlier, created by assigning the minimum values across all evaluation criteria, to test the method's sensitivity and robustness.

Overall, the results from both parts validated the effectiveness and practical applicability of the proposed approach in multicriteria decision analysis contexts.

GENERAL CONCLUSION

Summary of the results

The primary objective of this thesis is to develop a novel approach for detecting outliers in the MCDA domain, considering the multicriteria nature of the problem and providing a specific definition of an outlier in this context.

To this end, we present a contribution divided into two parts. In the first part, we introduce a measure called importance degree, which quantifies the strength of the preference relation between two alternatives. Each alternative is represented by its multicriteria preference indices as coordinates, and the importance of their relation is assessed using the Euclidean distance between these points.

In the second part, we propose an outlier detection approach based on a similarity measure derived from the importance degree. This new measure evaluates how similar each alternative is to the others. Outliers are then identified by applying statistical methods, specifically the IQR or SD, depending on the distribution of similarity scores.

Finally, we define what constitutes an outlier in the MCDA context by grounding the definition in the notion of similarity.

To validate our approach, we conducted experiments in two phases using two real-life case studies. In the first phase, we focused solely on evaluating the effectiveness of the degree of importance in the World Happiness Ranking problem. The results of our approach were compared with those of the method proposed in [18], both with and without incorporating the degree of importance. In the second phase, we assessed the overall effectiveness of our approach using the Human Development Index problem. We evaluated its performance under both outlier-present and outlier-absent scenarios, and it demonstrated robust results in both cases.

The results of the first phase demonstrated 100% accuracy across all scenarios when the degree of importance was applied, outperforming the use of the original preference relations. In the second phase, the approach did not identify any false outliers in the first scenario. In the second scenario, it successfully detected the artificially introduced outlier.

In conclusion, a key strength of our approach lies in its ability to account for the multicriteria nature of decision-making problems, which enables the accurate detection of outliers within the MCDA framework. Additionally, we have introduced a novel concept, namely, the degree of importance which offers a promising direction for improving the performance of existing and future MCDA methods through its integration.

However, our approach is based exclusively on preference relations generated by the PROMETHEE method. This represents a limitation in terms of generalizability, as the approach may not be directly applicable to other MCDA methods that use different forms of preference modeling or rely on alternative decision-making frameworks.

Perspectives

The proposed approach offers various opportunities for future research, as each of its phases allows for further development. One significant direction is to apply the degree of importance to other preference-based MCDA methods in areas such as clustering and outlier detection, where understanding the strength of preference relations is advantageous.

In the current formulation of our approach, the degree of importance is defined and applied exclusively to preference relations derived from the PROMETHEE method. However, we aim to broaden the applicability of this concept by extending it to other outranking methods. This expansion includes incorporating the degree of importance into preference structures generated by techniques such as ELECTRE, thereby enhancing the generality and versatility of our framework.

BIBLIOGRAPHY

[1] Roy B. Multicriteria methodology for decision aiding. vol. 12. Springer Science & Business Media; 1996.

[2] Yang X, Latecki LJ, Pokrajac D. Outlier detection with globally optimal exemplar-based GMM. Proc. 2009 SIAM Int. Conf. data Min., SIAM; 2009, p. 145–54.

[3] Dang TT, Ngan HYT, Liu W. Distance-based k-nearest neighbors outlier detection method in large-scale traffic data. 2015 IEEE Int. Conf. Digit. Signal Process., IEEE; 2015, p. 507–10.

[4] Breunig MM, Kriegel H-P, Ng RT, Sander J. LOF: identifying density-based local outliers. Proc. 2000 ACM SIGMOD Int. Conf. Manag. data, 2000, p. 93–104.

[5] MacQueen J. Some methods for classification and analysis of multivariate observations. Proc. fifth Berkeley Symp. Math. Stat. Probab., vol. 1, Oakland, CA, USA; 1967, p. 281–97.

[6] Moonesinghe HDK, Tan P-N. Outrank: a graph-based outlier detection framework using random walk. Int J Artif Intell Tools 2008;17:19–36.

[7] Nguyen HV, Ang HH, Gopalkrishnan V. Mining outliers with ensemble of heterogeneous detectors on random subspaces. Database Syst. Adv. Appl. 15th Int. Conf. DASFAA 2010, Tsukuba, Japan, April 1-4, 2010, Proceedings, Part I 15, Springer; 2010, p. 368–83.

[8] Pimentel T, Monteiro M, Veloso A, Ziviani N. Deep active learning for anomaly detection. 2020 Int. Jt. Conf. Neural Networks, IEEE; 2020, p. 1–8.

[9] Bolton RJ, Hand DJ. Unsupervised profiling methods for fraud detection. Credit scoring Credit Control VII 2001:235–55.

[10] Ding Q, Katenka N, Barford P, Kolaczyk E, Crovella M. Intrusion as (anti) social communication: characterization and detection. Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discov. data Min., 2012, p. 886–94.

[11] Tsou Y-L, Chu H-M, Li C, Yang S-W. Robust distributed anomaly detection using optimal weighted one-class random forests. 2018 IEEE Int. Conf. Data Min., IEEE; 2018, p. 1272–7.

[12] Roy B. Multicriteria methodology for decision aiding. vol. 12. Springer Science & Business Media; 2013.

[13] Vincke P. Multicriteria decision-aid. John Wiley & Sons; 1992.

[14] Korhonen P. Interactive methods. Mult criteria Decis Anal state art Surv 2005:641–61.

[15] Ishizaka A, Nemery P. Multi-attribute utility theory. Multi-Criteria Decis Anal Methods Softw 2013:81–113.

[16] Brans J-P, Vincke P, Mareschal B. How to select and how to rank projects: The PROMETHEE method. Eur J Oper Res 1986;24:228–38.

[17] De Smet Y, Hubinont J-PP, Rosenfeld J. A note on the detection of outliers in a binary outranking relation. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) 2017;10173 LNCS:151–9. https://doi.org/10.1007/978-3-319-54157-0_11.

[18] Rouba B, Nait-Bahloul S. Towards identifying multicriteria outliers: An outranking relation-based approach. Int J Decis Support Syst Technol 2018;10:27–38. https://doi.org/10.4018/IJDSST.2018070102.

[19] Rouba B. A net-flow based approach to detect outliers in multicriteria decision problems. Intell Decis Technol 2021;15:239–50. https://doi.org/10.3233/IDT-200046.

[20] Wang H, Bah MJ, Hammad M. Progress in outlier detection techniques: A survey. Ieee Access 2019;7:107964–8000.

[21] Aleskerov E, Freisleben B, Rao B. Cardwatch: A neural network based database mining system for credit card fraud detection. Proc. IEEE/IAFE 1997 Comput. Intell. Financ. Eng., IEEE; 1997, p. 220–6.

[22] Edgeworth FY. Xli. on discordant observations. london, edinburgh, dublin Philos Mag J Sci 1887;23:364–75.

[23] Ayadi A, Ghorbel O, Obeid AM, Abid M. Outlier detection approaches for wireless sensor networks: A survey. Comput Networks 2017;129:319–33.

[24] Grubbs FE. Procedures for detecting outlying observations in samples. Technometrics 1969;11:1–21.

[25] Boukerche A, Zheng L, Alfandi O. Outlier detection: Methods, models, and classification. ACM Comput Surv 2020;53:1–37.

[26] Smiti A. A critical overview of outlier detection methods. Comput Sci Rev 2020;38:100306.

[27] Dasgupta D, Nino F. A comparison of negative and positive selection algorithms in novel pattern detection. Smc 2000 Conf. proceedings. 2000 ieee Int. Conf. Syst. man Cybern. Evol. to Syst. humans, Organ. their complex Interact. no. 0, vol. 1, IEEE; 2000, p. 125–30.

[28] Atallah M, Szpankowski W, Gwadera R. Detection of significant sets of episodes in event sequences. Fourth IEEE Int. Conf. Data Min., IEEE; 2004, p. 3–10.

[29] Bhattacharyya S, Jha S, Tharakunnel K, Westland JC. Data mining for credit card fraud: A comparative study. Decis Support Syst 2011;50:602–13.

[30] Bolton RJ, Hand DJ. Statistical fraud detection: A review. Stat Sci 2002;17:235–55.

[31] Estévez PA, Held CM, Perez CA. Subscription fraud prevention in telecommunications using fuzzy rules and neural networks. Expert Syst Appl 2006;31:337–44.

[32] Perols J. Financial statement fraud detection: An analysis of statistical and machine learning algorithms. Audit A J Pract Theory 2011;30:19–50.

[33] Rabatel J, Bringay S, Poncelet P. Anomaly detection in monitoring sensor data for preventive maintenance. Expert Syst Appl 2011;38:7003–15.

[34] Van Capelleveen G, Poel M, Mueller RM, Thornton D, van Hillegersberg J. Outlier detection in healthcare fraud: A case study in the Medicaid dental domain. Int J Account Inf Syst 2016;21:18–31.

[35] Shu K, Sliva A, Wang S, Tang J, Liu H. Fake news detection on social media: A data mining perspective. ACM SIGKDD Explor Newsl 2017;19:22–36.

[36] Chenaoua K, Kurugollu F, Bouridane A. Data cleaning and outlier removal: Application in human skin detection. 2014 5th Eur. Work. Vis. Inf. Process., IEEE; 2014, p. 1–6.

[37] D'Urso C. EXPERIENCE: glitches in databases, how to ensure data quality by outlier detection techniques. J Data Inf Qual 2016;7:1–22.

[38] Ye N, Chen Q. An anomaly detection technique based on a chi-square statistic for detecting intrusions into information systems. Qual Reliab Eng Int 2001;17:105–12.

[39] Akaike H. Information theory and an extension of the maximum likelihood principle. Sel. Pap. hirotugu akaike, Springer; 1998, p. 199–213.

[40] Schwarz G. Estimating the dimension of a model. Ann Stat 1978:461–4.

[41] Rousseeuw PJ. Least median of squares regression. J Am Stat Assoc 1984;79:871–80.

[42] Huber PJ. Robust regression: asymptotics, conjectures and Monte Carlo. Ann Stat 1973:799–821.

[43] Box GEP, Jenkins GM. Times series analysis: forecasting and control. 2 e éd 1976.

[44] Tsay RS, Pena D, Pankratz AE. Outliers in multivariate time series. Biometrika 2000;87:789–804.

[45] Eskin E. Anomaly detection over noisy data using learned probability distributions. Proc. seventeenth Int. Conf. Mach. Learn., 2000, p. 255–62.

[46] Abraham B, Box GEP. Bayesian analysis of some outlier problems in time series. Biometrika 1979;66:229–36.

[47] Hickinbotham SJ, Austin J. Novelty detection in airframe strain data. Proc. 15th Int. Conf. Pattern Recognition. ICPR-2000, vol. 2, IEEE; 2000, p. 536–9.

[48] Pavlidou M, Zioutas G. Kernel density outlier detector. Top. nonparametric Stat. Proc.

first Conf. Int. Soc. nonparametric Stat., Springer; 2014, p. 241–50.

[49] Laurikkala J, Juhola M, Kentala E, Lavrac N, Miksch S, Kavsek B. Informal identification of outliers in medical data. Fifth Int. Work. Intell. data Anal. Med. Pharmacol., vol. 1, 2000, p. 20–4.

[50] Yamanishi K, Takeuchi J-I, Williams G, Milne P. On-line unsupervised outlier detection using finite mixtures with discounting learning algorithms. Proc. sixth ACM SIGKDD Int. Conf. Knowl. Discov. data Min., 2000, p. 320–4.

[51] Endler D. Intrusion detection. Applying machine learning to Solaris audit data. Proc. 14th Annu. Comput. Secur. Appl. Conf. (Cat. No. 98EX217), IEEE; 1998, p. 268–79.

[52] Ramaswamy S, Rastogi R, Shim K. Efficient algorithms for mining outliers from large data sets. Proc. 2000 ACM SIGMOD Int. Conf. Manag. data, 2000, p. 427–38.

[53] Eskin E, Arnold A, Prerau M, Portnoy L, Stolfo S. A geometric framework for unsupervised anomaly detection: Detecting intrusions in unlabeled data. Appl data Min Comput Secur 2002:77–101.

[54] Angiulli F, Pizzuti C. Fast outlier detection in high dimensional spaces. Eur. Conf. Princ. data Min. Knowl. Discov., Springer; 2002, p. 15–27.

[55] Zhang J, Wang H. Detecting outlying subspaces for high-dimensional data: the new task, algorithms, and performance. Knowl Inf Syst 2006;10:333–55.

[56] Knorr EM, Ng RT, Tucakov V. Distance-based outliers: algorithms and applications. VLDB J 2000;8:237–53.

[57] Bay SD, Schwabacher M. Mining distance-based outliers in near linear time with randomization and a simple pruning rule. Proc. ninth ACM SIGKDD Int. Conf. Knowl. Discov. data Min., 2003, p. 29–38.

[58] Angiulli F, Fassetti F. Very efficient mining of distance-based outliers. Proc. Sixt. ACM Conf. Conf. Inf. Knowl. Manag., 2007, p. 791–800.

[59] Vu NH, Gopalkrishnan V. Efficient pruning schemes for distance-based outlier detection. Mach. Learn. Knowl. Discov. Databases Eur. Conf. ECML PKDD 2009, Bled, Slov. Sept. 7-11, 2009, Proceedings, Part II 20, Springer; 2009, p. 160–75.

[60] Tang J, Chen Z, Fu AW-C, Cheung DW. Enhancing effectiveness of outlier detections for low density patterns. Adv. Knowl. Discov. data Min. 6th Pacific-Asia Conf. PAKDD 2002 Taipei, Taiwan, May 6–8, 2002 Proc. 6, Springer; 2002, p. 535–48.

[61] Papadimitriou S, Kitagawa H, Gibbons PB, Faloutsos C. Loci: Fast outlier detection using the local correlation integral. Proc. 19th Int. Conf. data Eng. (Cat. No. 03CH37405), IEEE; 2003, p. 315–26.

[62] Basu S, Bilenko M, Mooney RJ. A probabilistic framework for semi-supervised clustering. Proc. tenth ACM SIGKDD Int. Conf. Knowl. Discov. data Min., 2004, p. 59–68.

[63] Kaufman L, Rousseeuw PJ. Finding groups in data: an introduction to cluster analysis. John Wiley & Sons; 2009.

[64] Zahn CT. Graph-theoretical methods for detecting and describing gestalt clusters. IEEE Trans Comput 1971;100:68–86.

[65] Guha S, Rastogi R, Shim K. CURE: An efficient clustering algorithm for large databases. ACM Sigmod Rec 1998;27:73–84.

[66] Karypis G, Han E, Kumar V. A hierarchical clustering algorithm using dynamic modeling 1999.

[67] Ester M, Kriegel H-P, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. kdd, vol. 96, 1996, p. 226–31.

[68] Hinneburg A, Keim DA. An efficient approach to clustering in large multimedia databases with noise. vol. 98. Bibliothek der Universität Konstanz Konstanz, Germany; 1998.

[69] Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. J Comput Syst Sci 1997;55:119–39.

[70] Liu FT, Ting KM, Zhou Z-H. Isolation forest. 2008 eighth ieee Int. Conf. data Min., IEEE; 2008, p. 413–22.

[71] Zhao Y, Hryniewicki MK. Xgbod: improving supervised outlier detection with unsupervised representation learning. 2018 Int. Jt. Conf. Neural Networks, IEEE; 2018, p. 1–8.

[72] Zhao Y, Hryniewicki MK. DCSO: dynamic combination of detector scores for outlier ensembles. arXiv Prepr arXiv191110418 2019.

[73] Aggarwal CC. Outlier ensembles. Outlier Anal., Springer; 2016, p. 185–218.

[74] Das S, Wong W-K, Dietterich T, Fern A, Emmott A. Incorporating expert feedback into active anomaly discovery. 2016 IEEE 16th Int. Conf. Data Min., IEEE; 2016, p. 853–8.

[75] Pevný T. Loda: Lightweight on-line detector of anomalies. Mach Learn 2016;102:275–304.

[76] Veeramachaneni K, Arnaldo I, Korrapati V, Bassias C, Li K. AI^ 2: training a big data machine to defend. 2016 IEEE 2nd Int. Conf. big data Secur. cloud (BigDataSecurity), IEEE Int. Conf. high Perform. smart Comput. (HPSC), IEEE Int. Conf. Intell. data Secur., IEEE; 2016, p. 49–54.

[77] Zhang J, Jiang Y, Chang KH, Zhang S, Cai J, Hu L. A concept lattice based outlier mining method in low-dimensional subspaces. Pattern Recognit Lett 2009;30:1434–9.

[78] Dutta JK, Banerjee B, Reddy CK. RODS: Rarity based outlier detection in a sparse coding framework. IEEE Trans Knowl Data Eng 2015;28:483–95.

[79] Rubinstein R, Zibulevsky M, Elad M. Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit. Cs Tech 2008;40:1–15.

[80] Huang H, Mehrotra K, Mohan CK. Rank-based outlier detection. J Stat Comput Simul 2013;83:518–31.

[81] Müller E, Schiffer M, Seidl T. Statistical selection of relevant subspace projections for outlier ranking. 2011 IEEE 27th Int. Conf. data Eng., IEEE; 2011, p. 434–45.

[82] Kriegel H-P, Kröger P, Schubert E, Zimek A. Outlier detection in axis-parallel subspaces of high dimensional data. Adv. Knowl. Discov. Data Min. 13th Pacific-Asia Conf. PAKDD 2009 Bangkok, Thailand, April 27-30, 2009 Proc. 13, Springer; 2009, p. 831–8.

[83] Wang C, Gao H, Liu Z, Fu Y. A new outlier detection model using random walk on local information graph. IEEE Access 2018;6:75531–44.

[84] Gao J, Liang F, Fan W, Wang C, Sun Y, Han J. On community outliers and their efficient detection in information networks. Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discov. data Min., 2010, p. 813–22.

[85] Chalapathy R, Chawla S. Deep learning for anomaly detection: A survey. arXiv Prepr arXiv190103407 2019.

[86] Zong B, Song Q, Min MR, Cheng W, Lumezanu C, Cho Dほか. Deep autoencoding gaussian mixture model for unsupervised anomaly detection. Int. Conf. Learn. Represent., 2018.

[87] Chen J, Sathe S, Aggarwal C, Turaga D. Outlier detection with autoencoder ensembles. Proc. 2017 SIAM Int. Conf. data Min., SIAM; 2017, p. 90–8.

[88] Zhou C, Paffenroth RC. Anomaly detection with robust deep autoencoders. Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discov. data Min., 2017, p. 665–74.

[89] Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers. Found Trends® Mach Learn 2011;3:1–122.

[90] Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair Sほか. Generative adversarial nets. Adv Neural Inf Process Syst 2014;27.

[91] Schlegl T, Seeböck P, Waldstein SM, Schmidt-Erfurth U, Langs G. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) 2017;10265 LNCS:146–7. https://doi.org/10.1007/978-3-319-59050-9_12.

[92] Schlegl T, Seeböck P, Waldstein SM, Langs G, Schmidt-Erfurth U. f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks [J]. Med Image Anal 2019;54:30–44.

[93] Pang G, Cao L, Chen L, Liu H. Learning representations of ultrahigh-dimensional data for random distance-based outlier detection. Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discov. data Min., 2018, p. 2041–50.

[94] Schärlig A. Décider sur plusieurs critères: panorama de l'aide à la décision multicritère. vol. 1. EPFL Press; 1985.

[95] Roy B. A French-English decision aiding glossary. Newsl EWG MCDA (Supplement 1), Vol Ser 2000;3.

[96] Hajkowicz S, Collins K. A review of multiple criteria analysis for water resource planning and management. Water Resour Manag 2007;21:1553–66.

[97] Huang IB, Keisler J, Linkov I. Multi-criteria decision analysis in environmental sciences: Ten years of applications and trends. Sci Total Environ 2011;409:3578–94.

[98] Thokala P, Devlin N, Marsh K, Baltussen R, Boysen M, Kalo Zほか. Multiple criteria decision analysis for health care decision making—an introduction: report 1 of the ISPOR MCDA Emerging Good Practices Task Force. Value Heal 2016;19:1–13.

[99] Marsh K, IJzerman M, Thokala P, Baltussen R, Boysen M, Kaló Zほか. Multiple criteria decision analysis for health care decision making—emerging good practices: report 2 of the ISPOR MCDA Emerging Good Practices Task Force. Value Heal 2016;19:125–37.

[100] Tsoutsos T, Drandaki M, Frantzeskaki N, Iosifidis E, Kiosses I. Sustainable energy planning by using multi-criteria analysis application in the island of Crete. Energy Policy 2009;37:1587–600.

[101] Govindan K, Rajendran S, Sarkis J, Murugesan P. Multi criteria decision making approaches for green supplier evaluation and selection: a literature review. J Clean Prod 2015;98:66–83.

[102] AF Ferreira F, P. Santos S, SE Marques C, Ferreira J. Assessing credit risk of mortgage lending using MACBETH: a methodological framework. Manag Decis 2014;52:182–206.

[103] Angilella S, Mazzù S. The financing of innovative SMEs: A multicriteria credit rating model. Eur J Oper Res 2015;244:540–54.

[104] Ehrgott M, Klamroth K, Schwehm C. An MCDM approach to portfolio optimization. Eur J Oper Res 2004;155:752–70.

[105] Aouni B, Doumpos M, Pérez-Gladish B, Steuer RE. On the increasing importance of multiple criteria decision aid methods for portfolio selection. J Oper Res Soc 2018;69:1525–42.

[106] Bai C, Dhavale D, Sarkis J. Integrating Fuzzy C-Means and TOPSIS for performance evaluation: An application and comparative analysis. Expert Syst Appl 2014;41:4186–96.

[107] Lowe TJ, Wendell RE, Hu G. Screening location strategies to reduce exchange rate risk. Eur J Oper Res 2002;136:573–90.

[108] Kabašinskas A, Šutienė K, Kopa M, Lukšys K, Bagdonas K. Dominance-based decision rules for pension fund selection under different distributional assumptions. Mathematics 2020;8:719.

[109] Geoffrion AM, Dyer JS, Feinberg A. An interactive approach for multi-criterion optimization, with an application to the operation of an academic department. Manage Sci 1972;19:357–68.

[110] Zionts S, Wallenius J. An interactive programming method for solving the multiple criteria problem. Manage Sci 1976;22:652–63.

[111] Keeney RL, Raiffa H. Decisions with multiple objectives. 1976. John Wiley&Sons, New York 1976:34–8.

[112] Saaty TL. How to make a decision: the analytic hierarchy process. Eur J Oper Res 1990;48:9–26.

[113] Jacquet-Lagreze E, Siskos J. Assessing a set of additive utility functions for multicriteria decision-making, the UTA method. Eur J Oper Res 1982;10:151–64.

[114] Roy B. The outranking approach and the foundations of ELECTRE methods. Theory Decis 1991;31:49–73.

[115] Roy B, Bouyssou D. Aide multicritère à la décision: méthodes et cas. vol. 695. Economica Paris; 1993.

[116] Mareschal B, Brans JP, Vincke P. PROMETHEE: A new family of outranking methods in multicriteria analysis. ULB--Universite Libre de Bruxelles; 1984.

[117] Brans J-P, De Smet Y. PROMETHEE methods. Mult criteria Decis Anal state art Surv 2016:187–219.

[118] Figueira JJR, De Smet Y, Brans JP. MCDA methods for sorting and clustering problems: Promethee TRI and Promethee CLUSTER 2005.

[119] De Smet Y, Guzmán LM. Towards multicriteria clustering: An extension of the k-means algorithm. Eur J Oper Res 2004;158:390–8. https://doi.org/10.1016/j.ejor.2003.06.012.

[120] Mareschal B, Brans J-P. Geometrical representations for MCDA. Eur J Oper Res 1988;34:69–77. https://doi.org/https://doi.org/10.1016/0377-2217(88)90456-0.

[121] Dejaegere G, De Smet Y. Promethee γ: A new Promethee based method for partial ranking based on valued coalitions of monocriterion net flow scores. J Multi-Criteria Decis Anal 2023;30:147–60.

[122] Achir T, Rouba B. A Novel Weighted Preference Relation Approach to Detect Outliers in Multi-Criteria Decision Aid Context 日付なし.

[123] Barnett V, Lewis T. Outliers in statistical data. vol. 3. Wiley New York; 1994.

[124] Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika 1965;52:591–611.

[125] World Happiness Report 2015. https://www.kaggle.com/datasets/mathurinache/world-happiness-report?resource=download&select=2015.csv.

[126] De Smet Y, Nemery P, Selvaraj R. An exact algorithm for the multicriteria ordered clustering problem. Omega 2012;40:861–9.

APPENDICES

Appendix A. Ranking of 158 alternatives

Rank	Alternative	ϕ	ϕ^+	ϕ^-	Rank	Alternative	ϕ	ϕ^+	ϕ^-
1	New Zealand	0.7116	0.7208	0.0092	80	China	-0.072	0.2329	0.3049
2	Australia	0.7107	0.7219	0.0112	81	Myanmar	-0.0735	0.3367	0.4102
3	Canada	0.7003	0.7139	0.0135	82	Kyrgyzstan	-0.0743	0.2264	0.3007
4	Ireland	0.693	0.7119	0.0189	83	Bolivia	-0.0899	0.2175	0.3075
5	Norway	0.682	0.6998	0.0178	84	Lebanon	-0.0954	0.2114	0.3069
6	Netherlands	0.6816	0.6973	0.0157	85	Turkey	-0.097	0.2264	0.3235
7	Sweden	0.679	0.6922	0.0132	86	Azerbaijan	-0.0992	0.2264	0.3256
8	United Kingdom	0.6784	0.695	0.0165	87	Macedonia	-0.0994	0.1919	0.2913
9	Denmark	0.6719	0.6879	0.016	88	Russia	-0.1	0.2418	0.3418
10	Switzerland	0.6682	0.6886	0.0205	89	Peru	-0.1056	0.1867	0.2923
11	Luxembourg	0.6158	0.6429	0.0271	90	Algeria	-0.1083	0.2336	0.3419
12	Hong Kong	0.6148	0.6587	0.0439	91	Syria	-0.1124	0.3139	0.4263
13	Iceland	0.6029	0.6406	0.0377	92	Bulgaria	-0.1218	0.2039	0.3257
14	Qatar	0.5816	0.6361	0.0545	93	El Salvador	-0.1231	0.1913	0.3144
15	Finland	0.5711	0.6075	0.0363	94	Iran	-0.1285	0.2389	0.3674
16	Singapore	0.5615	0.6258	0.0644	95	Botswana	-0.1295	0.2087	0.3382
17	Austria	0.5448	0.5873	0.0425	96	Georgia	-0.1322	0.2777	0.41
18	Malta	0.544	0.5878	0.0438	97	Lithuania	-0.143	0.2405	0.3835
19	Germany	0.531	0.5767	0.0457	98	Montenegro	-0.1453	0.2	0.3452
20	United States	0.5108	0.5587	0.0479	99	Honduras	-0.1612	0.1679	0.3291
21	United Arab Emirates	0.5097	0.5657	0.056	100	Albania	-0.1631	0.1775	0.3405
22	Belgium	0.4775	0.5367	0.0592	101	Tajikistan	-0.1661	0.1951	0.3612
23	Oman	0.4103	0.5022	0.0919	102	Romania	-0.167	0.1834	0.3504
24	France	0.3961	0.498	0.1019	103	Ukraine	-0.1706	0.1855	0.3562
25	Uruguay	0.3915	0.476	0.0845	104	Serbia	-0.1887	0.1729	0.3615
26	Kuwait	0.3704	0.484	0.1136	105	Kenya	-0.1951	0.2007	0.3958
27	Japan	0.3422	0.4676	0.1254	106	Uganda	-0.1962	0.1791	0.3753
28	Bahrain	0.3262	0.4457	0.1195	107	South Africa	-0.2055	0.1666	0.3721
29	Slovenia	0.3076	0.4221	0.1145	108	Djibouti	-0.2323	0.2213	0.4537
30	Costa Rica	0.3016	0.4083	0.1067	109	Bosnia and Herzegovina	-0.2401	0.1801	0.4202
31	Thailand	0.3012	0.4533	0.1522	110	Greece	-0.2432	0.2176	0.4608
32	Israel	0.2891	0.4096	0.1205	111	Nepal	-0.2457	0.1661	0.4118
33	North Cyprus	0.2861	0.3971	0.1109	112	Zambia	-0.2462	0.1514	0.3976
34	Uzbekistan	0.2679	0.4645	0.1966	113	Kosovo	-0.2502	0.1616	0.4118
35	Chile	0.2678	0.3867	0.1189	114	Haiti	-0.2525	0.243	0.4955
36	Spain	0.2478	0.3872	0.1394	115	Iraq	-0.2548	0.1737	0.4285
37	Panama	0.222	0.343	0.121	116	Croatia	-0.2582	0.172	0.4302
38	Malaysia	0.2128	0.3514	0.1386	117	Moldova	-0.2633	0.1422	0.4055
39	Bhutan	0.1972	0.4024	0.2053	118	Gabon	-0.2665	0.1552	0.4217
40	Nicaragua	0.1823	0.3653	0.183	119	Palestinian Territories	-0.2786	0.1455	0.4241
41	Saudi Arabia	0.1724	0.3865	0.2142	120	Ivory Coast	-0.2788	0.1702	0.449
42	Paraguay	0.1705	0.3581	0.1876	121	Tanzania	-0.2866	0.1578	0.4443
43	Turkmenistan	0.1679	0.376	0.2081	122	Senegal	-0.3	0.1216	0.4216
44	Trinidad and Tobago	0.1668	0.3613	0.1944	123	Bangladesh	-0.3155	0.1355	0.451

45	Taiwan	0.1651	0.3274	0.1622	124	India	-0.3171	0.1412	0.4584
46	Mauritius	0.1632	0.3442	0.181	125	Mali	-0.3182	0.1183	0.4365
47	Poland	0.1627	0.3276	0.1649	126	Morocco	-0.3204	0.1348	0.4552
48	Dominican Republic	0.1597	0.3276	0.168	127	Niger	-0.3211	0.1471	0.4683
49	Brazil	0.1586	0.3356	0.177	128	Ethiopia	-0.3228	0.152	0.4749
50	Estonia	0.1534	0.3368	0.1834	129	Cameroon	-0.3279	0.115	0.4429
51	Sri Lanka	0.1498	0.3429	0.1931	130	Mozambique	-0.333	0.1365	0.4695
52	Mexico	0.1226	0.3325	0.2099	131	Mauritania	-0.3337	0.1468	0.4805
53	Cyprus	0.1211	0.3282	0.2071	132	Tunisia	-0.3429	0.1309	0.4738
54	Portugal	0.1192	0.3191	0.1999	133	Burkina Faso	-0.349	0.113	0.462
55	Belarus	0.1145	0.3273	0.2128	134	Sudan	-0.3494	0.1352	0.4846
56	Italy	0.1055	0.333	0.2275	135	Nigeria	-0.3501	0.1187	0.4688
57	Czech Republic	0.1	0.3097	0.2097	136	Swaziland	-0.3613	0.1078	0.4691
58	Laos	0.0975	0.4101	0.3126	137	Congo (Brazzaville)	-0.3663	0.1157	0.4821
59	Argentina	0.0941	0.2889	0.1949	138	Armenia	-0.3699	0.1213	0.4912
60	Colombia	0.0767	0.2808	0.2041	139	Egypt	-0.3702	0.1158	0.486
61	Libya	0.064	0.2592	0.1952	140	Lesotho	-0.3703	0.1032	0.4735
62	Indonesia	0.0629	0.3385	0.2756	141	Sierra Leone	-0.3725	0.0959	0.4684
63	Ecuador	0.0621	0.2952	0.2331	142	Ghana	-0.3737	0.1102	0.484
64	South Korea	0.0583	0.2915	0.2333	143	Pakistan	-0.3797	0.1577	0.5374
65	Somaliland region	0.0563	0.3936	0.3374	144	Comoros	-0.3966	0.1378	0.5343
66	Suriname	0.0543	0.2843	0.23	145	Zimbabwe	-0.4024	0.0838	0.4862
67	Mongolia	0.0479	0.3086	0.2607	146	Guinea	-0.42	0.112	0.532
68	Guatemala	0.0472	0.2756	0.2284	147	Malawi	-0.4238	0.1245	0.5482
69	Kazakhstan	0.0423	0.2617	0.2194	148	Benin	-0.4347	0.0871	0.5217
70	Slovakia	0.0412	0.2777	0.2365	149	Afghanistan	-0.444	0.1289	0.5729
71	Venezuela	0.0403	0.2794	0.2391	150	Yemen	-0.4635	0.0786	0.5421
72	Philippines	0.0385	0.2884	0.2499	151	Congo (Kinshasa)	-0.4674	0.0783	0.5457
73	Jamaica	-0.0088	0.2427	0.2515	152	Central African Republic	-0.4742	0.0891	0.5633
74	Vietnam	-0.0306	0.2402	0.2707	153	Liberia	-0.4806	0.0699	0.5504
75	Hungary	-0.0311	0.2416	0.2727	154	Angola	-0.4956	0.0702	0.5658
76	Jordan	-0.0377	0.2326	0.2704	155	Madagascar	-0.5083	0.0671	0.5754
77	Latvia	-0.0503	0.2138	0.2642	156	Togo	-0.5191	0.0581	0.5772
78	Rwanda	-0.0579	0.3203	0.3783	157	Chad	-0.5581	0.0383	0.5964
79	Cambodia	-0.0604	0.305	0.3654	158	Burundi	-0.6278	0.0366	0.6644

To simplify, we present the alternatives by their rank in the following appendices.

Appendix B. The weighted preference matrix of the top 20 alternatives and the lowest-ranked alternative.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	158
1	I	0.072P$^+$	0.060P$^+$	0.088P$^+$	0.089R	0.094P$^+$	0.079P$^+$	0.094P$^+$	0.083P$^+$	0.084R	0.116P$^+$	0.129R	0.112P$^+$	0.154R	0.116P$^+$	0.177R	0.180P$^+$	0.184P$^+$	0.186P$^+$	0.237P$^+$	0.654P$^+$
2	0.072P$^-$	I	0.061P$^+$	0.076P$^+$	0.077R	0.080P$^+$	0.065P$^+$	0.096P$^+$	0.085R	0.073R	0.105P$^+$	0.110R	0.101R	0.143R	0.099P$^+$	0.166R	0.163P$^+$	0.177P$^+$	0.177P$^+$	0.222P$^+$	0.639P$^+$
3	0.060P$^-$	0.061P$^-$	I	0.062R	0.105R	0.070P$^+$	0.080P$^-$	0.085R	0.096P$^-$	0.091P$^-$	0.100R	0.116R	0.098R	0.155R	0.094P$^+$	0.176R	0.154P$^+$	0.167P$^+$	0.165P$^+$	0.211P$^+$	0.627P$^+$
4	0.088P$^-$	0.076P$^-$	0.062R	I	0.101P$^-$	0.070R	0.092R	0.087R	0.101P$^-$	0.099P$^-$	0.093R	0.096R	0.088R	0.141R	0.086P$^+$	0.157R	0.146P$^+$	0.148P$^+$	0.159P$^+$	0.189P$^+$	0.605P$^+$
5	0.089R	0.077R	0.105R	0.101P$^-$	I	0.109P$^+$	0.075R	0.108P$^+$	0.091R	0.082P$^-$	0.103P$^+$	0.100P$^+$	0.100P$^+$	0.124R	0.107P$^+$	0.144R	0.168P$^+$	0.164P$^+$	0.176P$^+$	0.216P$^+$	0.625P$^+$
6	0.094P$^-$	0.080P$^-$	0.070P$^-$	0.070R	0.109P$^-$	I	0.094P$^-$	0.072R	0.105P$^-$	0.112P$^-$	0.099R	0.113R	0.082R	0.154R	0.084P$^+$	0.176R	0.128P$^+$	0.148P$^+$	0.138P$^+$	0.189P$^+$	0.609P$^+$
7	0.079P$^-$	0.065P$^-$	0.080P$^+$	0.092R	0.075R	0.094P$^+$	I	0.113R	0.077R	0.080R	0.098R	0.115R	0.100R	0.144R	0.099P$^+$	0.166R	0.152P$^+$	0.165P$^+$	0.164P$^+$	0.214P$^+$	0.632P$^+$
8	0.094P$^-$	0.096P$^-$	0.085R	0.087R	0.108P$^-$	0.072R	0.113R	I	0.119P$^-$	0.121P$^-$	0.097P$^+$	0.106R	0.084R	0.146R	0.100P$^+$	0.163R	0.149P$^+$	0.148P$^+$	0.146P$^+$	0.193P$^+$	0.600P$^+$
9	0.083R	0.085R	0.096P$^+$	0.101P$^+$	0.091R	0.105P$^+$	0.077R	0.119P$^+$	I	0.063P$^-$	0.096P$^+$	0.105P$^+$	0.107P$^+$	0.114R	0.102P$^+$	0.145R	0.164P$^+$	0.167P$^+$	0.177P$^+$	0.215P$^+$	0.631P$^+$
10	0.084R	0.073R	0.091P$^+$	0.099P$^+$	0.082P$^+$	0.112P$^+$	0.080R	0.121P$^+$	0.063P$^+$	I	0.106P$^+$	0.108P$^+$	0.116P$^+$	0.121R	0.111P$^+$	0.147R	0.180P$^+$	0.181P$^+$	0.191P$^+$	0.225P$^+$	0.636P$^+$
11	0.116P$^-$	0.105P$^-$	0.100R	0.093R	0.103P$^-$	0.099R	0.098R	0.097P$^-$	0.096P$^-$	0.106P$^-$	I	0.054R	0.084P$^-$	0.100R	0.076R	0.121R	0.125P$^+$	0.120P$^+$	0.127P$^+$	0.165P$^+$	0.574P$^+$
12	0.129R	0.110R	0.116R	0.096R	0.100P$^-$	0.113R	0.115R	0.106R	0.105P$^-$	0.108P$^-$	0.054R	I	0.089R	0.084R	0.091R	0.106R	0.136P$^+$	0.119P$^+$	0.138P$^+$	0.160P$^+$	0.567P$^+$
13	0.112P$^-$	0.101R	0.098R	0.088R	0.100P$^-$	0.082R	0.100R	0.084R	0.107P$^-$	0.116P$^-$	0.084P$^+$	0.089R	I	0.114R	0.081R	0.124R	0.121P$^+$	0.119P$^+$	0.128P$^+$	0.161P$^+$	0.575P$^+$
14	0.154R	0.143R	0.155R	0.141R	0.124R	0.154R	0.144R	0.146R	0.114R	0.121R	0.100R	0.084R	0.114R	I	0.121R	0.072R	0.168P$^+$	0.152P$^+$	0.177P$^+$	0.182P$^+$	0.564P$^+$
15	0.116P$^-$	0.099P$^-$	0.094P$^-$	0.086P$^-$	0.107P$^-$	0.084P$^-$	0.099P$^-$	0.100P$^-$	0.102P$^-$	0.111P$^-$	0.076P$^-$	0.091R	0.081R	0.121R	I	0.147R	0.105P$^+$	0.121P$^+$	0.121P$^+$	0.160P$^+$	0.574P$^+$
16	0.177R	0.166R	0.176R	0.157R	0.144R	0.176R	0.166R	0.163R	0.145R	0.147R	0.121R	0.106R	0.124R	0.072R	0.147R	I	0.189R	0.165P$^+$	0.197R	0.190P$^+$	0.550P$^+$
17	0.180P$^-$	0.163P$^-$	0.154P$^-$	0.146P$^-$	0.168P$^-$	0.128P$^-$	0.152P$^-$	0.149P$^-$	0.164P$^-$	0.180P$^-$	0.125P$^-$	0.136P$^-$	0.121P$^-$	0.168P$^-$	0.105P$^-$	0.189R	I	0.096R	0.066R	0.102R	0.530P$^+$
18	0.184P$^-$	0.177P$^-$	0.167P$^-$	0.148P$^-$	0.164P$^-$	0.148P$^-$	0.165P$^-$	0.148P$^-$	0.167P$^-$	0.181P$^-$	0.120P$^-$	0.119P$^-$	0.119P$^-$	0.152P$^-$	0.121P$^-$	0.165P$^-$	0.096R	I	0.087R	0.0109P$^+$	0.517P$^+$
19	0.186P$^-$	0.177P$^-$	0.165P$^-$	0.159P$^-$	0.176P$^-$	0.138P$^-$	0.164P$^-$	0.146P$^-$	0.177P$^-$	0.191P$^-$	0.127P$^-$	0.138P$^-$	0.128P$^-$	0.177P$^-$	0.121P$^-$	0.197R	0.066R	0.087R	I	0.106R	0.523P$^+$
20	0.237P$^-$	0.222P$^-$	0.211P$^-$	0.189P$^-$	0.216P$^-$	0.189P$^-$	0.214P$^-$	0.193P$^-$	0.215P$^-$	0.225P$^-$	0.165P$^-$	0.160P$^-$	0.161P$^-$	0.182P$^-$	0.160P$^-$	0.190P$^-$	0.102R	0.109P$^-$	0.106R	I	0.462P$^+$
158	0.654P$^-$	0.639P$^-$	0.627P$^-$	0.605P$^-$	0.625P$^-$	0.609P$^-$	0.632P$^-$	0.600P$^-$	0.631P$^-$	0.636P$^-$	0.574P$^-$	0.567P$^-$	0.575P$^-$	0.564P$^-$	0.574P$^-$	0.550P$^-$	0.530P$^-$	0.517P$^-$	0.523P$^-$	0.462P$^-$	I

Appendix C. The weighted preference matrix of the 20 alternatives ranked in the middle, along with both the first and last ranked alternatives.

	1	**70**	**71**	**72**	**73**	**74**	**75**	**76**	**77**	**78**	**79**	**80**	**81**	**82**	**83**	**84**	**85**	**86**	**87**	**88**	**89**	**158**
1	*I*	$0.384P^+$	$0.382P^+$	$0.402P^+$	$0.429P^+$	$0.435P^+$	$0.448P^+$	$0.442P^+$	$0.439P^+$	$0.453P^+$	$0.448P^+$	$0.449P^+$	$0.446P^+$	$0.452P^+$	$0.465P^+$	$0.458P^+$	$0.4454P$	$0.467P^+$	$0.479P^+$	$0.470P^+$	$0.482P^+$	$0.712P^+$
70	$0.384P^-$	*I*	$0.097R$	$0.166P^+$	$0.119P^+$	$0.160P^+$	$0.134P^+$	$0.143P^+$	$0.126P^+$	$0.220R$	$0.204P^+$	$0.140P^+$	$0.218R$	$0.165P^+$	$0.167P^+$	$0.174P^+$	$0.168P^+$	$0.172P^+$	$0.189P^+$	$0.196P^+$	$0.179P^+$	$0.421P^+$
71	$0.382P^-$	$0.097R$	*I*	$0.124P^+$	$0.099P^+$	$0.119P^+$	$0.120P^+$	$0.099P^+$	$0.100P^+$	$0.216R$	$0.197P^+$	$0.108P^+$	$0.212R$	$0.138P^+$	$0.148P^+$	$0.142P^+$	$0.131P^+$	$0.142P^+$	$0.151P^+$	$0.163P^+$	$0.146P^+$	$0.426P^+$
72	$0.402P^-$	$0.166P^-$	$0.124P^-$	*I*	$0.107P^+$	$0.078P^+$	$0.164P^+$	$0.124R$	$0.131R$	$0.172R$	$0.145R$	$0.118P^+$	$0.173R$	$0.114P^+$	$0.128P^+$	$0.135P^+$	$0.153P^+$	$0.139P^+$	$0.155P^+$	$0.155P^+$	$0.161P^+$	$0.397P^+$
73	$0.429P^-$	$0.119P^-$	$0.099P^-$	$0.107P^-$	*I*	$0.104R$	$0.103P^+$	$0.100R$	$0.091R$	$0.197R$	$0.169R$	$0.101P^+$	$0.193R$	$0.114P^+$	$0.130P^+$	$0.140P^+$	$0.145P^+$	$0.135P^+$	$0.155P^+$	$0.158P^+$	$0.148P^+$	$0.383P^+$
74	$0.435P^-$	$0.160P^-$	$0.119P^-$	$0.078P^-$	$0.104R$	*I*	$0.159R$	$0.112R$	$0.129R$	$0.165R$	$0.145R$	$0.116P^+$	$0.170R$	$0.108P^+$	$0.100P^+$	$0.130P^+$	$0.135P^+$	$0.122P^+$	$0.136P^+$	$0.156P^+$	$0.146P^+$	$0.378P^+$
75	$0.448P^-$	$0.134P^-$	$0.120P^-$	$0.164P^-$	$0.103P^-$	$0.159R$	*I*	$0.099P^-$	$0.073P^-$	$0.234R$	$0.210R$	$0.103R$	$0.219R$	$0.145R$	$0.166P^+$	$0.147P^+$	$0.145R$	$0.146P^+$	$0.153P^+$	$0.147R$	$0.153P^+$	$0.377P^+$
76	$0.442P^-$	$0.143P^-$	$0.099P^-$	$0.124R$	$0.100R$	$0.112R$	$0.099P^+$	*I*	$0.084P^-$	$0.211R$	$0.190R$	$0.090P^+$	$0.203R$	$0.113R$	$0.123P^+$	$0.120P^+$	$0.118R$	$0.112P^+$	$0.122P^+$	$0.138P^+$	$0.123P^+$	$0.390P^+$
77	$0.439P^-$	$0.126P^-$	$0.100P^-$	$0.131R$	$0.091R$	$0.129R$	$0.073P^+$	$0.084P^+$	*I*	$0.223R$	$0.201R$	$0.079P^+$	$0.215R$	$0.132R$	$0.151P^+$	$0.131P^+$	$0.132R$	$0.130P^+$	$0.138P^+$	$0.141P^+$	$0.136P^+$	$0.389P^+$
78	$0.453P^-$	$0.220R$	$0.216R$	$0.172R$	$0.197R$	$0.165R$	$0.234R$	$0.211R$	$0.223R$	*I*	$0.084R$	$0.190R$	$0.081R$	$0.153R$	$0.155R$	$0.178R$	$0.190R$	$0.175R$	$0.211R$	$0.189R$	$0.214R$	$0.300P^+$
79	$0.448P^-$	$0.204P^-$	$0.197P^-$	$0.145R$	$0.169R$	$0.145R$	$0.210R$	$0.190R$	$0.201R$	$0.084R$	*I*	$0.168R$	$0.092R$	$0.134R$	$0.137R$	$0.172R$	$0.181R$	$0.167R$	$0.200R$	$0.182R$	$0.199R$	$0.313P^+$
80	$0.449P^-$	$0.140P^-$	$0.108P^-$	$0.118P^-$	$0.101P^-$	$0.116P^-$	$0.103R$	$0.090P^-$	$0.079P^-$	$0.190R$	$0.168R$	*I*	$0.186R$	$0.097R$	$0.121P^+$	$0.094P^+$	$0.092R$	$0.095P^+$	$0.105R$	$0.111P^+$	$0.105R$	$0.357P^+$
81	$0.446P^-$	$0.218R$	$0.212R$	$0.173R$	$0.193R$	$0.170R$	$0.219R$	$0.203R$	$0.215R$	$0.081R$	$0.092R$	$0.186R$	*I*	$0.156R$	$0.163R$	$0.185R$	$0.189R$	$0.173R$	$0.218R$	$0.187R$	$0.214R$	$0.304P^+$
82	$0.452P^-$	$0.165P^-$	$0.138P^-$	$0.114P^-$	$0.114P^-$	$0.108P^-$	$0.145R$	$0.113R$	$0.132R$	$0.153R$	$0.134R$	$0.097R$	$0.156R$	*I*	$0.099R$	$0.095R$	$0.121P^-$	$0.111R$	$0.109R$	$0.108P^+$	$0.116R$	$0.338P^+$
83	$0.465P^-$	$0.167P^-$	$0.148P^-$	$0.128P^-$	$0.130P^-$	$0.100P^-$	$0.166P^-$	$0.123P^-$	$0.151P^-$	$0.155R$	$0.137R$	$0.121P^-$	$0.163R$	$0.099R$	*I*	$0.116P^-$	$0.123P^-$	$0.113R$	$0.128R$	$0.157R$	$0.125R$	$0.350P^+$
84	$0.458P^-$	$0.174P^-$	$0.142P^-$	$0.135P^-$	$0.140P^-$	$0.130P^-$	$0.147P^-$	$0.120P^-$	$0.131P^-$	$0.178R$	$0.172R$	$0.094P^-$	$0.185R$	$0.095R$	$0.116P^+$	*I*	$0.080P^-$	$0.080P^-$	$0.069R$	$0.086R$	$0.087R$	$0.349P^+$
85	$0.454P^-$	$0.168P^-$	$0.131P^-$	$0.153P^-$	$0.145P^-$	$0.135P^-$	$0.145R$	$0.118R$	$0.132R$	$0.190R$	$0.181R$	$0.092R$	$0.189R$	$0.121P^+$	$0.123P^+$	$0.080P^+$	*I*	$0.070P^+$	$0.079R$	$0.106P^+$	$0.084R$	$0.360P^+$
86	$0.467P^-$	$0.172P^-$	$0.142P^-$	$0.139P^-$	$0.135P^-$	$0.122P^-$	$0.146P^-$	$0.112P^-$	$0.130P^-$	$0.175R$	$0.167R$	$0.095P^-$	$0.173R$	$0.111R$	$0.113R$	$0.080P^-$	$0.070P^-$	*I*	$0.093R$	$0.104R$	$0.090R$	$0.348P^+$
87	$0.479P^-$	$0.189P^-$	$0.151P^-$	$0.155P^-$	$0.155P^-$	$0.136P^-$	$0.153P^-$	$0.122P^-$	$0.138P^-$	$0.211R$	$0.200R$	$0.105R$	$0.218R$	$0.109R$	$0.128R$	$0.069R$	$0.079R$	$0.093R$	*I*	$0.096R$	$0.069P^+$	$0.366P^+$
88	$0.470P^-$	$0.196P^-$	$0.163P^-$	$0.155P^-$	$0.158P^-$	$0.156P^-$	$0.147R$	$0.138P^-$	$0.141P^-$	$0.189R$	$0.182R$	$0.111P^-$	$0.187R$	$0.108P^-$	$0.157R$	$0.086R$	$0.106P^-$	$0.104R$	$0.096R$	*I*	$0.105R$	$0.333P^+$
89	$0.482P^-$	$0.179P^-$	$0.146P^-$	$0.161P^-$	$0.148P^-$	$0.146P^-$	$0.153P^-$	$0.123P^-$	$0.136P^-$	$0.214R$	$0.199R$	$0.105R$	$0.214R$	$0.116R$	$0.125R$	$0.087R$	$0.084R$	$0.090R$	$0.069P^-$	$0.105R$	*I*	$0.364P^+$
158	$0.712P^-$	$0.421P^-$	$0.426P^-$	$0.397P^-$	$0.383P^-$	$0.378P^-$	$0.377P^-$	$0.390P^-$	$0.389P^-$	$0.300P^-$	$0.313P^-$	$0.357P^-$	$0.304P^-$	$0.338P^-$	$0.350P^-$	$0.349P^-$	$0.360P^-$	$0.348P^-$	$0.366P^-$	$0.333P^-$	$0.364P^-$	*I*

Appendix D. The weighted preference matrix of the 20 alternatives with the lowest ranks along with the first ranked alternative.

	1	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158
1	*I*	0.485P^+	0.542P^+	0.569P^+	0.499P^+	0.510P^+	0.521P^+	0.553P^+	0.567P^+	0.594P^+	0.572P^+	0.587P^+	0.535P^+	0.618P^+	0.651P^+	0.578P^+	0.585P^+	0.583P^+	0.625P^+	0.622P^+	0.701P^+
139	0.485P^-	*I*	0.130*R*	0.139P^+	0.095*R*	0.113*R*	0.109*R*	0.150*R*	0.157P^+	0.187P^+	0.169*R*	0.178P^+	0.121*R*	0.175P^+	0.229P^+	0.148*R*	0.146P^+	0.151*R*	0.199P^+	0.181P^+	0.266P^+
140	0.542P^-	0.130*R*	*I*	0.071P^+	0.118P^-	0.104*R*	0.103P^-	0.089*R*	0.093P^+	0.122P^+	0.112P^+	0.107P^+	0.107P^-	0.125P^+	0.179P^+	0.101P^+	0.112P^+	0.099P^+	0.132P^+	0.124P^+	0.202P^+
141	0.569P^-	0.139P^-	0.071P^-	*I*	0.124P^-	0.124P^-	0.109P^-	0.084P^-	0.090*R*	0.111P^+	0.109*R*	0.099P^+	0.116P^-	0.105P^+	0.162P^+	0.097*R*	0.112*R*	0.092*R*	0.115P^+	0.107P^+	0.179P^+
142	0.499P^-	0.095*R*	0.118P^+	0.124P^+	*I*	0.091P^+	0.084P^+	0.113*R*	0.118P^+	0.158P^+	0.130P^+	0.141P^+	0.085P^+	0.165P^+	0.227P^+	0.120P^+	0.158P^+	0.133P^+	0.180P^+	0.169P^+	0.258P^+
143	0.510P^-	0.113*R*	0.104*R*	0.124P^+	0.091P^-	*I*	0.094*R*	0.118*R*	0.100P^+	0.130P^+	0.104*R*	0.109P^+	0.089*R*	0.159P^+	0.196P^+	0.116*R*	0.137P^+	0.128*R*	0.155P^+	0.152P^+	0.231P^+
144	0.521P^-	0.109*R*	0.103P^+	0.109P^+	0.084P^-	0.094*R*	*I*	0.073*R*	0.123P^+	0.160P^+	0.116P^+	0.127P^+	0.094P^+	0.156P^+	0.225P^+	0.116P^+	0.152P^+	0.121P^+	0.170P^+	0.151P^+	0.249P^+
145	0.553P^-	0.150*R*	0.089*R*	0.084P^+	0.113*R*	0.118*R*	0.073*R*	*I*	0.105P^+	0.134P^+	0.100P^+	0.105P^+	0.109*R*	0.137P^+	0.202P^+	0.105P^+	0.153*R*	0.110P^+	0.145P^+	0.136P^+	0.220P^+
146	0.567P^-	0.157P^-	0.093P^-	0.090*R*	0.118P^-	0.100P^-	0.123P^-	0.105P^-	*I*	0.070P^+	0.070*R*	0.075P^+	0.104P^-	0.107P^+	0.143P^+	0.099*R*	0.130*R*	0.110*R*	0.106P^+	0.119P^+	0.181P^+
147	0.594P^-	0.187P^-	0.122P^-	0.111P^-	0.158P^-	0.130P^-	0.160P^-	0.134P^-	0.070P^-	*I*	0.095P^-	0.088P^-	0.136P^-	0.109P^+	0.116P^+	0.111P^-	0.132*R*	0.123P^-	0.101P^+	0.128*R*	0.154P^+
148	0.572P^-	0.169*R*	0.112P^-	0.109*R*	0.130P^-	0.104*R*	0.116P^-	0.100P^-	0.070*R*	0.095P^+	*I*	0.062P^+	0.112P^-	0.118P^+	0.159P^+	0.106P^+	0.144*R*	0.118P^+	0.117P^+	0.120P^+	0.188P^+
149	0.587P^-	0.178P^-	0.107P^-	0.099P^-	0.141P^-	0.109P^-	0.127P^-	0.105P^-	0.075P^-	0.088P^+	0.062P^-	*I*	0.122P^-	0.121P^+	0.149P^+	0.103P^-	0.132*R*	0.115P^-	0.101P^+	0.114P^+	0.168P^+
150	0.535P^-	0.121*R*	0.107P^+	0.116P^+	0.085P^-	0.089*R*	0.094P^-	0.109*R*	0.104P^+	0.136P^+	0.112P^+	0.122P^+	*I*	0.143P^+	0.200P^+	0.099P^+	0.123P^+	0.099P^+	0.156P^+	0.134P^+	0.122P^+
151	0.618P^-	0.175P^-	0.125P^-	0.105P^-	0.165P^-	0.159P^-	0.156P^-	0.137P^-	0.107P^-	0.109P^-	0.118P^-	0.121P^-	0.143P^-	*I*	0.105P^+	0.094P^-	0.106*R*	0.124P^-	0.105P^+	0.082*R*	0.139P^+
152	0.651P^-	0.229P^-	0.179P^-	0.162P^-	0.227P^-	0.196P^-	0.225P^-	0.202P^-	0.143P^-	0.116P^-	0.159P^-	0.149P^-	0.200P^-	0.105P^-	*I*	0.156P^-	0.138P^-	0.183P^-	0.120P^-	0.138P^-	0.108*R*
153	0.578P^-	0.148*R*	0.101P^-	0.097*R*	0.120P^-	0.116*R*	0.116P^-	0.105P^-	0.099*R*	0.111P^+	0.106P^-	0.103P^+	0.099P^-	0.094P^+	0.156P^+	*I*	0.102*R*	0.086P^+	0.105P^+	0.083P^+	0.171P^+
154	0.585P^-	0.146P^-	0.112P^-	0.112*R*	0.158P^-	0.137P^-	0.152P^-	0.153*R*	0.130*R*	0.132*R*	0.144*R*	0.132*R*	0.123P^-	0.106*R*	0.138P^+	0.102*R*	*I*	0.101*R*	0.121*R*	0.099*R*	0.161P^+
155	0.583P^-	0.151*R*	0.099P^-	0.092*R*	0.133P^-	0.128*R*	0.121P^-	0.110P^-	0.110*R*	0.132P^+	0.118P^-	0.115P^+	0.099P^-	0.124P^+	0.183P^+	0.086P^-	0.101*R*	*I*	0.114P^+	0.090P^+	0.182P^+
156	0.625P^-	0.199P^-	0.132P^-	0.115P^-	0.180P^-	0.155P^-	0.170P^-	0.145P^-	0.106P^-	0.101P^-	0.117P^-	0.101P^-	0.156P^-	0.105P^-	0.120P^+	0.105P^-	0.121*R*	0.114P^-	*I*	0.087*R*	0.116P^+
157	0.622P^-	0.181P^-	0.124P^-	0.107P^-	0.169P^-	0.152P^-	0.151P^-	0.136P^-	0.119P^-	0.128*R*	0.120P^-	0.114P^-	0.134P^-	0.082*R*	0.138P^+	0.083P^-	0.099*R*	0.090P^-	0.087*R*	*I*	0.137P^+
158	0.701P^-	0.266P^-	0.202P^-	0.179P^-	0.258P^-	0.231P^-	0.249P^-	0.220P^-	0.181P^-	0.154P^-	0.188P^-	0.168P^-	0.224P^-	0.139P^-	0.108*R*	0.171P^-	0.161P^-	0.182P^-	0.116P^-	0.137P^-	*I*