Article

SUNFIT: A Machine Learning-Based Sustainable University Field Training Framework for Higher Education

Mohammed Gollapalli [1], Atta Rahman [2,*], Mariam Alkharraa [2], Linah Saraireh [3], Dania AlKhulaifi [2], Asiya Abdus Salam [1], Gomathi Krishnasamy [1], Mohammad Aftab Alam Khan [4], Mehwash Farooqui [4], Maqsood Mahmud [5] and Rehan Hatab [6]

[1] Department of Computer Information Systems (CIS), College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam 31441, Saudi Arabia

[2] Department of Computer Science (CS), College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam 31441, Saudi Arabia

[3] Department of Management Information System (MIS), College of Business Administration, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam 31441, Saudi Arabia

[4] Department of Computer Engineering (CE), College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam 31441, Saudi Arabia

[5] Business Analytic Program, Department of Management and Marketing, College of Business Administration, University of Bahrain, Sakhir 32038, Bahrain

[6] Department of Computer Science, The University of Jordan, Amman 11942, Jordan

* Correspondence: aaurrahman@iau.edu.sa

Abstract: With the rapid advances in Information Technology (IT), the focus on engaging computing students to gain practical experience in the IT industry before graduation is becoming increasingly complex without incorporating pedagogical strategies of success in curricula. The goal is to enable computing major students to gain in-depth knowledge and practical understanding of the IT working environment before graduating through essential industry-driven practical skills based on international standards and best practices. Unfortunately, tracking and analyzing students' practical skills performance during their IT field training programs, which are conducted primarily off-campus at various public and private organizations, before, during, and after the training period, is a daunting task for both the college instructors and the industry trainers. To overcome these challenges, this paper introduces a Sustainable University Field Training (SUNFIT) framework, which is a pedagogical approach towards mining the educational data using machine learning to integrate and measure the field training programs against the internationally recognized accreditation standards such as Accreditation Board for Engineering and Technology (ABET). The study employs machine learning models aimed at continuously measuring and monitoring international ABET accreditation requirements on computing major courses' academic data, elucidating student performance across various semesters, integrating best practices, and producing an evidence-based rationale approach for evaluating weak learning outcomes (LOs) with minimal manual intervention, as well as preventing faculty-specific portfolio errors. The proposed approach could be easily developed by academics, researchers, or even students, and for a variety of purposes, including enhancing poor student outcomes (SOs). In addition, various data mining and machine learning approaches have been investigated over field training assessment data for successful prediction in subsequent cycles. The results are promising, with Naïve Bayes obtaining the highest accuracy of 90.54% followed by J48 and PART algorithms at 87.83%.

Keywords: machine learning; computing programs; data management; higher education; curriculum development; practical training; communication skills; best practices; accreditation

1. Introduction

With the exponential increase in the number of public and private universities offering different computing programs, it is becoming an important factor and a long-term dream

worldwide to recognize universities as world leaders in delivering computing major students with excellent communication skills in higher education [1]. Moreover, a university degree is often measured based on the output of a university or college's young graduates over the years who have successfully taken on different market positions post-graduation. Problem solving in the real world is at the core of many theories and models in higher education [2]. Hence, the development of outstanding students with real-life IT experience has become the primary focus of the vision and mission of computing colleges around the globe. To achieve this goal, field training programs providing students with practical experience outside the classroom are becoming part of the core courses.

The Accreditation Board for Engineering and Technology (ABET) is an accrediting body recognized worldwide as a global leader, ensuring quality and innovations are well established and maintained in the fields of applied science, computer engineering, and information technology education [3]. To maintain educational standards at the highest levels, as highlighted by [4], colleges and departments with computing programs worldwide have recently begun to embrace ABET international accreditation along with national accreditation standards such as the National Commission for Academic Accreditation and Assessment [5], as well as departmental policies aimed at fulfilling the department's mission statement and the overall vision of the university. Unfortunately, the immensity of setting proper evaluation criteria and guidelines of best practices and cooperation among faculties, students, as well as industry staff without Information and Communication Technology (ICT) platforms is a daunting and time-consuming process [6,7]. In some of the core CS courses, for instance, "system analysis and design", apart from technical concepts, graduating students also need to be evaluated on soft skills including verbal and written communication skills to ensure that students are professionally and ethically developed as necessary for their lifelong careers, based on accreditation ontologies [8] and Blooms knowledge Taxonomy frameworks [9]. Identifying gaps between teaching practices at the faculty levels and improving student critical communication skills could ensure that the curriculum is delivered efficiently, accurately evaluated through rubrics, and continuously improved in line with the education standards and best practices [10]. Students do not achieve higher-order thinking skills and competency skills to the expected level in a traditional classroom setup for all learners, thus it is essential to motivate students with new pedagogy techniques and thereby improve learners' performance [11]. As highlighted in [12], there is little evidence now of a breakthrough in the application of 'modern' Artificial Intelligence (AI) specifically towards teaching and learning, in higher education, except for perhaps learning analytics. However, the mining process of student assessment data structures is a significant task, with the aim to acquire crucial fact-finding information that is not otherwise available, or that would require time-consuming and expensive manual procedures [13].

This paper formally introduces the research problem on the implementation of an educational data-mining framework using machine learning on how to collect, analyze, measure, and continuously improve students' industry-driven skills through assessment data based on ABET accreditation requirements and best practices. In this regard, most of the studies in the literature are limited to or mainly focused on the best practices in the education programs at a bigger level such as processes of assessment, evaluation, and continuous improvement (as given in the subsequent section "related work"). However, how these processes can be improved to ensure sustainability by ensuring the integrity of educational data is the main missing element, especially in computing program field training. This is where the proposed SUNFIT bridges this gap by providing the essence of data mining and machine learning in the best practices for field training in computing programs for sustainability. The scope of this research is limited to computing undergraduate and postgraduate program capstone courses such as final year "graduation project", and semi-capstone courses such as "cooperative field training", where skills are critically assessed as part of curricula inside the classrooms (by academic staff) and outside the classrooms (by industry staff). This research is particularly aimed to fulfill ABET accreditation

requirements of soft skills for computing students primarily in developing countries where English is the standard mode of academic teaching for non-English speaking students. Preparatory courses (normally the first 2 years of a 4-year undergraduate computing program before undertaking majors), technical courses (such as programming, networking, and mathematical courses), as well as theory-based and elective courses (such as business courses) with no practical assessment are considered outside the scope of this research. Further, the study investigates various data mining and machine learning algorithms on the duly-collected data for field training through various assessments to determine the best algorithm for prediction accuracy of the rubrics data. This is to highlight the best sustainable practices in the overall accreditation process through prediction assessments, rubrics, and key areas of improvement.

The rest of the paper is organized as follows: Section 2 presents the related work in literature. Section 3 sheds light on the proposed framework for field training. Section 4 presents the proposed educational data mining and machine framework, experimental results, and discussion, while Section 5 concludes the paper.

2. Related Work

Ref. [14] proposed machine learning algorithms using Weka data mining software to measure and monitor students' academic progress. The proposed study is primarily aimed at identifying weaker students and notifying instructors. However, the research fails to identify how to measure the criteria-based evaluation and how to integrate practical industry experience along with academic studies. The research also does not consider sets of courses that students undertake during each academic year, which are an important factor in students' gradual progress. Ref. [15] proposed a data mining-based approach on ABET criteria to discover relationships between program educational objectives (PEOs) and SOs in engineering programs. The study employed the Apriori algorithm to extract association rules and for decision making. Unfortunately, the study does not consider the core of ABET, i.e., the continuous improvement strategy. In addition, the study only considers at the SOs level and does not consider deep learning at the rubrics level, which is critical in measuring students' performance in different skill sets.

Ref. [16] proposed the integration of the agile Scrum Framework and Cooperative Learning guidance in response to the needs of ABET accreditation into one of the computer information systems (CIS) semi-capstone courses "systems analysis and design" to encourage collaboration, communication, and problem-solving skills while learning system analysis and method concepts. The research proposed two approaches, including "overlapped" and "delayed" methodology including theory-based applied learning best practices, context-based learning, modernized teaching methods, project-based learning in teams, and providing students with formative feedback. The research-recommended improved advice for CIS students could provide them with the ability to obtain input regarding their preliminary analysis before moving on to design and implementation aspects. Nevertheless, the methodology does not address how agile scrum framework implementation could affect the learning experience for CIS students both inside and outside the classroom. In addition, the analysis of data from the students' appraisal against best practices and metrics of performance has not even been discussed.

Ref. [11] proposed a model that incorporates Visual, Auditory, Read, and Kinesthetic (VARK) learning styles in a flipped classroom to improve students' higher order thinking skills. The research work was focused on developing a framework to successfully leverage the ABET learning outcomes and competencies in a systematic way when designing, delivering, or revising an undergraduate and postgraduate syllabus. The approach involved using Information and Communication Technology (ICT) cloud computing tools for education. The research recommended measuring competency skills such as communication skills using Mind-Mapping activity to determine higher-order thinking skills using the flipped classrooms. The outcomes of this research through the flipped classroom strategy enhanced student performance and proved to be a positive learning strategy for engineer-

ing courses. Fuzzy logic was used to analyze the performance of learners using MATLAB. However, it would have been interesting to see how in-class and out-class components could be experimented based on students' cognitive levels as suggested by [9].

Ref. [17] proposed a Course and Student Management System (CSMS) to address the course assessment matrix and help achieve department objectives on ABET requirements including teamwork, ethics, lifelong learning, oral, and communication skills. The research focused on facilitating means to assess courses based on course learning outcomes, student evaluation, and student tracking to fulfill ABET criteria. The research also aimed at helping faculties in identifying courses and student outcomes that need attention. However, the proposed CSMS was not a panacea towards the identification of students' performance data on communication skills and other practical skills without ABET assessment and improvement indicators. A deeper educational data mining is needed to analyze the breadth and depth of various forms of assessments and the evaluation of results for course continuous improvement process. Ref. [18] proposed Problem Based Learning (PBL) as a best methodology for students learning one of the CIS semi-capstone courses "system analysis and design", which has diverse areas of concepts from project management skills to communication, design, and implementation expertise. The research recommended that PBL could enhance learning soft skills including communication and teamwork and could be retained as part of lifelong learning. The outcomes of the research included students' positive feedback on changing five lectures to PBL-based classes and suggested incorporating PBL as a way forward and as part of fulfilling ABET criteria. The research recommended giving proper training to faculty members to use PBL effectively and transfer knowledge in their core courses. However, the proposed model does not involve any course measurement data collection strategy, which is critical for ABET for evidence. In addition, the PBL approach alone could be challenging and might not be suitable for courses where communication skills are measured by internal and external faculties such as in field-training cooperative programs while gaining practical industry experience as part of course curricula. Studies conducted in [19–25] are evident that field training cooperative courses are the pillars of professional programs especially in engineering, computer science and information technology curricula. Moreover, the role of IT as instructional technology, pedagogy and library information system has been instrumental.

Authors in [26] proposed a modern approach to smart education systems by developing a cloud-based collaborative filtering recommendation system using SVM and machine learning. The paper aims to improve students' learning efficiency, especially in cloud teaching by the classification and collection of the necessary knowledge for students. Another purpose of the presented algorithm was to increase the engagement between teachers and students in courses involving document writing and processing. A group of experiences has been conducted in this paper, concluding the improvement cloud computing can make in classroom environments. Authors in [27] proposed a sustainable quality assurance framework for outcome-based education (OBE) at the higher education level. Since higher education plays an important role in the life-long learning and other factors of a graduate, the authors suggested a set of guidelines and best practices to foster sustainable quality education; it was mainly based on their experience and the lessons learned from three programs accredited by the ABET. Stakeholders' (such as students, faculty, alumni, and industry partners) engagement was one the prominent guidelines in this regard. Saeed et al. [28] investigated sustainable program assessment practices under the umbrella of ABET and the National Center for Academic Accreditation and Evaluation (NCAAA); ABET is an international body and NCAAA is a local accreditation body in Saudi Arabia. The study contrasts in terms of sustainable program assessment practices by taking the case study of a computer information system program. In continuation to this, the authors further studied the importance and role of academic accreditation in sustainable quality education that instill due skills among the program graduates [29].

In [30], an algorithm for predicting classified training quality was introduced with the objective of the assessment and enhancement of active learning methods for English major

students. The presented algorithm uses SVM in combination with the Grey Wolf Optimizer (GWO) to generate a prediction model for classified training quality in English majors. The experimental results of this algorithm have shown superiority in comparison to previously proposed algorithms in terms of computational speed and prediction performance. In [31], authors proposed the Apriori algorithm as an educational data mining approach to investigate the best teaching and learning practices to enhance the overall learning environment at the higher education level. The study encompasses several aspects of students' interests such as preferred learning hours, days of the week, and their preferences towards various learning equipment. The study revealed quite interesting patterns of students' interests to enhance the overall teaching and learning environment at the higher education level. In [32], the authors have investigated the role of various data mining algorithms on the student success prediction at the secondary school. The algorithm includes Naïve Bayes, J48, and Random forests where Naïve Bayes outperformed the other algorithms. The data include the student educational outcomes along with human factors. From the literature review, the following can be concluded and justifies the current study:

1- SVM, J48, Apriori algorithm, and Naïve Bayes are among the most widely used algorithms in educational data mining.
2- Most of the studies focus on success prediction based on the course or the degree outcome in the form of percentage or grades.
3- Most of the studies focus on sustainable best practices in educational programs.
4- No studies focus on deeper levels such as the course/training level with the assessment, evaluation, and rubrics level to find out the granular-level issues.
5- Most of the studies focus on the taught courses rather than the field training courses that are governed partially by the academic and on-sight supervisors.

3. University Field Training

A successful assessment process leads to successful accreditation [4]. The proposed sustainable University Field Training (SUNFIT) is an educational data mining framework based on the pedagogical strategies of preparing, conducting, and assessing computing students' skills in courses involving practical industry engagement. The framework process involves workflow for collecting, classifying, and visualizing student practical skills measurement data as part of accreditation standards and the continuous improvement lifecycle. Field training also known as industry or cooperative training and internship (in different universities' nomenclature) facilitates by giving computing students an opportunity to become lifelong learners, with tools and techniques that cannot be taught out of a textbook, but only by being an integral part of the learning process [18]. Field training providers (trainers) are college-partnered public and private industry organizations where computing students undergo field training as part of their course curriculum for a specific IT work environment and for a specific training period usually spanned over 6–8 weeks. Field training-based practical courses not only allow computing students to be an active participant in the learning process but forces them to take an effective role by engaging themselves in a meaningful, thought-provoking way [18]. Unfortunately, for students from computing colleges, when dealing with different field training positions from different training partnered organizations, we are highly exposed to different levels of assessment as mandated by course syllabus requirements (academic perspective), IT industrial requirements (industry perspective), and standards requirements (accreditation perspective). Hence, successful IT field training assessment measurement is needed to quantify based on direct, indirect, quantitative, and qualitative rubric indicators appropriate to the outcomes of field training curricula [19].

The Accreditation Board for Engineering and Technology (ABET) is one such platform that is recognized internationally and is used in more than 40 countries and among nearly 895 universities [3]. Figure 1 visualizes the five layers of accreditation assessments identified by ABET for higher education. The first layer is the objectives of the educational program (PEOs) that include a set of goals that each student needs to achieve before graduating.

Each PEO includes a set of student outcomes (SOs) as a second layer that CIS students need to achieve to advance the PEOs. The third layer includes performance indicators (PIs) assessment processes to direct evaluation against SOs and for each of the evaluation criteria. The fourth layer includes the course evaluation objectives (Cos) for which different sets of rubrics are written to obtain the actual results data. Regarding the university field training, the following points describe the structure, feasibility, and the requirements.
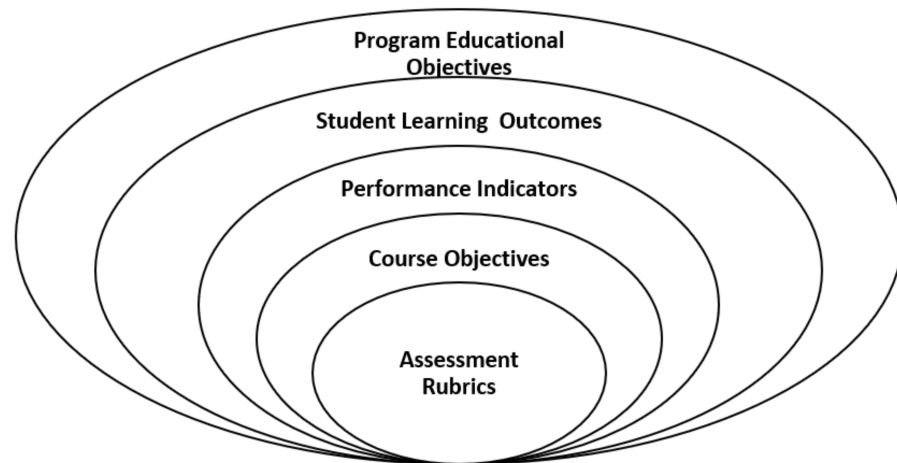

Figure 1. Five-Layer Accreditation Assessment.

Following are the common points followed in various universities:

1- The field training is a carefully designed course with the accreditation bodies guidelines and is an essential part of the computing curriculum whose purpose is to provide the students with practical hands-on experience by working real life industries/organizations.

2- Usually, the curriculum committee is responsible for the alignment of the course description, objective, and CLOs of the course with program SOs and their corresponding PIs and providing effective rubrics for measuring them.

3- Students should have passed minimum credit hours and exhibit hands-on skills needed for the industry. Field training (aka cooperative training, internship) is a course offered during summer for 6–8 weeks duration. Upon successful completion, the evaluation takes place during the following semester.

4- Field training opportunities are created by communicating with industry partners in different capacities. A university/college/department cell is dedicated to this task that continuously interacts with industry to find opportunities for the students.

5- Usually, there are two mentors/supervisors for each student, one on sight and the second academic mentor/supervisor. Proper communication is maintained between the two in terms of student's progress.

6- Mostly, one student is assigned exactly one field training. However, occasionally more than students may be working on same field training project, but they are evaluated separately.

7- Usually, one mentor can be assigned more than one mentee depending on the nature of the project and availability of the mentors.

8- Mentors are trained for mentoring/supervising field training projects in terms of rubrics and evaluation process.

9- The PI assessment processes include formative assessment from the on-sight supervisor and the academic supervisor. The summative assessments include the report evaluation and end of training presentation to two evaluators based on the provided rubrics for the report writing and the oral presentation, respectively. An aggregate score is then assigned as a grade for the students against the field training course.

4. Data Mining and Machine Learning Framework

According to accreditation bodies, assessment involves one or more procedures to identify the data and evidence gathered through evaluation processes. Assessment determines to what degree SOs are obtained. Unfortunately, carrying out assessment through different field training placements in the context of IT course requirements for thousands of computing students is a critical task that requires a complete analysis of each SO for proper data evaluation. Assessment results in decisions and actions requiring changes to the program [19]. Hence, to carry out analysis on field-training practical skills, data on the field training course need to be carefully collected and measured. The results should be analyzed through qualitative and quantitative data mining for which academic faculty as well as industry trainers conducting field training need to input different sets of data based on the results of the SOs. Once the summative and formative assessment data are obtained, a series of data analysis processes must be conducted at various evaluation stages.

Development of computing programs must also take into consideration the various views and needs of stakeholders (students, faculty, institution, and industry at large) to be successful [20]. To achieve this goal, performing data mining on education field-training data at different assessment levels is critical in providing evidence-based analysis, cross-validation, and continuous improvement life cycle as per ABET program and course level standards. Figure 2 officially demonstrates the proposed SUNFIT framework with different criteria layers for the collection, evaluation, and justification of field-training assessment data. The course mapping layer (top layer) includes the portfolio data of field training that must be prepared over the years in compliance with the course specifications of the ABET and the checklist of field-training skills expectations that needs to be analyzed throughout the semesters. The frequency of the field training data collection process (middle layer) depends on the three levels of assessments, which are the advanced level of introduced (I), reinforced (R), and emphasized/focused (E).
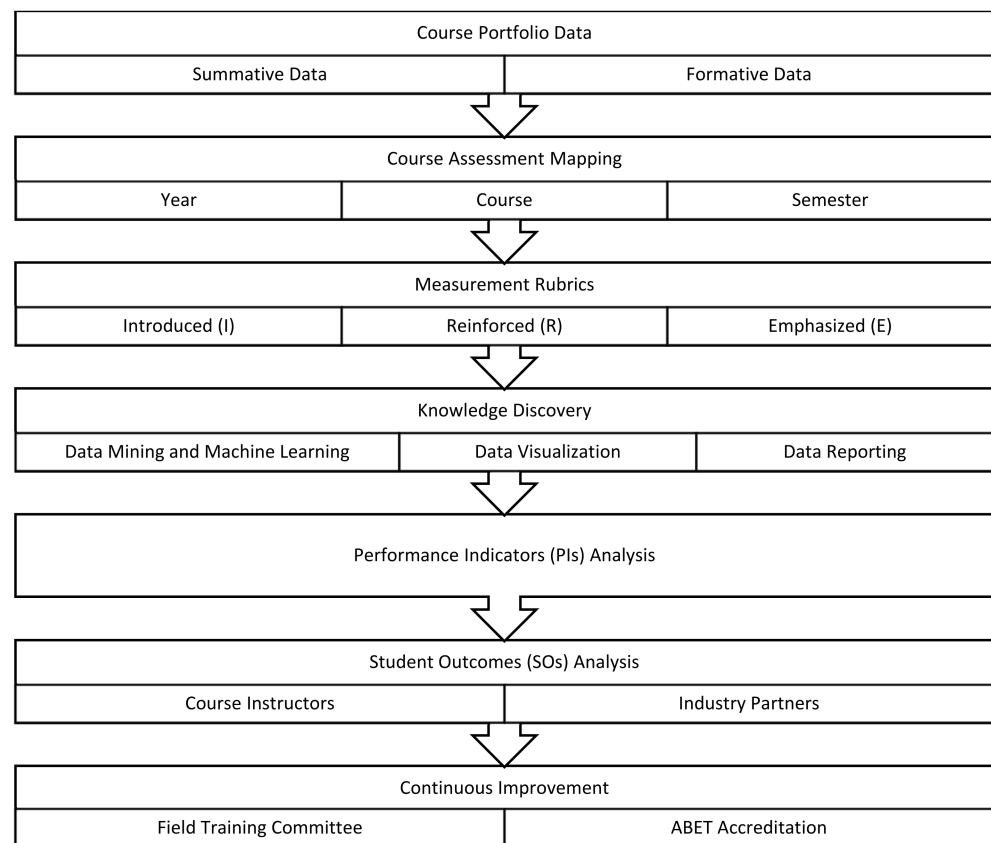

Figure 2. Proposed SUNFIT Framework.

The performance indicators (PIs) comprise a series of evaluation scales required at the end of each semester to evaluate student outcomes for field training. An important and crucial step towards accreditation is the evidence-based ABET assessment process for each of these layers, especially to achieve the overall educational SOs for field training. For each of these layers, therefore, field training courses must be specified with different assessment levels. SOs assist faculties in evaluating what computing students are expected to know and be able to do by graduation time. These relate to the knowledge, skills, and practical skills acquired by computing students as they progress through the training program [19]. Therefore, field training courses need to be carefully designed to ensure that as part of the practical field training industry experience, computing students could achieve the expected IT skills outcomes. While higher education institutions across the globe are broadly promoting the adoption and intensive use of diverse digital technology platforms for teaching, less data-driven evidence for teaching activities is occurring [21]. Moreover, the process of mining structures of databases is a significant task, with the aim to acquire crucial fact-finding information that is not otherwise available, or that would require time-consuming and expensive manual procedures [13]. The process of data mining as part of the proposed SUNFIT framework as visualized in Figure 2 includes the collection and analysis of field training data through different means to determine and measure the accomplishment of each of the field training SOs by the knowledge discovery process. This includes data mining and machine learning, data visualization, and data reporting. Consequently, the outcomes are used in corresponding PI analysis which is linked to the corresponding SO analysis and eventually contributes to the continuous improvement deemed to the field training committee and ABET accreditation committee, and this completes the cycle. Therefore, the proposed SUNFIT framework spans from data collection from the course portfolios to continuous improvement by utilizing the strength of data mining and machine learning in identification of the shortcoming in achieving the relevant indicators and rubrics. Each of these steps has been comprehensively explained and detailed below.

4.1. Data Preparation

An effective approach towards closing the gap between academia and industry in education and to best train students is for both the parties to work together on educational needs and goals [22]. Evaluation, as described by ABET for computer programs, includes one or more processes defining, gathering, and preparing data to determine SOs achievement. Efficient assessment uses applicable direct, indirect, quantitative, and qualitative indicators that are appropriate to the assessed outcome.

The dataset is hypothetically envisioned from the course portfolios collected from the coordinators and supervisors over the academic years in each cycle (usually comprised of two years). It is comprised of formative and summative assessment data from the whole discipline. The data are mainly in terms of percentage of students and fall in one of the four categories defined in the rubrics that are poor/non-existent (0–24%), developing (25–49%), developed (50–74%), and exemplary (75–100%). For the field training, the data from both mentors/supervisors are averaged for each rubric. The experiments have been conducted in Weka tool/environment.

As part of the ABET framework, the central problem that arises periodically (normally at the end of each semester) is the efficient performance of data handling, including the compilation, incorporation, and review of various computing student's assessment data. This issue of processing IT field training course data according to ABET accreditation requirements is collectively defined as a data management problem. ABET-based field data management is obviously not a new issue. However, a daunting task within the ABET data management process is the efficient and reliable alignment of field training data (also referred to as rubric-based assessment data) based on ABET criteria. Recently, great progress has been made in seeking alternative simpler ways to solve data gathering, handling, and data management issues due to the significant demands within the process [11,16,17].

However, the root of this problem lies in conducting summative and formative rubric evaluation to obtain student course outcomes. The main drawbacks with the current techniques have been the computational cost involved in performing the manual data collection process by field training faculties and evaluators (academic staff), as well as field training providers (IT industry staff) [10].

In the proposed SUNFIT framework as outlined in Figure 2, the rubrics data handling assessment process is categorized under two sets, namely data from the summative assessment (also known as quantitative analysis) and data from the formative assessment (also known as qualitative analysis). Summative assessment data are the collection of rubric-based assessment that is carefully defined and for which data evidence is mandatory. The summative assessments are performed periodically to assess what students know and do not know at a certain point in time [23]. Table 1 highlights the top five summative assessments measurements we have identified in most computing colleges have adopted in their field training programs. Summative assessments involve cumulative student performance results on various sets of rubrics. Table 2 shows the most frequently used industry trainer rubrics mapped to each of the five summative assessments identified in Table 1.

Table 1. Generic Summative Assessment.

Summative Assessment	
S1	Acquaintance with the real-world work environment.
S2	Prepare the students to transfer from the learning environment to work environment.
S3	Acquaintance with the applied work systems.
S4	Understand mechanism of different applications.
S5	Understand the attitude and the manner of the work.

Table 2. Generic Summative Assessment Rubrics.

Industry Trainer Assessment Rubrics	
R1(S1)	Dependability and reliability.
R2(S2)	Ability to learn and search for information.
R3(S5)	Judgment and decision making.
R4(S5)	Effective relations with his/her work colleagues.
R5(S2)	Ability of reporting and presenting his/her work.
R6(S5)	Attendance and punctuality.
R7(S4)	Initiative in taking lead towards task completion.
R8(S3)	Ability to deliver work with practical experience.

Unfortunately, when dealing with large volumes of rubrics data (numeric; categorical; string based; etc.) obtained from different sources, we are vulnerable to different types of 'data uncertainties' such as different formats, null values, length constraints, typographical errors, and shorthand notations, which may well be one of the biggest obstacles to performing successful data linkage [13]. Hence, the SUNFIT framework addresses this summative data collection process for field training courses through predefined data collection patterns using data mining tools for data smoothing and analysis purposes. The goal is to eliminate common human errors and coordinate the gathering of rubrics data among different faculty instructors and industry trainers from diverse organizations where the students are placed for field training.

Table 3 highlights the summative-based rubrics data we have experimented with as part of this study and collected from both industry trainers and academic instructors during students' field training semester. Table 3 highlights the field training industry trainer rubrics data (R1–R5) for the top five students (S. Id), whose overall attendance rates (Att.) and whether they were certified during their training (Cert.) have been highlighted. Table 4 highlights similar academic instructor rubrics data (R6–R10) that assessed the students (provided as a standard across all field training organizations). Tables 5 and 6 are other

assessment rubrics data on field training students reports (R11–R16) and presentation (R17–R22) rubrics data. Table 7 displays the remaining rubrics data (R23–R28) on students' self-assessment of training.

Table 3. Assumed industry trainer data of students' progress.

Form 1			Industry Trainer Rubrics Data				
S. Id	Att.	Cert.	R1	R2	R3	R4	R5
1	82.5	0 (N)	1.0	3.0	3.0	5.0	5.5
2	91.4	1 (Y)	3.0	2.5	3.0	5.0	5.5
3	95.0	1 (Y)	5.0	3.0	3.0	5.0	5.5
4	79.5	0 (N)	3.0	3.0	3.0	5.0	5.5
5	85.5	0 (N)	3.0	3.0	3.0	6.0	6.0

Table 4. Assumed course instructor data of students' progress.

Form 2			Instructor Rubrics Data				
S. Id	GPA	PGPA	R6	R7	R8	R9	R10
1	2.53	2.0	2.0	4.0	3.0	6.0	6.0
2	1.42	2.5	3.0	3.0	2.5	6.0	6.0
3	3.20	3.5	3.0	3.0	2.5	6.0	6.0
4	1.84	3.8	3.0	3.0	2.5	6.0	6.0
5	1.97	1.5	3.0	2.5	2.5	6.0	6.0

Table 5. Assumed evaluators data on student's report.

Form 3		Examiners Report Rubrics Data					
S. Id	GPA	R11	R12	R13	R14	R15	R16
1	2.53	1.2	1.2	1.2	1.5	1.5	1.5
2	1.42	1.2	1.2	1.2	1.5	1.5	1.5
3	3.20	1.2	1.2	1.2	1.5	1.5	1.5
4	1.84	1.3	1.4	1.4	1.5	1.3	1.5
5	1.97	1.3	1.4	1.4	1.5	1.3	1.5

Table 6. Assumed evaluators data on student's presentations.

Form 4		Examiners Presentation Rubrics Data					
S. Id	GPA	R17	R18	R19	R20	R21	R22
1	2.53	4.7	4.3	5.0	5.0	4.3	4.7
2	1.42	4.7	4.3	5.0	5.0	4.3	4.7
3	3.20	5.0	4.3	5.0	5.0	4.3	5.0
4	1.84	5.0	4.3	5.0	5.0	4.3	5.0
5	1.97	4.8	4.7	4.7	4.7	4.8	5.0

Table 7. Students' data on industry training.

Form 5			Students Training Rubrics Data				
S. Id	IE	R23	R24	R25	R26	R27	R28
1	82.5	0 (N)	1.0	3.0	3.0	5.0	5.5
2	91.4	1 (Y)	3.0	2.5	3.0	5.0	5.5
3	95.0	1 (Y)	5.0	3.0	3.0	5.0	5.5
4	79.5	0 (N)	3.0	3.0	3.0	5.0	5.5
5	85.5	0 (N)	3.0	3.0	3.0	6.0	6.0

Formative assessment is an informal, collaborative, ongoing evaluation process requiring instructional changes and feedback [23]. Formative Evaluation (also known as qualitative appraisal) data include the cumulative input received at the end of the semester by the field training course instructors. Formative assessment provides the overall information required (no data collection is needed) to change the teaching and learning styles as it is taking place. Formative assessment provides information about students understanding to both instructors and students at a time when timely improvements can be made [23]. Although data for formative assessment courses are not collected, the instructors' feedback assists in the process of quality improvement. The data collected from faculty members are thus composed of tightly mapped summative assessments with assessment rubrics for proper assessments and improvement cycle.

4.2. ML Classification Algorithms

The current use of learning analytics and artificial intelligence in the field of education is only at a preliminary stage, mainly due to lack of demand from educational institutions [12]. The data mining of field training rubrics data collected from different assessment PIs on course rubrics (summative and formative data) is one of the most important components in recognizing and predicting overall field training experience and student's skills achievements across different IT organizations because of its ability to synthesize large amounts of academic data into effective databased graphic visualizations [24]. Hence, through the systematic representation of data analytics, this research considers the issue of generating evidence-based ABET criteria measurement of students' practical industry skills during field training programs. In this research, we employed Weka data mining software and compared Naïve bayes (NB), J48, PART, Support Vector Machine (SVM), and Logistic Regression (LR). These algorithms were carefully chosen based on their frequency of use we identified through the literature review in educational data mining (EDM) [32]. The collected hypothetical dataset (randomly generated by using statistical functions based on intuition) from 998 student instances from the assessment forms over several years was split into 80:20 ratio as training (799) and testing (199) sets with grades (A–D) as the class attribute and rubrics as the other/non-class attributes. The experiments were conducted using the default parameters in the Weka environment. Based on the given rubrics feature set, the class attribute is predicted using various algorithms. The performance of each of these algorithms is highlighted in Table 8 showing the correct classification, Kappa statistics, and the time of execution (in seconds). As can be seen, the NB classification algorithm outperformed the other algorithms and hence was chosen for measuring and reporting on the rubrics data for the following ABET cycle.

Table 8. Performance of EDM algorithm for envisioned field training data.

Classifier	Correctly Classified	Kappa Statistic	Time (s)
NB	90.54%	0.8724	00.00
J48	87.83%	0.8366	00.00
PART	87.83%	0.8364	00.00
SVM	85.13%	0.7979	00.17
LR	82.43%	0.7624	00.55

4.3. Measurement and Reporting

The performance of the learners can be measured easily using a variety of in-class and out-class evaluation criteria [11]. Hence, we further employed the most effective identified classifier Naïve Bayes (NB) to analyze the implementation of the data visualization process needed in line with ABET-based standards for quantitative data analysis and generate data-driven visualized charts based on the 10 guidelines outlined in [24] for effective rubrics data analysis and predictions. The sensitivity, also known as recall, was used to analyze the proportion of true positive (TP) values over the entire number of positive cases, and

the specificity, also known as the negative ratio, was used to determine all the correctly predicted negative values (TN) as per [33–35].

$$\text{Sensitivity} = \text{TP}/\,(\text{TP} + \text{FN}),$$
$$\text{Specificity} = \text{TN}/\,(\text{TN} + \text{FP}). \tag{1}$$

The accuracy of the classifiers was measured using the performance metric utilized in this study. These accuracy measurements were used to represent the predicted measure of cases as positive, whereas the F1 score was utilized to compare the weighted average accuracy and recall.

$$\text{Precision}\,(\Lambda) = \text{TP}/\,(\text{TP} + \text{FP}), \tag{2}$$

Table 9 illustrates NB classifier results on different rubrics (R1–R7), such as the mean (M), standard deviation (σ), and variances (Λ), identifying related information, including formative and summative course-mapping data and student performance in terms of grades (A+, A, B+, B, C+, C, D+, D etc.) under various rubrics. Unlike the top-down approach, the SUNFIT framework is intended as part of analyzing data from the bottom-up approach, i.e., from rubrics to the program objectives. The key assumption of this methodology is that the colleges and ABET committee who analyze data through various data files would not have knowledge of the meaningful results of field training students experience levels until relationships have been established and related information is visualized between PEOs, SOs, PIs, COs, and the rubrics.

Table 9. Naïve Bayes Classifier Data Analysis.

	R1	R2	R3	R4	R5	R6	R7
Λ	0.33	0.5	0.25	0.5	1.0	0.25	0.5
A+ (0.35)							
M	4.95	1.37	1.42	2	3	2.94	2.92
σ	0.11	0.21	0.17	0.08	0.16	0.15	0.17
D (0.11)							
M	4.95	1.18	1.25	1.81	3	2.93	3
σ	0.11	0.24	0.25	0.24	0.16	0.16	0.08
D+ (0.04)							
M	4.5	1.5	1.5	1.75	3	3	3
σ	0.5	0.08	0.04	0.25	0.16	0.04	0.08
A (0.27)							
M	4.6	1.4	1.3	1.9	2.9	2.9	2.8
σ	0.39	0.19	0.19	0.19	0.21	0.08	0.32
C (0.05)							
M	5	1.3	1.3	2	3	3	3
σ	0.05	0.23	0.08	0.16	0.04	0.08	0.08
C+ (0.04)							
M	5	1.5	1.5	2	3	3	3
σ	0.05	0.08	0.04	0.08	0.16	0.04	0.08
B+ (0.05)							
M	4.2	1.5	1.4	2	3	2.9	2.8
σ	0.3	0.08	0.11	0.08	0.16	0.11	0.23
B (0.1)							
M	4.8	1.5	1.5	2	3	2.7	3
σ	0.16	0.08	0.04	0.08	0.16	0.24	0.08

The indicators of performance measurements could describe the attitudes, skills, and behavior of the CIS student's ability to perform during the industry field-training period [17]. For instance, Figure 3 illustrates the means of field training rubrics R1–R8 for students who secured A+ grades. As can be seen, even though the students secure top grades, the students' performance has been minimal in R2 and R3, which need attention. Similarly, Figure 4 displays the performance of different students who secured different

grades (A–D). Again, as can be seen, almost all the students, irrespective of their grades, did very well in R1; however, most students did little in R2.
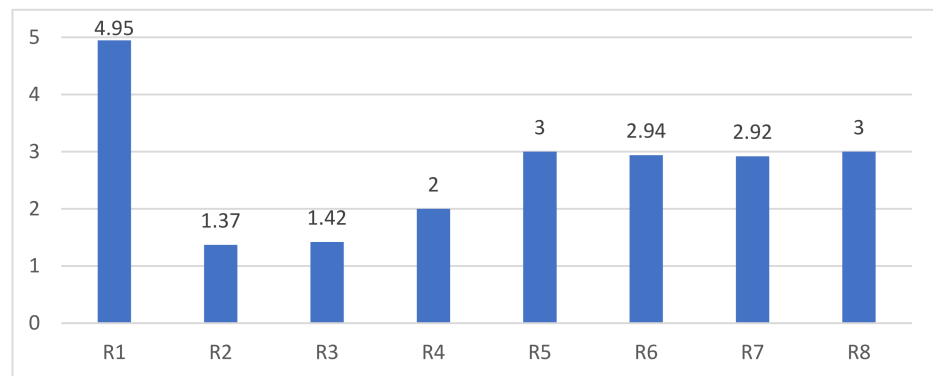

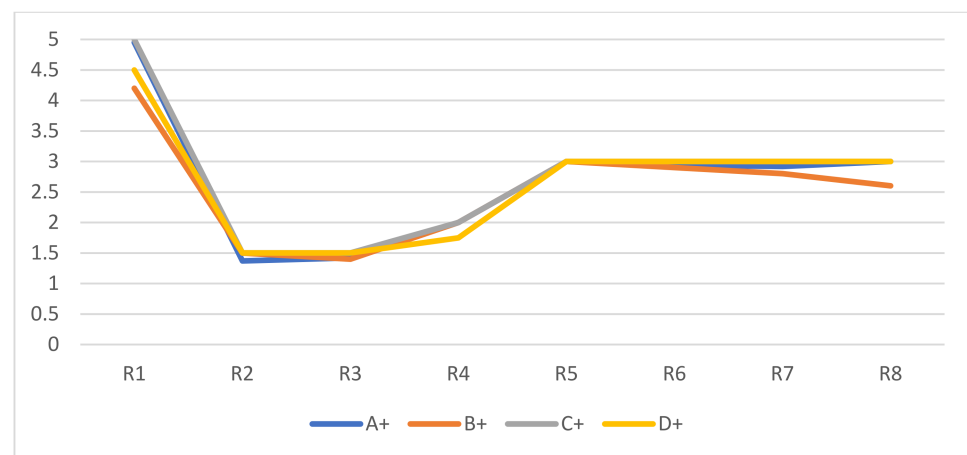**Figure 3.** Means of rubrics (R1–R8) on students scoring A+.

Figure 4. Rubrics (R1–R8) consistency across grades.

As computing education progresses in the 21st century, rapid technological changes will affect both what we teach and how we teach [4]. To understand these changes in teaching and learning, we proposed generating a hierarchy tree of students performance from Year 1 (prep years) to Year 4 (graduation year). Figure 5 shows one instance of the tree of courses for students majoring in programming stream courses as part of the ABET accreditation continuous improvement lifecycle.
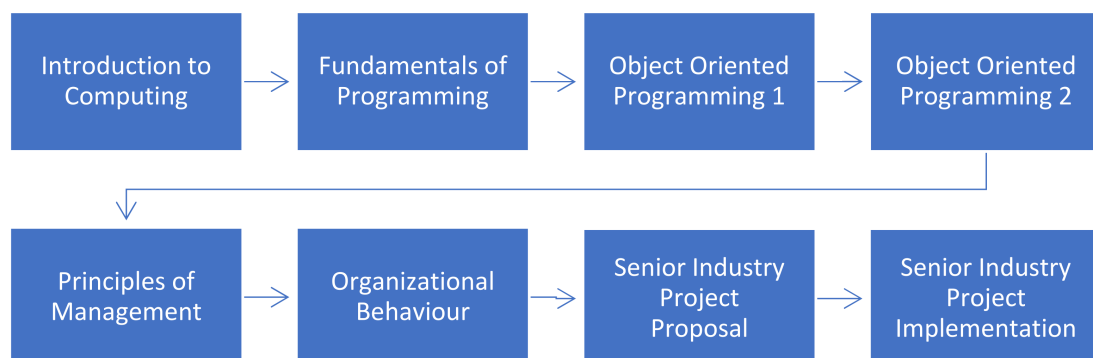
Figure 5. Rubrics Measurement across Course Hierarchies.

The hierarchy performance of students using the SUNFIT data mining technique is also analyzed in Table 10. As can be seen, the student's attendance rates (Att.), if the student

is certified (Cert.) during the field training, GPA, and the previous GPA (PGPA), along with Best (BR), Average (AR), and Worst (WR) rubric rates could give the college ample opportunity to understand and constantly improve the field training course as well as students' assessment criterions.

Table 10. Naïve Bayes Classifier Data Analysis using SUNFIT.

			Student Performance Tree Data				
S. Id	Att.	Cert.	GPA	PGPA	BR	AR	WR
1	82.5	0 (N)	2.53	2.0	R1	R15	R2
2	91.4	1 (Y)	1.42	2.5	R3	R8	R8
3	95.0	1 (Y)	3.20	3.5	R5	R6	R5
4	79.5	0 (N)	1.84	3.8	R1	R24	R25
5	85.5	0 (N)	1.97	1.5	R2	R19	R15

4.4. Performance Cycle

One of the long-term goals of computing colleges is to develop and continuously improve students' performance standards that measure their academic levels and skills [18]. The ABET accreditation process involves the tracking of computing student performance against overall PEOs for each assessment cycle, with each cycle comprising academic years of (quantitative and qualitative assessment) course data. In line with ABET, the SUNFIT framework proposes the following four performance analysis requirements at the end of each field training course semester.

- Performance comparison of students for field training courses against PEOs, SOs, PIs, and course-specific rubrics to analyze the overall progress of the student's outcomes.
- Performance comparison of faculty for field training course against ABET criteria and identification of data anomalies to analyze the overall performance of faculties in handling the course.
- Performance comparison of field training courses through formative course mapping (qualitative analysis data) against summative courses (summative analysis data) to analyze and constantly improve the course structure.
- Performance comparison of programs against self-study report (SSR) checklist based on ABET questionnaires [25] to measure fulfilment of ABET criteria across different cycles of data for gap analysis, continuous improvement process, and closing the loop.

4.5. Continuous Improvement

Assessment is the process that determines that true learning is taking place at all levels of the curriculum, and that discussions are always present for continuous improvement [4]. Continuous assessment of field training courses SOs are critical to accreditation by ABET. Because ABET is an ongoing process for developing PEOs, this process needs to be semi-automated with the help of the SUNFIT framework to produce the most needed evidence-based summarized computing programs data and improvements that could take place on a periodic basis based on the SUNFIT data mining framework. Using this approach, we believe computing colleges could take specific decisions and provide data-driven evidence on course achievements and course improvements to the ABET accreditation body. More importantly, computing students' performance could be re-measured after the improvements are applied to see if the intervention was truly an improvement leading to a cycle of continuous improvement.

The SUNFIT framework is intended to assist in the implementation of best practices for continuous field training course improvement processes. As illustrated in Figure 6, SUNFIT recommends connecting various constituencies which define the Program Educational Objectives (PEOs) with the Program Outcomes (POs), which are further mapped with the Course Learning Outcomes (CLOs) having different sets of assessment rubrics as part of the continuous improvement cycle. This connection could be easily achieved through visualizing and analyzing field training data using data mining approaches as explained

in the above section. The results of continuous improvement become self-motivating as faculty could see their efforts leading to successful student achievements and improving student learning outcomes.
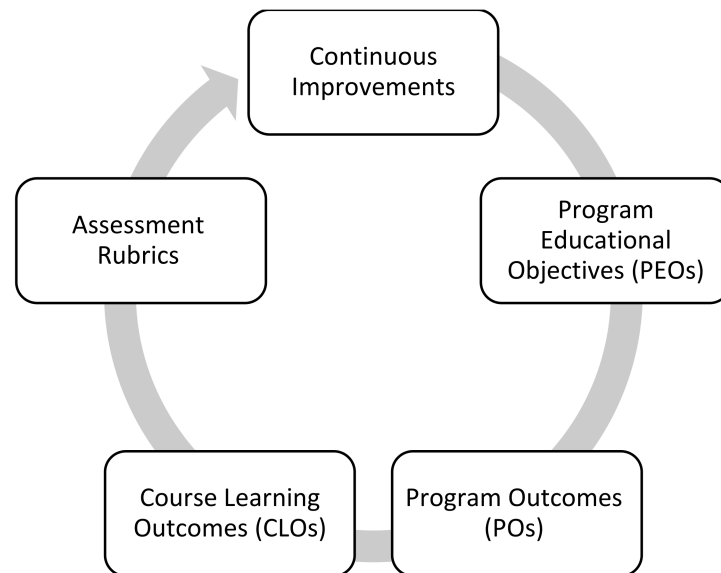

Figure 6. Students Continuous Improvement Lifecycle.

5. Conclusions

The first contribution of this study deals with an important approach on understanding computing student's field training assessment data measurement gaps among the stakeholders such as colleges, field training providers, students, and supervisors as well as field training course evaluators due to complex ABET data collection, data-driven analysis, and continuous course management process. Current practices in the processing of field training relevant data from different organizations are often noisy with irregularities. To provide an effective training program and analyze the performance, this study covers open-ended challenges in both academic and industry field training-based courses on how to structure the field training courses in a way that simplifies building students' communication skills through the incorporation of best practices.

The second contribution to this study is the development of the SUNFIT framework. SUNFIT promotes and facilitates best practices in the management of multi-layer complex field training within each organization based on the context of each training opportunity. The SUNFIT framework further supports facilitating in fulfilling quality education accreditation requirements of ABET, which is globally adopted by some of the world's top universities. While the focus in this paper is primarily made on IT students in computing colleges in the region for measuring field-training experience, SUNFIT can easily be incorporated into other colleges' programs course learning outcomes (CLOs) and performance indicators (PIs).

The third contribution we made in this paper is the alignment of SUNFIT with ABET through multiple stages of education data mining process using machine learning, providing a highly effective and efficient approach for the collection and integration of various sets of field training evaluation results data from academic faculties, industry trainees/trainers, and other participating bodies in line with ABET-based best practice standards and as a blueprint for the accreditation process. The SUNFIT framework provides a variety of approaches for evaluating various sets of statistical data for current and historical SOs to promote an evidence-based cycle of continuous improvement. The SUNFIT framework results could also support the development of a best practice digital library, based on the criteria of the ABET data repository. Other colleges at different universities (such as engineering and medical colleges) who also provide similar field training programs and are

keen to develop the soft skills of their students could also adopt recommended framework structure and best sustainable practices.

In conclusion, exploration of SUNFIT and its broader applications could form part of the future work as a potential new research problem. The proposed approach could be easily developed by academics, researchers, or even students, and for a variety of purposes, including enhancing poor SOs, since it does not require an expensive ABET specialist or detailed knowledge of assessment. The series of measures undertaken as part of the SUNFIT approach is aimed at monitoring ABET requirements unique to academic data, elucidating computing students' performance across various semesters, integrating best practices, and producing an evidence-based rationale approach for evaluating learning outcomes with minimal manual intervention, as well as preventing faculty-specific portfolio errors. In other words, the proposed framework in this research explicitly addresses the incorporation of best practices criteria into higher education standards through a multi-layer strategy that significantly reduces the amount of faculty work required for field training assessment at each stage and enables clusters of PEO outcomes by creating data-driven course-mapping data facts. In addition, the proposed approach is based on the use of rubrics between sets of summative and formative course data, which individual universities or colleges could easily define, measure, monitor, and continuously improve.

As part of the future work, this research study could be enhanced for aiming at implementing the concept of preliminary filtration for the comprehensive mapping and automation of the triple student-organization-supervisory filtration process that takes place before students engage in the field training cycles. In carrying out this move, the future direction of this research could provide the means to tightly incorporate field training positions based on eligibility criteria as well as ABET best practice criteria. The research direction also seeks to implement a clear standardized performance live monitoring model that can effectively tag the training outcomes of the students and field training providers during the training period. The future work could also consider building mobile and web platforms based on different stages of the proposed SUNFIT framework.

References

1. Ragonis, N.; Hazzan, O.; Har-Shai, G. Students' Awareness and Embracement of Soft Skills by Learning and Practicing Teamwork. *J. Inf. Technol. Educ. Innov. Pract.* **2020**, *19*, 185–201. [CrossRef]
2. Mills, R.; Hauser, K.; Pratt, J. A Software Development Capstone Course and Project for CIS Majors. *J. Comput. Inf. Syst.* **2008**, *48*, 1–14.
3. ABET Accreditation Program. Accreditation Board for Engineering and Technology. Available online: https://www.abet.org/ (accessed on 3 March 2022).
4. Lending, D.; Mitri, M.; Dillon, T. Ingredients of a High-Quality Information Systems Program in a Changing IS. *J. Inf. Syst. Educ.* **2019**, *30*, 266–286.
5. National Center for Academic Accreditation and Evaluation (NCAAA). Education and Training Evaluation Commission. Available online: https://etec.gov.sa/en (accessed on 19 January 2022).
6. Kulturel-Konak, S.; Konak, A.; Esparragoza, I.E.; Kremer, G.E.O. Measuring global awareness interest development of engineering and information technology students. In Proceedings of the 2016 IEEE Frontiers in Education Conference (FIE), Erie, PA, USA, 12–15 October 2016. [CrossRef]
7. Davis, C.E.; Sluss, J.J.; Landers, T.L.; Pulat, P.S. Innovative practices for engineering professional development courses. In Proceedings of the 2013 IEEE Frontiers in Education Conference (FIE), Oklahoma City, OK, USA, 23–26 October 2013. [CrossRef]
8. Abdel-Haq, M.S. Conceptual Framework for Developing an ERP Module for Quality Management and Academic Accreditation at Higher Education Institutions: The Case of Saudi Arabia. *Int. J. Adv. Comput. Sci. Appl.* **2020**, *11*, 144–152. [CrossRef]
9. Gottipati, S.; Shankararaman, V. Competency analytics tool: Analyzing curriculum using course competencies. *Educ. Inf. Technol.* **2017**, *23*, 41–60. [CrossRef]
10. Saito, D.; Kaieda, S.; Washizaki, H.; Fukazawa, Y. Rubric for Measuring and Visualizing the Effects of Learning Computer Programming for Elementary School Students. *J. Inf. Technol. Educ. Innov. Pract.* **2020**, *19*, 203–227. [CrossRef] [PubMed]
11. Priyaadharshini, M.; Sundaram, B.V. Evaluation of higher order thinking skills using learning style in an undergraduate engineering in flipped classroom. *Comput. Appl. Eng. Educ.* **2018**, *26*, 2237–2254. [CrossRef]
12. Bates, T.; Cobo, C.; Mariño, O.; Wheeler, S. Can artificial intelligence transform higher education? *Int. J. Educ. Technol. High. Educ.* **2020**, *17*, 42. [CrossRef]
13. Gollapalli, M.; Li, X.; Wood, I. Automated discovery of multi-faceted ontologies for accurate query answering and future semantic reasoning. *Data Knowl. Eng.* **2013**, *87*, 405–424. [CrossRef]
14. Khan, I.; Ahmad, A.R.; Jabeur, N.; Mahdi, M.N. An artificial intelligence approach to monitor student performance and devise preventive measures. *Smart Learn. Environ.* **2021**, *8*, 17. [CrossRef]
15. Yahya, A.A.; Osman, A. A data-mining-based approach to informed decision-making in engineering education. *Comput. Appl. Eng. Educ.* **2019**, *27*, 1402–1418. [CrossRef]
16. Magana, A.; Seah, Y.; Thomas, P. Fostering Cooperative Learning with Scrum in a Semi-Capstone Systems Analysis and Design Course. *J. Inf. Syst. Educ.* **2018**, *29*, 75–92.
17. El Rahman, S.A.; Shabanah, S.S. Course and Student Management System Based on ABET Computing Criteria. *Int. J. Inf. Eng. Electron. Bus.* **2016**, *8*, 1–10. [CrossRef]
18. Fatima, S.; Abdullah, S. Improving Teaching Methodology in System Analysis and Design using Problem Based Learning for ABET. *Int. J. Mod. Educ. Comput. Sci.* **2013**, *5*, 60–68. [CrossRef]
19. AlGhamdi, R. Fostering information technology students' internship program. *Educ. Inf. Technol.* **2019**, *24*, 2727–2739. [CrossRef]
20. Shah, V.; Kumar, A.; Smart, K. Moving Forward by Looking Backward: Embracing Pedagogical Principles to Develop an Innovative MSIS Program. *J. Inf. Syst. Educ.* **2018**, *29*, 139–156.
21. Rahman, A.; Alhaidari, F. The Digital Library and the Archiving System for Educational Institutes. *Pak. J. Inf. Manag. Libr.* **2019**, *20*, 94–117. [CrossRef]
22. Garousi, V.; Giray, G.; Tüzün, E.; Catal, C.; Felderer, M. Aligning software engineering education with industrial needs: A meta-analysis. *J. Syst. Softw.* **2019**, *156*, 65–83. [CrossRef]
23. Fidalgo-Blanco, Á.; Sein-Echaluce, M.L.; García-Peñalvo, F.J.; Conde, M.Á. Using Learning Analytics to improve teamwork assessment. *Comput. Hum. Behav.* **2015**, *47*, 149–156. [CrossRef]
24. McKenzie, F.D.; Mielke, R.R.; Leathrum, J.F. A Successful EAC-ABET Accredited Undergraduate Program in Modeling and Simulation Engineering (M&SE). In Proceedings of the 2015 Winter Simulation Conference (WSC), Huntington Beach, CA, USA, 6–9 December 2015. [CrossRef]
25. Al-Shagran, A.; Sahraoui, A.-E.-K. Assessment of E-learning Systems: A Systems Engineering Approach System. *Int. J. Comput. Sci. Softw. Eng.* **2017**, *6*, 173–179.
26. Li, J.; Zhou, Z. Matching Teaching Content and Strategy of Practical Document Writing and Processing Courses Based on Wisdom Education. *Mob. Inf. Syst.* **2022**, *2022*, 4282141. [CrossRef]
27. Almuhaideb, A.M.; Saeed, S. Fostering Sustainable Quality Assurance Practices in Outcome-Based Education: Lessons Learned from ABET Accreditation Process of Computing Programs. *Sustainability* **2020**, *12*, 8380. [CrossRef]
28. Saeed, S.; Almuhaideb, A.M.; Bamarouf, Y.A.; Alabaad, D.A.; Gull, H.; Saqib, M.; Iqbal, S.Z.; Salam, A.A. Sustainable Program Assessment Practices: A Review of the ABET and NCAAA Computer Information Systems Accreditation Process. *Int. J. Environ. Res. Public Health* **2021**, *18*, 12691. [CrossRef]

29. Almurayh, A.; Saeed, S.; Aldhafferi, N.; Alqahtani, A.; Saqib, M. Sustainable Education Quality Improvement Using Academic Accreditation: Findings from a University in Saudi Arabia. *Sustainability* **2022**, *14*, 16968. [CrossRef]

30. Prediction, F.C. Monitoring, and Management of the Classified Training Quality of English Majors Based on Support Vector Machine. *Int. J. Emerg. Technol. Learn. (IJET)* **2022**, *17*, 233–248.

31. Rahman, A.; Sultan, K.; Aldhafferi, N.; Alqahtani, A. Educational data mining for enhanced teaching and learning. *J. Theor. Appl. Inf. Technol.* **2018**, *96*, 4417–4427.

32. Alghamdi, A.S.; Rahman, A. Data Mining Approach to Predict Success of Secondary School Students: A Saudi Arabian Case Study. *Educ. Sci.* **2023**, *13*, 293. [CrossRef]

33. Alqarni, A.; Rahman, A. Arabic Tweets-Based Sentiment Analysis to Investigate the Impact of COVID-19 in KSA: A Deep Learning Approach. *Big Data Cogn. Comput.* **2023**, *7*, 16. [CrossRef]

34. Dash, S.; Luhach, A.K.; Chilamkurti, N.; Baek, S.; Nam, Y. A Neuro-fuzzy approach for user behaviour classification and prediction. *J. Cloud Comp.* **2019**, *8*, 17. [CrossRef]

35. Rahman, A.; Nasir, M.U.; Gollapalli, M.; Zubair, M.; Saleem, M.A.; Mehmood, S.; Khan, M.A.; Mosavi, A. Advance Genome Disorder Prediction Model Empowered with Deep Learning. *IEEE Access* **2022**, *10*, 70317–70328. [CrossRef]