

Research Article

The Design of Academic Programs Using Rough Set Association Rule Mining

Mofreh A. Hogo ^{1,2}

¹Department of Electrical Engineering, Benha Faculty of Engineering, Benha University, Benha, Egypt

²Department of Computer Science, University College in Umluj, University of Tabuk, Tabuk, Saudi Arabia

Correspondence should be addressed to Mofreh A. Hogo; mhogo@ut.edu.sa

Received 30 May 2022; Revised 1 July 2022; Accepted 2 July 2022; Published 30 August 2022

Academic Editor: Manikandan Ramachandran

Copyright © 2022 Mofreh A. Hogo. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Program accreditation is important for determining whether or not a program or institution meets quality standards. It helps employers to evaluate the programs and qualifications of their graduates as well as to achieve its strategic goals and its continuous improvement plans. Preparing for accreditation requires extensive effort. One of the required documents is the program's self-study report (SSR), which includes the PEO-SO map (which allocates the program's educational objectives (PEOs) to student learning outcomes (SOs)). It influences program structure design, performance monitoring, assessment, and continuous improvement. Professionals in each academic engineering program have designed their PEO-SO maps in accordance with their experiences. The problem with the incorrect design of map design is that the SOs are either missing altogether or cannot be assigned to the correct PEOs. The objective of this work is to use a hybrid data mining approach to design the correct PEO-SO map. The proposed hybrid approach utilizes three different data mining techniques: classification to find the similarities between PEOs, crisp association rules to find the crisp rules for the PEO-SO map, and rough set association rules to find the coarse association rules for the PEO-SO map. The work collected 200 SSRs of accredited engineering programs by the ABET-EAC. The paper presents the different phases of the work, such as data collection and preprocessing, building of three data mining models (classification, crisp association rules, and rough set association rules), and analysis of the results and comparison with related work. The validation of the obtained results by different fifty specialists (from the academic engineering field) and their recommendations were also presented. The comparison with other related works proved the success of the proposed approach to discover the correct PEO-SO maps with higher performance.

1. Introduction

In designing academic programs, more emphasis is placed on improving students' knowledge and skills; this is accomplished in undergraduate programs through a series of courses in the subject area. General courses provide core foundational knowledge, and each course in these courses provides a range of activities to enhance students' skills in knowledge, cognitive skills, communication, leadership, teamwork, presentation, technical writing, and psychomotor skills domains. The life cycle of the program includes various phases such as design and specification, implementation, and continuous updating as shown by the authors in [1–3]. PEOs are specified by various stakeholders and components and are defined as a set of skills that relate to knowledge,

skills, and attitudes that learners are likely to demonstrate. PEOs are then mapped to the expected student learning outcomes in each course. Figure 1 illustrates a hierarchical structure of academic program design. The design of the PEO-SO map is a core phase of academic program design. In terms of student performance, PEOs are assessed 4–5 years after graduation based on the academic model presented by the authors in [4]. The ABET-EAC (American Board for Engineering and Technology-Engineering Accreditation Commission) has accredited 380 computer engineering programs by 2019. The ABET-EAC criteria for engineering programs are related to the knowledge, skills, and student behaviors acquired during the program. The PEO-SO direct and indirect assessments presented by the authors in [5] are used to ensure the achievement of program objectives.

The guide to accreditation policies and procedures of ABET describes the details of the PEOs in Criterion 2 and the student learning outcomes in Criterion 3 (see Figure 2, Appendix A) and shows the ABET SOs in their old and new updated versions. One of the most common problems encountered in engineering program design is the incorrect mapping of PEOs and SOs. This influences program design, implementation, assessment, and accreditation processes. It also increases the burden of the accreditation process and decreases the quality of the program. Therefore, a robust design of the PEO-SO map is a critical issue to avoid these problems and minimizes the effort required to prepare the accreditation documentation. On the other hand, the proper design of the PEO-SO map improves the SO selection and allocation, improves student performance, enhances graduate skills, and increases the satisfaction of the professional community, society, creativity, professionals, ethical issues, goal achievement of academic programs, economic issues, and local and international competition.

The work in this paper aims to introduce a hybrid of data mining techniques (three different data mining models) to discover the correct PEO-SO map that avoids the problems of incorrectly mapping PEOs to their SOs. The paper uses the rough set association rule mining technique to eliminate the confusion in the association of PEOs to SOs and minimize the number of SOs associated with the corresponding PEOs and eliminate the ambiguous association. Only the rules that are certain to result in the correct mapping of PEOs and SOs are found.

This in turn minimizes the effort required to design academic engineering programs and prepare them for accreditation. In this work, a dataset of 200 SSRs of academic engineering programs was used to develop and validate the proposed model. The remainder of this paper is presented as follows: Section 2 provides a summary of related work using data mining techniques in education or higher education. Section 3 presents the proposed approach. Section 4 presents the different stages of data acquisition, preprocessing and presentation, experimental design, results obtained, discussion, and analysis. Finally, Section 5 is reserved for the conclusion and future work.

2. Related Work

Many researchers have used artificial intelligence and machine learning algorithms, as well as statistical theories and techniques, in their work to discover key patterns in educational datasets. Their goals are to support academic program design and assessment, accreditation and reaccreditation processes, and decision-making processes and to improve program performance in higher education institutions. Data mining techniques help remove difficulties and impurities to produce a good analysis of large datasets and discover the hidden knowledge in datasets captured by various information systems. Various researchers have worked on developing data mining models to summarize, classify, cluster data, and develop association rules and other features to be applied to various datasets in educational data mining (EDM). The data mining techniques in [6] detect the

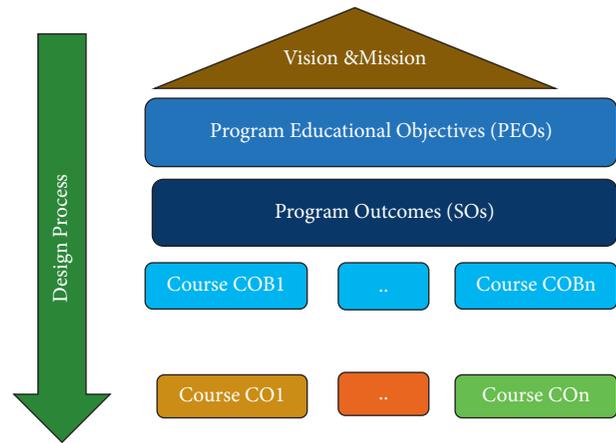


FIGURE 1: The design process for PEOs and SOs.

correspondence between course content learning objects and program level. In [7], an analysis of two different datasets of academic courses is presented using graphical, statistical, and quantitative techniques to select an appropriate ensemble learner from a combination of six potential learning algorithms. In [8], several relevant studies on computer-supported learning analytics (CSLA), computer-supported predictive analytics (CSPA), computer-supported behavior analysis (CSBA), and computer-supported visualization analysis (CSVA) from 2000 to 2017 are presented. Predicting student performance using educational data mining was presented in [9], where the base classifiers were random tree, J48, KNN, and naive Bayes. The use of decision trees and hierarchical linear models using data from the Spanish PISA 2015 and high and low effectiveness schools is presented in [10]. In [11], an assessment framework for capstone courses is presented to evaluate student's performance and project quality by assessing student learning outcomes. In [12], a methodology for collecting information and measuring student learning outcomes for the ABET accreditation preparation process is presented to assist in the completion of ABET SSR. The Apriori technique for identifying association rules was used to create the PEO-SO map that describes the relationship between the program's educational goals and student learning outcomes, as shown in [13]. The Apriori algorithm was used to create association rules to describe the mapping of PEOs and SOs. The problem with this approach is that the number of rules is very high and not specific. In the work presented in [14], a tool for computational methods of information was introduced that includes rough sets, incompleteness of information, data mining, granular computation and extraction of association rules, and new mathematical frameworks. The tool is named RNIA (rough nondeterministic information analysis). In [15], an assessment and evaluation strategy for ABET student outcomes (SOs) of computer science and computer information systems programs is presented, where the assessment is developed through direct and indirect methods. Quality of education through accreditation, teaching, and learning in nursing programs is presented in [16]. A framework and work phase is developed to collect and document the selection for assessment of ABET curriculum requirements.

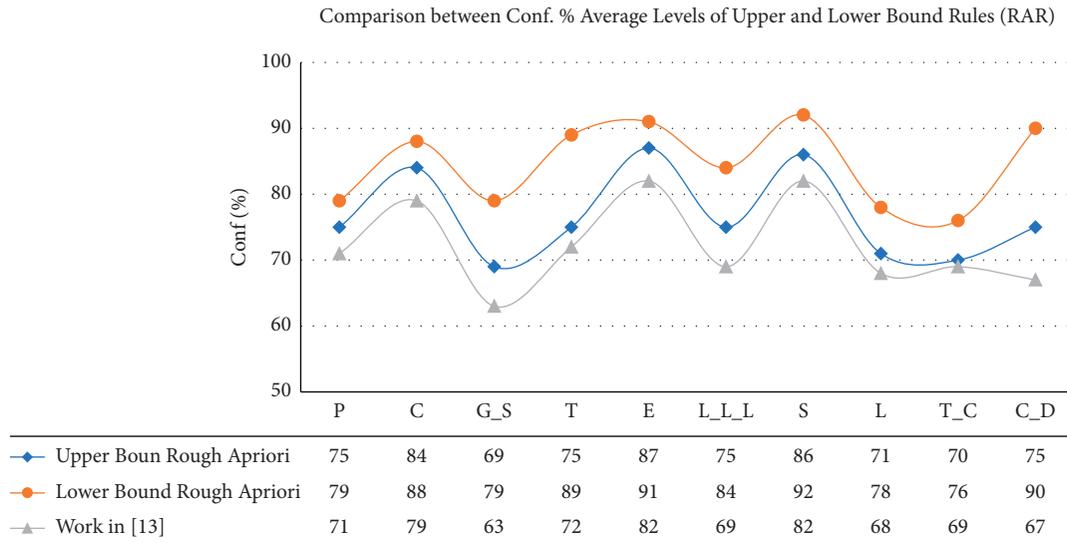


FIGURE 2: Comparison between Conf. percentage Average Levels for Upper and Lower Bound Rules at Work in [13].

The work developed a tool called ABETAS to automate this framework [17], which helps the institution prepare for ABET accreditation to minimize the burden of assessing student outcomes. The relationship between teaching quality and accreditation is presented in [18], illustrating how accreditation can assist the program in maintaining quality and programs in improving or achieving educational quality through the accreditation process. The approach used rough set theory to adapt the association rule model to discover customer favorites, and its analysis was presented in [19]. A modification mechanism for the attributes and association rules of the rough set was presented, and this proposal was applied to e-commerce platforms to categorize the rough recommendations. An investigation of mining class association rules using the rough set approach is presented in [20]. In [21], an algorithm for finding the finest class rules is presented that uses the adaptation of the Apriori association rule algorithm based on rough set theory to compute the support and confidence of the elementary set of lower approximation concepts. Rough set theory is used in association rule discovery. It simplifies the process of traditional association rule mining and avoids redundant rules introduced in [22] to determine rough set rules. In the work presented in [23], a collection methodology was presented for mapping PEOs to SOs derived from the SSRs of 32 engineering programs accredited by ABET. It minimizes the effort and time-consuming processes. An association rule mining algorithm based on the properties of rough set theory is said to improve the Apriori algorithm for association rule mining based on a decision table. Assessment methods for the ABET-CAC accreditation criteria for computer science undergraduate programs are presented in [24]. They adopted set of student outcomes for the computer science program to meet the ABET program outcomes and PEOs. Using the theory of rough sets and their properties to discover information in a simpler way than the normal Apriori association rule mining method presented in [25] minimizes the

attributes in the dataset and develops a simpler data mining model.

In the above review, many research papers were presented that focused on the accreditation process and its requirements. They made good contributions in different areas related to the design of PEOs and SOs, but they did not pay attention to the size of the rules governing the relationship between PEOs and SOs in the PEO-SO maps. They presented the discovery of PEO-SO with high ambiguity and large scale with low confidence. This leads to many errors in program design and accreditation processes.

The idea of the approach proposed in this paper is to discover the correct PEO-SO map that generates minimum size and accurate rules for the relationship between PEOs and SOs. The proposed approach uses three different data mining techniques as follows:

- (i) The decision tree (J48) classifier is very popular to represent the data in a tree form similar to rules. Its goal is to discover the similarities and dissimilarities between different PEO categories. It discovers the confusion among different PEOs, which assists in eliminating this confusion during the PEO-SO map design phase. Its results can be used as a guide for PEO-SO map design.
- (ii) The Apriori algorithm for association rules is used to determine the clear rules describing the relationship between PEOs and SOs.
- (iii) The adaptation of the rough set theory for Apriori association rules is used to determine the association rules of the rough set. It consists of lower bound rules (describing the safe region, i.e., high conf.% rules) and upper bound rules (describing the uncertain region, i.e., low and high conf.% rules). The goal is to select the rules with the lower conf.% and avoid redundancy, eliminate the ambiguity between different PEO rules, and simplify the association

rule mining with the correct rule sets that govern the relationship between PEOs and SOs.

The resulting PEO-SO maps are evaluated by 50 independent academic professionals to obtain their assessment and feedback. Figure 3 illustrates the complete structure of the proposed approach and shows the different phases of the work.

3. The Proposed Approach

This section presents the theoretical basis of the machine learning algorithms utilized in this paper to develop the proposed hybrid approach. We start with the bagged J48 decision tree algorithm for machine learning, then the Apriori algorithm for association rules, and finally the rough set Apriori algorithm for rough association rules (upper and lower bounds).

3.1. Bagged J48 Machine Learning Algorithm.

Classification is a process of recognizing, understanding, and grouping objects into predefined classes using training datasets. Machine learning software uses different types of algorithms to classify future elements of datasets into correct categories. A decision tree is a classification algorithm that uses a divide and conquer algorithm, which consists of decision nodes and leaf nodes. The decision node identifies a test over one of the attributes and the leaf node represents the class value [26]. The classification error is the percentage of misclassified cases [27]. In practice, the training datasets are usually large, which leads to a larger number of branches and layers in the generated decision tree. When there are more class categories in the decision tree, the classification accuracy decreases significantly. There are various decision tree generation algorithms such as ID3, J48, FT, BF Tree, LMT, and many more. The performance is evaluated using the F-measure. By using machine learning algorithms, the proposed work is automated which increases the accuracy of the result [28, 29]. The J48 algorithm was proposed and developed by Quinlan in 1993. In this work, the objectives of using the J48 classifier are as follows: first, we validate the process of data collection and representation, and second, we discover how the classifier is confounded with different PEO categories of the different SSR reports in different academic programs. In this work, we used the J48 algorithm because it has higher accuracy. To increase the accuracy of the classifier, an ensemble technique can be used; the classification performance is greatly improved by combining the decisions of different classifiers into a single classifier. This J48 algorithm uses two ensemble learning approaches, bagging and boosting, which are applied to five traditional classifiers.

3.2. *Apriori Association Rule.* The association rules are formally described as introduced in [22, 30]. Let $Z = \{Z_1, Z_2, \dots, Z_m\}$ be the set of attributes and T be an instance in the dataset S , where $T \subseteq Z$. Each instance in S is identified by TID. If the set of objects satisfies $X \subseteq Z$, $Y \subseteq Z$, and $X \cap Y = \varnothing$, the implication $X \longrightarrow Y$ is defined as an association rule, and

if $s\%$ in D matches $X \longrightarrow Y$, then the support of the rule $X \longrightarrow Y$ is $s\%$, which is computed by $s\% = \text{support}(X \longrightarrow Y) = P(X|Y)$. If the instance contains X and Y , then the confidence of rule $X \longrightarrow Y$ is $c\%$, which is calculated by $c\% = \text{support}(X \longrightarrow Y) = P(X|Y) = P(XY)/P(X)$. Minsup and minconf if $s\% \geq \text{minsup}$ and $c\% \geq \text{minconf}$, which is defined as a strong association rule. The Apriori algorithm is updated in [31, 32]. To reduce bias, a lift judgment measure is defined in [33, 34] and defined by the formula $\text{lift} = c(X \longrightarrow Y)/s(Y)$. Table 1 (see Appendix B) illustrates an explanatory example for the calculation of support % and conf. % using the Apriori algorithm.

3.3. *Rough Set Theory.* Rough set theory can make a proper measurable analysis of vague, unpredictable, and imperfect information [35, 36]. The universe (all instances in a dataset) is divided into units of imperceptible objects, which are defined as basic sets. This imperceptibility is related to the outcome and granularity of the information [37, 38]. Pawlak's rough set model is the basis for formal reasoning and data analysis and self-directed decision-making [39]. Rough set theory classifies uncertain information expressed in terms of experience data. A set of similar objects is called an elementary set, which is a fundamental atom of knowledge. Any union of elementary sets is called a crisp set, while other sets are rough sets. Each rough set has boundary line elements that can belong to the set or its complement, as shown in [40, 41]. There are three types of approximation in rough set theory; the different regions that represent the approximation properties of rough set theory are the upper boundary BU, the lower boundary (BL), and the boundary region (BU-BL), which are shown in Figure 4.

3.4. *Rough Set Association Rules (RSARs).* The algorithm for generating the rough set association rule is denoted by R_Apriori and was presented in [18]. It modifies the Apriori algorithm for generating association rules, which generates frequent rules. It consists of three phases. The first phase computes the support of each element with one rule (which contains only one element), while the second phase computes the support of each element with two rules (which contains two elements), and the third phase computes the support of each element with three rules (which contains three elements). The steps of the RSAR algorithm are described with an explanatory example illustrated in Table 2 (see Appendix C). The example uses a dataset from the instances of the collected SSR dataset for the PEO-SO map.

4. Datasets and Experimental Result Analysis

4.1. *Dataset Description and Preprocessing.* The raw datasets used in this work were collected manually using the Google search engine and accredited academic programs' websites. From 200 SSRs of accredited engineering programs, the work aims to discover the robust and correct PEO-SO maps. Therefore, we focus only on the map section in the SSR documents. Each academic engineering program should illustrate the mapping between PEOs and a set of SOs (11

Old ABET 11 Outcomes a to k	Updated ABET 7 Outcomes
a. an ability to apply knowledge of mathematics, science, and engineering. b. an ability to design and conduct experiments, as well as to analyze and interpret data. c. an ability to design a system, component, or process to meet desired needs within realistic constraints such as economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability. d. an ability to function on multidisciplinary teams e. an ability to identify, formulate, and solve engineering problems. f. an understanding of professional and ethical responsibility. g. an ability to communicate effectively (3g1 orally, 3g2 written). h. the broad education necessary to understand the impact of engineering solutions in a global, economic, environmental, and societal context. i. a recognition of the need for, and an ability to engage in lifelong learning. j. a knowledge of contemporary issues. k. an ability to use the techniques, skills, and modern engineering tools necessary for engineering practice.	1. an ability to identify, formulate, and solve complex engineering problems by applying principles of engineering, science, and mathematics. 2. an ability to apply engineering design to produce solutions that meet specified needs with consideration of public health, safety, and welfare, as well as global, cultural, social, environmental, and economic factors. 3. an ability to communicate effectively with a range of audiences. 4. an ability to recognize ethical and professional responsibilities in engineering situations and make informed judgments, which must consider the impact of engineering solutions in global, economic, environmental, and societal contexts. 5. an ability to function effectively on a team whose members provide leadership, create a collaborative and inclusive environment, establish goals, plan tasks, and meet objectives. 6. an ability to develop and conduct appropriate experimentation, analyze and interpret data, and use engineering judgment to conclude. 7. an ability to acquire and apply new knowledge as needed, using appropriate learning strategies.

FIGURE 3: A List of ABET 11 Old SOs (11 from a to k) and New Updated 7 SOs (7 from 1 to 7).

TABLE 1: Examples of support measure.

ID	Element	$s \% = \text{Support} = \text{Occurrence} / \text{Total instances}$	$c\% = s (Y \cup X) / s (X)$
1	ABD	Total number of instances = 6 $\text{Support}(\{AB\}) = s \% = 4/6 = 66.66\%$ $\text{Support}(\{BC\}) = s \% = 1/6 = 33.33\%$	i) $c\% (A \rightarrow B) = 4/5 = 80\%$ ii) $c\% (B \rightarrow C) = 1/6 = 16.6\%$
2	AB		
3	ABB		
4	BCD		
5	BAC		
6	ABA		

outcomes from *a* to *k*, see Figure 2, Appendix A). This is one of the requirements of ABET-EAC, presented in subsection B of Section 3 (outcome subsection). Therefore, the dataset used was selected from these subsections in the SSRs of the entire dataset. Then, we create a table for 200 PEO-SO maps extracted from 200 SSRs. Each PEO-SO map consists of a predefined number of entries, which are first collected in the form of symbols and words and then converted into numerical entries (on average, the total entries are $200 \times 8 = 1600$ entries, and each entry is represented as 11 student learning outcomes. Some entries were omitted because they were not complete or were missing). The representation of the data is performed in several steps. First, we encode the set of PEOs with a set of symbols similar to those in [13], and Table 3 shows a symbolic form of the common set of PEOs for all engineering programs. PEOs were presented in text form, and in many cases, each PEO was confusingly prepared (not specific and/or merged with two or more PEOs), e.g., “excel in industrial or graduate work in computer engineering and related fields” represents two PEOs as graduate studies and career development. To solve this problem, an essential word processing step was performed by implementing a software program similar to the Word2Vec model to convert each

PEO into 11-dimensional word vectors (each word represents a single PEO) using the collected SSR dataset described above, and Table 4 shows an example of the output. The third step is the conversion, normalization, and presentation of PEO-SO map datasets, which are essential for the further data mining steps; it strongly affects the output of the data mining model, and Tables 5–8 show examples of these preprocessing steps. The following remarks are intended to illustrate these steps:

- (i) Table 3 shows an example of the PEO-SO map for academic program A. PEO1 is divided into two PEO categories: career development and graduate studies, which are symbolically represented as C_D and G_S, respectively. The last column refers to the assigned student outcomes for these two PEO categories, where “*x*” means that this SO is unassigned for this PEO, while “ $\sqrt{}$ ” means that it is assigned.
- (ii) Table 4 shows another example of the assignment of PEOs and SOs in academic program B. It illustrates the different representations of PEOs from one program to another and the difference in the assignment of SOs.

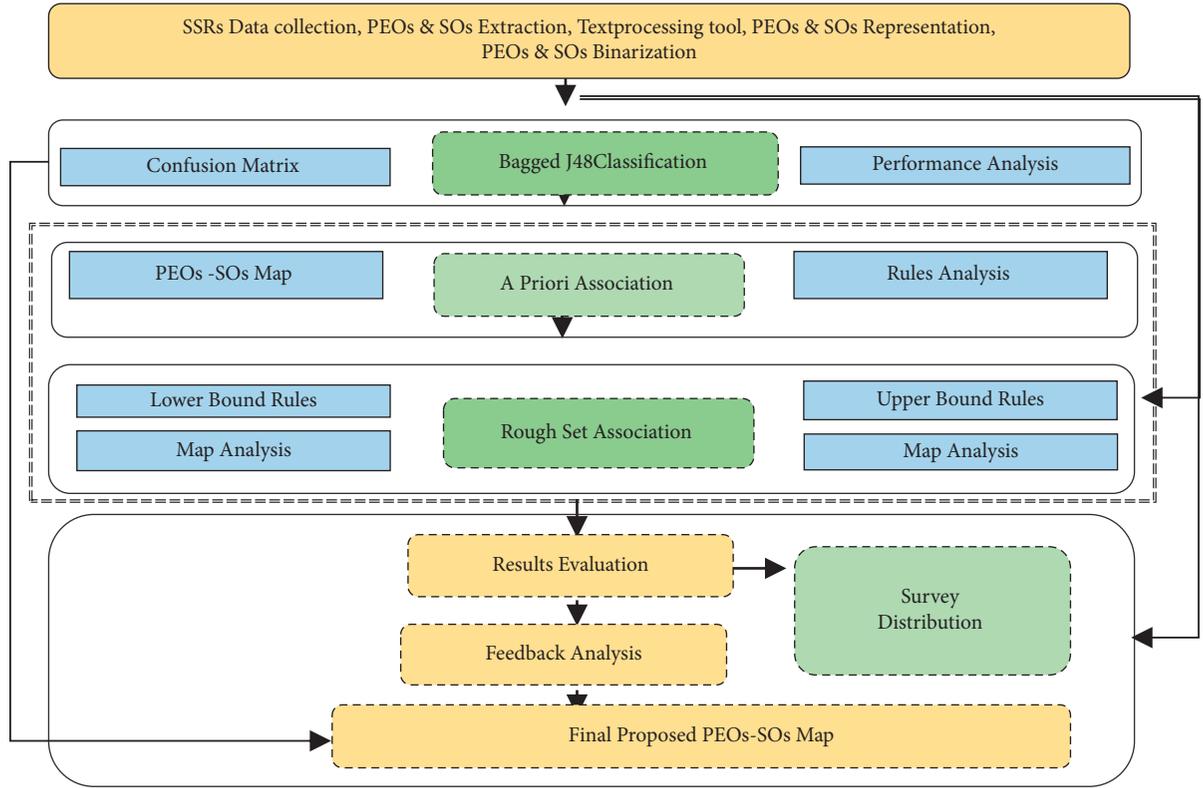


FIGURE 4: The Structure of the proposed approach for PEOs-SOs Map Design.

TABLE 2: Support and confidence measures using Apriori and rough Apriori.

ID	PEO	a	b	c	d	e	f	G	H	i	J	k
1	C_D	1	1	1	1	1	1	1	1	1	1	1
2	G_S	1	1	1	0	1	1	0	0	1	0	0
3	P	0	0	1	0	1	0	0	0	1	1	1
4	E	0	0	1	0	0	1	0	0	0	0	0
5	T_C	0	0	0	0	0	0	0	1	1	1	1
6	L	0	0	0	1	0	1	1	0	0	1	0
7	K_C	1	1	1	1	0	0	0	1	0	1	0
8	L_L_L	0	0	0	0	0	1	1	0	1	1	0
9	S	0	0	0	0	1	1	1	1	1	1	0
10	L	0	0	1	1	0	0	1	0	0	0	0
11	C_D	1	0	1	1	0	1	0	0	1	1	1
12	T	0	1	0	1	0	0	1	0	0	0	0
13	P	0	1	1	0	1	0	1	1	1	1	1
14	E	0	0	1	0	0	1	0	0	1	0	0
15	C	0	0	1	1	0	0	1	0	0	0	0

Apriori association rule algorithms $s\% = S = \text{occurrence}/\text{total trans}$, total number of trans = 15, $s(\{ab\}) = 3/15 = 60\%$, $s(\{bc\}) = 4/15 = 40\%$, and $\text{conf}(a \Rightarrow b) = 3/4 = 75\%$, $\text{conf}(b \Rightarrow c) = 4/5 = 80\%$.

Rough set Apriori association rule algorithm RSAR, $s(\{c\}) = 2$, $\{\text{decSet}\} = \{\{e\}d\} = \{P\} = \{3,13\}$ and $s(\{e\}d) = 2$, $\{\text{condSet}\} = \{\{e\}c\} = \{c, e\} = \{3,13\}$ and $s(\{e\}c) = 2$, and s of RSAR = $s(\{\text{condSet} \cup \text{decSet}\}) = s(\{e\}c, \{e\}d) = s(\{3,13\}) = 2$, $|\Omega| = 15$. Then, s of RSAR is computed as $(\{e\}d)/|\Omega| = 2/15 = 13\%$, and the confidence of RSAR is $S(\text{RSAR})/s(\{e\}c) = 2/2 = 100\%$.

- (iii) Table 5 illustrates the conversion of PEOs, and the PEOs shown in Tables 3 and 4 into binary forms are suitable for input to different data mining models in the next steps.

TABLE 3: PEO symbolic representations.

PEO#	PEO class	Symbolic PEOs
1	Lifelong learning	L_L_L
2	Career development	C_D
3	Communication	C
4	Ethics	E
5	Graduate studies	G_S
6	Knowledge competency	K_C
7	Leadership	L
8	Professionalism	P
9	Social	S
10	Team	T
11	Technical competency	T_C

- (iv) Table 6 illustrates the final binary representation of the mapping of PEOs and SOs. Each row contains 13 features such as the pattern ID, the PEO class, and 11 student outcomes from $a \dots k$.

4.2. *Result Analysis of Applying the Bagged J48 Machine Learning Algorithm.* This section presents the interpretation and analysis of the results obtained by applying the bagged J48 classifier to the dataset. The obtained results are presented in Table 9, Table 10, and Table 11 (see Appendix C). The results illustrate the confusion matrix for the classifier outputs. It is important to illustrate the misattribution and misunderstanding of PEOs to their SOs when designing PEO-SO maps for PEO categories:

TABLE 4: Examples of PEO-SO mapping for computer engineering program A.

Program educational objectives	<i>C_D</i>	<i>E</i>	<i>P</i>	<i>C</i>	<i>L_L_L</i>	<i>G_S</i>	<i>S</i>	<i>T</i>	<i>T_C</i>	<i>K_C</i>	<i>L</i>
PEO1. Excel in industrial or graduate work in computer engineering and related fields	1	0	0	0	0	1	0	0	0	0	0
PEO2. Practice computer engineering in both national and international settings, by forming ethical values and environmental friendly policies	0	1	1	0	0	0	0	0	0	0	0
PEO3. Have sufficient professional competence allowing them to assume leadership roles in their organization	0	0	0	0	0	0	0	0	1	1	0
PEO4. Successfully adapt to new technologies and stay current with their professions	0	0	0	0	1	0	0	0	0	0	0
PEO5. Make suggestions for improving the processes used by their institutions and take the initiative to make autonomous decisions	0	0	0	0	0	0	1	0	0	0	0

TABLE 5: Examples of PEO-SO mappings for computer engineering program A.

PEO categories	Symbol	Program educational objectives (PEOs)	Student outcomes										
			<i>a</i>	<i>B</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>F</i>	<i>g</i>	<i>h</i>	<i>i</i>	<i>j</i>	<i>K</i>
Career development, graduate studies	<i>C_D, G_S</i>	PEO6. Excel in industrial or graduate work in computer engineering and related fields	√	√	√	√	√	√	√	√	√	√	√
		PEO7. Practice computer engineering in both national and international settings, by forming ethical values and environmental friendly policies	√	√	√	√	√	√	√	√	√	√	√
Technical and knowledge competence, leadership	<i>T_C, K_C, L</i>	PEO8. Have sufficient professional competence allowing them to assume leadership roles in their organization	<i>x</i>	<i>X</i>	<i>x</i>	√	√	√	√	√	√	√	√
Longlife learning	<i>L_L_L</i>	PEO9. Successfully adapt to new technologies and stay current with their professions	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	√	√	<i>x</i>	√	√	<i>x</i>
Society service	<i>S</i>	PEO10. Make suggestions for improving processes used by their institutions and take initiatives to make autonomous decisions	<i>x</i>	<i>x</i>	<i>x</i>	<i>x</i>	√	√	√	√	√	√	<i>x</i>

TABLE 6: Examples of PEO-SO mappings for computer engineering program B.

PEO categories	Symbol	Program educational objectives (PEOs)	Student outcomes										
			<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>F</i>	<i>g</i>	<i>h</i>	<i>i</i>	<i>j</i>	<i>k</i>
Technology development, leadership, career development	<i>T, L, C_D</i>	PEO1. Prepare graduates with technical knowledge and skills	√	√	√	√	<i>x</i>	√	√	√	√	√	√
Professionalism, ethics, communication	<i>P, C, E</i>	PEO2. Prepare graduates to establish themselves in professional careers	<i>x</i>	<i>x</i>	<i>x</i>	√	√	<i>X</i>	√	√	√	√	√

- (1) Category *C_D* has 100 instances, and 84 of them are classified correctly as *L_L_L* and 16 instances are not, which means that specialists may get confused in assigning PEOs to SOs for these two categories.
- (2) Category *E* has 86 instances, and six of them are classified correctly as *L_L_L* and two of them are not.
- (3) Category *P* has 84 instances, and 70 of them are classified correctly, while 14 instances are not.
- (4) Category *C* has 44 instances, and 24 of them are classified correctly, and 8 instances are incorrectly classified as *E*, while 12 instances are incorrectly classified as *T*, which means that there may be confusion among the specialists in assigning PEOs to SOs for these three categories.
- (5) The *L_L_L* category has 204 instances, and 196 of them are classified correctly, while 8 instances are incorrectly classified as *T*, implying that the specialists might have made mistakes in assigning PEOs to SOs for *L_L_L* and *T* categories.
- (6) The *G_S* category has 40 instances, and 24 of them are classified correctly, eight instances are incorrectly classified as *L_L_L*, two instances are incorrectly classified as *S* category, and two are classified as *T* category, which means that the specialists may get confused when assigning PEOs to SOs for these four categories.
- (7) The *S* category has 112 instances, and 100 of them are classified correctly. Eight instances are incorrectly classified as *L_L_L* and 4 instances are incorrectly classified as *P* category, which means that specialists may get confused in assigning PEOs to SOs for these three categories.
- (8) The *T_C* category has 40 instances, and 24 of them are classified correctly, 8 instances are incorrectly classified as *S* category, 4 instances are incorrectly classified as *P* category, and 4 are classified as *T* category, which means that the specialists may get confused in assigning PEOs to SOs for these four categories.

TABLE 7: Convert the PEO-SO mapping for computer engineering program A&B to a suitable form.

PEOs categories	Symbol	Program educational objectives (PEOs)	Student outcomes													
			<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>F</i>	<i>g</i>	<i>h</i>	<i>i</i>	<i>j</i>	<i>k</i>			
Career development	<i>C_D</i>	PEO1. Excel in industrial work in computer engineering and related fields	1	1	1	1	1	1	1	1	1	1	1	1	1	
Graduate studies	<i>G_S</i>	PEO2. Excel in graduate work in computer engineering and related fields	1	1	1	0	1	1	0	0	1	0	0	1	0	0
Professionalism	<i>P</i>	PEO3. Practice computer engineering in both national and international settings	0	0	1	0	1	0	0	0	0	0	1	1	1	1
Ethics	<i>E</i>	PEO4. Forming ethical values and environmental friendly policies	0	0	1	0	0	1	0	0	0	0	0	0	0	0
Technical competence	<i>T_C</i>	PEO5. Have sufficient professional competence allowing them to assume	0	0	0	0	0	0	0	0	0	1	1	1	1	1
Leadership	<i>L</i>	PEO6. Leadership roles in their organization	0	0	0	1	0	1	1	1	0	0	1	0	1	0
Knowledge competence	<i>K_C</i>	PEO7. Have sufficient professional competence allowing them to assume	1	1	1	1	0	0	0	0	1	0	1	0	1	0
Longlife learning	<i>L_L_L</i>	PEO8. Successfully adapt to new technologies and stay current with their professions	0	0	0	0	0	1	1	0	1	1	0	1	1	0
Society service	<i>S</i>	PEO9. Make suggestions for improving processes used by their institutions and take initiatives to make autonomous decisions	0	0	0	0	1	1	1	1	1	1	1	1	1	0
Leadership	<i>L</i>	PEO10. Prepare graduates with technical knowledge	0	0	1	1	0	0	1	0	0	0	0	0	0	0
Technology development	<i>C_D</i>	PEO11. Prepare graduates with technical skills	1	0	1	1	0	1	0	0	1	1	1	1	1	1
Team	<i>T</i>	PEO12. Prepare graduates with technical knowledge and skills	0	1	0	1	0	0	1	0	0	0	0	0	0	0
Professionalism	<i>P</i>	PEO13. Prepare graduates to establish themselves in professional careers	0	1	1	1	1	1	1	1	1	1	1	1	1	1
Ethics	<i>E</i>	PEO14. Prepare graduates to establish themselves in professional careers	0	0	1	0	0	1	0	0	1	0	0	1	0	0
Communication	<i>C</i>	PEO15. Prepare graduates to establish themselves in professional careers	0	0	1	1	0	0	1	0	0	1	0	0	0	0

TABLE 8: The final representation of the PEO-SO maps.

ID	PEOs	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>F</i>	<i>g</i>	<i>h</i>	<i>I</i>	<i>j</i>	<i>k</i>
1	<i>C_D</i>	1	1	1	1	1	1	1	1	1	1	1
2	<i>G_S</i>	1	1	1	0	1	1	0	0	1	0	0
3	<i>P</i>	0	0	1	0	1	0	0	0	1	1	1
4	<i>E</i>	0	0	1	0	0	1	0	0	0	0	0
5	<i>T_C</i>	0	0	0	0	0	0	0	1	1	1	1
6	<i>L</i>	0	0	0	1	0	1	1	0	0	1	0
7	<i>K_C</i>	1	1	1	1	0	0	0	1	0	1	0
8	<i>L_L_L</i>	0	0	0	0	0	1	1	0	1	1	0
9	<i>S</i>	0	0	0	0	1	1	1	1	1	1	0
10	<i>L</i>	0	0	0	1	0	0	1	0	0	0	0
11	<i>C_D</i>	1	0	1	1	0	1	0	0	1	0	1
12	<i>T</i>	0	1	0	1	0	0	1	0	0	0	0
13	<i>P</i>	0	1	1	1	1	1	1	1	1	1	1
14	<i>E</i>	0	0	1	0	0	1	0	0	1	0	0
15	<i>C</i>	0	0	1	1	0	0	1	0	0	0	0

TABLE 9: Confusion matrix for bagged J48.

	<i>C_D</i>	<i>E</i>	<i>P</i>	<i>C</i>	<i>L_L_L</i>	<i>G_S</i>	<i>S</i>	<i>T</i>	<i>T_C</i>	<i>K_C</i>	<i>L</i>
<i>C_D</i>	84	0	0	0	16	0	0	0	0	0	0
<i>E</i>	0	78	0	0	6	0	0	2	0	0	0
<i>P</i>	0	0	70	0	14	0	0	0	0	0	0
<i>C</i>	0	8	0	24	0	0	0	12	0	0	0
<i>L_L_L</i>	0	0	0	0	196	0	0	8	0	0	0
<i>G_S</i>	0	0	0	0	8	24	4	4	0	0	0
<i>S</i>	0	0	4	0	8	0	100	0	0	0	0
<i>T</i>	0	0	4	0	8	0	0	120	0	0	0
<i>T_C</i>	0	0	4	0	0	0	8	0	28	0	0
<i>K_C</i>	0	0	0	8	0	0	4	0	0	0	0
<i>L</i>	0	0	8	8	8	8	0	32	0	0	56

TABLE 10: Performance summary for bagged J48.

Correctly classified instances	380
Incorrectly classified instances	86
Kappa statistic	0.78
Mean absolute error	0.11
Root mean squared error	0.22
Relative absolute error	72%
Root relative squared error	75%
Total number of instances	466

- (9) Category *T* has 132 instances, and 120 of them are classified correctly, 8 instances are incorrectly classified as *L_L_L*, and 4 instances are incorrectly classified as *P* category, which means that specialists may get confused when assigning PEOs to SOs for these three categories.
- (10) Category *K_C* has 12 instances, but all are misclassified, which means that this class should be removed from the proposed 11 PEO categories.
- (11) Category *L* has 120 instances, and 56 of them are classified correctly, and 32 instances are incorrectly classified as *T*. Eight instances are incorrectly classified as category *P*. Eight instances are incorrectly classified as *C*. Eight instances are incorrectly classified as *L_L_L*. Eight instances are incorrectly classified as *G_S*. This means that specialists could get confused in assigning PEOs to SOs for these six categories.
- (12) Table 10 (see Appendix D) shows the summary of the results obtained using the bagged J48 decision tree classifier, which performs the highest for the datasets and provides more details, while Table 11 (see Appendix D) shows the detailed accuracy of each PEO class. It shows the percentage of correct

TABLE 11: Detailed accuracy by PEO category.

Class	TP	Rate	FP rate	Precision	Recall	F-measure	MCC area	ROC area
<i>C_D</i>	0.75	0	1.00	0.75	0.86	0.86	1.00	0.83
<i>P</i>	0.83	0.02	0.90	0.83	0.86	0.84	0.981	0.91
<i>G_S</i>	0.60	0.004	0.86	0.60	0.71	0.71	0.98	0.77
<i>S</i>	0.89	0.02	0.89	0.89	0.89	0.88	1.00	0.96
<i>L_L_L</i>	0.96	0.10	0.73	0.96	0.83	0.78	0.96	0.81
<i>L</i>	0.48	0	1.00	0.43	0.61	0.65	0.97	0.68
<i>E</i>	0.90	0	0.90	0.90	0.90	0.89	0.99	0.96
<i>T</i>	0.90	0.06	0.70	0.90	0.78	0.75	0.96	0.69
<i>C</i>	0.54	0	0.86	0.55	0.67	0.67	0.97	0.70
<i>T_C</i>	0.70	0	1	0.70	0.82	0.83	0.99	0.84
<i>K_C</i>	0	0	0.6	0.60	0.60	0.70	0.70	0.40
Weight average	0.80	0.03	0.8	0.80	0.80	0.78	0.98	0.82

TABLE 12: The upper bound rough set rules for the PEO-SO map.

PEOs	<i>a</i>	<i>b</i>	<i>C</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>G</i>	<i>H</i>	<i>I</i>	<i>j</i>	<i>k</i>	Conf. interval	Avg. conf
<i>P</i>	0/1	0/1	1	0	0/1	0/1	0	0/1	0/1	1	1	0.72–0.77	0.75
<i>C</i>	0	0	0/1	1	0/1	0/1	0/1	0/1	0	0/1	0/1	0.77–0.91	0.84
<i>G_S</i>	0/1	1	0/1	0/1	1	0	0/1	0/1	1	0/1	0/1	0.60–0.77	0.69
<i>T</i>	0	0	0	1	0	0/1	0/1	0	0	0	0/1	0.68–0.82	0.75
<i>E</i>	0	0	0/1	0/1	0	1	0	0/1	0	0	0/1	0.84–0.89	0.87
<i>L_L_L</i>	0	0	0	0	0/1	0/1	0/1	0/1	1	1	0/1	0.65–0.85	0.75
<i>S</i>	0	0	0/1	0/1	0/1	0/1	0/1	1	0/1	0/1	0/1	0.81–0.90	0.86
<i>L</i>	0	0	0	1	0/1	0/1	0/1	0/1	0/1	0/1	0/1	0.65–0.76	0.71
<i>T_C</i>	0	0/1	0/1	0/1	1/0	0/1	0	0	0/1	0/1	1	0.66–0.74	0.70
<i>C_D</i>	0/1	0/1	1	0/1	1	0/1	0/1	0/1	0/1	0/1	1	0.78–0.71	0.75

P rule represents the uncertain range of a rough set (may or may not), which means that PEO with “0/1” SO entries, and *a*, *b*, *e*, *f*, and *h* SOs mean that these SOs may or may not exist in the rule.

TABLE 13: Sample of the upper bound rough set rules for PEOs - SOs map.

<i>P</i>	<i>A</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>E</i>	<i>F</i>	<i>g</i>	<i>H</i>	<i>I</i>	<i>j</i>	<i>k</i>
0/1	0/1	1	0	0/1	0/1	0	0/1	0/1	0/1	1	1

TABLE 14: Rules interpretation for the upper bound rough set rules for PEOs - SOs map.

<i>P</i>	<i>A</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>G</i>	<i>H</i>	<i>i</i>	<i>J</i>	<i>K</i>
Rule 1	0	0	1	0	0	0	0	0	0	1	1
Rule 2	1	1	1	0	1	1	0	1	1	1	1

TABLE 15: The lower bound rough set rules for PEO-SO maps.

PEOs	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>	<i>i</i>	<i>j</i>	<i>k</i>	Conf. interval	Avg. conf
<i>P</i>	1	0	1	0	1	1	0	0	1	1	1	0.78–0.8	0.79
<i>C</i>	0	0	1	1	1	0	1	0	0	0	0	0.85–0.91	0.88
<i>G_S</i>	1	1	1	1	1	0	0	1	1	1	0	0.70–0.88	0.79
<i>T</i>	0	0	0	1	0	0	1	0	0	0	0	0.86–0.92	0.89
<i>E</i>	0	0	1	0	0	1	0	1	0	0	0	0.88–0.93	0.91
<i>L_L_L</i>	1	1	0	0	0	1	0	0	1	1	1	0.74–0.93	0.84
<i>S</i>	0	0	0	0	1	0	0	1	1	1	0	0.90–0.94	0.92
<i>L</i>	0	0	0	1	0	1	1	0	1	1	0	0.72–0.84	0.78
<i>T_C</i>	0	0	0	0	0	0	0	0	1	1	1	0.70–0.81	0.76
<i>C_D</i>	0	0	0	0	1	1	1	0	1	1	1	0.88–0.92	0.9
Weight	3	2	4	4	5	5	4	3	7	7	4	0.8–0.9	0.85

TABLE 16: The alignment of ABET 11 old SOs (a–k) to 7 new updated SOs (1–7).

New outcomes 1–7 (7)	Old outcomes a–k (11)										
	<i>a</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>e</i>	<i>F</i>	<i>G</i>	<i>h</i>	<i>I</i>	<i>J</i>	<i>k</i>
#1		√			√						
#2			√								
#3								√			
#4									√	√	
#5				√							
#6		√									√
#7										√	√

PEO-SO mappings for each PEO category and indicates that the *K_C* category performs the worst in both the number of instances and percentage accuracy, so this category is eliminated as it may be implicitly included in another PEO category. This classification model is considered the first step in the analysis of the assignment of PEOs and SOs. It is very useful for the academic specialist to avoid the incorrect assignment of PEOs to corresponding SOs during the design of the academic program.

4.3. Result Analysis of Applying the Apriori Rough Set Association Rule Mining Algorithm

4.3.1. Upper Bound Apriori Rough Set Association Rule Mining Algorithm. This section presents the results of the

TABLE 17: The alignment of lower bound PEO-SO map rules with the new updated SOs.

PEOs	New updated SOs							Avg. conf
	#1	#2	#3	#4	#5	#6	#7	
<i>P</i>	1	1	1	1	0	0	1	0.79
<i>C</i>	1	1	1	0	1	0	0	0.88
<i>G_S</i>	1	1	0	0	1	1	1	0.79
<i>T</i>	0	0	1	0	1	0	0	0.89
<i>E</i>	0	1	0	1	0	0	0	0.91
<i>L_L_L</i>	1	0	0	1	0	0	1	0.84
<i>S</i>	1	0	0	1	0	1	1	0.92
<i>L</i>	0	0	1	1	1	1	0	0.78
<i>T_C</i>	0	0	0	0	0	0	1	0.76
<i>C_D</i>	1	0	1	1	0	0	1	0.9
Weight	6	4	5	6	4	3	6	0.85

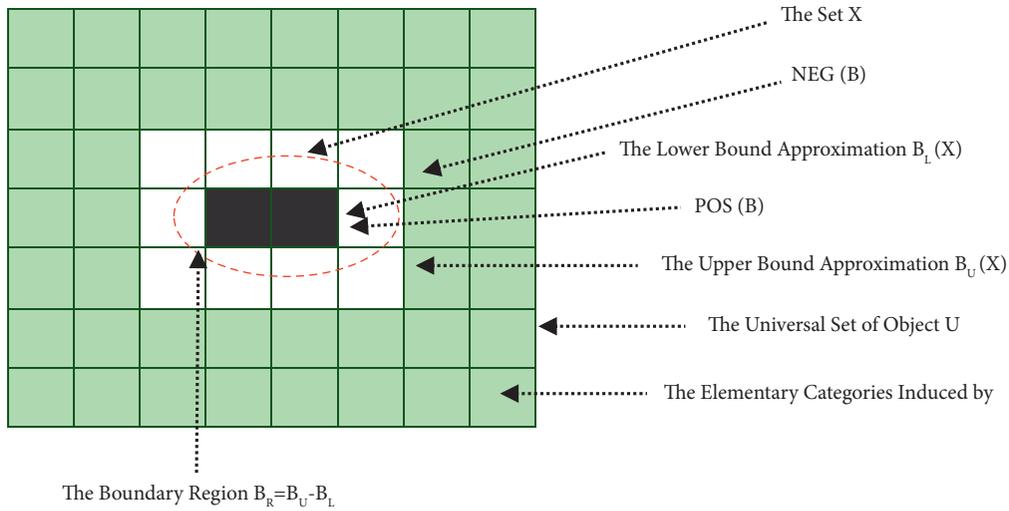


FIGURE 5: The Structure of the proposed approach for PEOs-SOs Map Design.

TABLE 18: Sample of the lower bound rough set rules for PEOs - SOs map.

	#1	#2	#3	#4	#5	#6	#7
<i>P</i>	1	1	1	1	0	0	1

TABLE 19: Rules interpretation for the lower bound rough set rules for PEOs - SOs map.

<i>P</i>	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>H</i>	<i>i</i>	<i>j</i>	<i>k</i>
	1	0	1	0	1	1	0	0	1	1	1

interpretation and analysis of the upper bound rules by applying the RSAR algorithm to the dataset. The obtained results are presented in Table 12, which shows the upper bound rough set association rules that govern the relationship between PEOs and SOs. To interpret the rules shown in Table 12, we consider the first row of *P* in the PEO-SO map as follows.

If PEO has only “1” SO entry, it means that this SO exists, or only “0,” which means that this SO does not exist. The upper bound rule for *P* can be processed using the gray area of results that may or may not exist with an average confidence

of 0.75. The obtained upper bound rules are shown in Table 12 and can be interpreted as follows: *E*, *S*, and *C* have the highest confidence level, which means that these PEOs are well defined in the proposed PEO-SO map, while *G_S* and *T_C* have the lowest confidence level, which means that these two PEOs are unclear in their mapping. Each PEO is defined with a rule associating all groups of student outcome SOs, and each association rule should be defined twice because each rule contains some of the student outcomes with two values of either 1 or 0. The upper bound rules are shown with SOs, which may or may not be present in the low confidence PEO-SO rules (see Tables 13 and 14).

4.3.2. Lower Bound Apriori Rough Set Association Rule Mining Algorithm. This section presents the results of the interpretation and analysis of the lower bound Apriori rough sets association rule (safe region) obtained by applying the (RSAR) algorithm to the dataset. Table 15 illustrates the lower bound rule sets of RSAR that govern the relationship between PEOs and SOs. To interpret the rules shown in

TABLE 20: Questionnaire results for the proposed PEO-SO maps.

	Questions	Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Average	Percentage	Result
Q1	The proposed PEO-SO mapping satisfies objectives correctly	47	0	2	1	0	0.98	98%	Agree
Q2	The proposed PEO-SO mapping helps to correctly allocate courses for academic engineering programs	42	2	4	1	1	0.96	96%	Agree
Q3	The proposed PEO-SO mapping is clear to faculty members	46	4	0	0	0	1	100%	Agree
Q4	The proposed PEO-SO mapping eliminates the ambiguity	43	3	2	1	1	0.96	96%	Agree
Q5	The proposed PEO-SO mapping satisfies the ABET requirements correctly	47	2	1	0	0	1	100%	Agree
Q6	The proposed PEO-SO mapping satisfies scientific and professional requirements	43	3	2	1	1	0.94	94%	Agree
Q7	The proposed PEO-SO mapping satisfies market and employers' requirements	45	2	3	0	0	1	100%	Agree
Q8	The proposed PEO-SO mapping can maximize chances	40	1	5	1	3	0.92	92%	Agree
Q9	The proposed PEO-SO mapping can be used as a standard reference for academic engineering programs	44	1	2	1	2	0.94	94%	Agree

Table 15 that govern the PEOs with 11 old SOs, we consider the first row P in the PEO-SO maps as follows.

P is the lower bound (safe region) rule that includes only the SOs with “1” or “0” entries, which means that the SO either exists or does not exist in rules such as $a, c, e, f, i, j,$ and k SOs, while SOs $b, d, g,$ and h do not exist in rule P . Therefore, the lower bound rule for P can be defined with only one rule, not two rules with an average confidence of 0.79, which is higher than that of the upper bound rule with an average confidence of 0.75.

The rules obtained can be summarized as follows: the lower bound RSAR is more concise and robust in its description; each rule can be described once, but not twice as in the upper bound RSAR. The comparison between upper bound RSAR conf. % and lower bound RSAR conf. % is shown in Figure 5. From this, it can be seen that the average confidence level for the lower bound rules is higher than for the upper bound rules for all PEO domains, proving their concreteness and robustness to the Apriori association rules.

4.4. Result Mapping to the New Updated ABET-EAC SOs. In this section, we present the mapping of the results obtained with the proposed approach to the new updated ABET-EAC SOs listed in Figure 2 (see Appendix A). We have mapped 11 (a–k) old SOs listed in Figure 2 (see Appendix A) with the new updated SOs as shown in Table 16. The mapping between old and updated SOs is created using a software program similar to the Word2Vec model that maps 11 old SOs to seven new updated SOs, as shown in Table 16.

Each new SO is assigned to 11 old SOs; e.g., outcome “#” is assigned to both old outcomes “a” and “e,” while “#5” is assigned to only one old outcome “d.” Table 17 illustrates the assignment of outcomes from the old to the new updated ABET-EAC SOs. The interpretation of the rules is explained

in Table 17, which governs the PEOs with the new seven SOs, as follows.

For example, we consider the first line of P in PEO-SO maps, which is the lower bound of the rule (safe area). It contains only the SOs with “1” or “0” entries, which means that the SO either exists or does not exist in this rule as SOs #1, #2, #3, #4, and #5 exist in rule P , while SOs #5 and #6 do not exist in rule P (see Tables 18 and 19).

4.5. Questionnaires, Analysis, and Feedback about the Obtained Results. This section is reserved for the validation and evaluation of the obtained results by 50 independent professionals and experts (lecturers and professors) from different academic engineering professions. The evaluation was performed by distributing a survey among these experts and obtaining their feedback for further analysis. The survey template was designed to determine the level of satisfaction with the results obtained from the proposed approach to PEO-SO maps. The survey and results of the analysis are presented in Table 20 and Figure 6 (see Appendix E); it includes 9 questions and five satisfaction levels: strongly agree, agree, neutral, disagree, and strongly disagree. The average column is calculated by taking the average of the 3 boxes: strongly agree, agree, and neutral. The survey showed a 97% agreement with the results obtained with the proposed approach.

4.6. Comparison with the Related Work Presented. In this section, we present the comparison of the proposed approach with the work presented in [13], which uses the Apriori association rule mining technique to discover the rules describing the relationship between PEOs and SOs, while our approach uses the rough set Apriori association rule mining technique. Our proposed approach uses the

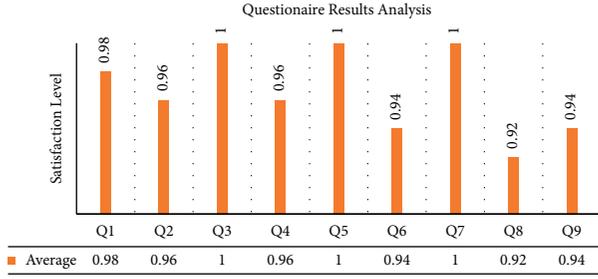


FIGURE 6: The feedback of survey analysis on the PEO-SO map.

bagged J48 classifier to detect and resolve the confusion in the association of different PEO-SOs. Our approach uses the text processing software similar to the Word2V model to perform automatic preprocessing of the merged PEOs into PEO-SO maps, while this was performed manually in [13]. Figure 5 illustrates a graphical comparison of the obtained results as lower bound and upper bound rules conf. % of the Apriori association results in [13]. It shows that our proposed approach performs better for the upper bound and lower bound rules with higher conf. %. It concludes that our proposed approach develops a PEO-SO map with a higher conf. % and a smaller number of rules with a small size, as well as greater robustness. The work presented in [13] describes the PEO-SO map with 22 rules and ours with only 11 rules (lower bound 50% reduction in the number of rules). The work uses a questionnaire survey to evaluate the obtained results by 50 different specialists which is not the case in [13].

5. Conclusion and Future Extensions

This paper presents the framework for the proposed hybrid approach to support the design and accreditation process in academic engineering programs, which uses three different data mining techniques to determine the best design for PEO-SO. It can minimize the effort required to design a PEO-SO map and the time required to prepare the accreditation processes. The proposed approach has succeeded in designing a correct assignment PEO-SO map with a minimal set of rules and small rule sizes. The paper presented different phases for the development of the proposed approach, including dataset collection and preprocessing, data mining construction, modeling, analysis, interpretation, and evaluation of the obtained results. The main feedback from the questionnaires proved the validity of the proposed approach with 97% agreement in the design and minimizing the effort required for accreditation requirements. The comparison with related methods showed that our proposed approach performs better with a high percentage of performance. The limitation of the presented work is the choice of the confidence rate threshold, which distinguishes between lower and upper bounds. Future extensions of this work will be the use of natural language processing (NLP) techniques for the preprocessing phase of PEOs. The second extension will be the generalization of the proposed approach to all academic courses. The third extension will be the development of an interactive web-based system for the

proposed approach. Finally, additional datasets will be collected and preprocessed for use as a benchmark dataset.

Appendix

A. A List of ABET 11 Old SOs (11 from a to k and New Updated 7 SOs (7 from 1 to 7)

The list of ABET is given in the following figure.

B. Examples of Support and Confidence Measure calculations

The measure support examples are as follows.

C. An Explanatory Example for the Rough Sets Association Rules (RSAR) for PEO-SO Map

This section illustrates an explanatory example to illustrate the steps for modifying the Apriori association rules using the rough set theory technique as shown in Table 1, which presents the comparison between the association rule and the rough set association rule algorithms in the way of computing the support and confidence percentage. This section uses a sample from the collected preprocessed dataset PEO-SO maps as follows: the value of an attribute is described as $\varepsilon = (a, \alpha)$, $a \in A$, $\alpha \in \alpha$, and $[\varepsilon]$ describes a universe unit in the U set (the set of all samples). Each attribute has α value and unit, and their s (support) can be estimated as follows:

- (i) $[\varepsilon]_1: [\{a\}] = \{1, 2, 7, 11\}$ and $s([\varepsilon]_1) = 4$.
- (ii) $[\varepsilon]_2: [\{b\}] = \{1, 2, 7, 12, 13\}$ and $s([\varepsilon]_2) = 5$.
- (iii) $[\varepsilon]_3: [\{c\}] = \{1, 2, 3, 4, 7, 10, 11, 13, 14, 15\}$ and $s([\varepsilon]_3) = 10$.
- (iv) $[\varepsilon]_4: [\{d\}] = \{1, 6, 7, 10, 11, 12, 15\}$ and support $([\varepsilon]_4) = 7$.
- (v) $[\varepsilon]_5: [\{e\}] = \{1, 2, 3, 9, 13\}$ and $s([\varepsilon]_5) = 5$.
- (vi) $[\varepsilon]_6: [\{f\}] = \{1, 2, 4, 6, 8, 9, 11, 14\}$ and $s([\varepsilon]_6) = 8$.
- (vii) $[\varepsilon]_7: [\{g\}] = \{1, 6, 8, 9, 10, 11, 12, 13, 15\}$ and $s([\varepsilon]_7) = 9$.
- (viii) $[\varepsilon]_8: [\{h\}] = \{1, 5, 7, 9, 13\}$ and $s([\varepsilon]_8) = 5$.
- (ix) $[\varepsilon]_9: [\{i\}] = \{1, 2, 3, 5, 8, 9, 11, 13\}$ and $s([\varepsilon]_9) = 8$.
- (x) $[\varepsilon]_{10}: [\{j\}] = \{1, 3, 5, 6, 7, 8, 9, 11, 13\}$ and $s([\varepsilon]_{10}) = 9$.
- (xi) $[\varepsilon]_{11}: [\{k\}] = \{1, 3, 5, 11, 13\}$ and $s([\varepsilon]_{11}) = 5$.

For example, to combine 6 attribute values (c, d, f, i, j , and k), the elementary set of $B = \{c, d, f, i, j$, and $k\}$ is defined as follows:

First step:

- (i) $[\varepsilon]_3: [\{c\}] = \{1, 2, 3, 4, 7, 10, 11, 13, 14, 15\}$ and $s([\varepsilon]_3) = 10$.
- (ii) $[\varepsilon]_4: [\{d\}] = \{1, 6, 7, 10, 11, 12, 15\}$ and $s([\varepsilon]_4) = 7$.
- (iii) $[\varepsilon]_6: [\{f\}] = \{1, 2, 4, 6, 8, 9, 11, 14\}$ and $s([\varepsilon]_6) = 8$.
- (iv) $[\varepsilon]_9: [\{i\}] = \{1, 2, 3, 5, 8, 9, 11, 13\}$ and $s([\varepsilon]_9) = 8$.
- (v) $[\varepsilon]_{10}: [\{j\}] = \{1, 3, 5, 6, 7, 8, 9, 11, 13\}$ and $s([\varepsilon]_{10}) = 9$.
- (vi) $[\varepsilon]_{11}: [\{k\}] = \{1, 3, 5, 11, 13\}$ and $s([\varepsilon]_{11}) = 5$.

Second step:

$\{[\varepsilon]_3, [\varepsilon]_4, [\varepsilon]_6, [\varepsilon]_9, [\varepsilon]_{10}, \text{ and } [\varepsilon]_{11}\} = \{c, d, f, i, j, \text{ and } k\} = \{1, 11\}$ and $s([\varepsilon]_3, [\varepsilon]_4, [\varepsilon]_6, [\varepsilon]_9, [\varepsilon]_{10}, \text{ and } [\varepsilon]_{11}) = 2$. $s\%$ and $c\%$ of class association rules are computed directly, and S and C are calculated by the following two equations:

$$S = \frac{B_L(i) \cup B_L(y)}{\Omega}, \quad (1)$$

$$C = \frac{|B_L(i) \cup B_L(y)|}{B_L(i)}. \quad (2)$$

B_L is defined as the lower bound approximation in rough set theory, $B_L(i) \cup B_L(y)$ indicates the number of i occurs in conjunction with y across the dataset in the table, and $|\Omega|$ is the number of all the samples in the dataset table.

Then, the rough set association rule is as follows: $RSAR = \{c, d, f, i, j, k \Rightarrow CD\}$. The elementary of B' set in the condition of the rule contains two attributes and is defined as follows:

- (i) $\{\text{condSet}\} = \{[\varepsilon]_c\} = \{c, d, f, i, j, k\} = \{1, 11, 13\}$, and $s([\varepsilon]_c) = 2$.
- (ii) $\{\text{decSet}\} = \{[\varepsilon]_d\} = \{C_D\} = \{1, 11\}$ and $s([\varepsilon]_d) = 2$.
- (iii) $S\%$ of $RSAR = s\{\text{condSet decSet}\} = s\{[\varepsilon]_c, [\varepsilon]_d\} = s\{1, 11\} = 2$.
- (iv) $|\Omega| = 15$.

Then, $S\%$ of $(RSAR)$ is $([\varepsilon]_d)/|\Omega| = 2/15 = 13\%$. $c\%$ of $RSAR$ is $S(RSAR)/s([\varepsilon]_c) = 2/2 = 100\%$.

D. Detailed Accuracy by PEO Category

The performance summary and detailed accuracy are given as follows.

E. Questionnaire Results for the Proposed PEO-SO Map

The questionnaire results are as follows.

Data Availability

The unavailability of the data set because it is collected manually and preprocessed to be suitable for further processing.

Conflicts of Interest

The author declares no conflicts of interest.

References

- [1] S. Subbaraman, R. R. Jagtap, and S. S. Shinde, "Outcome-based learning: a case study," in *Proceedings of the 2013 IEEE International Conference in MOOC, Innovation and Technology in Education (MITE)*, Jaipur, India, 2013.
- [2] S. Balachandran and L. Balachandran, *Guidelines for Assessment and Evaluation of Engineering and Engineering Technology Programs*, 2007, <https://dokumen.tips/documents/guidelines-for-assessment-and-evaluation-of-engineering-ilinaseeorgconference2007programpapersconference.html?page=1>.
- [3] S. I. Malik, "Assessing the teaching and learning process of an introductory programming course with Bloom's taxonomy and Assurance of Learning (AOL)," *International Journal of Information and Communication Technology Education (IJICTE)*, vol. 15, no. 2, pp. 130–145, 2019.
- [4] F. Bouslama, A. Lansari, A. Mahmoud Al-Rawi, and A. Abonamah, "A novel outcome-based educational model and its effect on student learning, curriculum development, and assessment," *Journal of Information Technology Education: Research*, vol. 2, no. 1, pp. 203–214, 2003.
- [5] M. Iqbal Khan, S. M. Mourad, and W. M. Zahid, "Developing and qualifying civil engineering programs for ABET accreditation," *Journal of King Saud University-Engineering Sciences*, vol. 28, no. 1, pp. 1–11, 2016.
- [6] S. M. A. Ghaly, "An indirect evaluation of program educational objectives and student outcomes for engineering programs: a case study," *Engineering, Technology & Applied Science Research*, vol. 10, no. 5, 2020.
- [7] M. Injadat, A. Moubayed, A. B. Nassif, and A. Shami, "Systematic ensemble model selection approach for educational data mining," *Knowledge-Based Systems*, vol. 200, Article ID 105992, 2020.
- [8] H. Aldowah, H. Al-Samarraie, and W. M. Fauzy, "Educational data mining and learning analytics for 21st century higher education: a review and synthesis," *Telematics and Informatics*, vol. 37, pp. 13–49, 2019.
- [9] M. Ashraf, M. Zaman, and M. Ahmed, "An intelligent prediction system for educational data mining based on ensemble and filtering approaches," *Procedia Computer Science*, vol. 167, 2020.
- [10] F. Martínez-Abad, A. Gamazo, and M. J. Rodríguez-Conde, "Educational Data Mining: Identification of factors associated with school effectiveness in PISA assessment," *Studies in Educational Evaluation*, vol. 66, Article ID 100875, 2020.
- [11] G. Akçapınar and M. Kokoç, "Analyzing the relationship between student's assignment submission behaviors and course achievement through process mining analysis," *Türk Bilgisayar Ve Matematik Eğitimi Dergisi*, vol. 11, no. 2, pp. 386–401, 2020.
- [12] A. Shafi, S. Saeed, Y. A. Bamarouf, S. Z. Iqbal, N. Min-Allah, and M. A. Alqahtani, "Student outcomes assessment methodology for ABET accreditation: a case study of computer science and computer information systems programs," *IEEE Access*, vol. 7, 2019.
- [13] A. A. Yahya and A. Osman, "Using data mining techniques to guide academic programs design and assessment," *Procedia Computer Science*, vol. 163, pp. 472–481, 2019.
- [14] J. Lin, H. Pu, Y. Li, and J. Lian, "Intelligent recommendation system for course selection in smart education," *Procedia Computer Science*, vol. 129, pp. 449–453, 2018.
- [15] J. Alhiyafi and I. Abuaqel, "An automated system for collection and analysis of ABET curriculum requirements," in *Proceedings of the 2018 International Conference on Computational Science and Computational Intelligence (CSCI)*, Las Vegas, NV, USA, 2018.
- [16] N. Ard, S. Beasley, and K. Nunn-Ellison, "Quality education through accreditation," *Teaching and Learning in Nursing*, vol. 12, no. 2, pp. 85–87, 2017.
- [17] J. Miller, "Minimizing effort for ABET Student Outcomes Assessment while maintaining effective results," in *Proceedings of the 2016 International Conference on Computational*

- Science and Computational Intelligence (CSCI)*, Las Vegas, NV, USA, 2016.
- [18] T. Slimani, "Class association rules mining-based rough set method," 2015, <https://arxiv.org/ftp/arxiv/papers/1509/1509.05437.pdf>.
- [19] X. Zheng and J. Xu, "Studies on the application of rough set analysis in mining of association rules and the realization in provincial road transportation management information system," in *Proceedings of the 2012 International Conference on Industrial Control and Electronics Engineering*, Xi'an, China, 2012.
- [20] A. Naimpally, H. Ramachandran, and C. Smith, "Accreditation of engineering programs and their relationship to lifelong learning," *Lifelong Learning for Engineers and Scientists in the Information Age*, vol. 11, 19 pages, 2012.
- [21] A. Osman, A. A. Yahya, and M. B. Kamal, "A dataset of program educational objectives mapped to ABET outcomes: data cleansing, exploratory data analysis, and modeling," *International Journal of Computer and Information Engineering*, vol. 12, no. 6, pp. 479–485, 2018.
- [22] J. Xun, L. C. Xu, and L. Qi, "Association rules mining algorithm based on a rough set," *Proc International Symposium on Information Technologies in Medicine and Education*, vol. 1, pp. 361–364, 2012.
- [23] Y. L. Chen and C. H. Weng, "Mining fuzzy association rules from questionnaire data," *Knowledge-Based Systems*, vol. 22, no. 1, pp. 46–56, 2009.
- [24] A. K. Sigurdardottir, H. Jonsdottir, and R. Benediktsson, "Outcomes of educational interventions in type 2 diabetes: WEKA data-mining analysis," *Patient Education and Counseling*, vol. 67, no. 1-2, pp. 21–31, 2007.
- [25] R. S. Baker, "Data mining for education," *International Encyclopedia of Education*, vol. 7, no. 3, pp. 112–118, 2010.
- [26] S. Raj S and M. Nandhini, "Ensemble human movement sequence prediction model with Apriori-based probability tree classifier (APTC) and Bagged J48 on machine learning," *Journal of King Saud University-Computer and Information Sciences*, vol. 33, no. 4, pp. 408–416, 2021.
- [27] S. Ruggieri, "Efficient C4.5 classification algorithm," *IEEE Transactions on Knowledge and Data Engineering*, vol. 14, no. 2, pp. 438–444, 2002.
- [28] S. Sahu and B. M. Mehtre, "Network intrusion detection system using J48 Decision Tree," in *Proceedings of the 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Kochi, India, 2015.
- [29] S. Banitaan, M. Azzeh, and A. B. Nassif, "User movement prediction: the contribution of machine learning techniques," in *Proceedings of the 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Anaheim, CA, USA, 2016.
- [30] B. Xiao, Q. F. Xu, Z. Q. Lin, J. Guo, and C. G. Li, "Credible association rule and its mining algorithm based on maximum clique," *Journal of Software*, vol. 19, no. 10, 2008.
- [31] R. Agrawal, T. Imieliński, and A. Swami, "Mining association rules between sets of items in large databases," *ACM SIGMOD Record*, vol. 22, no. 2, pp. 207–216, 1993.
- [32] R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, and A. I. Verkamo, "Fast discovery of association rules," *Advances in Knowledge Discovery and Data Mining*, vol. 12, no. 1, pp. 307–328, 1996.
- [33] Q. Wang, M. J. Rowan, and R. Anwyl, " β -amyloid-mediated inhibition of NMDA receptor-dependent long-term potentiation induction involves activation of microglia and stimulation of inducible nitric oxide synthase and superoxide," *Journal of Neuroscience*, vol. 24, no. 27, 2004.
- [34] S. Y. Kim, J. E. Park, Y. J. Lee et al., "Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity," *Journal of Clinical Epidemiology*, vol. 66, no. 4, pp. 408–414, 2013.
- [35] Z. Pawlak, "Rough sets," *International Journal of Computer and Information Sciences*, vol. 11, no. 5, pp. 341–356, 1982.
- [36] Z. Pawlak, "Rough sets and intelligent data analysis," *Information Sciences*, vol. 147, no. 1-4, pp. 1–12, 2002.
- [37] S. Greco, B. Matarazzo, and R. Slowinski, "Rough sets theory for multicriteria decision analysis," *European Journal of Operational Research*, vol. 129, no. 1, pp. 1–47, 2001.
- [38] S. S. Dey, A. K. Singh, D. Chandel, and T. K. Behera, "Genetic diversity of bitter melon (*Momordica charantia* L.) genotypes revealed by RAPD markers and agronomic traits," *Scientia Horticulturae*, vol. 109, no. 1, pp. 21–28, 2006.
- [39] G. Xue-Jie, W. Mei-Li, and F. Giorgi, "Climate change over China in the 21st century as simulated by BCC_CSM1.1-RegCM4.0," *Atmospheric and Oceanic Science Letters*, vol. 6, no. 5, pp. 381–386, 2013.
- [40] H. Takiguchi, S. Kaneiwa, H. Kudo, C. Sakane, and T. Yoshida, "European patent application No. 19880300330," 1993, <https://www.epo.org/searching-for-patents.html>.
- [41] D. Parmar, T. Wu, and J. Blackhurst, "MMR: An algorithm for clustering categorical data using rough set theory," *Data and Knowledge Engineering*, vol. 63, no. 3, pp. 879–893, 2007.