IDENTIFYING KEY FACTORS FOR ADOPTING ARTIFICIAL INTELLIGENCE-ENABLED AUDITING TECHNIQUES BY JOINT UTILIZATION OF FUZZY-ROUGH SET THEORY AND MRDM TECHNIQUE

Kuang-Hua HU[1], Fu-Hsiang CHEN[2*],
Ming-Fu HSU[3], Gwo-Hshiung TZENG[4]

[1]*School of Accounting, Finance and Accounting Research Center,
Nanfang College of Sun Yat-sen University, Guangzhou, China*
[2]*Department and Graduate School of Accounting, Chinese Culture University, Taipei, Taiwan*
[3]*English Program of Global Business, Chinese Culture University, Taipei, Taiwan*
[4]*Graduate Institute of Urban Planning, College of Public Affairs,
National Taipei University, Taipei, Taiwan*

Abstract. In today's big-data era, enterprises are able to generate complex and non-structured information that could cause considerable challenges for CPA firms in data analysis and to issue improper audited reports within the required period. Artificial intelligence (AI)-enabled auditing technology not only facilitates accurate and comprehensive auditing for CPA firms, but is also a major breakthrough in auditing's new environment. Applications of an AI-enabled auditing technique in external auditing can add to auditing efficiency, increase financial reporting accountability, ensure audit quality, and assist decision-makers in making reliable decisions. Strategies related to the adoption of an AI-enabled auditing technique by CPA firms cover the classical multiple criteria decision-making (MCDM) task (i.e., several perspectives/criteria must be considered). To address this critical task, the present study proposes a fusion multiple rule-based decision making (MRDM) model that integrates rule-based technique (i.e., the fuzzy rough set theory (FRST) with ant colony optimization (ACO)) into MCDM techniques that can assist decision makers in selecting the best methods necessary to achieve the aspired goals of audit success. We also consider potential implications for articulating suitable strategies that can improve the adoption of AI-enabled auditing techniques and that target continuous improvement and sustainable development.

Keywords: artificial intelligence (AI), audit, certified public accountant (CPA), multiple criteria decision making (MCDM), multiple rule-based decision making (MRDM), fuzzy rough set theory (FRST).

JEL Classification: M15, M42, M42, O14, O33.

---

*Corresponding author. E-mail: *chenfuhsiang1@gmail.com*

Introduction

Artificial intelligence (AI) has experienced several "winters" and "springs", so to speak. With real progress now being made, AI is gaining traction and companies are witnessing enormous benefits from the expansion in the spectrum of human cognitive and functional outcomes. Although AI applications in accounting and auditing are not new (Keenoy, 1958), its impact on these fields is expected to continue to increase over the next few years due to further advancements in information technology. With powerful data processing and analytical ability, AI is dramatically changing traditional audit practices (Thibodeau, 2003). The massive amount of structured, semi-structured, and non-structured data is rendering the assessment of corporate financial and non-financial performances more challenging (Lam, 2004; Baldwin et al., 2006). In view of this, auditing is more suitable for AI applications. By applying AI technology, CPA firms/auditors acquire and process data and audit reviews of the entire population of their audited entities, greatly increasing audit efficiency and effectiveness (Baldwin et al., 2006). CPA firms are embracing these advanced technologies, and AI-enabled auditing techniques are creating a new auditing paradigm to provide better services to clients (Bizarro & Dorian, 2017).

The nature of audit work has dramatically changed with the evolution of computerization, automation, and new audit technologies that have ushered in a new era of performance milestones in the audit process. Compared with traditional auditing techniques, AI-enabled auditing techniques are more efficient and comprehensive, particularly in the two aspects of repetitive automation work and professional judgment of different types of results. Automation is the best tool for structural tasks, such as substantive tests (verification, footing, and vouching) when performing automated tasks to crunch data and provide instantaneous verification, so that the actual values correspond to the advanced analysis results (Kokina & Davenport, 2017). Furthermore, AI technology can be used to analyze big data, by collecting and summarizing data from multiple sources, providing auditors with sufficient evidence and information that can be included in judgments, supplementing their judgment capability, and making more informed decisions to provide clients with higher levels of assurance. More importantly, AI-enabled auditing techniques are prominent in data analysis, such as data extraction, comparison, and validation (Hsu & Lin, 2016; Lin, 2017), which means that AI-based technology can extract textual information from complex electronic documents (Deloitte, 2015). Thus, auditors can spend more time in areas that require higher levels of judgment and present greater insights to businesses.

Recognizing the extraordinary potential in AI and its incredible advantages, the Big 4 CPA firms (i.e., Deloitte, E&Y, KPMG, and PwC) are constantly developing AI techniques and investing heavily in technological innovation (Kokina & Davenport, 2017). Deloitte has applied cognitive computing (a synonym for artificial intelligence) using IBM's Watson technology that provides foundational capabilities in natural language processing (NLP) during the course of the audit. Such profound cognitive capabilities and advanced analytical solutions can help clients make judicious decisions and create and capture critical information (Deloitte, 2015; Meskovic et al., 2018). Similarly, KPMG is also focused on using cognitive computing for audits, because AI-enabled auditing techniques can automatically extract key information from copious datasets for clients, thus unleashing the real value of AI to acceler-

ate their decision-making (KPMG, 2016; Meskovic et al., 2018). To drive audit efficiencies, PwC uses "Halo" computing and processing with the aim of providing customers guaranteed services and service models that are closer to the digital age (M2 Presswire, 2016). Ernst and Young (EY) has already adopted AI technology for automating routine tasks in auditing, using its own proprietary Robotic Process Automation (RPA) system for delivering more accurate and effective auditing for its clients (Faggella, 2018).

Although existing research on AI applications is not rare, there are relatively few studies in the area of auditing, and those that do lack empirical and systematical analysis. Therefore, to remedy this lack of literature and provide a proper analytical procedure for CPA firms, we have to realize the dimensions and criteria and their complex interactive influencing relationship in the adoption of AI-enabled auditing techniques (Figure 1), so as to enhance audit quality as well as prevent audit failures. The motivation for this paper is to explore more in-depth the key dimensions and criteria in the application of AI-enabled auditing techniques by CPA firms and to premeditate the decision making of interactive influencing relationships among multiple criteria/attributes (Liou & Tzeng, 2012; Chen, 2015; Hu et al., 2018). As for the complex and influencing relationship for the adoption of AI-enabled auditing techniques, we take a fusion of multi-relational data mining in multiple rule-based decision making (MRDM) (i.e., the joint utilization of soft computing technique and MRDM) as the most appropriate approach for analyzing this issue.

To apply ranking and selection and to establish a performance improvement strategy based on the characteristics of a finite number of attributes, sorting/selecting should be shifted by existing systematic classification methods to find the best way for improving the performance and achieve the business level. In other words, "a systematic approach to exhaustively
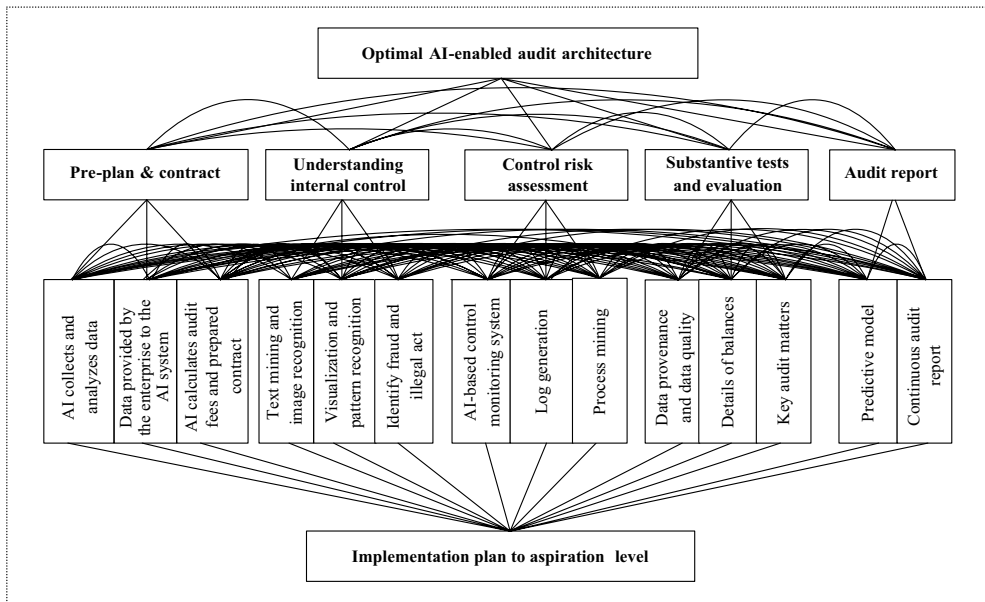

Figure 1. Analytic architecture of AI-enabled auditing mutual influence relationship

address AI-enabled auditing technique adoption issues for CPA firms' performance improvement is what we need" to avoid "just taking expedient measures" (Peng & Tzeng, 2013; Chen & Chi, 2015; Hu et al., 2017, 2018).

The contributions of this study are highlighted as follows: (1) A rule-based technique (i.e., FRST+ACO) is used to determine the essential features from a large amount of datasets without deteriorating the model's forecasting quality as well as speeding up the data processing procedure. (2) The selected features were then fed into D-ANP (DEMATEL-based ANP (Analytic Network Process)) to determine the most influential dimension and criteria. By doing so the dependency and feedback relationship for adopting AI-enabled auditing techniques can be clearly represented. (3) A modified-VIKOR approach assesses the performance gap of each dimension and each criterion for CPA firms to facilitate improvement toward achieving a targeted level. Moreover, it aims to reduce the performance gap for reaching the best improvement strategy. The interactive influential network relationship map (IINRM) can help reach this level. Furthermore, the modified-VIKOR (Tzeng & Shen, 2017; Peng & Tzeng, 2019; Lo et al., 2019; Shen & Tzeng, 2018; Lin et al., 2019; Hu & Tzeng, 2019; Qu et al. 2019; Huang et al., 2019) can be adopted to rank and select criteria as well as to improve performance gaps in AI-based applications in external audits. (4) Sorting the influencing factors can be used as a reference for selecting key criteria for CPA firms to exploit AI when conducting audit procedures. At the same time, the hybrid MRDM model can advance performance improvement from simple sorting and ranking for selecting the best solutions. The research flows in this study are shown in Figure 2.

The remainder of this study is organized as follows. Section 1 reviews the existing literature on the adoption of AI-enabled auditing techniques by CPA firms. Section 2 proposes the research design and methodologies and presents the empirical results. Last Section concludes.

1. Framework for adopting AI-enabled auditing techniques

Efficiency and effectiveness are attributes in an AI-based system that have been introduced in the audit process in lieu of traditional audit trails by large CPA firms, so that formalization of an overall audit framework of AI technology can be implemented in the near future (Meskovic et al., 2018). For enhancing their audit judgments, some CPA firms have already adopted AI to evaluate important clues in contracts, bookkeeping processes, and predictive analysis. With AI, automation technology can unimpeachably search semi-structured and unstructured data across multiple databases, thereby allowing CPA firms (auditors) to focus on providing more in-depth professional services to their clients. To help CPA firms achieve their AI-based performance indicators, this study captures the practical perspectives based on existing research and investigates deeper into the extreme issues. According to the past literature (American Institution of Certified Public Accountants [AICPA], 2015a, 2015b, 2015c, 2015d; Issa et al., 2016; Kokina & Davenport, 2017), interviews with domain experts and brainstorming, and the adoption of AI-enabled auditing techniques, we identify five dimensions as the evaluation framework: (1) pre-plan and contract, (2) understanding internal control, (3) control risk assessment, (4) substantive tests and evaluation, and (5) audit report. Each dimension is outlined in detail below.
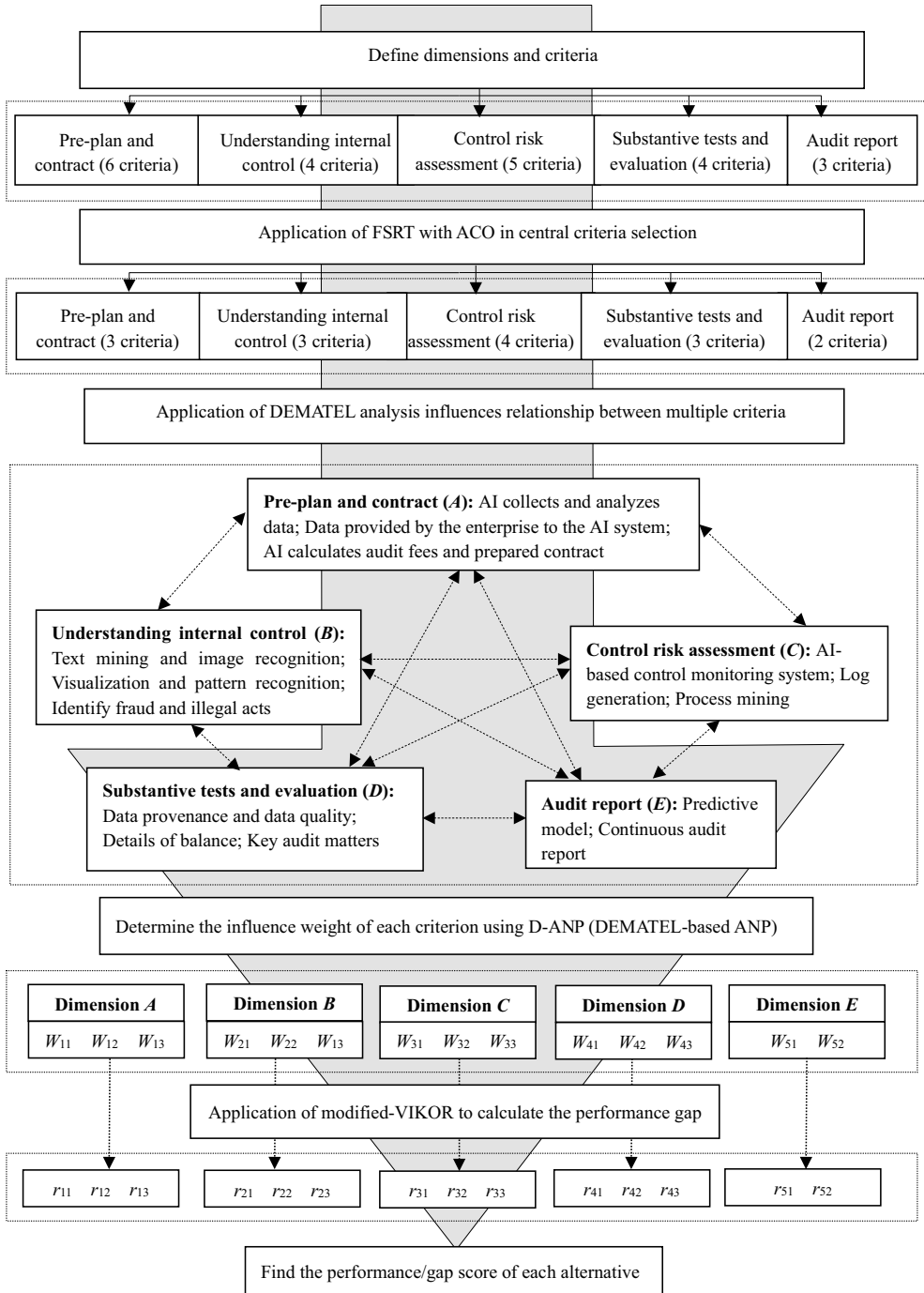
Define dimensions and criteria

| Pre-plan and contract (6 criteria) | Understanding internal control (4 criteria) | Control risk assessment (5 criteria) | Substantive tests and evaluation (4 criteria) | Audit report (3 criteria) |

Application of FSRT with ACO in central criteria selection

| Pre-plan and contract (3 criteria) | Understanding internal control (3 criteria) | Control risk assessment (4 criteria) | Substantive tests and evaluation (3 criteria) | Audit report (2 criteria) |

Application of DEMATEL analysis influences relationship between multiple criteria

Pre-plan and contract (*A*): AI collects and analyzes data; Data provided by the enterprise to the AI system; AI calculates audit fees and prepared contract

Understanding internal control (*B*): Text mining and image recognition; Visualization and pattern recognition; Identify fraud and illegal acts

Control risk assessment (*C*): AI-based control monitoring system; Log generation; Process mining

Substantive tests and evaluation (*D*): Data provenance and data quality; Details of balance; Key audit matters

Audit report (*E*): Predictive model; Continuous audit report

Determine the influence weight of each criterion using D-ANP (DEMATEL-based ANP)

| **Dimension *A*** | **Dimension *B*** | **Dimension *C*** | **Dimension *D*** | **Dimension *E*** |
| W_{11} W_{12} W_{13} | W_{21} W_{22} W_{13} | W_{31} W_{32} W_{33} | W_{41} W_{42} W_{43} | W_{51} W_{52} |

Application of modified-VIKOR to calculate the performance gap

| r_{11} r_{12} r_{13} | r_{21} r_{22} r_{23} | r_{31} r_{32} r_{33} | r_{41} r_{42} r_{43} | r_{51} r_{52} |

Find the performance/gap score of each alternative

Figure 2. Research flows of AI-enabled auditing

1.1. Pre-plan and contract

A pre-plan and contract make up the first phase in the process of adopting an AI-assisted auditing technique, which describes the acquisition of initial knowledge for clients. During the pre-planning and contracting phase, data acquisition is at the core of AI-enabled auditing techniques. Before the enterprise undertakes an official audit, a CPA firm has to ensure the accuracy and alignment of all the information (Davenport & Raphael, 2017). External data structure and accounting and financial systems provided by organizations reliably feed into the AI system. Next, a CPA firm creates a set of communicable APIs (Application Programming Interfaces) to transform a wide range of document types for use in an AI system (Wang et al., 2014; Davenport & Raphael, 2017). As a result, an effective analysis procedure and a correct audit report can be generated. More importantly, to confirm the rights and obligations between the CPA firm and the client, a contract for the adoption of the AI-enabled auditing technique is signed, and audit fees are decided based on the information provided by the client and evaluation of AI (Issa et al., 2016). The CPA firm then submits the audit report with the highest standard requested by the client and in accordance with the service contract content.

1.2. Understanding internal control

Establishing considerable internal control systems, a sound and effective internal operational mechanism not only can be helpful in reducing business risks, but also can deliver value to the business (Cangemi & Taylor, 2018). Most CPA firms have a responsibility for assessing and inspecting the impact of AI on the design and operation of internal control systems on a continuous basis by identifying its adequacy and effectiveness after the introduction of AI. However, owing to a huge amount of unstructured data (i.e., textual data), such as annual reports of public-listed companies, CPA firms can employ considerable and promising "minerals" text mining technology to extract meaningful information from the documents that help supplement the lack of structured data (Faria et al., 2016). In addition, the use of still and moving images is increasing exponentially in everyday life, and the capability to leverage image recognition technology for retrieving meaningful data from these images is turning into a progressively important application (Burgess, 2017). The application of data visualization technology or pattern recognition in auditing provides multiple functionality to help determine illegal acts or any anomalies not easily identified by statistical models, thus gaining better insights for clients (Roscoe & Howorth, 2009; Rehman & Saba, 2014; Sutton et al., 2016).

1.3. Control risk assessment

This dimension of CPA firms is to examine if the design and implementation of the client's internal control system conform to specifications. Continuous auditing automatically focuses on risk assessment and the execution of an internal control. In the current era of big data, a refinement of an entire population of records is the clear trend, and the function of an AI-based continuous control and monitoring system for conducting risk assessment by internal auditors is indispensable (Issa & Kogan, 2014). It is thus crucial to prioritize the level of

riskiness detected if a large number of outliers is detected (Issa et al., 2016). Consequently, corporates can benefit from the use of AI-based continuous auditing technology for improving audit efficiency and effectiveness. To prevent the tampering of audit evidence during the audit process and preclude the user from accessing chronological record of corporate activities, an automatically generated log provides a robust basis for the going-concern audit opinion. However, the process mining technique can be applied to extract pivotal information and identify anomalous outliers from event logs recorded in an audit context (Jans et al., 2014; Whitehouse, 2015) and provide CPA firms with insights on any inconsistencies between a firm's operations and design processes. Therefore, AI-based process mining on the entire sample should be necessary in the current big data environment.

1.4. Substantive tests and evaluation

A substantive testing method is used to examine account balances, and complete and valid test results offer evidence to support the financial records of a company (Quick & Henrizi, 2018). In this perspective, the most significant difference between AI-enabled auditing approaches and traditional auditing approaches is that the former can conduct a review of data provenance and data quality of the population on a continuous basis in real time and improve the reliability of the audit results (Appelbaum et al., 2018). To test the details of balances, AI-based continuous and comprehensive testing technology is appropriate for reducing the likelihood that anomalous records are not discovered (Hooda et al., 2018; Kang et al., 2015). In particular, for a large hybrid dataset, relevant information can be excavated by examining the details of the financial statement balances through AI-assisted auditing techniques. Audit evidence from third parties (such as CPA firms) is often more reliable than audit evidence provided directly by a company (Kang et al., 2015). A substantive test can thus prioritize identified exceptions to draw auditors to pay attention to more suspicious abnormal events (Issa & Kogan, 2014) and propose preventive measures. In order to improve user satisfaction with audit reports, the International Auditing Standards Board (IASB) set up new auditor reporting standards in 2015, in which CPAs have to communicate key audit matters (KAMs) after discovering major deficiencies in the audit risk assessment or verification process. The adoption of IT-related KAMs not only have improved the quality and efficiency of information available for investors or other financial statement users, but also have had positive effects on relevant disclosures and may help professionals interpret complex financial statements when there is information overload (Sneller et al., 2016; Sirois et al., 2018).

1.5. Audit report

CPA firms have proposed AI-enabled auditing techniques to help design and develop a prediction model as a decision-support framework for the smooth operation of the audit process. Auditors may run continuous tests with different models to anticipate the risk assessment identified with continuing client activities (Sinclair, 2015). The prime function of the prediction model is to execute audit risk prediction, which involves performance evaluation of the prediction model, prediction of a risk class, and fraud prediction. On the completion of an AI-enabled auditing technique adoption, CPA firms deliver a continuous audit report

called a verdict to a company comprising details of all the findings in the final step (PwC, 2014). The issuance of a traditional audit report is divided into three categories: clean, qualified, and adverse. However, the audit report adopted by AI can be continuous, and the grade could range from 1 to 100, such that the actual situation of the enterprise can be confirmed more clearly and precisely. Several research studies suggest that the audit report results obtained from using AI technologies, achieving a higher level of assurance in an avalanche of datasets, are better versus traditional audit techniques and basic statistical methods.

2. Research design and methodology

China is the second-largest economy in the world and plays a pivotal role in international financial markets. Its stock markets are some of the most important capital markets for global investors. Financial statements issued by enterprises cannot be fully trusted by shareholders due to the presence of information asymmetry. Therefore, the fundamental issue is how to improve the quality of financial statements and protect investors, thus enhancing shareholders' trust in a company. Most CPA firms conduct an external audit, which is widely considered to be the most effective supervisory form of financial reports. However, in this era of big data, CPA firms that perform limited audit procedures through existing audit technology find it difficult to discriminate among the indispensable information given the over-abundance of data. Chinese CPA firms recognize and support the merits of AI-enabled auditing, based on attributes such as facilitating better decision making under an anticipated risk level and reducing the possibility of audit failure and accounting fraudulence. To evaluate the performance of AI-enabled auditing in CPA firms, this research considers the Big 4 CPA firms in China that have adopted AI as an auditing technology over a long period. The performance gap of each criterion and the total performance gap of AI-enabled auditing for these evaluated firms can be clearly understood from the results.

This study uses the hybrid MRDM (Multiple Rule-based Decision Making) method (Tzeng & Shen, 2017; Shen et al., 2019) to evaluate how the external auditing architecture under AI application is effectively implemented and enhanced to offer a more accurate and all-around corporate message for clients and investors. However, the decision on the adoption of an AI-enabled auditing technique is complex, involving many criteria that can easily confuse decision-makers (auditors); thus, data pre-processing (i.e., feature selection) is an inevitable process before model construction. In addition, the pre-processing technique is performed by a rule-based architecture and the inherent decision logics of decision rules can be examined by users so as to increase it applicable capabilities.

Rough set theory (RST) has demonstrated its usefulness in discovering data dependencies, reducing the number of features contained in a dataset, using data without requiring additional information, and handling vague, imprecise, and uncertain information. However, the classical RST cannot work effectively on a real-valued dataset (i.e., it only accepts discrete/crisp data). Thus, many extension models of RST have been introduced to cope with this problem. The fuzzy rough set theory (FRST) (Dubois & Prade, 1980, 1990), which encapsulates the related but distinct concepts of vagueness (fuzzy set theory) and indiscernibility (rough set theory), appropriately handles datasets both with fuzziness and vagueness with

continuous features. Many previous studies have implemented the hill-climbing algorithm to determine the best feature subsets for FRST (Cheng, 2018), but this algorithm often leads to the discovery of non-optimal feature subsets, both in terms of the resulting dependency measure and the subset size (Jensen & Shen, 2005; Jensen et al., 2014). Ant colony optimization (ACO) is a population-based algorithm that relies on the foraging behavior of ants for finding optimal paths. It has been applied to a large number of difficult combinatorial tasks successfully, such as the traveling salesman problem (TSP), scheduling problems, and optimization tasks. Thus, we conduct FRST with ACO to handle the task of essential feature subset selection (i.e., the essential criteria).

The selected criteria are then fed into the DEMATEL (decision-making trial and evaluation laboratory) technique to establish an interactive influential network relationship map (IINRM) to address AI-enabled auditing issues. The DEMATEL technique can determine the influence relationship under the interactive degree of factors/criteria, and the effect is more significant under the practical experience of experts (Gopal et al., 2018). Our paper integrates the concept of ANP (Saaty, 1996) into the DEMATEL technique to determine the influence relation matrix and to construct IINRM and identify the influential weights (called "global weights") for the D-ANP (DEMATEL-based analytic network process) dimensions and criteria. Accordingly, D-ANP can effectively and accurately measure the major influencing factors (performance weights) of an audit work for CPA firms in AI applications by submitting the influential weight of each criterion derived by D-ANP into the modified-VIKOR method. We use the "aspiration-worst" of the modified-VIKOR technique to replace the traditional "max-min" as a benchmark for reducing the performance gap ratio and estimating the integrated performance values in all individual criteria. We use the gaps in each criterion to achieve the aspiration level by diminishing the difference between current situation and target situation and to determine the optimal AI-enabled auditing architecture for CPA firms toward achieving the aspiration level of "avoiding a bunch of rotten apples to find the best apple". Figure 3 illustrates the details of how each technique is synthesized.

2.1. Data collection

This study designed a three-stage questionnaire survey to collect data, including dimensions and criteria determination, a pre-test questionnaire, and an official questionnaire. The first-stage acknowledged the 5 dimensions and 21 criteria based on the related literature and professional in-depth reviews by domain experts as shown in Table 1. To ensure consistency and effectiveness of the pairwise comparisons, Saaty (1996) recommended a limited number of factors within a single dimension. In the second stage, these criteria were preliminarily formulated based on Table 1 to create a pre-test questionnaire. The pre-test questionnaire survey consisted of interviews with 15 CPAs and 15 senior engineers from CPA firms have that incorporated AI techniques into their auditing business task in Guangzhou and Shanghai. In the pre-test questionnaire, the respondents were asked to rate the influence of the 21 criteria on AI-based auditing implementation using a 10-point scale (0 extremely unimportant; 5 neutral; and 10 extremely important). Providing a crisp number on the degree of influence between criteria by experts (i.e., CPAs and engineers from CPA firms) is a sophisticated

and complicated task. Due to the increasing complexity of decision architecture, experts are accustomed to employing confidence levels/linguistic values in their assessments (Si et al., 2017; Ding & Liu, 2018; Salarpour et al., 2019). In addition, not all of the given criteria are familiar to experts due to their limited working experience and knowledge. To combat this, the confidence level is incorporated into a 10-point scale – that is, the experts provided two values: one is for a 10-point scale (i.e., from 0 to 10), and the other is for the confidence level (i.e., from 1% to 100%). For example, if the expert fills the criteria "A1: AI collects and analyzes data" with values 8 and 9, and also provides the confidence level with value 8 at 80% and value 9 at 20 %, then the final score will be 8.2 (i.e., 8*80% + 9*20%). Compared to other crisp decision models, the proposed architecture not only can increase the flexibility of the decision process, but also prevent the data from exhibiting distortion (Wang et al., 2017; Du et al., 2019; Çolak et al., 2020).

High dimensionalities (i.e., too many criteria) will confuse users and lead to improper judgments (Hsu, 2019a; Lin et al., 2019; Hsu et al., 2020). To combat this, feature selection (FS) aims at determining the minimal feature subsets from a problem domain while retaining appropriate precision in representing the original data semantics, and thus we consider it in the second stage. FRST has proven its success in FS applications (Lin & Hsu, 2018). Thus, we employ FRST to handle the FS task. Determining the minimal feature subset in FRST is a NP-hard problem (Skowron & Rauszer, 1992). Inspired by swarm intelligences, ACO (one type of swarm intelligence algorithm) has superior ability at handling the combinatorial optimization task, and thus we use it. The FRST-ACO method used herein was introduced by Jensen
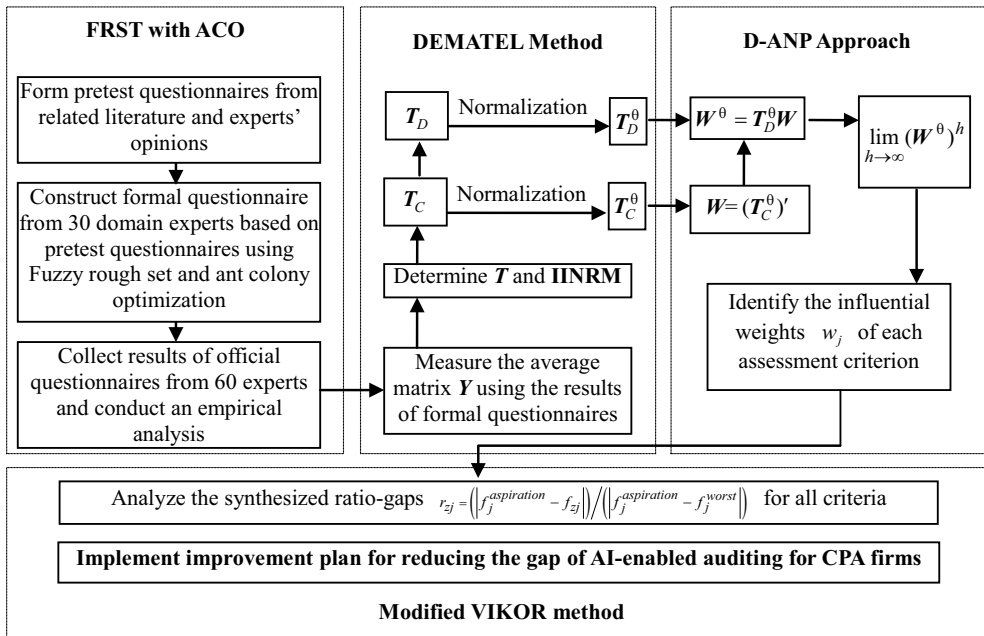

Figure 3. Integrated techniques and methods of a fusion MRDM model

and Shen (2005) and is a supervised FS technique. Thus, before implementing FRST-ACO to identify the essential features, the decision attribute has to be provided in advance. In accordance with the study by Thangavel et al. (2005), we apply the K-means algorithm to determine the decision attribute. The main purpose of K-means is to partition all observations into K clusters in which each observation belongs to the cluster with the nearest mean (that is, it minimize the within-cluster sum of squared errors), serving as a prototype of the cluster. The critical issue is how to determine the suitable number of clusters (i.e., K). The trial-and-error strategy was performed. Here, K is set from 1 to 5, and the aggregation of test accuracy and rule coverage (ATARC) is taken as an assessment measure. According to the experiments (see Table 2), K is set to 2 in order to achieve optimal performance (i.e., when K is set to 2, the value of ATARC is the highest among all comparisons). Thus, the number of clusters of this study is set to 2. Table 2 presents the selected criteria/attributes. Table 3 provides detailed descriptions of the selected criteria.

Table 1. Criteria of AI-enabled auditing for the pre-test questionnaire

Dimensions	Criteria
Pre-plan and contract (A)	A1: AI collects and analyzes data
	A2: Data provided by the enterprise to the AI system
	A3: Client's initial risk level
	A4: The number of hours the engagement
	A5: AI calculates audit fees and prepared contract
Understanding internal control (B)	B1: Text mining and image recognition
	B2: AI analyzes the generated video
	B3: Visualization and pattern recognition
	B4: Identify fraud and illegal acts
Control risk assessment (C)	C1: AI-based control monitoring system
	C2: Identified violations of level of riskiness
	C3: Logs are automatically generated
	C4: Process mining
	C5: Response time
Substantive tests and evaluation (D)	D1: Data provenance and data quality
	D2: Details of transactions
	D3: Details of balances
	D4: Key audit matters
Audit report (E)	E1: Review
	E2: Predictive model
	E3: Continuous audit report

Table 2. The selected criteria by FRST-ACO

Criteria Subset	Testing accuracy	Rule coverage	ATARC*
Cluster numbers: K = 2			
Subset 1: A1, A2, B1, B3, B4, C1, C2, C3, D1, D3, E1, E2	0.85	0.89	1.74
Subset 2: A1, A2, A4, B2, B3, B4, C1, C3, C4, C5, D1, D4, E1, E3	0.87	0.91	1.78
Subset 3: A1, A2, A5, B1, B3, B4, C1, C3, C4, D1, D3, D4, E2, E3	***0.93***	***0.92***	***1.85***
Subset 4: A1, A3, A4, B2, B4, C1, C3, C4, C5, D2, D3, D4, E2, E3	0.90	0.87	1.77
Subset 5: A1, A2, A4, B2, B4, C2, C4, C5, D1, D4, E1, E3	0.88	0.86	1.74
Cluster numbers: K = 3			
Subset 1: A1, A3, A4, B2, B3, C4, C5, E3	0.75	0.79	1.54
Subset 2: A2, A3, B4, C3, C5, D2, D3, E3	0.78	0.81	1.59
Subset 3: A1, A4, B4, C2, C4, C5, D4, E2, E3	0.81	0.8	1.61
Subset 4: A2, A3, B4, C2, C4, D3, D4, E1, E3	0.82	0.82	1.64
Subset 5: A3, A5, B4, C1, C5, D1, D4, E1, E3	0.79	0.79	1.58
Cluster numbers: K = 4			
Subset 1: A1, B1,C4, C5, D4, E1, E2	0.69	0.72	1.41
Subset 2: A1, A5, B3, D1, D4, E1, E3	0.71	0.74	1.45
Subset 3: A2, B1, B4, C1, C5, D3, E2	0.7	0.73	1.43
Cluster numbers: K = 5			
Subset 1: A2, B3, C1, D3, E1	0.65	0.68	1.33
Subset 2: A1, B1, C1, C4, D4, E1	0.68	0.7	1.38
Subset 3: A1, B2, B4, C3, C5, D4, E3	0.66	0.67	1.33

Note: *ATARC: The aggregation of testing accuracy and rule coverage.

Table 3. The descriptions of selected criteria

Dimensions/Criteria	Descriptions	Sources
Pre-plan and contract (A)		
A1: AI collects and analyzes data (a_1)	AI can receive big data across multiple-source databases.	Davenport and Raphael (2017)
A2: Data provided by the enterprise to the AI system (a_2)	AI can incorporate any information and operating system provided by the enterprise.	Wang et al. (2014); Davenport and Raphael (2017)
A5: AI calculates audit fees and prepared contract (a_3)	AI calculates audit fees, and CPA firms and client signed an AI-prepared contract.	Issa et al. (2016)
Understanding internal control (B)		
B1:Text mining and image recognition (b_1)	Text mining and image recognition technology of AI can analyze client-provided semi-structured and structured data.	Faria et al. (2016); Burgess (2017)

Dimensions/Criteria	Descriptions	Sources
B3:Visualization and pattern recognition (b_2)	AI utilizes visualization and pattern recognition to determine risk factors.	Roscoe and Howorth, (2009); Sutton et al. (2016)
B4: Identify fraud and illegal acts (b_3)	AI applied aggregated information to identify fraud and illegal act.	Rehman and Saba (2014); Meskovic et al. (2018)
Control risk assessment (*C*)		
C1:AI-based control monitoring system (c_1)	An AI-based continuous control monitoring system can effectively review the level of riskiness and assess the priority of its violations.	Issa and Kogan (2014); Issa et al. (2016)
C3:Log generation (c_2)	Logs are generated automatically to ensure the integrity of a chronological work record of the audit process and to prevent unauthorized tampering such as audit evidence and audit report.	Jans et al. (2014); Whitehouse (2015)
C4:Process mining (c_3)	Big data mining using the AI running process not only ensures that the internal control system is properly designed but can also be correctly configured and implemented.	Jans et al. (2014)
Substantive tests and evaluation (*D*)		
D1:Data provenance and data quality (d_1)	AI examined continuously to ensure reliability of data sources and data quality in real time.	Appelbaum et al. (2018); Hooda et al. (2018)
D3: Details of balances (d_2)	Reduce undetected abnormal conditions with continuous and comprehensive test of details of balances.	Appelbaum et al. (2018); Hooda et al. (2018); Quick and Henrizi (2018); Kang et al. (2015)
D4: Key audit matters (d_3)	CPAs provide a professional judgment, including disclosing company-specific information and considerable risks of material misstatements, in the audit report.	Sneller (2016); Sirois (2018)
Audit report (*E*)		
E2: Predictive model (e_1)	AI uses predictive model to evaluate all the risks identified based on collected evidence.	Sinclair (2015)
E3: Continuous audit report (e_2)	Issuance of continuous audit report based on the complete audit process is under AI.	PwC (2014)

The third stage constructed the formal questionnaire based on Table 3 and collected the results of the experts' reviews. The respondents were invited to complete comparison levels of 14 criteria and their satisfaction level of the criteria at this stage. The interviewed population consisted of 32 CPAs and 28 senior engineers belonging to the information technology (IT) department of the Big 4 CPA firms in Guangzhou and Shanghai. Respondents included CPAs who are a senior partner or a partner and senior engineers with at least 10 years of work experience in the IT department. The questionnaires were administered through face-to-face interviews lasting over 90 minutes between March and December 2018. The score assessment on the influence of a pair-comparison was conducted on a scale of 0 (absolutely no influence) to 4 (very high influence) based on the experts' thoughts and opinions. In addition, the performance questionnaire was scored from "very bad/extreme dissatisfaction (0)" to "very good/extreme satisfaction (10) by domain experts. Each domain expert simultaneously evaluated the performance values of each alternative (each Big 4 CPA firm). Finally, 60 completed expert forecasting questionnaires were submitted for the research methods depicted herein as the basis for empirical analysis.

2.2. Creation of an IINRM by DEMATEL

This study found 14 relatively important criteria in an AI-enabled auditing evaluation framework of CPA firms using a pre-test significance. The DEMATEL method was used to construct IINRM, which illustrated the mutual relationships among the 14 criteria and 5 dimensions. The influential weights of D-ANP were determined by relaxing the AHP assumption that the factors presented in the hierarchical structure are independent; however, this assumption is inappropriate in the dynamic environment. The ANP proposed by Saaty (1996) addresses the dependence and feedback issues between the dimensions/clusters or criteria (inner dimensions/clusters) through a diagonal matrix (Zhou et al., 2016). The ANP is based on the assumption that independence of factors of AHP can be eliminated and that the measurement of the dependency among the factors helps in the decision making process (Jing et al., 2018). Therefore, we propose a hybrid MRDM model that involves FRST with ACO, the DEMATEL technique, D-ANP (DEMATEL-based ANP) in IINRM, and the modified-VIKOR method for improving the performance toward achieving the targeted level. Briefly, FRST is used with ACO to screen important variables from the pre-test questionnaire to create an official questionnaire. Next, the DEMATEL technique was applied to construct IINRM of the influential relationship matrix from an official questionnaire survey of domain experts. The D-ANP influential weights based on the basic concepts of ANP are better able to deal with real-world problems than AHP. Therefore, according to the processes of constructing an average matrix and finding the direct-effect matrix (Appendix A, stage 1), we determine the IINRM and D-ANP influential weights. Figure 4 illustrates the process of empirical analysis.

The 60 domain experts from the Big 4 CPA firms in Guangzhou and Shanghai were invited to score the cause-effect relationship of each criterion on the other criteria, and so the initial influence matrix was created by a pairwise comparison. To identify the reliability of the sample collection, a random selection of 59 questionnaires was used to analyze the consensus level with an average gap ratio of 1.01% < 5% (i.e., consensus ratio is 98.99%, or
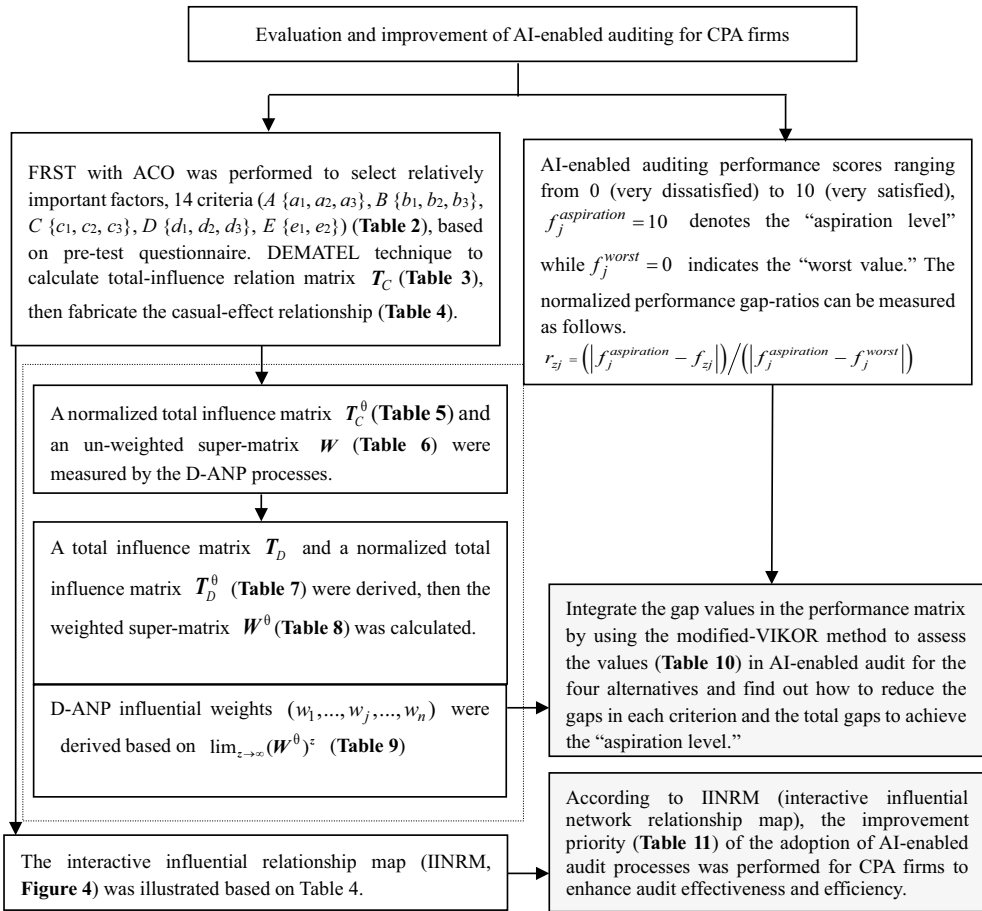
Figure 4. Process of the empirical case

more than 95% confidence), indicating consensus (see Note in Table 4). After the average matrix is normalized, a direct-influence relation matrix can be measured.

Using Eqs (A.4) (Appendix A, stage 1), the total impact relationship matrix T (Table 4) is calculated, and the construction of IINRM among the dimensions and criteria is accomplished. The IINRM of the structural relation influence matrix is composed of the relationship between the total influence row vector $r = (\ldots, r_i, \ldots)$ and the column vector $s = (\ldots, s_i, \ldots)$. The vector $r = (\ldots, r_i, \ldots)$ was used to determine the extent of the direct and indirect effects of the criterion/factor i on other criteria/factors. The vector $s = (\ldots, s_i, \ldots)$ is the measurement of the total direct and indirect influences of criterion j by other criteria. Therefore, the vector $(r_i + s_i)$ shows the degree of the major between the sum total of influences of criterion i and other criteria; the vector $(r_i - s_i)$ offers a causal cluster among the criteria (or called factors). If the vector $(r_i - s_i)$ is positive, then criterion (dimension) i is part of the causal cluster affecting other criteria, and if the vector $(r_i - s_i)$ is negative, then criterion (dimension) i is an influenced cluster affected by the other criteria. The results of the total influence (direct and indirect) relation matrix T are listed in Table 4.

Table 4. Normalized direct influence relationship matrix T through Eq. (4)

Criteria	a_1	a_2	a_3	b_1	b_2	b_3	c_1	c_2	c_3	d_1	d_2	d_3	e_1	e_2
a_1	0.081	0.135	0.101	0.116	0.091	0.121	0.098	0.129	0.136	0.133	0.141	0.089	0.097	0.100
a_2	0.127	0.095	0.104	0.132	0.109	0.150	0.100	0.129	0.128	0.120	0.136	0.103	0.131	0.125
a_3	0.206	0.213	0.094	0.171	0.143	0.167	0.127	0.157	0.173	0.163	0.174	0.125	0.148	0.126
b_1	0.146	0.156	0.115	0.087	0.108	0.142	0.107	0.145	0.148	0.125	0.145	0.114	0.123	0.121
b_2	0.176	0.183	0.132	0.183	0.087	0.198	0.135	0.178	0.175	0.156	0.179	0.137	0.146	0.139
b_3	0.121	0.129	0.100	0.111	0.086	0.079	0.093	0.123	0.138	0.115	0.126	0.088	0.107	0.105
c_1	0.188	0.207	0.151	0.176	0.151	0.192	0.093	0.203	0.207	0.180	0.191	0.125	0.161	0.157
c_2	0.103	0.111	0.084	0.090	0.074	0.098	0.081	0.068	0.112	0.106	0.120	0.088	0.094	0.093
c_3	0.100	0.096	0.078	0.095	0.083	0.107	0.087	0.107	0.068	0.091	0.097	0.082	0.087	0.083
d_1	0.132	0.138	0.103	0.122	0.100	0.124	0.107	0.135	0.142	0.081	0.142	0.101	0.122	0.116
d_2	0.089	0.096	0.077	0.073	0.060	0.072	0.055	0.088	0.086	0.077	0.057	0.070	0.078	0.078
d_3	0.212	0.228	0.168	0.191	0.160	0.203	0.151	0.194	0.195	0.214	0.223	0.100	0.183	0.171
e_1	0.182	0.189	0.160	0.181	0.152	0.192	0.155	0.180	0.180	0.169	0.180	0.129	0.105	0.143
e_2	0.223	0.236	0.178	0.210	0.177	0.223	0.168	0.210	0.230	0.210	0.215	0.168	0.235	0.119

Note: Average gap ratio (%) $= \dfrac{1}{n \times (n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\left| \overline{y}_{ij}^{60} - \overline{y}_{ij}^{59} \right| \Big/ \overline{y}_{ij}^{60} \right) \times 100\% = 1.01\% < 5\%$. This result indicates that significant confidence of consensus is 98.99%, where $\overline{y}_{ij}^{59}$ and $\overline{y}_{ij}^{60}$ are the average scores of the experts for 59 and 60, respectively; n is number of critical criteria (factors), here $n = 14$ and $n \times n$ matrix.

As Table 5 shows, of the five dimensions, dimension A (understanding internal control) had the greatest degree of influence ($r_i - s_i = 0.036$), whereas dimensions A (pre-plan and contract, $r_i - s_i = -0.007$) and C (control risk assessment, $r_i - s_i = -0.034$) are influenced the least among all the dimensions. The strongest influence relationship ($r_i + s_i$) with other criteria is measured for criterion a_3 (AI calculates audit fees and the prepared contract) at 0.812, whereas the value of e_1 (predictive model) at 0.406 has the weakest relationship. In the criteria evaluation of AI-enabled auditing factors, criteron d_3 (key audit matters) has the highest ($r_i - s_i$) influence, indicating it is the most important criterion from the perspective of CPA firms. Criterion d_2 (details of balances) has the lowest ($r_i - s_i$) influence among all the criteria, representing that CPA firms are concerned by it the least.

We construct IINRM by measuring the degree of mutual influence and the important degree of AI-enabled auditing among the five dimensions and 14 criteria using the DEMATEL technique, as shown in Figure 5. The horizontal axis ($r_i + s_i$) and the vertical axis ($r_i - s_i$) in Figure 5 represent the degree of relationship between the criteria and the degree of causality between the variables, respectively. The IINRM of dimensions and criteria in this study allow us to clearly pinpoint the interdependence in influential relationship between the criteria for the adoption of AI-enabled auditing for CPA firms. Dimensions A and C are below the horizontal axis, indicating that they are influencing dimensions to CPA firms. However, Dimensions B, E, and D are above the horizontal axis, representing that they affect other dimensions. Dimension B (understanding internal control) acknowledges the direct effect on other dimensions; Dimension E (Audit report) also has a strong influence on dimensions D (substantive tests and evaluation), A (pre-plan and contract), and C (control risk assessment). In other words, based on the degree of impact of the dimensions on IINRM, the recommend-

Table 5. Sum of cause (r_i)-effect (s_i) relationships among the key dimensions (phases) and criteria (steps)

Dimensions/Criteria	Row sum (r_i)	Column sum (s_i)	$r_i + s_i$	$r_i - s_i$
Pre-plan and contract (*A*)	**0.524**	**0.531**	**1.055**	**–0.007**
AI collects and analyzes data (a_1)	0.318	0.414	0.732	–0.096
Data provided by the enterprise to the AI system (a_2)	0.326	0.443	0.768	–0.117
AI calculates audit fees and prepared contract (a_3)	0.513	0.300	0.812	0.213
Understanding internal control (*B*)	**0.530**	**0.494**	**1.024**	**0.036**
Text mining and image recognition (b_1)	0.337	0.381	0.717	–0.044
Visualization and pattern recognition (b_2)	0.468	0.281	0.749	0.187
Identify fraud and illegal acts (b_3)	0.275	0.419	0.695	–0.144
Control risk assessment (*C*)	**0.477**	**0.511**	**0.988**	**–0.034**
AI-based control monitoring system (c_1)	0.503	0.261	0.764	0.242
Log generation (c_2)	0.261	0.378	0.639	–0.117
Process mining (c_3)	0.262	0.387	0.649	–0.125
Substantive tests and evaluation (*D*)	**0.507**	**0.502**	**1.009**	**0.005**
Data provenance and data quality (d_1)	0.325	0.373	0.698	–0.048
Details of balances (d_2)	0.205	0.423	0.627	–0.218
Key audit matters (d_3)	0.537	0.271	0.808	0.266
Audit report (*E*)	**0.750**	**0.725**	**1.475**	**0.025**
Predictive model (e_1)	0.247	0.159	0.406	0.088
Continuous audit report (e_2)	0.354	0.200	0.554	0.154

ed improvements should be prioritized as *B*, *E*, *D*, *A*, and *C*. Analysis of the straightforward interactive influential relationship among criteria within the dimensions recommends that criteria a_3 (AI calculates audit fees and prepared contract), b_2 (visualization and pattern recognition), c_1 (AI-based control monitoring system), d_3 (key audit matters), and e_2 (continuous audit report) have a significant effect on each dimension. These five criteria are the core of each dimension in the system having a powerful effect on the sustainability and affirmation of the criteria in their dimensions. Consequently, this result demonstrates that a healthy and strong understanding of internal control of AI-enabled auditing is the most conclusive aspect for the sustainability of audit tasks in CPA firms.

2.3. Evaluation of influential weights by the D-ANP method

The basic assumption in AHP is that the relationship among criteria (inner clusters) and dimensions (outer clusters) is independent; however, ANP commits to the relationship and correlation between dimensions and between criteria through a diagonal matrix until it converges to a null matrix (independent) or an identity matrix (diag (1,1,…,1)). The influential weights (called "global weights") among the criteria are calculated by the D-ANP technology on the DEMATEL basis ANP can indicate. The specific process refers to Stage 2 in Appendix A.
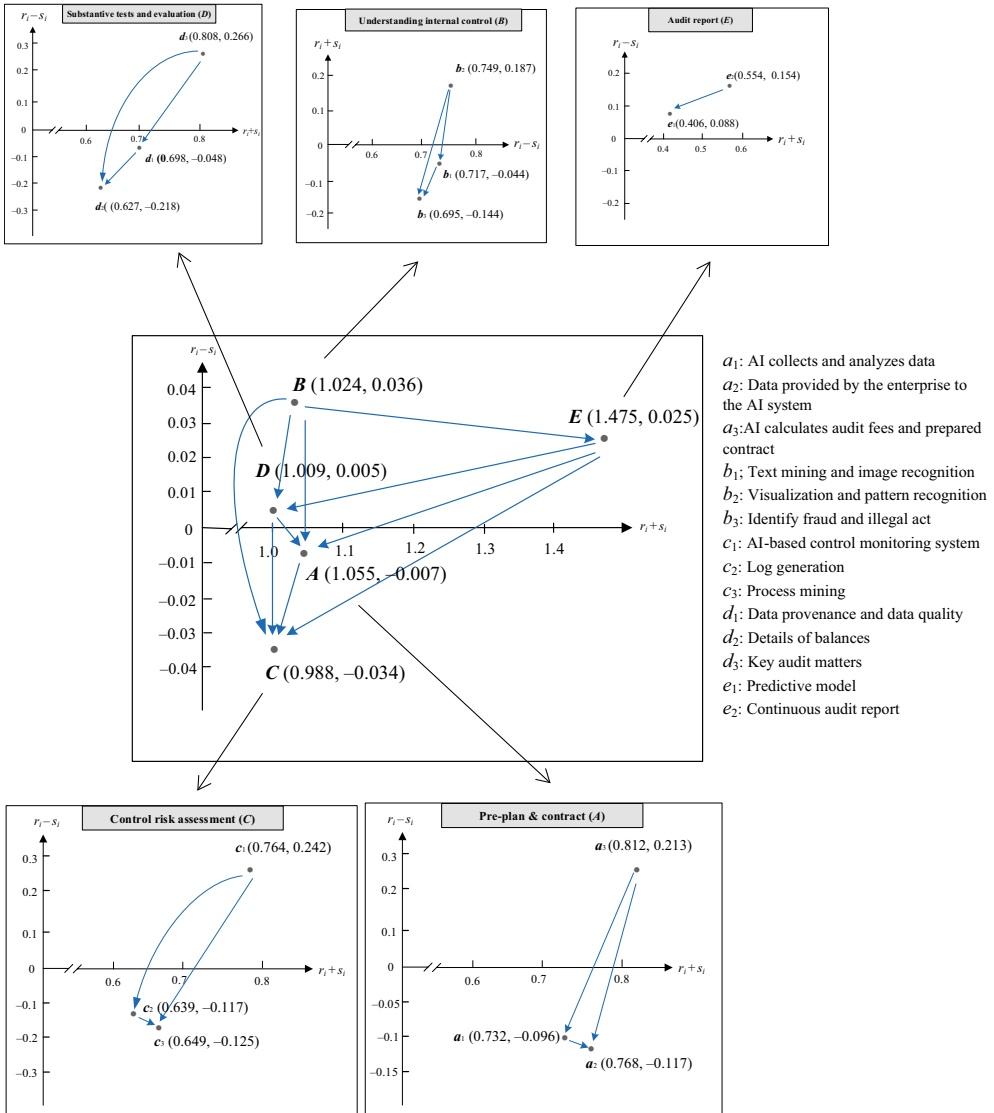
Figure 5. The IINRM of influence relationships based on DEMATEL
within AI-enabled auditing adoption

Using Eqs (A.5) and (A.6) (Stage 2 in Appendix A), the total influence relationship matrix T_C and the normalized total influence matrix T_C^θ for the criteria are obtained. Using Eq. (A.7), the unweighted super-matrix W is estimated by transposing the normalized total matrix T^θ. Total influence relationship matrix T_D and the normalized total matrix T_D^θ for the dimensions are demonstrated using Eqs (A.8) and (A.9). Next, the weighted super-matrix W^θ can be derived through Eq. (A.10). Subsequently, through self-multiplication of the super-weighted matrix among the criteria, the infinite power of the limit is taken until it converges and emerges as a long-term stable pattern (Table 6), which is the D-ANP influential weights.

Table 6. D-ANP influential weights (Global weights) of system factors by $\lim_{h\to\infty}(W^{\theta})^{h}$.

Criteria	a_1	a_2	a_3	b_1	b_2	b_3	c_1	c_2	c_3	d_1	d_2	d_3	e_1	e_2
a_1	0.079	0.083	0.064	0.073	0.060	0.078	0.060	0.077	0.080	0.073	0.080	0.059	0.069	0.065
a_2	0.079	0.083	0.064	0.073	0.060	0.078	0.060	0.077	0.080	0.073	0.080	0.059	0.069	0.065
a_3	0.079	0.083	0.064	0.073	0.060	0.078	0.060	0.077	0.080	0.073	0.080	0.059	0.069	0.065
b_1	0.079	0.083	0.064	0.073	0.060	0.078	0.060	0.077	0.080	0.073	0.080	0.059	0.069	0.065
b_2	0.079	0.083	0.064	0.073	0.060	0.078	0.060	0.077	0.080	0.073	0.080	0.059	0.069	0.065
b_3	0.079	0.083	0.064	0.073	0.060	0.078	0.060	0.077	0.080	0.073	0.080	0.059	0.069	0.065
c_1	0.079	0.083	0.064	0.073	0.060	0.078	0.060	0.077	0.080	0.073	0.080	0.059	0.069	0.065
c_2	0.079	0.083	0.064	0.073	0.060	0.078	0.060	0.077	0.080	0.073	0.080	0.059	0.069	0.065
c_3	0.079	0.083	0.064	0.073	0.060	0.078	0.060	0.077	0.080	0.073	0.080	0.059	0.069	0.065
d_1	0.079	0.083	0.064	0.073	0.060	0.078	0.060	0.077	0.080	0.073	0.080	0.059	0.069	0.065
d_2	0.079	0.083	0.064	0.073	0.060	0.078	0.060	0.077	0.080	0.073	0.080	0.059	0.069	0.065
d_3	0.079	0.083	0.064	0.073	0.060	0.078	0.060	0.077	0.080	0.073	0.080	0.059	0.069	0.065
e_1	0.079	0.083	0.064	0.073	0.060	0.078	0.060	0.077	0.080	0.073	0.080	0.059	0.069	0.065
e_2	0.079	0.083	0.064	0.073	0.060	0.078	0.060	0.077	0.080	0.073	0.080	0.059	0.069	0.065

The D-ANP influential weights, among the dimensions/criteria, are transformed into the local weights of the criteria for applying to the modified-VIKOR and to evaluate the performance gap ratio r_{Zj} of each criterion, as well as making an overall assessment.

2.4. Measurement of performance using the modified-VIKOR

The AI-enabled auditing model was scored using 14 filtered criteria, with each criterion's rating scales ranging from 0 to 10 based on a formal questionnaire: 0 = very dissatisfied, 10 = very satisfied. The new benchmark for the modified-VIKOR method using "aspiration-worst" is thus set as $f_j^{worst}=0$ for the worst value and $f_j^{aspiration}=10$ as the aspiration level in criterion j, where $j=1,2,\ldots,n$. The normalized performance matrix (values) $[f_{zj}]_{Z\times n}$ is used to yield the performance ratio-gap $[r_{zj}]_{Z\times n}$, thus avoiding the choice of a seemingly best solution among the inferior options/alternatives (Opricovic & Tzeng, 2004), as measured based on Appendix A, Stage 3 Eq. (A.11).

The minimal gap (called group utility) S_z and the maximal gap (called maximal regret) Q_z are derived by Eqs (A.12) and (A.13). The comprehensive performance indicator R_z is then measured using S_Z, Q_Z and $S_Z^{aspiration}=Q_Z^{aspiration}=0$, $S_z^{aspiration}=Q_z^{aspiration}=0$ (the aspiration level), where $S_Z^{worst}=Q_Z^{worst}=1$ (the worst value). Accordingly, Eq. (A.14) for R_Z can be re-written as $R_z=vS_z+(1-v)Q_z$, where $0\le v\le 1$. The average gap can be minimized when the weight $v=1$; the maximum gap can be determined when the weight $v=0$ for improvement priority. In most cases, $v=0.5$ is employed; however, it can be adjusted dependent on the AI-enabled auditing experts. The details of the procedures for the modified-VIKOR method are illustrated in Stage 3 of Appendix A.

According to the above description and Stage 3 of Appendix A, the performance values of the Big 4 CPA firms can be assessed through the modified-VIKOR method, as shown in Table 7. The decision maker defines the aspiration level (zero gaps) as a benchmark in this approach, and the gap-weighted measurements indicate the improvement between the Big 4 CPA firms (alternatives) and the benchmark (aspiration levels). The results can help CPA firms to capture the gaps between actual performance and aspiration levels, determine the improvement goals and priorities, and enhance their auditing efficiency and competitiveness.

Our results clearly indicate that CPA firm A_1 has the lowest total gap (0.215) among the Big 4 CPA firms, indicating that this firm has the best outcome for the adoption of AI-enabled auditing as unanimously advocated by the panel of experts. This also means that its performance is the easiest to achieve over other companies. For the dimension, the results of the indicated gaps demonstrate that "Audit report (E)" has the lowest performance gap for

Table 7. Gap-ratio value of AI-enabled audit for Big 4 CPA firms by modified-VIKOR

Dimensions/ Criteria	Local weight	Global weight	Alternative A_1		Alternative A_2		Alternative A_3		Alternative A_4	
			Score	Gap	Score	Gap	Score	Gap	Score	Gap
A	**0.225**		**7.870**	**0.213**	**7.830**	**0.217**	**7.610**	**0.239**	**7.340**	**0.266**
a_1	0.349	0.079	8.429	0.271	8.429	0.157	7.714	0.229	7.286	0.271
a_2	0.369	0.083	7.143	0.286	7.571	0.243	7.857	0.214	7.429	0.257
a_3	0.282	0.064	8.143	0.186	7.429	0.257	7.143	0.286	7.000	0.300
B	**0.212**		**7.640**	**0.236**	**7.730**	**0.227**	**7.450**	**0.255**	**8.300**	**0.170**
b_1	0.346	0.073	8.143	0.186	7.714	0.229	8.429	0.157	8.143	0.186
b_2	0.285	0.060	7.286	0.271	6.857	0.314	6.857	0.314	7.571	0.243
b_3	0.368	0.078	7.429	0.257	8.429	0.157	7.000	0.300	9.000	0.100
C	**0.217**		**7.740**	**0.226**	**7.680**	**0.232**	**6.920**	**0.308**	**8.100**	**0.190**
c_1	0.275	0.060	8.000	0.200	6.286	0.371	6.714	0.329	8.143	0.186
c_2	0.357	0.077	7.714	0.229	8.143	0.186	7.429	0.257	7.429	0.257
c_3	0.368	0.080	7.571	0.243	8.286	0.171	6.571	0.343	8.714	0.129
D	**0.213**		**7.960**	**0.204**	**7.990**	**0.231**	**8.050**	**0.195**	**7.450**	**0.255**
d_1	0.344	0.073	8.143	0.186	7.429	0.257	8.857	0.114	7.286	0.271
d_2	0.376	0.080	7.857	0.214	7.286	0.271	7.571	0.243	7.286	0.271
d_3	0.279	0.059	7.857	0.214	8.571	0.143	7.714	0.229	7.857	0.214
E	**0.134**		**8.150**	**0.185**	**8.000**	**0.200**	**8.000**	**0.200**	**8.010**	**0.199**
e_1	0.513	0.069	8.429	0.157	7.857	0.214	8.143	0.186	8.286	0.171
e_2	0.487	0.065	7.857	0.214	8.143	0.186	7.857	0.214	7.714	0.229
Total performance			7.850	–	7.770	–	7.570	–	7.800	–
Total gap (S_z)			–	0.215	–	0.223	–	0.243	–	0.220

Note: For example, CPA firm A_1, Audit report (E): 0.185 = (0.513×0.157) + (0.487×0.214), and total gap-ratio value: 0.215 = (0.225×0.253) + (0.212×0.236) + (0.217×0.226) + (0.213×0.204) + (0.134×0.185). The gap-ratio S_z is calculated by $L_z^{h=1} = S_z = \sum_{j=1}^{n} w_j r_{zj} = \sum_{j=1}^{n} w_j (|f_j^{aspiration} - f_{zj}|)/(|f_j^{aspiration} - f_j^{worst}|)$ for CPA firms (alternatives) $z = 1, 2, ..., m$ and criteria $j = 1, 2, ..., n$.

firms A_1 and A_2, while firms A_3 and A_4 are the "Substantive tests and evaluation (D)" and "Understanding internal control (B)," respectively. From the assessment criteria, the Big 4 CPA firms present different results, with firms A_1, A_2, A_3, and A_4 for "Data provided by the enterprise to the AI system (a_2)," "AI-based control monitoring system (c_1)," "Process mining (c_3)," and "AI calculates audit fees and prepares contract (a_3)" showing the highest gaps, respectively. The results also reveal that the firms have already adopted AI-enabled auditing for corporates and have different weakest aspects in experts' opinions. Accordingly, our model can also aid companies to achieve the given targeted/desired level for each criterion and for improving overall performance.

2.5. Discussions and implications

This study has examined the information congregated from the experiences and knowledge of domain experts using the DEMATEL method to construct IINRM (Figure 5). The cause-effect relationships among the systems (dimensions) and sub-systems (criteria) are illustrated in Figure 5 for assessing AI-enabled auditing for CPA firms. According to the magnitude of the influence, the priority for improvement of dimensions is: understanding internal control (B), audit report (E), substantive tests and evaluation (D), pre-plan and contract (A), and control risk assessment (C). The results also indicate that understanding internal control (B) reveals the most important and immediate influence relationship on other dimensions – that is, if the CPA firms choose "understanding internal control" as a priority improvement objective, then this will have a multiplication effect on audit quality. The main purpose of internal control is to ameliorate the capacity of risk prevention, and it requires involvement of top-level management in enterprises, thereby creating and realizing operational efficiencies and financial performance. Effective internal controls can strengthen the risk prevention and control mechanisms by identifying, assessing, managing, and controlling risks and by managing corporate risks within the defined risk capacity and risk tolerance, thereby reducing risk losses without eroding corporate profit. To avoid audit risks that can emerge via major financial misstatements, underreporting, or false statements made by a company, it is essential for CPA firms/auditors to understand the internal control systems of the enterprise (the audited entity).

A wide range of non-traditional data sources is available that have a profound impact on the value of enterprises. However, the identification of risk factors is the most important part of internal control. For CPA firms, traditional audit techniques are unable to collect complete audit evidence beyond traditional data, conduct comprehensive audit reviews, or control audit risk to an acceptable level. According to COSO (Committee of Sponsoring Organizations) of the Treadway Commission's Internal Control-Integrated Framework (2013), we can understand that the internal control system of an enterprise is complex and uncontrollable. Therefore, CPA firms have a clear understanding of the internal control system when they sign an audit engagement with the audited entity, which can effectively prevent high levels of fraud attempts. It is the most important dimension in AI-enabled auditing. The criteria also have the same network influence effect on individual dimensions, such as AI calculating audit fees and the prepared contract (a_3), visualization and pattern recognition (b_2), AI-based control monitoring system (c_1), key audit matters (d_3), and continuous audit report (e_2).

Among all the criteria, key audit matters (d_3) has the highest importance and influence on other criteria. It indicates that when AI-enabled auditing is built and adopted by CPA firms, KAMs (key audit matters) are the most important during the audit procedures. According to the Auditing Standard ASA 701 released by International Auditing and Assurance Standards Board [IAASB] (2015), CPA firms are required to communicate KAMs in the auditor's report attached to the company's financial statements. Highly relevant matters are first identified in the audit report, such as significant misrepresentation of risks, accounting estimates of high estimation uncertainty, and events or transactions that have a significant impact on the audit, which are listed as KAMs. However, due to the increasing complexity of financial information, corporate information disclosure is growing (KPMG, 2011), and it is difficult for users to extract key information from the excess information available in financial statements (Public Company Accounting Oversight Board [PCAOB], 2014).

Based on understanding the industry and the company, the planning process for audit risk assessment, and major flaws found in the verification process, from their professional judgment CPAs put forward KAMs that are unique to the company and the user. Accordingly, the CPA's audit report is no longer immutable and can provide customized information content for clients, thus echoing the new trend of the external audit report in recent years. In addition, compared with past audit reports, the audit report from joining KAMs can apply the responsibility of the company (client) and the CPA to the text and clearly define the responsibilities for each other. Thus, KAMs provide better reliability and conciseness than the disclosure of underlying financial statements. This can increase users' attention to prominent information and utilize it as a substitute for corporate information disclosure (IAASB, 2012). With KAMs, users can still have better navigation capabilities even for complex financial reports. The finding is consistent with the previous studies. Furthermore, CPA firms can consider the relationships among multiple criteria and initiate improvement strategies (see Table 8) to advance an integrated AI-enabled audit framework.

Table 8. AI-enabled auditing implementation improvement plan for CPA firms based on DEMATEL method

Formula	Strategy (the priority of improvement targets)
F1: Interactive influential relationship of dimensions	$B\_E\_D\_A\_C$
F2: Interactive influential relationship of criteria within individual dimensions	B: $(b_2)\_(b_1)\_(b_3)$ E: $(e_2)\_(e_1)$ D: $(d_3)\_(d_1)\_(d_2)$ A: $(a_3)\_(a_1)\_(a_2)$ C: $(c_1)\_(c_2)\_(c_3)$

Note: 1. Dimensions: Pre-plan and contract (*A*), Understanding internal control (*B*), Control risk assessment (*C*), Substantive tests and evaluation (*D*), Audit report (*E*). 2. Criteria (factors): AI collects and analyzes data (a_1), Data provided by the enterprise to the AI system (a_2), AI calculates audit fees and prepared contract (a_3), Text mining and image recognition (b_1), Visualization and pattern recognition (b_2), Identify fraud and illegal acts (b_3), AI-based control monitoring system (c_1), Log generation (c_2), Process mining (c_3), Data provenance and data quality (d_1), Details of balances (d_2), Key audit matters (d_3), Predictive model (e_1), Continuous audit report (e_2).

Conclusions

This research proposes a synthesized improvement strategy for AI-enabled external auditing that can serve as a reference when CPA firms adopt AI-enabled auditing techniques. A promising decision model formulated herein integrates FRST with ACO, DEMATEL, D-ANP, and modified-VIKOR methods, so as to address the inter-relationships on IINRM between the influenced criteria/factors. FRST with ACO not only has the ability to handle data with uncertainty and vagueness and to determine the optimal feature subset without impeding its discriminant ability, but also can prevent loss of information encountered by RST as well as provide suitable suggestions for users to form an appropriate judgment. The selected core criteria are fed into the DEMATEL technique to construct IINRM, which can be used to navigate the evaluation and improvement direction based on the influential network relation map of the AI-enabled auditing framework. The D-ANP (DEMATEL-based ANP) approach simultaneously considers the problems of interdependence and feedback relationships among the criteria to measure the mutual influential weights of the criteria. The modified-VIKOR method measures the total performance and criteria gaps for the adoption of AI-enabled auditing techniques by China's Big 4 CPA firms, with the expectation for how to reduce the gaps between actual and aspiration performance when targeting a goal level based on the proposed technique.

For dimensions, the magnitude of this finding is based on diverse expert opinions, and the improvement priorities include understanding internal control, audit reports, substantive tests, evaluation, pre-plan and contract, and control risk assessment. The synthetic methodology described herein can be used to resolve and evaluate complex dynamic tasks related to the adoption of AI-enabled auditing techniques in China. The research findings not only provide implications for senior supervisors, but their insights can also help in creating pragmatic problem-solving strategies and structures in the era of AI, by assisting through effective, efficient, and intelligent audit approaches. This generalized model can be extended to other countries to aid CPA firms in constructing their best AI-based auditing architecture.

While this study has built a successful evaluation practice-audit model for CPA firms, there are still some interesting ideas worth developing for future research. The evaluation framework proposed herein is based on the American Institute of CPA (AICPA) reports, and the selection of experts is not comprehensive enough. It is recommended that future research consider the choice of more samples and examples (top 10 vs. other CPA firms) so that the development of AI can become broader and more mature to provide reliable and versatile results. Future research should make comparisons of this approach with other ranking methods not only as a means to see how it outperforms against other rankings, but also in terms of identifying the appropriateness of the AI-based audit problem discussed here for reaching a target level.

Acknowledgements

References

American Institution of Certified Public Accountants. (2015a). *Audit Data Standards-Based Standard.*

American Institution of Certified Public Accountants. (2015b). *Audit Data Standards-General Ledger Standard.*

American Institution of Certified Public Accountants. (2015c). *Audit Data Standards-Order to Cash Sub-ledger Standard.*

American Institution of Certified Public Accountants. (2015d). *Audit Data Standards – Procure to Pay Standard.*

Appelbaum, D. A., Kogan, A., & Vasarhelyi, M. A. (2018). Analytical procedures in external auditing: A comprehensive literature survey and framework for external audit analytics. *Journal of Accounting Literature*, *40*, 83–101. https://doi.org/10.1016/j.acclit.2018.01.001

Baldwin, A. A., Brown, C. E., & Trinkle, B. S. (2006). Opportunities for artificial intelligence development in the accounting domain: The case for auditing. *Intelligent Systems in Accounting, Finance and Management*, *14*, 77–86. https://doi.org/10.1002/isaf.277

Bizarro, P. A., & Dorian, M. (2017). Artificial intelligent: the future of auditing. *Internal Auditing*, (September/October), 21–26.

Bonabeau, E., Dorigo, M., & Theraulez, G. (1999). *Swarm intelligence: From natural to artificial systems*. Oxford University Press Inc.

Burgess, A. (2017). *The executive guide to artificial intelligence-How to identify and implement applications for AI in your organization*. Palgrave Macmillan.

Cangemi, M. P., & Taylor, P. (2018). Harnessing artificial intelligence to deliver real-time intelligence and business process improvements. *EDPACS*, *57*(4), 1–6. https://doi.org/10.1080/07366981.2018.1444007

Chen, F. S. (2015). Application of a hybrid dynamic MCDM to explore the key factors for the internal control of procurement circulation. *International Journal of Production Research*, *53*(10), 2951–2969. https://doi.org/10.1080/00207543.2014.961210

Chen, M., Zhang, S., Zhang, W., & Lin, J. (2019). Collaborative vehicle routing problem with rough location using extended ant colony optimization algorithm. *Journal of Intelligent* & *Fuzzy Systems*, *37*(2), 2385–2402. https://doi.org/10.3233/JIFS-182715

Cheng, Y. (2018). Dynamic maintenance of approximations under fuzzy rough sets. *International Journal of Machine Learning and Cybernetics*, *9*, 2011–2026. https://doi.org/10.1007/s13042-017-0683-7

Chen, F. S., & Chi, D. J. (2015). Application of a new DEMATEL to explore key factors of China's corporate social responsibility: evidence from accounting experts. *Quality* & *Quantity*, *49*(1), 135–154. https://doi.org/10.1007/s11135-013-9978-2

Çolak, M., Kaya, İ., Özkan, B., Budak, A., & Karaşan, A. (2020). A multi-criteria evaluation model based on hesitant fuzzy sets for blockchain technology in supply chain management. *Journal of Intelligent* & *Fuzzy Systems*, *38*(1), 935–946. https://doi.org/10.3233/JIFS-179460

Davenport, T. H., & Raphael, J. (2017). Creating a cognitive audit. *CFO.Com.*

Deloitte. (2015). *Disruption ahead: Deloitte's point of view on IBM Watson*. https://www2.deloitte.com/content/dam/Deloitte/us/Documents/about-deloitte/us-ibm-watson-client.pdf

Dubois, D., & Prade, H. (1980). *Fuzzy sets and systems: Theory and applications*. Academic Press.

Dubois, D., & Prade, H. (1990). Rough fuzzy sets and fuzzy rough sets. *International Journal of General Systems*, *17*(2–3), 191–209. https://doi.org/10.1080/03081079008935107

Du, S., Li, H., & Sun, B. (2019). Hybrid Kano-fuzzy-DEMATEL model based risk factor evaluation and ranking of cross-border e-commerce SMEs with customer requirement. *Journal of Intelligent* & *Fuzzy Systems*, *37*(6), 8299–8315. https://doi.org/10.3233/JIFS-190830

Ding, X. F., & Liu, H. C. (2018). A 2-dimension uncertain linguistic DEMATEL method for identifying critical success factors in emergency management. *Applied Soft Computing*, *71*, 386–395. https://doi.org/10.1016/j.asoc.2018.07.018

Faria, E. R., Gonçalves, I. J. C. R., de Carvalho, A. C. P. L. F., & Gama, J. (2016). Novelty detection in data streams. *Artificial Intelligence Review*, *45*(2), 235–269. https://doi.org/10.1007/s10462-015-9444-8

Faggella, D. (2018). *AI in the accounting big four – comparing Deloitte, PwC, KPMG, and EY*. Emerj.

Gupta, S., Upadhyaya, V., Singh, A., Varshney, P., & Srivastava, S. (2018). Modeling of fractional order chaotic systems using artificial bee colony optimization and ant colony optimization. *Journal of Intelligent & Fuzzy Systems*, *35*(5), 5337–5344. https://doi.org/10.3233/JIFS-169816

Gopal, J., Sangaiah, A. K., Basu, A., & Gao, X. Z. (2018). Integration of fuzzy DEMATEL and FMCDM approach for evaluating knowledge transfer effectiveness with reference to GSD project outcome. *International Journal of Machine Learning and Cybernetics*, *9*, 225–241. https://doi.org/10.1007/s13042-015-0370-5

Hooda, N., Bawa, S., & Rana, P. S. (2018). Fraudulent firm classification: A case study of an external audit. *Applied Artificial Intelligence*, *32*, 48–64. https://doi.org/10.1080/08839514.2018.1451032

Hsu, Y. S., & Lin, S. J. (2016). An emerging hybrid mechanism for information disclosure forecasting. *International Journal of Machine Learning and Cybernetics*, *7*, 943–952. https://doi.org/10.1007/s13042-014-0295-4

Hsu, M. F. (2019a). A fusion mechanism for management decision and risk analysis. *Cybernetics and Systems*, *50*(6), 497–515. https://doi.org/10.1080/01969722.2018.1541596

Hsu, M. F. (2019b). Integrated multiple-attribute decision making and kernel-based mechanism for risk analysis and evaluation. *Journal of Intelligent & Fuzzy Systems*, *36*(3), 2895–2905. https://doi.org/10.3233/JIFS-171366

Hsu, M. F., Chang, T. M., & Lin, S. J. (2020). News-based soft information as a corporate competitive advantage. *Technological and Economic Development of Economy*, *26*(1), 48–70. https://doi.org/10.3846/tede.2019.11328

Hu, K. H., Wei, J., & Tzeng, G. H. (2017). Risk Factor Assessment improvement for China's cloud computing auditing using a new hybrid MADM model. *International Journal of Information Technology and Decision Making*, *16*(3), 737–777. https://doi.org/10.1142/S021962201750016X

Hu, K. H., Wei, J., & Tzeng, G. H. (2018). Improving China's regional financial center modernization development using a new hybrid MADM model. *Technological and Economic Development of Economy*, *24*(2), 429–466. https://doi.org/10.3846/20294913.2016.1213195

Huang, J. Y., Shen, K. Y., Shieh, J. C. P., & Tzeng, G. H. (2019). Strengthen financial holding companies' business sustainability by using a hybrid corporate governance evaluation model. *Sustainability*, *11*(3), 582. https://doi.org/10.3390/su11030582

Hu, S. K., & Tzeng, G. H. (2019). A hybrid multiple-attribute decision-making model with modified promethee for identifying optimal performance-improvement strategies for sustainable development of a better life. *Social Indicators Research*, *144*, 1021–1053. https://doi.org/10.1007/s11205-018-2033-x

International Auditing and Assurance Standards Board. (2012). *Invitation to Comment: Improving the Auditor's Report*. International Federation of Accountants.

International Auditing and Assurance Standards Board. (2015). *International Standard on Auditing (ISA) 701 (NEW), Communicating Key Audit Matters in the Independent Auditor's Report*.

Issa, H., Sun, T., & Vasarhelyi, M. A. (2016). Research ideas for artificial intelligence in auditing: The formalization of audit and workforce supplementation. *Journal of Emerging Technologies in Accounting*, *13*(2), 1–20. https://doi.org/10.2308/jeta-10511

Issa, H., & Kogan, A. (2014). A predictive ordered logistic regression model as a tool for quality review of control risk assessments. *Journal of Information Systems*, *28*(2), 209–229. https://doi.org/10.2308/isys-50808

Jans, M., Alles, M., & Vasarhelyi, M. (2014). A field study on the use of process mining of event logs as an analytical procedure in auditing. *Accounting Review*, *89*(5), 1751–1773. https://doi.org/10.2308/accr-50807

Jensen, R., & Shen, Q. (2005). Fuzzy-rough data reduction with ant colony optimization. *Fuzzy Sets and Systems*, *149*(1), 5–20. https://doi.org/10.1016/j.fss.2004.07.014

Jensen, R., & Shen, Q. (2009). New approaches to fuzzy-rough feature selection. *IEEE Transactions on Fuzzy Systems*, *17*(4), 824–838. https://doi.org/10.1109/TFUZZ.2008.924209

Jensen, R., Tuson, A., & Shen, Q. (2014). Finding rough and fuzzy-rough set reducts with SAT. *Information Sciences*, *255*, 100–120. https://doi.org/10.1016/j.ins.2013.07.033

Jing, M., Jie, Y., Shou-yi, L., & Lu, W. (2018). Application of fuzzy analytic hierarchy process in the risk assessment of dangerous small-sized reservoirs. *International Journal of Machine Learning and Cybernetics*, *9*, 113–123. https://doi.org/10.1007/s13042-015-0363-4

Kang, M., Kim, J. W., Lee, H. Y., & Lee, M. G. (2015). Financial statement comparability and audit efficiency: Evidence from South Korea. *Applied Economic*, *47*(4), 358–373. https://doi.org/10.1080/00036846.2014.972543

Keenoy, C. L. (1958). The impact of automation on the field of accounting. *Accounting Review*, *33*(2), 230–236.

Ke, L., Feng, Z., & Ren, Z. (2008). An efficient ant colony optimization approach to attribute reduction in rough set theory. *Pattern Recognition Letters*, *29*(9), 1351–1357. https://doi.org/10.1016/j.patrec.2008.02.006

Kokina, J., & Davenport, T. H. (2017). The emergence of artificial intelligence: how automation is changing auditing. *Journal of Emerging Technologies in Accounting*, *14*(1), 115–122. https://doi.org/10.2308/jeta-51730

KPMG. (2011). *Disclosure overload and complexity: hidden in plain sight*. Financial Executives Research Foundation, Inc.

KPMG. (2016). *Game changer: The impact of cognitive technology on business and financial reporting.*

Lam, M. (2004). Neural network techniques for financial performance prediction: integrating fundamental and technical analysis. *Decision Support Systems*, *37*(4), 567–581. https://doi.org/10.1016/S0167-9236(03)00088-5

Lin, P. J, Shiue, Y. C., Tzeng, G. H., & Huang, S. L. (2019). Developing a sustainable long-term ageing health care system using the DANP-mV model: Empirical case of Taiwan. *International Journal of Environmental Research and Public Health*, *16*(8), 1349. https://doi.org/10.3390/ijerph16081349

Lin, S. J. (2017). Integrated artificial intelligence-based resizing strategy and multiple criteria decision making technique to form a management decision in an imbalanced environment. *International Journal of Machine Learning and Cybernetics*, *8*, 1981–1992. https://doi.org/10.1007/s13042-016-0574-3

Lin, S. J., & Hsu, M. F. (2018). Decision making by extracting soft information from CSR news report. *Technological and Economic Development of Economy*, *24*(4), 1344–1361. https://doi.org/10.3846/tede.2018.3121

Lin, S. J., Chang, T. M., & Hsu, M. F. (2019). An emerging online business decision making architecture in a dynamic economic environment. *Journal of Intelligent* & *Fuzzy Systems*, *37*(2), 1893–1903. https://doi.org/10.3233/JIFS-179251

Liou, J. J. H., & Tzeng, G. H. (2012). Comments on "Multiple criteria decision making (MCDM) methods in economics: an overview". *Technological and Economic Development of Economy*, *18*(4), 672–695. https://doi.org/10.3846/20294913.2012.753489

Lo, H. W, Liou, J. J. H, & Tzeng, G. H. (2019). Comments on "Sustainable recycling partner selection using fuzzy DEMATEL-AEW-FVIKOR: A case study in small-and-medium enterprises". *Journal of Cleaner Production*, *228*, 1011–1012. https://doi.org/10.1016/j.jclepro.2019.04.376

M2 Presswire. (2016). *PwC Wins "Audit Innovation of the Year" at the accountant & international accounting bulletin awards*. https://www.m2.com/m2/web/story.php/20166219039

Meskovic, E., Garrison, M., Ghezal, S., & Chen, Y. (2018). Artificial intelligence: Trends in business and implications for the accounting profession. *Internal Auditing*, (May/June), 5–11.

Opricovic, S., & Tzeng, G. H. (2004). Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS. *European Journal of Operational Research*, *156*, 445–455. https://doi.org/10.1016/S0377-2217(03)00020-1

Parthaláin, N. M., & Jensen, R. (2010). Measures for unsupervised fuzzy-rough feature selection. *International Journal of Hybrid Intelligent Systems*, *7*, 249–259. https://doi.org/10.3233/HIS-2010-0118

Parthaláin, N. M., & Jensen, R. (2013). Unsupervised fuzzy-rough set-based dimensionality reduction. *Information Sciences*, *229*, 106–121. https://doi.org/10.1016/j.ins.2012.12.001

Pawlak, Z. (1982). Rough sets. *International Journal of Computer and Information Sciences*, *11*(5), 341–356. https://doi.org/10.1007/BF01001956

Peng, K. H., & Tzeng, G. H. (2013). A hybrid dynamic MADM model for problems-improvement in economics and business. *Technological and Economic Development of Economy*, *19*(4), 638–660. https://doi.org/10.3846/20294913.2013.837114

Peng, K. H., & Tzeng, G. H. (2019). Exploring heritage tourism performance improvement for making sustainable development strategies using the hybrid-modified MADM model. *Current Issues in Tourism*, *22*(8), 921–947. https://doi.org/10.1080/13683500.2017.1306030

Public Company Accounting Oversight Board. (2014). *Transcript of the Public Company Accounting Oversight Board's April 2, 2014 Public Meeting on the Auditor's Reporting Model* [*unofficial*]. Public Company Accounting Oversight Board, 381, Washington, D.C.

PwC. (2014). *Sizing the prize – What's the real value of AI for your business and how can you capitalise*?

Qu, G. B., Zhao, T. Y., Zhu, B. W., Tzeng, G. H., & Huang, S. L. (2019). Use of a modified DANP-mV model to improve quality of life in rural residents: The empirical case of xingshisi village, China. *International Journal of Environmental Research and Public Health*, *16*(1), 153. https://doi.org/10.3390/ijerph16010153

Quick, R., & Henrizi, P. (2018). Review of managerial science experimental evidence on external auditor reliance on the internal audit. *Review of Managerial Science*, *13*, 1143–1176. https://doi.org/10.1007/s11846-018-0285-0

Radzikowska, A. M., & Kerre, E. E. (2002). A comparative study of fuzzy rough sets. *Fuzzy Sets and Systems*, *126*(2), 137–155. https://doi.org/10.1016/S0165-0114(01)00032-X

Rehman, A., & Saba, T. (2014). Evaluation of artificial intelligent techniques to secure information in enterprises. *Artificial Intelligence Review*, *42*, 1029–1044. https://doi.org/10.1007/s10462-012-9372-9

Roscoe, P., & Howorth, C. (2009). Identification through technical analysis: A study of charting and UK non-professional investors. *Accounting, Organizations and Society*, *34*(2), 206–221. https://doi.org/10.1016/j.aos.2008.05.003

Saaty, T. L. (1996). *Decision making with dependence and feedback: Analytic network process*. RWS Publications.

Salarpour, H., Ghodrati A. G., & Meysam, M. (2019). A hierarchical group decision approach based on DEMATEL and dynamic hesitant fuzzy sets to evaluate sustainability criteria for strategic management of housing market problem. *Journal of Intelligent & Fuzzy Systems*, *37*(1), 821–833. https://doi.org/10.3233/JIFS-181482

Shen, K. Y., & Tzeng, G. H. (2018). Advances in multiple criteria decision making for sustainability: Modeling and applications. *Sustainability*, *10*(5), 1600. https://doi.org/10.3390/su10051600

Shen, K. Y., Sakai, H., & Tzeng, G. H. (2019). Comparing two novel hybrid MRDM approaches to consumer credit scoring under uncertainty and fuzzy judgments. *International Journal of Fuzzy Systems*, *21*(1), 194–212. https://doi.org/10.1007/s40815-018-0525-0

Shen, K. Y., Zavadskas, E. K., & Tzeng, G. H. (2018). Updated discussions on 'Hybrid multiple criteria decision-making methods: a review of applications for sustainability issues'. *Economic Research-Ekonomska Istraživanja*, *31*(1), 1437–1452. https://doi.org/10.1080/1331677X.2018.1483836

Skowron, A., & Rauszer, C. (1992). The discernibility matrices and functions in information systems. In R. Slowinski (Ed.), *Intelligent decision support: Handbook of applications and advances of rough sets theory* (pp. 331–362). Kluwer Academic Publishers. https://doi.org/10.1007/978-94-015-7975-9_21

Siemiński, A., & Kopel, M. (2019). Solving dynamic TSP by parallel and adaptive ant colony communities. *Journal of Intelligent & Fuzzy Systems*, *37*(6), 7607–7618. https://doi.org/10.3233/JIFS-179366

Sinclair, N (2015, October 27). How KPMG is using Formula 1 to transform audit. *CA Today*.

Sirois, L. P., Bédard, J., & Bera, P. (2018). The informational value of key audit matters in the auditor's report: evidence from an Eye-tracking study. *Accounting Horizons*, *32*(2), 141–162. https://doi.org/10.2308/acch-52047

Si, S. L., You, X. Y., Liu, H. C., & Huang, J. (2017). Identifying key performance indicators for holistic hospital management with a modified DEMATEL approach. *International Journal of Environmental Research and Public Health*, *14*(8), 934. https://doi.org/10.3390/ijerph14080934

Sneller, L., Bode, R., & Klerkx, A. (2016). Do IT matters matter? IT-related key audit matters in Dutch annual reports. *International Journal of Disclosure and Governance*, *14*(2), 139–151. https://doi.org/10.1057/s41310-016-0017-0

Sutton, S. G., Holt, M., & Arnold, V. (2016). The reports of my death are greatly exaggerated – Artificial intelligence research in accounting. *International Journal of Accounting Information Systems*, *22*, 60–73. https://doi.org/10.1016/j.accinf.2016.07.005

Thangavel, K., Karnan, M., & Pethalakshmi, A. (2005). Performance analysis of rough reduct algorithms in mammogram. *International Journal on Global Vision and Image Processing*, *5*(8), 13–21.

Thibodeau, J. C. (2003). The development and transferability of task knowledge. *Auditing: A Journal of Practice and Theory*, *22*(1), 47–67. https://doi.org/10.2308/aud.2003.22.1.47

Tzeng, G. H., & Shen, K. Y. (2017). *New concepts and trends of hybrid multiple criteria decision making*. CRC Press, Taylor & Francis Group, Chapman & Hall Book. https://doi.org/10.1201/9781315166650

Uthayakumar, J., Metawa, N., Shankar, K., & Lakshmanaprabu, S. K. (2018). Financial crisis prediction model using ant colony optimization. *International Journal of Information Management*, *50*, 538–556. https://doi.org/10.1016/j.ijinfomgt.2018.12.001

Wang, X. Y., Choi, T. M., Liu, H. K., & Yue, X. H. (2018). A novel hybrid ant colony optimization algorithm for emergency transportation problems during post-disaster scenarios. *IEEE Transactions on Systems Man Cybernetics-Systems*, *48*(4), 545–556. https://doi.org/10.1109/TSMC.2016.2606440

Wang, W., Guyet, T., Quiniou, R., Cordier, M. O., Masseglia, F., & Zhang, X. (2014). Autonomic intrusion detection: Adaptively detecting anomalies over unlabeled audit data streams in computer networks. *Knowledge-Based Systems*, *70*, 103–117. https://doi.org/10.1016/j.knosys.2014.06.018

Wang, J. Q., Cao, Y. X., & Zhang, H. Y. (2017). Multi-criteria decision-making method based on distance measure and Choquet integral for linguistic z-numbers. *Cognitive Computation*, *9*(6), 827–842. https://doi.org/10.1007/s12559-017-9493-1

Whitehouse, T. (2015, December 1). The technology transforming your annual audit. *Compliance Week*.

Zhou, J., Wang, Q., Tsai, S. B., Xue, Y. Z., & Dong, W. W. (2016). How to evaluate the job satisfaction of development personnel. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *47*(11), 2809–2816. https://doi.org/10.1109/TSMC.2016.2519860

APPENDIX

A fusion multiple rule-based decision making (MRDM) model

Stage 1: Fuzzy rough set theory with ant colony optimization: FRST-ACO

It is often desirable to collect as much information as possible for any given domain, so as to realize the underlying realities of that domain (Parthaláin & Jensen, 2013; Hsu, 2019b). However, most information from annual reports or narrative documents may be contaminated by some degree of errors, which in turn can confuse users and lead to biased decisions when performing AI techniques for information retrieval. Feature selection (FS) (an information retrieval technique) looks to determine the feature subsets that are rich in information and to retain the original semantics of the features following the reduction. Removing noisy, irrelevant, or redundant information improves data quality and leads to a more reliable and precise outcome. Introduced by Pawlak (1982), rough set theory (RST) is a mathematical approach that handles data with uncertain, vague, and imprecise characteristics and can deal with the FS task, which has gained so much success in numerous domains. A brief description of RST runs as follows.

We let $H = (\mathfrak{R}, N)$ be an information system, $\mathfrak{R}$ denotes a non-empty set of finite objects, and N denotes a non-empty finite set of attributes, such that $b : \mathfrak{R} \to V_b$ for every $b \in N$. For any attribute set $P \subseteq N$, the associated indiscernibility relation can be defined by:

$$IND(P) = \left\{ (x, y) \in \mathfrak{R}^2 \,\middle|\, \forall b \in P, b(x) = b(y) \right\}. \tag{A.1}$$

Based on $IND(\mathrm{P})$, the partition of $\mathfrak{R}$ can be determined as $\mathfrak{R} / IND(P)$. If $(x, y) \in IND(\mathrm{P})$, then x and y are indispensable by attribute from P. We denote an equivalence class of P–indiscernibility relation containing x as $\left[x \right]_P$. The indiscernibility relation is the fundamental concept of RST (Parthaláin & Jensen, 2013).

RST has two important operations: lower and upper approximations. Given an arbitrary set $X \subseteq \mathfrak{R}$, we define the P-lower approximation of X as $\underline{P}X$, which is the set of all elements of $\mathfrak{R}$ that can be precisely discriminated as elements of X relying on the attribute set P. The P-upper approximation of X, denoted as $\overline{P}X$, is the set of all elements of $\mathfrak{R}$ that can be possibly discriminated as elements of X by relying on the attribute set P (Jensen & Shen, 2005). The mathematical formulations of these two concepts are as follows.

$$\underline{P}X = \left\{ x \,\middle|\, [\mathrm{x}]_P \subseteq X \right\}; \tag{A.2}$$

$$\overline{P}X = \left\{ x \,\middle|\, [\mathrm{x}]_P \cap X \neq \phi \right\}. \tag{A.3}$$

Let P and G be the equivalence relation over $\mathfrak{R}$. Thus, the positive region can be determined.

$$POS_P(\mathrm{G}) = \bigcup_{X \in \mathfrak{R}/G} \underline{P}X. \tag{A.4}$$

The positive region contains all objects of $\mathfrak{R}$ that can be accurately classified to the classes of $\mathfrak{R} / G$ by means of the information in attribute P. Based on this concept, we determine the degree of dependency between P and G.

$$k = r_p(\mathrm{G}) = \left| POS_p(\mathrm{G}) \right| / \left| \mathfrak{R} \right|. \tag{A.5}$$

Attribute reduction (or feature selection) can be achieved through a comparison of the degree of dependency generated by sets of attributes.

Although RST has numerous benefits (such as it requires no external information and no human input), it still comes with some weakness: (1) it cannot handle real-valued data, and (2) it incurs considerable computational costs to reach the minimal reduct (that is, minimal feature subsets) (Ke et al., 2008; Jensen et al., 2014). In order to tackle these problems, a number of extensions to RST has emerged. To prevent information loss due to data discretization, fuzzy rough set theory (FRST) encapsulates the membership function from fuzzy set theory (FST) in RST in order to extend its flexibility in handing uncertainty (Jensen and Shen, 2005). Radzikowska and Kerre (2002) identified the fuzzy lower and upper approximations, and a τ-transitive fuzzy similarity relation is then applied to depict the fuzzy concept X.

$$\underline{\theta R_p} X(x) = \underset{y \in \mathfrak{R}}{Inf}\, \pi(\theta_{R_p}(x,y), \theta_X(y)); \tag{A.6}$$

$$\overline{\theta R_p} X(x) = \underset{y \in \mathfrak{R}}{Sup}\, \tau(\theta_{R_p}(x,y), \theta_X(y)). \tag{A.7}$$

The fuzzy implicator and t-norm are denoted here as π and τ, respectively. On the basis of features P, the fuzzy similarity relation can be defined as R_P.

$$R_P(x,y) = \tau_{a \in P}\left\{\theta_{R_a}(x,y)\right\}. \tag{A.8}$$

In the same way as conventional RST, the fuzzy positive region is represented in Eq. (9) (Jensen & Shen, 2009).

$$\theta_{pos_p(B)}(x) = \underset{X \in \mathfrak{R}/B}{\sup}\, \theta_{R_p X}(x). \tag{A.9}$$

From attribute subset P, we express the fuzzy-rough degree of dependency of B as:

$$r'_P(B) = \frac{\sum_{x \in \mathfrak{R}} \theta_{pos_p(B)}(x)}{\left|\mathfrak{R}\right|}. \tag{A.10}$$

Based on Eq. (10), we construct the minimal subset of features (i.e., minimal reduct) by performing the hill-climbing method. As with the original method (i.e., hill-climbing method), for a dimensionality of n, the worst case dataset results in $(n^2 + n)/2$ assessments of the dependency function. This method is not appropriate for minimal feature subset selection (i.e., feature selection) when the dataset's dimensionality is very large (Jensen & Shen, 2005). If the minimal feature subset selection can be transformed into a combinatorial optimization problem, then the meta-heuristic algorithm can be applied to combat this.

This study executed a meta-heuristic algorithm, ant colony optimization (ACO), due to its superior searching capability and considerable generalization ability (Gupta et al., 2018; Chen et al., 2019; Siemiński & Kopel, 2019). We take the dependency measure as an assessing measurement to determine which attribute subsets are more informative than the others. We then utilize the heuristic desirability of traversal and edge pheromone levels to form probabilistic transition rules (Bonabeau et al., 1999), representing the probability of an ant at attribute i choosing to travel to attribute k at time t.

$$q_{ij}^k(t) = \frac{[\omega_{ij}(t)]^\alpha \cdot [\varepsilon_{ij}]^\beta}{\sum_{\sigma \in J_i^k} [\omega_{i\sigma}(t)]^\alpha \cdot [\varepsilon_{i\sigma}]^\beta}. \tag{A.11}$$

In the above equation, k denotes the number of ants, J_i^k denotes the set of ant k's unvisited attributes, ε_{ij} denotes the heuristic desirability of selecting attribute j when at attribute i, and $\omega_{ij}(t)$ denotes the amount of pheromone on edge (i,j). We decide α and β experimentally. The pheromone on each edge is updated based on the following equation.

$$\omega_{ij}(t+1) = (1-\xi) \cdot \omega_{ij}(t) + \Delta\omega_{ij}(t); \tag{A.12}$$

Here, $\Delta\omega_{ij}(t) = \sum_{k=1}^n (r'(D^k)/|D^k|)$.

If edge (i,j) has been traversed, then $\Delta\omega_{ij}(t)$ is 0. Moreover, ξ denotes a decay constant applied to simulate the evaporation of the pheromone, and D^k is the attribute subset found by ant k. Through joint utilization of FRST and ACO, the most important attributes can now be determined. More detailed illustrations of FRST and ACO can be seen in Jensen and Shen (2005), Ke et al. (2008), Parthalain and Jensen (2010), and Parthalain and Jensen (2013).

Stage 2: Measuring the interrelations among criteria using the DEMATEL technique

Step 1: Constructing an average matrix Y. The scoring ranges from 0 (absolutely no influence) to 4 (very high influence) and determine the degree of mutual influence using pairwise comparison. Then, an average matrix Y (called the initial direct relation matrix) was directly constructed through 60 domain experts in AI-enabled audit, as shown in Eq. (A.13):

$$Y = \begin{bmatrix} y_{11} & \cdots & y_{1j} & \cdots & y_{1n} \\ \vdots & & \vdots & & \vdots \\ y_{i1} & \cdots & y_{ij} & \cdots & y_{in} \\ \vdots & & \vdots & & \vdots \\ y_{n1} & \cdots & y_{nj} & \cdots & y_{nn} \end{bmatrix}. \tag{A.13}$$

Step 2: Finding the direct-effect matrix G. The normalized direct influence relationship G can be acquired from the average matrix $Y = [y_{ij}]_{n \times n}$ by using Eqs (A.14) and (A.15):

$$G = \pi \cdot Y; \tag{A.14}$$

$$\pi = \min\left\{ \frac{1}{\max_{1 \le i \le n} \sum_{j=1}^n y_{ij}}, \frac{1}{\max_{1 \le j \le n} \sum_{i=1}^n y_{ij}} \right\}. \tag{A.15}$$

Step 3: Calculate the total-influence matrix T. The total-influence relationship matrix $T = \left[t_{ij} \right]_{n \times n}$ is obtained using Eq. (A.16).

$$T = G + G^2 + \ldots + G^l = G(I - G), \text{ when } \lim_{l \to \infty} G^l = \left[\mathbf{0} \right]_{n \times n}, \tag{A.16}$$

where $G = [g_{ij}]_{n \times n}$, $0 \le g_{ij} < 1$, $0 < \sum_{i=1}^n g_{ij} \le 1$, and $0 < \sum_{j=1}^n g_{ij} \le 1$.

Accordingly, the result of the total-influence relation matrix T is obtained and presented in Table 3.

Step 4: Illustrating the IINRM based on the total-influence matrix T. The sum of each column vector $r = (r_i)_{n \times 1}$ ($\left[\sum_{j=1}^{n} t_{ij} \right]_{n \times 1} = (...,r_i,...)'$) and row vector $s = (s_j)_{n \times 1}$ ($\left[\sum_{i=1}^{n} t_{ij} \right]'_{1 \times n} = (...,s_j,...)'$) in the total-influence (direct and indirect) matrix T was applied to examine the relationship among criteria, when $i = j$ the IINRM is constructed.

The results of IINRM are shown in Table 4 and illustrated in Figure 5.

Stage 3: Deriving the influence weights of criteria/dimensions by the D-ANP

Step 1: Determining the un-weighted super-matrix W. First, the total influence relation matrix T_C is constructed in terms of each dimension (perspective or cluster), as shown in Eq. (A.17).

$$T_C = \begin{array}{c} \\ D_1 \\ \\ \\ D_i \\ \\ \\ D_m \end{array} \begin{array}{c} c_{11} \\ c_{12} \\ \vdots \\ c_{1m_1} \\ c_{i1} \\ c_{i2} \\ \vdots \\ c_{im_i} \\ c_{m1} \\ c_{m2} \\ \vdots \\ c_{mm_m} \end{array} \begin{bmatrix} T_C^{11} & ... & T_C^{1j} & ... & T_C^{1m} \\ \vdots & & \vdots & & \vdots \\ T_C^{i1} & ... & T_C^{ij} & ... & T_C^{im} \\ \vdots & & \vdots & & \vdots \\ T_C^{m1} & ... & T_C^{mj} & ... & T_C^{mm} \end{bmatrix}_{m_j \times m_j \mid m < n, \sum_{j=1}^{m} m_j = n} , \quad \text{(A.17)}$$

where T_C^{θ} is the submatrix of the total influence relation matrix. Next, the total influence relation matrix T_C is normalized, and the normalized total influence relation matrix T_C^{θ} with respect to the degree of total influence among criteria/dimensions can be derived from Eq. (A.18).

$$T_C^{\theta} = \begin{array}{c} \\ D_1 \\ \\ \\ D_i \\ \\ \\ D_m \end{array} \begin{array}{c} c_{11} \\ c_{12} \\ \vdots \\ c_{1m_1} \\ c_{i1} \\ c_{i2} \\ \vdots \\ c_{im_i} \\ c_{m1} \\ c_{m2} \\ \vdots \\ c_{mm_m} \end{array} \begin{bmatrix} T_C^{\theta 11} & ... & T_C^{\theta 1j} & ... & T_C^{\theta 1m} \\ \vdots & & \vdots & & \vdots \\ T_C^{\theta i1} & ... & T_C^{\theta ij} & ... & T_C^{\theta im} \\ \vdots & & \vdots & & \vdots \\ T_C^{\theta m1} & ... & T_C^{\theta mj} & ... & T_C^{\theta mm} \end{bmatrix}_{m_j \times m_j \mid m < n, \sum_{j=1}^{m} m_j = n} . \quad \text{(A.18)}$$

Then, by transposing the normalized total influence relation matrix T_C^{θ}, the un-weighted super-matrix $W = (T_C^{\theta})'$ is defined, as shown in Eq. (A.19).

$$W = (T_C^{\theta})' = \begin{array}{c} \\ D_1 \\ \\ \\ D_j \\ \\ \\ D_m \end{array} \begin{array}{c} c_{11} \\ c_{12} \\ \vdots \\ c_{1m_1} \\ c_{j1} \\ c_{j2} \\ \vdots \\ c_{jm_j} \\ c_{m1} \\ c_{m2} \\ \vdots \\ c_{mm_m} \end{array} \begin{bmatrix} W^{11} & ... & W^{i1} & ... & W^{m1} \\ \vdots & & \vdots & & \vdots \\ W^{1j} & ... & W^{ij} & ... & W^{mj} \\ \vdots & & \vdots & & \vdots \\ W^{1m} & ... & W^{im} & ... & W^{mm} \end{bmatrix}_{n \times n \mid m < n, \sum_{j=1}^{m} m_j = n} . \quad \text{(A.19)}$$

Step 2: Constructing the weighted super-matrix W^θ. First, the total influence relation matrix T_D is derived based on the dimensions using the DEMATEL technique, as shown in Eq. (A.20).

$$T_D = \begin{bmatrix} t_D^{11} & \cdots & t_D^{1j} & \cdots & t_D^{1m} \\ \vdots & & \vdots & & \vdots \\ t_D^{i1} & \cdots & t_D^{ij} & \cdots & t_D^{im} \\ \vdots & & \vdots & & \vdots \\ t_D^{m1} & \cdots & t_D^{mj} & \cdots & t_D^{mm} \end{bmatrix}_{m\times m}. \tag{A.20}$$

The normalizing the influential dimension matrix T_D^θ is obtained through each element divided by $d_i = \sum_{j=1}^{n} t_{ij}^{D}$ in this matrix, as shown as Eq. (A.21).

$$T_D^\theta = \begin{bmatrix} t_D^{11}/d_1 & \cdots & t_D^{1j}/d_1 & \cdots & t_D^{1m}/d_1 \\ \vdots & & \vdots & & \vdots \\ t_D^{i1}/d_i & \cdots & t_D^{ij}/d_i & \cdots & t_D^{im}/d_i \\ \vdots & & \vdots & & \vdots \\ t_D^{m1}/d_m & \cdots & t_D^{mj}/d_m & \cdots & t_D^{mm}/d_m \end{bmatrix}_{m\times m} = \begin{bmatrix} t_{11}^{\theta D} & \cdots & t_{1j}^{\theta D} & \cdots & t_{1m}^{\theta D} \\ \vdots & & \vdots & & \vdots \\ t_{i1}^{\theta D} & \cdots & t_{ij}^{\theta D} & \cdots & t_{im}^{\theta D} \\ \vdots & & \vdots & & \vdots \\ t_{m1}^{\theta D} & \cdots & t_{mj}^{\theta D} & \cdots & t_{mm}^{\theta D} \end{bmatrix}_{m\times m}. \tag{A.21}$$

Then, the weighted super-matrix W^θ for normalization can be easily obtained as in $T_D^\theta W$ Eq. (A.22):

$$W^\theta = T_D^\theta W = \begin{bmatrix} t_{11}^{\theta D} \times W^{11} & \cdots & t_{i1}^{\theta D} \times W^{i1} & \cdots & t_{m1}^{\theta D} \times W^{m1} \\ \vdots & & \vdots & & \vdots \\ t_{1j}^{\theta D} \times W^{1j} & \cdots & t_{ij}^{\theta D} \times W^{ij} & \cdots & t_{mj}^{\theta D} \times W^{mj} \\ \vdots & & \vdots & & \vdots \\ t_{1m}^{\theta D} \times W^{1m} & \cdots & t_{im}^{\theta D} \times W^{im} & \cdots & t_{mm}^{\theta D} \times W^{mm} \end{bmatrix}, \tag{A.22}$$

where $t_{ij}^{\theta D} = t_D^{ij}\big/d_i$ is a scalar and $\sum_{j=1}^{m} m_j = n$.

Step 3: Finding D-ANP influential weights (global weights). Increase a sufficient power q through self-multiplication of the weighted super-matrix until a stable super-matrix appeared to weigh the global vectors, the D-ANP influence weights ($w_1,...,w_j,...,w_n$), thus yielded from $\lim_{h\to\infty} (W^\theta)^h$ are shown as Table 5.

**Stage 4: Comparing the performance gap values
with a modified-VIKOR technique.**

Step 1: Normalizing the initial rating matrix. The "aspiratiosn-worst" of the modified-VIKOR approach was applied to measure the performance matrix of AI-enabled audit of alternative (company) k for criterion j. The performance scale as $f_j^{aspiration} = 10$ of the positive ideal (called aspiration level) and the worst value as $f_j^{worst} = 0$ (called the lowest score)

in this study was defined. Accordingly, by a normalized performance matrix $[f_{zj}]_{Z \times n}$, the performance gap ratios $[r_{zj}]_{Z \times n}$ can be obtained as shown in Eq. (A.23).

$$[r_{zj}]_{Z \times n} = [(|f_j^{aspiration} - f_{zj}|)/(|f_j^{aspiration} - f_j^{worst}|)]_{Z \times n}, \tag{A.23}$$

where $\boldsymbol{f}^{aspiration} = (f_1^{aspiration}, ..., f_j^{aspiration}, ..., f_n^{aspiration})$ is the aspiration level (desired level) and $\boldsymbol{f}^{worst} = (f_1^{worst}, ..., f_j^{worst}, ..., f_n^{worst})$ is the worst value.

Step 2: Determining the means of group utility S_z and the maximal regret gap Q_z. These gap values can be explored using Eqs (A.24) and (A.25).

$$L_z^{h=1} = S_z = \sum_{j=1}^{n} w_j r_{zj} = \sum_{j=1}^{n} w_j (|f_j^{aspiration} - f_{zj}|)/(|f_j^{aspiration} - f_j^{worst}|); \tag{A.24}$$

$$L_z^{h=\infty} = Q_z = \max_j (r_{zj} | j = 1, 2, ..., n), \tag{A.25}$$

where S_z represents the average gap-ratios of the distance between aspiration level $f_j^{aspiration}$ and real performance value f_{zj} in each criterion j of each alternative (company) z; and w_j is the influential weight of the D-ANP. The main purpose of this study is examining how to reduce the performance gap-ratio $(r_{zj}, j = 1, 2, ..., n)$ of alternative p, and the results are presented in Table 6.

Step 3: Finding the comprehensive performance indicator R_z. The integrated values are measured by Eq. (A.26).

$$R_z = v(S_z - S^{aspiration})/(S^{worst} - S^{aspiration}) + (1 - v)(Q_z - Q^{aspiration})/(Q^{worst} - Q^{aspiration}). \tag{A.26}$$

Using R_z from performance gap matrix, the improvement priority can be determined.