

## Short Text Mining for Classifying Educational Objectives and Outcomes

Yousef Asiri\*

College of Computer Science and Information Systems, Najran University, Najran, 61441, Saudi Arabia

\*Corresponding Author: Yousef Asiri. Email: yasiri@nu.edu.sa

Received: 09 May 2021; Accepted: 11 June 2021

**Abstract:** Most of the international accreditation bodies in engineering education (e.g., ABET) and outcome-based educational systems have based their assessments on learning outcomes and program educational objectives. However, mapping program educational objectives (PEOs) to student outcomes (SOs) is a challenging and time-consuming task, especially for a new program which is applying for ABET-EAC (American Board for Engineering and Technology the American Board for Engineering and Technology—Engineering Accreditation Commission) accreditation. In addition, ABET needs to automatically ensure that the mapping (classification) is reasonable and correct. The classification also plays a vital role in the assessment of students' learning. Since the PEOs are expressed as short text, they do not contain enough semantic meaning and information, and consequently they suffer from high sparseness, multidimensionality and the curse of dimensionality. In this work, a novel associative short text classification technique is proposed to map PEOs to SOs. The datasets are extracted from 152 self-study reports (SSRs) that were produced in operational settings in an engineering program accredited by ABET-EAC. The datasets are processed and transformed into a representational form appropriate for association rule mining. The extracted rules are utilized as delegate classifiers to map PEOs to SOs. The proposed associative classification of the mapping of PEOs to SOs has shown promising results, which can simplify the classification of short text and avoid many problems caused by enriching short text based on external resources that are not related or relevant to the dataset.

**Keywords:** ABET accreditation; association rule mining; educational data mining; engineering education; program educational objectives; student outcomes; associative classification

### 1 Introduction

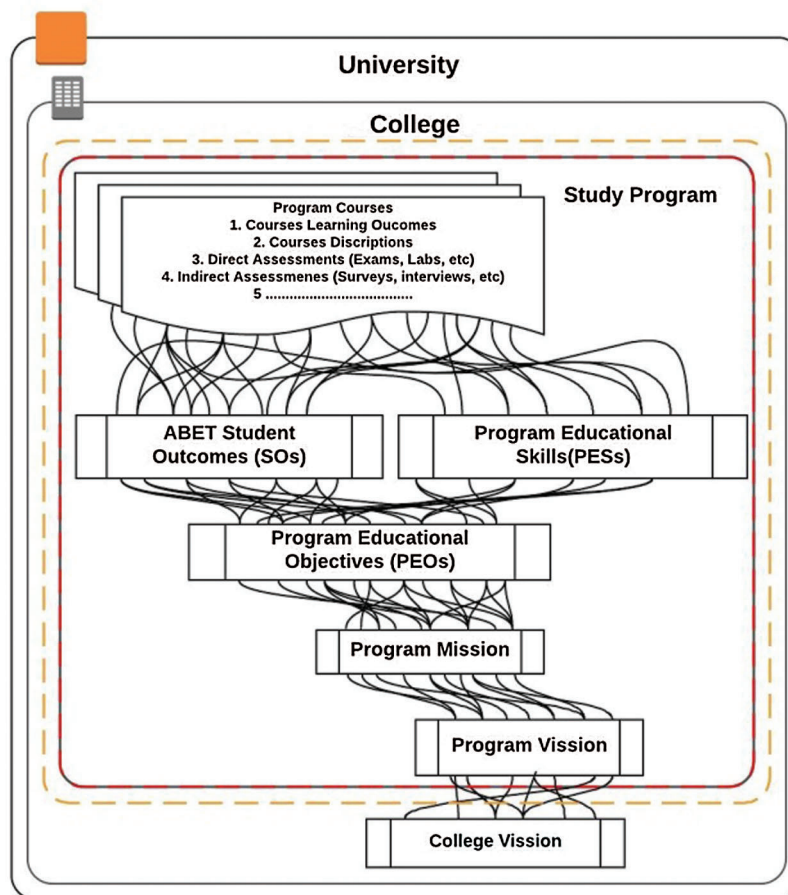
The Engineering Technology Accreditation Commission (ETAC) of the ABET outcome-based accreditation organization assures quality education of engineering programs. All programs applying for this accreditation must demonstrate that they meet eight general criteria governing students, program educational objectives (PEOs), student outcomes (SOs), continuous improvement, curriculum, faculty, facilities and institutional support. PEOs are broad statements that describe the career and professional



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

accomplishments (program education skills) that the program is preparing graduates to achieve. The programs must publish PEOs which are consistent with the mission of the institution, the needs of the program's various constituencies, and the criteria of ETAC. In addition, the programs must have documented and publicly stated and adopted ABET SOs that map to the PEOs. SOs describe what students are expected to know and be able to do by the time they graduate. The programs must submit their self-study reports (SSRs) to ABET. SSRs are the primary documents that programs use to demonstrate their compliance with all applicable ABET criteria and policies. SSRs are the foundation for the review team's judgement of whether a program meets ABET criteria for accreditation. It addresses all paths to completion of the degree, all methods of instructional delivery used for the program, and all remote location offerings [1–4].

Fig. 1 shows the importance of PEOs at any program, department, college or university. It also shows the relationship between PEOs, SOs and Program Educational Skills (PESs) as well as their central roles in any educational program, college or university. The relationship between the three entities forms an undirected cyclic graph, and the relationship between each pair of the entities is a many-to-many relationship. The most common form of curriculum mapping, which shows courses at the program level, makes it visible how courses align to the learning outcomes towards which that curriculum strives.



**Figure 1:** The relationships between PEOs, PESs and SOs and other related components of an educational institution

All PEOs presented in SSRs of engineering programmes are manually mapped to the ABET SOs. They represent one of the most significant parts of SSRs because they are considered the ultimate purpose by which the quality of programmes is judged. They are directly mapped to the SOs and PESs and indirectly through SOs to all courses of a programme, including course learning outcomes, as depicted in Fig. 1. Every course learning outcome (CLO) must be directly mapped to at least one SO. The mapping between PEOs and SOs can be considered a bridge which links the core components of a programme (e.g., courses) on one side to the department, college and university on the other side (e.g., missions and visions). With this mapping in hand, programme administrators can then develop a programme map; this will help them to adjust the alignment of courses within a degree programme with greater ease and efficiency.

The original concept of outcome-based education should include cyclic continual improvement with meaningful revision of teaching and learning strategies, delivery and assessment methods. The complexity is also increased during the creation of the mapping between PEOs and SOs [5]. Mapping PEOs to SOs is a challenging and time-consuming task, especially for new programmes applying for ABET accreditation. In addition, ABET must also ensure that the mapping (classification) makes sense and is correct. The classification also plays a vital role in the assessment of student learning [6].

The contribution of this work is to use a novel associative-short text classification technique to map PEOs to SOs using PESs as delegate classifiers. In the proposed approach, association rule mining (ARM) techniques [7–9] are utilized to inductively interrogate a set of PES→SO mapping data extracted from 152 self-study reports (SSRs) of engineering programmes accredited by EAC-ABET. Since PEOs are short texts, they face many difficulties in enriching or expanding the dataset, and a set of rules extracted from mappings of SOs to PESs are employed as delegate classifiers to map PEOs to SOs.

The structure of this paper is as follows: chapter 2 shows related work, chapter 3 describes the proposed model, chapter 4 describes the results and analysis, chapter 5 presents the discussion and chapter 6 presents the conclusion and discusses future work.

## 2 Related Work

Short text streams do not contain enough semantic meaning and information because they are too short and suffer from high sparseness and multidimensionality. It is quite simple to bring about the curse of dimensionality, which makes it difficult to obtain statistically sound and reliable results because the amount of data required to support the result often grows exponentially with the dimensionality. Short text data can change easily over time, which may result in poor performance of the predictive models. This is known as concept drift in the field of machine learning. Therefore, it is very difficult to directly implement conventional classification techniques for short text streams, such as support vector machines [10] and random forest [11]. To tackle the sparseness and multidimensionality problems of short texts in classification, there are two main approaches adopted in the literature [12]:

- a. Enriching or expanding short texts with additional related data extracted from external sources of knowledge used for short text topic modelling, before training a standard topic model. For example, [13] trained a topic model on massive external data collected from Wikipedia or MEDLINE and then predicted the topic distribution of short texts using latent Dirichlet allocation (LDA) [14]. Reference [15] proposed a dual-LDA model that learns topics over both the original short text and related long documents.
- b. Expanding short texts by using rules and statistics hidden in short texts, such as by applying distributional semantic models [16,17]. For instance, Kim et al. [18] proposed a novel kernel for short text classification, called a language independent kernel based on semantic annotation.

Although both of the above-mentioned approaches to some extent contribute to improving the efficiency of short text classification, there are some serious drawbacks which need addressing. Enriching short text by employing external resources adds to the original text or replaces it with unrelated and noisy information, especially when naively expanding the data.

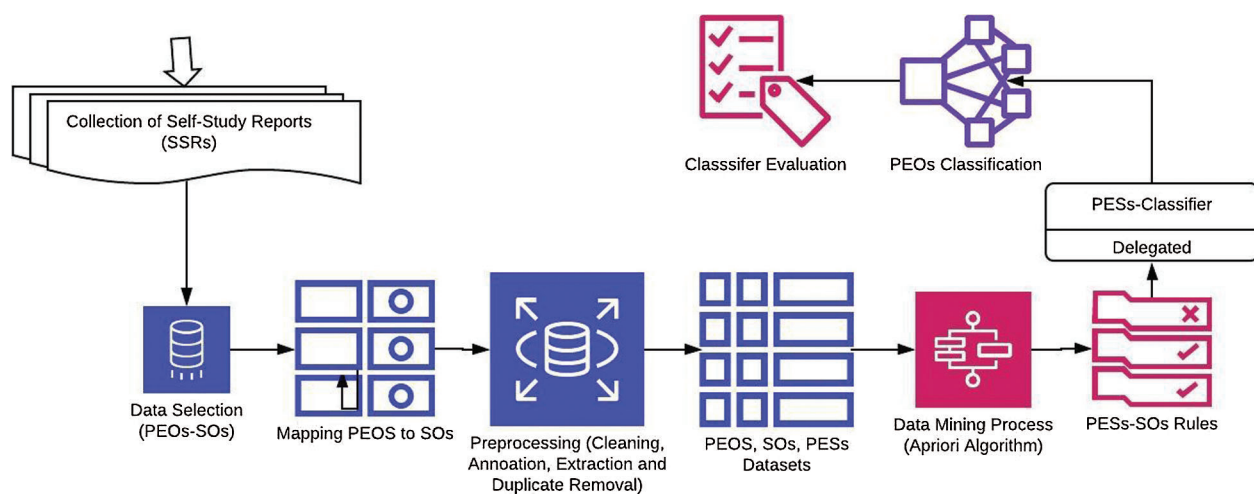
The basic idea of these kinds of models is to learn the topics from domain-related datasets and assume that each text is a multinomial distribution over these topics [12]. As the number of topics is relatively small, the dimensionality of each text becomes lower, and the vector space of texts is no longer sparse. In real applications, however, a text may be bound up with a small number of topics and often has no relations with others at all. Applying only the topic distribution has obvious limitations, especially for dealing with short texts [19]. In addition, external resources cannot reliably provide the most suitable new data. The features extracted using the second approach do not necessarily have the appropriate level of generalization to overcome the sparseness of the short text. In the two approaches, the data expansion process creates new terms and/or phrases that do not appear clearly in the original text [16].

The objective of this study is to overcome the above-mentioned drawbacks, we propose a novel approach called classification based on associations by delegation (CBAD), in which the association rules of program educational skills (PESs) and SOs are combined with a classifier and then used as delegates to classify PEOs with SOs.

### 3 The Proposed Model

The process of merging association rules and classification in data mining, called associative classification (AC), has attracted attention as a promising research discipline with great potential [20]. A limited number of studies have explored the topic of classification combined with another task in data mining, such as association or finding sequential patterns. Several algorithms that combine classification with other data mining tasks are classification based on associations (CBA), which combines classification with association rules, and the classify by sequence (CBS) algorithm, which combines classification with sequential patterns [21].

In this paper, we propose the classification based on associations by delegation (CBAD) model, as shown in Fig. 2.



**Figure 2:** Proposed model-classification based on associations by delegation (CBAD)

The CBAD technique aims to overcome the problems faced by enriching and expanding the short text by employing the association rules created from mapping PESs to SOs and using these rules to build a classifier to classify the PEOs with SOs by delegation. The aim of the association-rules technique is to reveal a correlation or association between attributes, while classification is the prediction of a categorical variable, which is known as a class label, based on training data. Delegation allows the same classifier to predict the class labels of instances of another dataset that are directly related to the classifier and the dataset used in the training process. For example, in a rule such as  $Pes1, Pes2 \rightarrow C1$ ,  $C1$  must be a class attribute, while  $Pes1$  and  $Pes2$  are attribute values. This rule can be interpreted as follows: if the  $Pes1$  and  $Pes2$  attribute values appear together for any instance, this instance can be classified as  $C1$  [20,22,23]. In our approach, when implementing the delegation process, the rule  $(Peo1 \rightarrow Pes1) \& (Pes1 \rightarrow C1)$ ,  $Pes1 \rightarrow C1$  is used and can be interpreted as follows: if the attribute  $Peo1$  is mapped to attribute  $Pes1$  and  $Peo1$  is classified as  $C1$ , then  $Peo1$  can be classified as  $C1$ .

### 3.1 The Methodology

In this section, we describe the datasets used in the research and the proposed associative classification based on delegation.

#### 3.1.1 Data Collection

The datasets used in this paper were collected from SSRs that were produced in operational settings at some ABET accredited engineering programs under procedures managed and operated by specialists and experts in the area. The PEOs and SOs together with their mapping were extracted from these SSRs. Most of the information related to the datasets used in this work have been used in previously published research [6,24,25].

#### 3.1.2 Data Pre-Processing

Data pre-processing has a substantial impact on improving the quality of the experimental datasets of any machine learning model because unreliable samples may lead to erroneous outputs. The pre-processing step involves substantial verification and validation of the content of the datasets. First, the datasets were cleaned, and some spelling errors and duplicates were eliminated. Eventually, a set of PESs were developed and used as labels to annotate each PEO instance with one or multiple PESs. Based on the requirements of the research, five datasets were extracted from these SSRs, namely, PEOs, SOs, PESs, PEO-PES, and PES-SO datasets. Tab. 1 shows the statistical aspects of the PES-SO dataset.

**Table 1:** Statistical aspects of PES-SO data [6,25]

Statistical aspect	Value
No. of instances	667
No. of programmes	152
Max. no. of textual PESs in a programme	17
Min. no. of textual PESs in a programme	2
Avg. no. of textual PESs in a programme	4.39

### 3.1.3 PES Labels

A set of common PES labels has been proposed for this work, as shown in [Tab. 2](#). The second column shows the 12 categories of the PESs used in this work. The third column shows the abbreviations of the PES categories (labels), and the last column shows the frequency of each label in the dataset.

**Table 2:** PES labels [6,25]

No.	PES category	PES label	Frequency
1	Life-long learning	LL	130
2	Communication	C	82
3	Leadership	L	62
4	Teaming	T	98
5	Ethical conduct	EC	81
6	Professionalism	P	137
7	Social and community	SC	78
8	Career success	CS	127
9	Technical competency	TC	212
10	Knowledge competency	KC	96
11	Graduate studies	GS	64
12	Others	O	29

## 3.2 The Algorithm

As mentioned previously, the objective of this research is to extract association rules between PESs and SOs and employ these rules to classify PEOs. Our proposed model has three main stages: association-rule generation, pruning and classification by delegation.

### 3.2.1 Association-Rule Generation and Pruning

A widely used Apriori algorithm is the ARM method for discovering interesting relationships between variables in a dataset, and it is adopted in this work. It is intended to identify strong rules extracted from datasets using a measurement of interestingness. The method adopted in this work is based on the concept of strong rules [26]. The rule extraction process is performed in two stages:

1. Mining frequent itemsets.
2. Discovering rules from these frequent itemsets.

Frequent itemsets are a form of frequently occurring pattern or item in a dataset. The classical procedure of finding frequent itemsets is simple, but it is very time consuming because of the large number of potential combinations to be processed. When they have been extracted, the production of rules is a very easy task, which is based on the following premise: “all sub-itemsets of a frequent itemset must also be frequent”.

Using this premise, the Apriori algorithm prunes a huge number of examinations of itemsets since it is certain that they are not frequent. Frequent sub-itemsets are extended one item at a time (candidate generation), and groups of candidates are examined. The algorithm terminates when no further extensions are found. In other words, the Apriori algorithm generates candidate itemsets of length  $L$  from itemsets of length  $L-1$  and then prunes the candidates that have a non-frequent sub-itemset. Thus, it keeps only the

frequent itemsets among the candidates. Mathematically, the Apriori algorithm works by iteratively generating the frequent  $k$ -itemsets whose count is equal to or greater than a pre-specified minimum support count. The support is the fraction of the total number of instances in which the itemset occurs, which is calculated by Eq. (1) [6,24].

$$\text{support}(\{\text{PESs}\} \rightarrow \{\text{SOs}\}) = \frac{\text{count}(\{\text{PESs}\} \cup \{\text{SOs}\})}{N} \quad (1)$$

This means that the support ( $\{\text{PESs}\} \rightarrow \{\text{SOs}\}$ ) is calculated by dividing the total number of instances that contain both PESs and SOs by the total number of instances ( $N$ ).

Thereafter, the process of joining and pruning is applied to the current frequent  $k$ -itemsets to create frequent  $(k + 1)$ -itemsets. The frequent  $k$ -itemsets are combined, and the  $k$ -itemsets that cannot reach or exceed the minimum support count are pruned; these are considered infrequent itemsets. The process of joining and pruning is repeated until the frequent itemsets become empty. Then, the algorithm halts, and the process of association rules generation begins. The confidence level of the rule must satisfy a confidence threshold pre-specified by the user. The confidence level of the rules is calculated using Eq. (2).

$$\text{confidence}(\{\text{PESs}\} \rightarrow \{\text{SOs}\}) = \frac{\text{count}(\{\text{PESs}\} \cup \{\text{SOs}\})}{\text{Ncount}(\{\text{PESs}\})} \quad (2)$$

The Apriori algorithm's work is shown by the following pseudo code:

- Join Step:  $C_k$  is generated by joining  $L_{k-1}$  with itself
- Prune Step: Any  $(k - 1)$ -itemset that is not frequent cannot be a subset of a frequent  $k$ -itemset
- Pseudo-code:

```

 $C_k$ : Candidate itemset of size  $k$ 
 $L_k$ : frequent itemset of size  $k$ 
 $L_1 = \{\text{frequent items}\};$ 
for ( $k = 1$ ;  $L_k \neq \emptyset$ ;  $k++$ ) do begin
     $C_{k+1} = \text{candidates generated from } L_k;$ 
    for each instance  $i$  in dataset do
        increment the count of all candidates in  $C_{k+1}$  that are contained in  $i$ 
     $L_{k+1} = \text{candidates in } C_{k+1} \text{ with min\_support}$ 
end
return  $\cup_k L_k;$ 

```

### 3.2.2 Associative Classification by Delegation

The problems of enriching the short text (PEOs) can be solved by extracting data from courses such as exams, tutorials, and lab materials indirectly through SOs, which in this case are called relative external datasets, instead of using related datasets from external resources. They are relative because there is a link or relationship between them. However, the main difficulty of adopting this approach is that the data are not available, and even if they were available, they are not publicly accessible; and so, the mapping between the PEOs and SOs is missed. Therefore, since the three entities form an undirected cyclic graph, which can be considered a circular channel to transfer information among them, our proposed method utilizes the association rules created from the SOs and PESs as delegates to classify the PEOs using SOs as class labels; in this paper, we call our method CBAD. Delegation is assignment of any authority to

another person (normally from a manager to a subordinate to carry out specific activities). However, the person who delegated the work remains accountable for the outcome of the delegated work. In our case, from Fig. 1, we can consider the PEOs as a manager.

The general description of the CBAD algorithm is depicted in Fig. 1, and details are given in the following subsections.

1. Input: Datasets PESs, SOs and PESs, Support and Confidence Thresholds, Number of Top Association Rules (n)
2. Find the candidate single itemset PEO→ SOs.
3. Find the support for each candidate (using Eq. (1)).
4. For each itemset PEO→ SOs:
5. If the support (PEO→ SOs) < minimum support
6. Remove from the list
7. Else if confidence (PEO→ SOs) >= minimum confidence (calculated using Eq. (1))
8. Add to Class Association Rules (CARs)
9. Else
10. Leave it in the itemset PEO→ SOs.
11. If all items visited
12. Find next candidate itemset PEO→ SOs
13. If the itemset PEO→ SOs is not empty
14. Go to step 2
15. Else
16. For each PEO→ SOs
17. Select the Top n Association Rules
18. For each Selected Rule (PEO→ SOs)
19. If PEO → SOs and PEO → PES
20. Class of PES is SOs
21. end

## 4 Analysis and Results

We conducted an extensive study to evaluate the accuracy and efficiency of CBAD. In this section, we will show the creation of the rules and use them for the process of associative classification of delegates.

The Apriori algorithm has been implemented as shown in the previous section, and it has resulted in a very large number of rules. These rules have been filtered to concentrate on two types of rules: the PES-SO mapping rules and the SO correlation rules. The implementation of the method of filtering the rules and the interpretation of the relevant insights extracted are presented in the following subsections.

### 4.1 PES Correlation

A commonly used type of rule can be expressed as follows: IF condition THEN conclusion, where the IF part is called the “antecedent” or “condition”, and the THEN part is called the “consequent” or “conclusion”. It means that if the condition of the rule is satisfied, we can infer or deduce the conclusion. As such, we can also write the rule in the following format: condition  $\Rightarrow$  conclusion. The condition typically consists of one or more feature tests (e.g., feature 1 > value 2, feature 5 = value 3) connected using logical operators (i.e., and, or, not) [27].



The generated rules have been filtered out to extract rules of the form  $SO_x \Rightarrow SO_y$  and  $\overline{SO_x} \Rightarrow \overline{SO_y}$ , where  $SO_x$  and  $SO_y$  are two previously specified SOs. This process has been performed to discover the correlations among ABET-EAC SOs. The extracted rules together with their confidence values are shown in [Tab. 3](#). Using these rules, the correlation between  $SO_x$  and  $SO_y$  is defined in terms of the confidence in their equivalency. Formally, this can be described as follows:

$$\text{Correlation}(\text{PEO}_x, \text{PEO}_y) = \text{Confidence}(\text{PEO}_x \Leftrightarrow \text{PEO}_y) \quad (3)$$

$$= \text{Confidence} ((SO_x \wedge SO_y) \vee (\overline{SO_x} \wedge \overline{SO_y})) \quad (4)$$

$$= P((SO_x \wedge SO_y) \vee (\overline{SO_x} \wedge \overline{SO_y})) \quad (5)$$

$$= P(SO_x \wedge SO_y) + P(\overline{SO_x} \wedge \overline{SO_y}) \quad (6)$$

$$= P(SO_x | SO_y) \times P(SO_y) + P(\overline{SO_x} | \overline{SO_y}) \times P(\overline{SO_y}) \quad (7)$$

$$= \text{Confidence} (SO_x \Rightarrow SO_y) \times P(SO_y) + \text{Confidence} (\overline{SO_x} \Rightarrow \overline{SO_y}) \times P(\overline{SO_y}) \quad (8)$$

The correlations between each pair ( $SO_x, SO_y$ ) were calculated using [Eq. \(8\)](#), as shown in [Tab. 4](#).

#### 4.2 PES Correlation

In this step, the PEO-SO mapping rules were extracted by filtering the generated rules. The salient feature of these rules is that a specific PES is shown in the premise of the rule, a single SO or multiple SOs in its conclusion and the confidence on top of the mapping arrow (e.g.,  $SC \xrightarrow{0.81} \bar{b} \wedge \bar{k}$ ). The top 10 rules were selected based on their confidence, as shown in [Tab. 5](#). Clearly, most of the rules shown in the table have a single SO in their conclusions as an indicator of the possibility of mapping or not mapping a PES in the premise to the SO or SOs in the conclusion. The over-bar SOs indicate the possibility of not mapping the PES to the SO with the confidence level shown (e.g.,  $SC \xrightarrow{0.81} \bar{b} \wedge \bar{k}$ , meaning that the PES S is not recommended to be mapped to SOs b and k with 81% confidence). It is also obvious that as the confidence of a rule declines, more SOs appear in the conclusion of the rule. [Tab. 6](#) shows more generic representations of the PES-SO mapping employing the top 10 rules.

In [Tab. 6](#), the presence of a given SO in the conclusion of a  $\text{PES} \Rightarrow \text{SO}$  rule is represented by the symbol  $\checkmark$ , its absence by the symbol  $\times$ , and an ambiguous state by  $?$ . It can be observed that the Ethical Conduct and Social and Community PEOs (EC and SC) have the highest average confidence, whereas Knowledge Competency and Graduate Studies PES (KC and GS) have the lowest. Moreover, it can be observed that the Social and Community and Knowledge Competency PESs (SC and KC) are not dependent on the presence of any SOs. They mainly depend on the absence of different combinations of SOs. The Lifelong Learning, Communication, Ethical Conduct, and Professionalism (LL, C, EC, P) PESs depend on the presence of a single SO, which indicates a one-to-one mapping between the graduate attributes of the PES and the skills of the SO.

Interestingly, two PESs, Leadership and Teamwork (L and T), depend on the presence of the same SOs (d and g) and on the absence of almost the same combination of SOs. This suggests a correlation between these two graduate attributes. Similarly, the PESs Career Success and Technical Competency (CS and TC) depend on the presence of the same combination of technical skill SOs (a, b, c, e, and k) but differ in their dependence on the soft skill SOs (d, f, g, h, i, and j), which suggests that they are somewhat correlated. The PES Knowledge Competency (KC) does not show dependence on the presence of any SOs; however, it depends on the absence of seven SOs.

**Table 3:** Confidence (SOx ⇒ SOy) and Confidence (SOy ⇒ SOx) [25]

$\frac{conf}{LL} \Rightarrow \frac{PEOy}{C}$	$\frac{conf}{C} \Rightarrow \frac{PEOy}{C}$	$\frac{conf}{L} \Rightarrow \frac{PEOy}{L}$	$\frac{conf}{T} \Rightarrow \frac{PEOy}{T}$	$\frac{conf}{EC} \Rightarrow \frac{PEOy}{EC}$	$\frac{conf}{P} \Rightarrow \frac{PEOy}{P}$	$\frac{conf}{SC} \Rightarrow \frac{PEOy}{SC}$	$\frac{conf}{CS} \Rightarrow \frac{PEOy}{CS}$	$\frac{conf}{TC} \Rightarrow \frac{PEOy}{TC}$	$\frac{conf}{KC} \Rightarrow \frac{PEOy}{KC}$	$\frac{conf}{GS} \Rightarrow \frac{PEOy}{GS}$
$\frac{0.07}{LL} \Rightarrow \frac{0.81}{C}$	$\frac{0.13}{C} \Rightarrow \frac{0.68}{L}$	$\frac{0.05}{L} \Rightarrow \frac{0.67}{L}$	$\frac{0.02}{T} \Rightarrow \frac{0.67}{L}$	$\frac{0.09}{EC} \Rightarrow \frac{0.65}{L}$	$\frac{0.35}{P} \Rightarrow \frac{0.75}{L}$	$\frac{0.04}{SC} \Rightarrow \frac{0.65}{L}$	$\frac{0.30}{CS} \Rightarrow \frac{0.71}{L}$	$\frac{0.16}{TC} \Rightarrow \frac{0.67}{L}$	$\frac{0.22}{KC} \Rightarrow \frac{0.69}{L}$	$\frac{0.62}{GS} \Rightarrow \frac{0.77}{L}$
$\frac{0.03}{LL} \Rightarrow \frac{0.78}{L}$	$\frac{0.23}{C} \Rightarrow \frac{0.85}{L}$	$\frac{0.23}{L} \Rightarrow \frac{0.86}{C}$	$\frac{0.57}{T} \Rightarrow \frac{0.69}{C}$	$\frac{0.11}{EC} \Rightarrow \frac{0.83}{C}$	$\frac{0.05}{P} \Rightarrow \frac{0.78}{C}$	$\frac{0.06}{SC} \Rightarrow \frac{0.82}{C}$	$\frac{0.11}{CS} \Rightarrow \frac{0.83}{C}$	$\frac{0.13}{TC} \Rightarrow \frac{0.84}{C}$	$\frac{0.15}{KC} \Rightarrow \frac{0.84}{C}$	$\frac{0.02}{GS} \Rightarrow \frac{0.82}{C}$
$\frac{0.09}{LL} \Rightarrow \frac{0.73}{T}$	$\frac{0.79}{C} \Rightarrow \frac{0.89}{T}$	$\frac{0.55}{L} \Rightarrow \frac{0.85}{T}$	$\frac{0.42}{T} \Rightarrow \frac{0.91}{L}$	$\frac{0.09}{EC} \Rightarrow \frac{0.82}{L}$	$\frac{0.06}{P} \Rightarrow \frac{0.78}{L}$	$\frac{0.13}{SC} \Rightarrow \frac{0.83}{L}$	$\frac{0.11}{CS} \Rightarrow \frac{0.82}{L}$	$\frac{0.13}{TC} \Rightarrow \frac{0.83}{L}$	$\frac{0.05}{KC} \Rightarrow \frac{0.81}{L}$	$\frac{0.02}{GS} \Rightarrow \frac{0.81}{L}$
$\frac{0.07}{LL} \Rightarrow \frac{0.71}{EC}$	$\frac{0.17}{C} \Rightarrow \frac{0.76}{EC}$	$\frac{0.13}{L} \Rightarrow \frac{0.75}{EC}$	$\frac{0.04}{T} \Rightarrow \frac{0.74}{EC}$	$\frac{0.13}{EC} \Rightarrow \frac{0.76}{T}$	$\frac{0.09}{P} \Rightarrow \frac{0.71}{T}$	$\frac{0.09}{SC} \Rightarrow \frac{0.75}{T}$	$\frac{0.09}{CS} \Rightarrow \frac{0.74}{T}$	$\frac{0.13}{TC} \Rightarrow \frac{0.76}{T}$	$\frac{0.09}{KC} \Rightarrow \frac{0.75}{T}$	$\frac{0.00}{GS} \Rightarrow \frac{0.75}{T}$
$\frac{0.46}{LL} \Rightarrow \frac{0.66}{P}$	$\frac{0.11}{C} \Rightarrow \frac{0.57}{P}$	$\frac{0.14}{L} \Rightarrow \frac{0.58}{P}$	$\frac{0.16}{T} \Rightarrow \frac{0.56}{P}$	$\frac{0.67}{EC} \Rightarrow \frac{0.71}{P}$	$\frac{0.41}{P} \Rightarrow \frac{0.88}{EC}$	$\frac{0.60}{SC} \Rightarrow \frac{0.86}{EC}$	$\frac{0.05}{CS} \Rightarrow \frac{0.71}{EC}$	$\frac{0.07}{TC} \Rightarrow \frac{0.71}{EC}$	$\frac{0.15}{KC} \Rightarrow \frac{0.75}{EC}$	$\frac{0.00}{GS} \Rightarrow \frac{0.73}{EC}$
$\frac{0.03}{LL} \Rightarrow \frac{0.73}{SC}$	$\frac{0.08}{C} \Rightarrow \frac{0.78}{SC}$	$\frac{0.16}{L} \Rightarrow \frac{0.77}{SC}$	$\frac{0.08}{T} \Rightarrow \frac{0.77}{SC}$	$\frac{0.52}{EC} \Rightarrow \frac{0.50}{SC}$	$\frac{0.25}{P} \Rightarrow \frac{0.83}{SC}$	$\frac{0.47}{SC} \Rightarrow \frac{0.64}{P}$	$\frac{0.26}{CS} \Rightarrow \frac{0.58}{P}$	$\frac{0.10}{TC} \Rightarrow \frac{0.52}{P}$	$\frac{0.23}{KC} \Rightarrow \frac{0.58}{P}$	$\frac{0.31}{GS} \Rightarrow \frac{0.61}{P}$
$\frac{0.26}{LL} \Rightarrow \frac{0.75}{CS}$	$\frac{0.19}{C} \Rightarrow \frac{0.73}{CS}$	$\frac{0.18}{L} \Rightarrow \frac{0.70}{CS}$	$\frac{0.11}{T} \Rightarrow \frac{0.70}{CS}$	$\frac{0.05}{EC} \Rightarrow \frac{0.68}{CS}$	$\frac{0.18}{P} \Rightarrow \frac{0.70}{CS}$	$\frac{0.03}{SC} \Rightarrow \frac{0.69}{CS}$	$\frac{0.02}{CS} \Rightarrow \frac{0.74}{SC}$	$\frac{0.08}{TC} \Rightarrow \frac{0.76}{SC}$	$\frac{0.24}{KC} \Rightarrow \frac{0.81}{SC}$	$\frac{0.00}{GS} \Rightarrow \frac{0.77}{SC}$
$\frac{0.14}{LL} \Rightarrow \frac{0.78}{TC}$	$\frac{0.32}{C} \Rightarrow \frac{0.75}{TC}$	$\frac{0.21}{L} \Rightarrow \frac{0.76}{TC}$	$\frac{0.16}{T} \Rightarrow \frac{0.76}{TC}$	$\frac{0.08}{EC} \Rightarrow \frac{0.68}{TC}$	$\frac{0.07}{P} \Rightarrow \frac{0.62}{TC}$	$\frac{0.10}{SC} \Rightarrow \frac{0.67}{TC}$	$\frac{0.35}{CS} \Rightarrow \frac{0.71}{TC}$	$\frac{0.34}{TC} \Rightarrow \frac{0.78}{CS}$	$\frac{0.16}{KC} \Rightarrow \frac{0.72}{CS}$	$\frac{0.38}{GS} \Rightarrow \frac{0.77}{CS}$

**Table 4:** Confidence (SOx, SOy) [6,25]

$\overset{\text{conf}}{a} \Leftrightarrow \text{SOy}$	$\overset{\text{conf}}{b} \Leftrightarrow \text{SOy}$	$\overset{\text{conf}}{c} \Leftrightarrow \text{SOy}$	$\overset{\text{conf}}{d} \Leftrightarrow \text{SOy}$	$\overset{\text{conf}}{e} \Leftrightarrow \text{SOy}$	$\overset{\text{conf}}{f} \Leftrightarrow \text{SOy}$	$\overset{\text{conf}}{g} \Leftrightarrow \text{SOy}$	$\overset{\text{conf}}{h} \Leftrightarrow \text{SOy}$	$\overset{\text{conf}}{i} \Leftrightarrow \text{SOy}$	$\overset{\text{conf}}{j} \Leftrightarrow \text{SOy}$
$\overset{0.84}{a} \Leftrightarrow b$	$\overset{0.83}{b} \Leftrightarrow c$	$\overset{0.61}{c} \Leftrightarrow d$	$\overset{0.59}{d} \Leftrightarrow e$	$\overset{0.46}{e} \Leftrightarrow f$	$\overset{0.60}{f} \Leftrightarrow g$	$\overset{0.60}{g} \Leftrightarrow h$	$\overset{0.60}{h} \Leftrightarrow i$	$\overset{0.71}{i} \Leftrightarrow j$	$\overset{0.49}{j} \Leftrightarrow k$
$\overset{0.79}{a} \Leftrightarrow c$	$\overset{0.60}{b} \Leftrightarrow d$	$\overset{0.81}{c} \Leftrightarrow e$	$\overset{0.61}{d} \Leftrightarrow f$	$\overset{0.60}{e} \Leftrightarrow g$	$\overset{0.69}{f} \Leftrightarrow h$	$\overset{0.54}{g} \Leftrightarrow i$	$\overset{0.74}{h} \Leftrightarrow j$	$\overset{0.52}{i} \Leftrightarrow k$	
$\overset{0.59}{a} \Leftrightarrow d$	$\overset{0.82}{b} \Leftrightarrow e$	$\overset{0.51}{c} \Leftrightarrow f$	$\overset{0.79}{d} \Leftrightarrow g$	$\overset{0.53}{e} \Leftrightarrow h$	$\overset{0.60}{f} \Leftrightarrow i$	$\overset{0.58}{g} \Leftrightarrow j$	$\overset{0.52}{h} \Leftrightarrow k$		
$\overset{0.84}{a} \Leftrightarrow e$	$\overset{0.48}{b} \Leftrightarrow f$	$\overset{0.57}{c} \Leftrightarrow g$	$\overset{0.59}{d} \Leftrightarrow h$	$\overset{0.51}{e} \Leftrightarrow i$	$\overset{0.66}{f} \Leftrightarrow j$	$\overset{0.57}{g} \Leftrightarrow k$			
$\overset{0.48}{a} \Leftrightarrow f$	$\overset{0.58}{b} \Leftrightarrow g$	$\overset{0.55}{c} \Leftrightarrow h$	$\overset{0.53}{d} \Leftrightarrow i$	$\overset{0.50}{e} \Leftrightarrow j$	$\overset{0.45}{f} \Leftrightarrow k$				
$\overset{0.55}{a} \Leftrightarrow g$	$\overset{0.52}{b} \Leftrightarrow h$	$\overset{0.50}{c} \Leftrightarrow i$	$\overset{0.59}{d} \Leftrightarrow j$	$\overset{0.80}{e} \Leftrightarrow k$					
$\overset{0.49}{a} \Leftrightarrow h$	$\overset{0.53}{b} \Leftrightarrow i$	$\overset{0.51}{c} \Leftrightarrow j$	$\overset{0.57}{d} \Leftrightarrow k$						
$\overset{0.50}{a} \Leftrightarrow i$	$\overset{0.49}{b} \Leftrightarrow j$	$\overset{0.74}{c} \Leftrightarrow k$							
$\overset{0.48}{a} \Leftrightarrow j$	$\overset{0.78}{b} \Leftrightarrow k$								
$\overset{0.78}{a} \Leftrightarrow k$									

From Tab. 4, it can be observed that some PESs are clearer than others in their mapping to SOs. This can be measured by counting the number of ? Symbols in their rows. Since the Technical Competency (TC) has only one ?, it has the clearest mapping to SOs. However, the most ambiguous PES in mapping to SOs is Social and Community (SC), with seven ?. By the same token, the SOs can also be distinguished based on their clarity in the PES-SO mapping by counting the number of ?s in their columns. In this manner, SO b is the clearest one, whereas h and j are the most ambiguous. Moreover, a distinction between SOs can be observed based on their contributions to the PES-SO mapping by counting the number of ✓’s in their columns. In this manner, SOs a, b, e, and g have the highest contribution, while h has the lowest contribution.

### 4.3 Associative Classification by Delegation

This is the final stage of the CBAD algorithm, in which the rules set has already been processed, and a final classifier is built and utilized as a delegate classifier. A linear pass is processed by means of the set of association rules that have been specified as strong rules. As the set is processed, information is maintained about how many elements from the original data set were correctly and incorrectly classified by the set of already-processed rules.

The process is performed employing the rules of PES-SO to build a classifier that is then delegated to classify PEOs with SOs. For example, the PEO “Apply modern technology tools, such as software and test equipment, to analyze, simulate, design and improve electrical systems” was actually mapped to PES number 9 (TC) in the dataset. This PES was mapped to SOs a and c and not to b, j d, e, f, g, h, i, j, and k. After the classifier was implemented on the dataset, the predicted multi-class labels (SOs) of the PESs employing the CBAD classifier were, in turn, considered as the multi-classes for the PESs; they are shown in Tab. 7.

Table 5: Top 10 PES-SO mapping rules [6,25]

Life-long learning	Communication	Leadership	Teamwork	Ethical conduct	Professionalism	Social & community	Career success	Technical compet.	Knowledge compet.	Graduate study
$LL \Rightarrow i$	$C \Rightarrow g$	$L \Rightarrow d$	$T \Rightarrow d$	$EC \Rightarrow b$	$P \Rightarrow b$	$SC \Rightarrow b$	$CS \Rightarrow e$	$TC \Rightarrow a$	$KC \Rightarrow d$	$GS \Rightarrow i$
$LL \Rightarrow d$	$C \Rightarrow i$	$L \Rightarrow b$	$T \Rightarrow a$	$EC \Rightarrow k$	$P \Rightarrow a$	$SC \Rightarrow a$	$CS \Rightarrow a$	$TC \Rightarrow c$	$KC \Rightarrow g$	$GS \Rightarrow a$
$LL \Rightarrow b$	$C \Rightarrow b$	$L \Rightarrow a$	$T \Rightarrow b$	$EC \Rightarrow a$	$P \Rightarrow c$	$SC \Rightarrow a \wedge b$	$CS \Rightarrow k$	$TC \Rightarrow e$	$KC \Rightarrow i$	$GS \Rightarrow e$
$LL \Rightarrow e$	$C \Rightarrow a$	$L \Rightarrow k$	$T \Rightarrow a \wedge b$	$EC \Rightarrow e$	$P \Rightarrow e$	$SC \Rightarrow k$	$CS \Rightarrow c$	$TC \Rightarrow f$	$KC \Rightarrow f$	$GS \Rightarrow k$
$LL \Rightarrow c$	$C \Rightarrow g \wedge i$	$L \Rightarrow g$	$T \Rightarrow i$	$EC \Rightarrow a \wedge b$	$P \Rightarrow b \wedge c$	$SC \Rightarrow e$	$CS \Rightarrow b$	$TC \Rightarrow i$	$KC \Rightarrow d \wedge g$	$GS \Rightarrow b$
$LL \Rightarrow g$	$C \Rightarrow f$	$L \Rightarrow i$	$T \Rightarrow c$	$EC \Rightarrow f$	$P \Rightarrow d$	$SC \Rightarrow b \wedge k$	$CS \Rightarrow a \wedge e$	$TC \Rightarrow b$	$KC \Rightarrow k$	$GS \Rightarrow a \wedge e$
$LL \Rightarrow a$	$C \Rightarrow a \wedge b$	$L \Rightarrow a \wedge k$	$T \Rightarrow e$	$EC \Rightarrow b \wedge e$	$P \Rightarrow f$	$SC \Rightarrow a \wedge k$	$CS \Rightarrow b \wedge e$	$TC \Rightarrow k$	$KC \Rightarrow g \wedge i$	$GS \Rightarrow d$
$LL \Rightarrow b \wedge c$	$C \Rightarrow b \wedge g$	$L \Rightarrow b \wedge k$	$T \Rightarrow g$	$EC \Rightarrow b \wedge k$	$P \Rightarrow b \wedge e$	$SC \Rightarrow b \wedge e$	$CS \Rightarrow a \wedge b$	$TC \Rightarrow j$	$KC \Rightarrow h$	$GS \Rightarrow f$
$LL \Rightarrow b \wedge e$	$C \Rightarrow b \wedge i$	$L \Rightarrow c$	$T \Rightarrow b \wedge c$	$EC \Rightarrow a \wedge k$	$P \Rightarrow a \wedge b$	$SC \Rightarrow a \wedge b \wedge k$	$CS \Rightarrow a \wedge c$	$TC \Rightarrow g$	$KC \Rightarrow d \wedge i$	$GS \Rightarrow j$
$LL \Rightarrow d \wedge i$	$C \Rightarrow a \wedge g$	$L \Rightarrow a \wedge b$	$T \Rightarrow k$	$EC \Rightarrow a \wedge e$	$P \Rightarrow g$	$SC \Rightarrow a \wedge e$	$CS \Rightarrow b \wedge c$	$TC \Rightarrow h$	$KC \Rightarrow b$	$GS \Rightarrow a \wedge b$

**Table 6:** PES-SO mapping using the top 10 rules [25]

No.	PES	SOs											Conf. interval	Avg. conf.
		a	b	c	d	e	f	g	h	i	j	k		
1	LL	×	×	×	×	×	?	×	?	✓	?	?	0.65–0.85	0.693
2	C	×	×	?	?	?	×	✓	?	×	?	?	0.77–0.91	0.792
3	L	×	×	×	✓	?	?	✓	?	×	?	×	0.65–0.76	0.679
4	T	×	×	×	✓	×	?	✓	?	×	?	×	0.66–0.82	0.726
5	EC	×	×	?	?	×	✓	?	?	?	?	×	0.81–0.89	0.827
6	P	×	×	×	×	×	✓	×	?	?	?	?	0.68–0.8	0.711
7	SC	×	×	?	?	×	?	?	?	?	?	×	0.81–0.9	0.826
8	CS	✓	✓	✓	?	✓	?	?	?	?	?	✓	0.69–0.74	0.678
9	TC	✓	✓	✓	?	✓	×	×	×	×	×	✓	0.65–0.74	0.691
10	KC	?	×	?	×	?	×	×	×	×	?	×	0.53–0.75	0.62
11	GS	✓	✓	?	×	✓	×	?	?	✓	✓	✓	0.56–0.77	0.63

Tab. 7 shows that the SOs a and c are correctly classified with the PES mentioned above, with a confidence of 71% for both of them, and not f with a confidence of 71%, not g with a confidence of 65%, not h with a confidence of 65%, and not i with a confidence of 71%. The classifier incorrectly classified the PEO with SO e with a confidence of 71%. As mentioned previously, the PEO-SO dataset is a multi-label dataset, which means that one PEO can be mapped to more than one SO; the same is true for the PES-SO and PEO-PES datasets. The rules created by the CBAD algorithm have been implemented in all instances of the PES-SO, PEO-SO and PEO-PES datasets. The process is as follows:

1. The multi-classes for every instance of the PEO-PES dataset are specified (e.g.,  $Peo_{15} \rightarrow \{CS, GS\}$ )
2. The association rules are implemented to specify which class labels (SOs) will be mapped to these PESs (CS and GS) separately and then combined to provide the final SOs for the PEO. In this case, the SOs d, f, g, h, i, and j are assigned as class labels for  $Peo_{15}$ .

**Table 7:** Mapping of a PES to SOs using multi-class labels [25]

a	c	e	$\bar{f}$	$\bar{i}$	b	k	$\bar{j}$	$\bar{g}$	$\bar{h}$
0.74	0.71	0.71	0.71	0.71	0.68	0.68	0.67	0.65	0.65

The analysis of the accuracy of PEO classification is based on PES classes (e.g., LL and C) separately, as shown in Tab. 8. It should be noted that in those datasets, a very large number of rules are required to cover all the classes.

**Table 8:** Accuracy of mapping PEOs to SOs using the CBAD classifier [25]

No.	PESs	SO									
1	LL	i	$\bar{d}$	$\bar{b}$	$\bar{e}$	$\bar{c}$	$\bar{g}$	$\bar{a}$	$\bar{b} \ \& \ \bar{c}$	$\bar{b} \ \& \ \bar{e}$	$\bar{d} \ \& \ i$
		40.15%	23.69%	23.44%	25.14%	24.04%	23.59%	22.91%	23.33%	26.10%	55.33%
2	C	g	$\bar{i}$	$\bar{b}$	$\bar{a}$	$g \ \& \ \bar{i}$	$\bar{f}$	$\bar{a} \ \& \ \bar{b}$	$\bar{b} \ \& \ \bar{g}$	$\bar{b} \ \& \ \bar{i}$	$\bar{a} \ \& \ \bar{b}$
		27.55%	18.32%	17.96%	19.14%	43.71%	16.97%	20.30%	40.88%	25.00%	42.21%
3	L	d	$\bar{b}$	$\bar{a}$	$\bar{k}$	g	$\bar{i}$	$\bar{a} \ \& \ \bar{k}$	$\bar{b} \ \& \ \bar{k}$	$\bar{c}$	$\bar{a} \ \& \ \bar{b}$
		17.67%	10.97%	11.59%	11.85%	8.50%	10.43%	14.29%	14.29%	9.29%	12.12%
4	T	d	$\bar{a}$	$\bar{b}$	$\bar{a} \ \& \ \bar{b}$	$\bar{i}$	$\bar{c}$	$\bar{e}$	G	$\bar{b} \ \& \ \bar{c}$	$\bar{b}$
		30.08%	21.29%	19.70%	22.42%	17.30%	18.31%	18.51%	22.79%	9.63%	19.70%
5	EC	$\bar{b}$	$\bar{k}$	$\bar{a}$	$\bar{e}$	$\bar{a} \ \& \ \bar{b}$	F	$\bar{e}$	$\bar{b} \ \& \ \bar{k}$	$\bar{a} \ \& \ \bar{k}$	$\bar{a} \ \& \ \bar{e}$
		17.96%	19.28%	18.60%	18.51%	20.30%	23.74%	20.75%	21.43%	9.76%	20.52%
6	P	$\bar{b}$	$\bar{a}$	$\bar{c}$	$\bar{e}$	$\bar{b} \ \& \ \bar{c}$	$\bar{d}$	f	$\bar{a} \ \& \ \bar{b}$	$\bar{a} \ \& \ \bar{b}$	$\bar{g}$
		27.43%	26.95%	27.05%	27.35%	29.81%	23.69%	34.17%	27.43%	8.18%	24.93%
7	SC	$\bar{b}$	$\bar{a}$	$\bar{a} \ \& \ \bar{b}$	$\bar{k}$	$\bar{e}$	$\bar{b} \ \& \ \bar{k}$	$\bar{a} \ \& \ \bar{k}$	$\bar{b} \ \& \ \bar{e}$	$\bar{a} \ \& \ \bar{b} \ \& \ \bar{k}$	$\bar{a} \ \& \ \sigma\epsilon$
		17.46%	18.33%	20.61%	18.18%	17.40%	20.45%	21.60%	19.50%	15.04%	19.87%
8	CS	e	a	k	c	B	$\bar{a} \ \& \ \bar{e}$	$\bar{b} \ \& \ \bar{e}$	$\bar{a} \ \& \ \bar{b}$	$\bar{a} \ \& \ \bar{c}$	$\bar{b} \ \& \ \bar{c}$
		30.82%	30.41%	12.50%	29.24%	32.71%	9.12%	8.71%	9.70%	6.04%	10.25%
9	TC	a	c	e	$\bar{f}$	$\bar{i}$	B	k	$\bar{j}$	$\bar{g}$	$\bar{h}$
		53.04%	49.83%	49.18%	38.56%	38.17%	54.14%	47.37%	37.47%	32.44%	35.84%
10	KC	$\bar{d}$	$\bar{g}$	$\bar{i}$	$\bar{f}$	$\bar{d} \ \& \ \bar{g}$	$\bar{k}$	$\bar{g} \ \& \ \bar{i}$	$\bar{h}$	$\bar{d} \ \& \ \bar{i}$	$\bar{b}$
		17.96%	12.87%	16.54%	15.94%	19.49%	15.43%	22.73%	13.77%	12.35%	12.72%
11	GS	i	a	e	k	B	a & e	$\bar{d}$	$\bar{f}$	i	a & b
		17.88%	15.54%	14.10%	13.82%	15.04%	15.77%	9.23%	9.25%	10.22%	16.00%

The prediction of the SOs by the algorithm based on the PESs is not homogeneous, meaning that every PES is mapped to a different SO. For example, the PEOs related to LL are mapped to SOs i, not d, not b, not e, etc. and the PEOs related to C are mapped to SOs g, not i, not b, etc. As shown in [Tab. 8](#), some PEOs obtained good results; for example, the classification of PEOs using the PES LL is mapped to SOs not d and i with an accuracy of 55.33%, which is the highest accuracy in the figure. The classifier did not obtain good results in some cases since it optimizes the confidence measure in isolation; that is, it generates very specific classifiers that are not able to correctly predict unseen examples.

## 5 Conclusion and Future Work

The proposed associative mapping of PEOs to SOs by delegation has shown promising results; it can simplify the classification of short text and avoid many problems faced by enriching short text with external resources. The mappings should have a purpose, helping to improve an educational programme so that it achieves objectives at every level. Targeted use of programme educational objectives, programme educational skills and student outcome mapping can benefit educational institutions that want to develop and offer coherent and effective degree programmes and plan to apply for ABET accreditation.

In addition, it can be beneficial to ABET accreditation bodies to check the consistency and efficiency of the mappings. The outcomes of this work can be applied to other short text-based analysis in various fields such as healthcare, social media and management. In fact, it can serve as a basis for the creation of tools, systems and applications for facilitating the current technique to enhance short text analysis, data mining and machine learning. In future work, we plan to enrich the datasets by including more programme educational objectives and testing the technique in different educational fields with larger datasets than engineering. Moreover, we would like to enrich the datasets with relative datasets by including more unstructured data extracted from the other components of the self-study reports and curriculum as well as other educational components (e.g., exams, tutorials and labs). In addition, the classification accuracy was tested by a hybrid classifier combining the associative classification using delegation and classical classifiers. Relevant educational data mining techniques for short texts will be implemented in future work to compare the accuracy of the results in different areas.

**Funding Statement:** The author received no specific funding for this study.

**Conflicts of Interest:** The author declares that they have no conflicts of interest to report regarding the present study.

## References

- [1] ABET, "ABET strategic plan," in *Accreditation Board for Engineering and Technology, Inc.*, ABET, United States, 1997.
- [2] ABET, "Criteria for Accrediting Engineering Programs Effective for Reviews during the 2016–2017 Accrediting Cycle," 415 N. Charles Street Baltimore, MD 21201, United States of America, ABET, 2017.
- [3] EAC-ABET, "Criteria for Accrediting Engineering Programs Effective for Review during the 2015–2016 Accreditation Cycle," 415 N. Charles Street Baltimore, MD 21201, United States of America, ABET, 2014.
- [4] J. Crichigno, I. L. Hurtado, R. R. Peralta and A. J. Perez, "From conception to accreditation: The path of an engineering technology program," in *Proc. ASEE Annual Conf. & Exposition*, American Society for Engineering Education, Seattle WA, USA, pp. 14–17, 2015.
- [5] N. Rajae, E. Junaidi, S. N. Taib, S. F. Salleh and M. A. Munot, "Issues and challenges in implementing outcome based education in engineering education," *International Journal for Innovation Education and Research*, vol. 1, no. 4, pp. 1–9, 2013.
- [6] A. Osman, A. A. Yahya and M. B. Kamal, "A benchmark collection for mapping program educational objectives to ABET student outcomes: Accreditation," in *Proc. the 5th Int. Symposium on Data Mining Applications. Advances in Intelligent Systems and Computing*, pp. 46–60, 2018.
- [7] C. C. Aggarwal, "*Data Mining: The Textbook*," Switzerland: Springer, 2015.
- [8] M. Dunham, "*Data Mining: Introductory and Advanced Topics*," Singapore: Prentice Hall, 2003.
- [9] Y. H. Jiang and L. Golab, "On competition for undergraduate co-op placements: A graph mining approach," in *Proc. the 9th Int. Conf. of Educational Data Mining*, North Carolina, USA, pp. 394–399, 2016.
- [10] V. N. Vapnik, "An overview of statistical learning theory," *IEEE Transactions on Neural Networks*, vol. 10, no. 5, pp. 988–999, 1999.
- [11] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [12] L. Yang, C. Li, Q. Ding and L. Li, "Combining lexical and semantic features for short text classification," in *Proc. the 7th Int. Conf. in Knowledge Based and Intelligent Information and Engineering Systems*, Oxford, UK, vol. 22, no. 1, pp. 78–86, 2013.
- [13] X. H. Phan, C. T. Nguyen, D. T. Le, L. M. Nguyen, S. Horiguchi *et al.*, "A hidden topic-based framework toward building applications with short web documents," *IEEE Transactions on Knowledge & Data Engineering*, vol. 23, no. 7, pp. 961–976, 2011.
- [14] D. M. Blei, A. Y. Ng and M. I. Jordan, "Latent dirichlet allocation," *Machine Learning Research Archive*, vol. 3, no. 1, pp. 993–1022, 2003.

- [15] O. Jin, N. N. Liu, K. Zhao, Y. Yu and Q. Yang, "Transferring topical knowledge from auxiliary long texts for short text clustering," in *Proc. of the 20th ACM Int. Conf. on Information and Knowledge Management*, Glasgow Scotland, UK, pp. 775–784, 2011.
- [16] M. Kozłowski and H. Rybinski, "Clustering of semantically enriched short texts," *Journal of Intelligent Information Systems*, vol. 53, no. 1, pp. 69–92, 2019.
- [17] V. S. Tseng and C. H. Lee, "Effective temporal data classification by integrating sequential pattern mining and probabilistic induction," *Journal of Expert System with Application*, vol. 36, no. 5, pp. 9524–9532, 2009.
- [18] K. Kim, B. S. Chung, Y. Choi, S. Lee, J. Y. Jung *et al.*, "Language independent semantic kernels for short-text classification," *Expert Systems with Applications*, vol. 41, no. 2, pp. 735–743, 2014.
- [19] Y. Lili, L. Chunping, D. Qiang and L. Li, "Combining lexical and semantic features for short text classification," in *Proc. of the 17th Int. Conf. in Knowledge Based and Intelligent Information and Engineering Systems*, Kitakyushu, Japan, pp. 78–86, 2013.
- [20] N. Abdelhamid, A. Ayesh and W. Hadi, "Multi-label rules algorithm based associative classification," *Parallel Processing Letters*, vol. 24, no. 1, pp. 01–21, 2014.
- [21] N. Lesh, M. J. Zaki and M. Ogiharra, "Mining features for sequence classification," in *Proc. of the 5th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, San Diego, California, USA, pp. 31–36, 1999.
- [22] B. Liu, W. Hsu, and Y. Ma, "Integrating classification and association rule mining," in *Proc. of the 4th Int. Conf. on Knowledge Discovery and Data Mining*, New York, pp. 80–86, 1998.
- [23] N. Abdelhamid, A. Ayesh and F. Thabtah, "Emerging trends in associative classification data mining," *International Journal of Electronics and Electrical Engineering*, vol. 3, no. 1, pp. 50–53, 2015.
- [24] A. A. Yahya and A. Osman, "Using data mining techniques to guide academic programs design and assessment," in *Proc. of the 6th Int. Learning & Technology Conf.*, New York, NY, USA, pp. 472–481, 2019.
- [25] A. Yahya and A. Osman, "A data-mining-based approach to informed decision-making in engineering education," *Computer Applications in Engineering Education*, vol. 27, no. 6, pp. 1402–1418, 2019.
- [26] R. Agrawal and S. Ramakrishnan, "Fast algorithms for mining association rules in large databases," in *Proc. 20th Int. Conf. on Very Large Data Bases VLDB Conf.*, Santiago, Chile, USA, pp. 487–499, 1994.
- [27] L. Xiao-Li and L. Bing, "Data Classification: Algorithms and Applications," in *Taylor & Francis Group, Charu C. Aggarwal IBM T. J. Watson Research Center Yorktown Heights*, New York, USA, CRC Press, 2015.