IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received May 17, 2018, accepted June 25, 2018, date of current version August 7, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2851790

E-Learning: Challenges and Research Opportunities Using Machine Learning & Data Analytics

ABDALLAH MOUBAYED[1], MOHAMMADNOOR INJADAT[1], ALI BOU NASSIF[2], HANAN LUTFIYYA[3], AND ABDALLAH SHAMI[1]

[1]Electrical and Computer Engineering Department, Western University, London, ON N6A 3K7, Canada
[2]Electrical and Computer Engineering Department, University of Sharjah, Sharjah, United Arab Emirates
[3]Computer Science Department, Western University, London, ON N6A 3K7, Canada

Corresponding author: Abdallah Moubayed (amoubaye@uwo.ca)

ABSTRACT With the proliferation of technology, the field of e-learning has garnered significant attention in recent times. This is because it has allowed users from around the world to learn and access new information. This has added to the growing amount of collected data that is already being generated through different devices and sensors employed around the world. This has led to the need to analyze collected data and extract useful information from it. Machine learning (ML) and data analytics (DA) are proposed techniques that can help extract information and find valuable patterns within the collected data. In this paper, the field of e-learning is investigated in terms of definitions and characteristics. Moreover, the various challenges facing the different participants within this process are discussed. In addition, some of the works proposed in the literature to tackle these challenges are presented. Then, a brief survey about some of the most popular ML and DA techniques is given. Finally, some of the research opportunities available that employ such techniques are proposed to give insights into the areas that merit further exploration and investigation.

INDEX TERMS E-learning, machine learning, data analytics.

I. INTRODUCTION

E-learning is one of the more recent fields that is contributing to the great amount of generated data. E-learning can be defined to be the use of electronic devices and technology for learning new information and skills. The proliferation of technology throughout the world and the boom in information access has made distance learning more popular in recent times. Distance learning is one component of the e-learning process as it allows people to share knowledge despite geographical boundaries and limitations. In layman's terms, e-learning can be defined as the access to educational curriculum outside of a traditional classroom by utilizing electronic technologies [1]. However, despite the available resources and benefits of such a learning process, there are some challenges facing it such content transmission and delivery as well as the enabling technologies [2].

One of the sources of data growth is the data being generated through online learning websites and learning management systems (LMSs) as part of e-Learning environments. Statistics show that online course websites such as Coursera, edX, and Udacity have more than 78 million students combined with around 10 thousand courses being offered by more than 600 universities [3]. Moreover, it has been estimated that there are almost one thousand event log entries per student every month and around 60,000 course visits every month for online courses [4] with the increased size and variety of the collected data, a need to analyze and extract useful information has risen. To this end, machine learning (ML) and data analytics (DA) have been proposed to perform these tasks. ML techniques aim to "teach" computers to complete specific tasks without being explicitly programmed using the collected data. This is important in many fields as such techniques and algorithms have several applications including house pricing prediction, spam filtering, and autonomous driving [5]. On the other hand, DA aims to analyze and draw conclusions from the raw collected data to be able to make better decisions [6]. Such techniques are also crucial since they can be applied in different fields such as business, marketing, and education [6]. Hence, ML and DA techniques can become useful tools for researchers to tackle the challenges found in the e-learning field. For example, DA techniques can be used to get better insights into the performance of

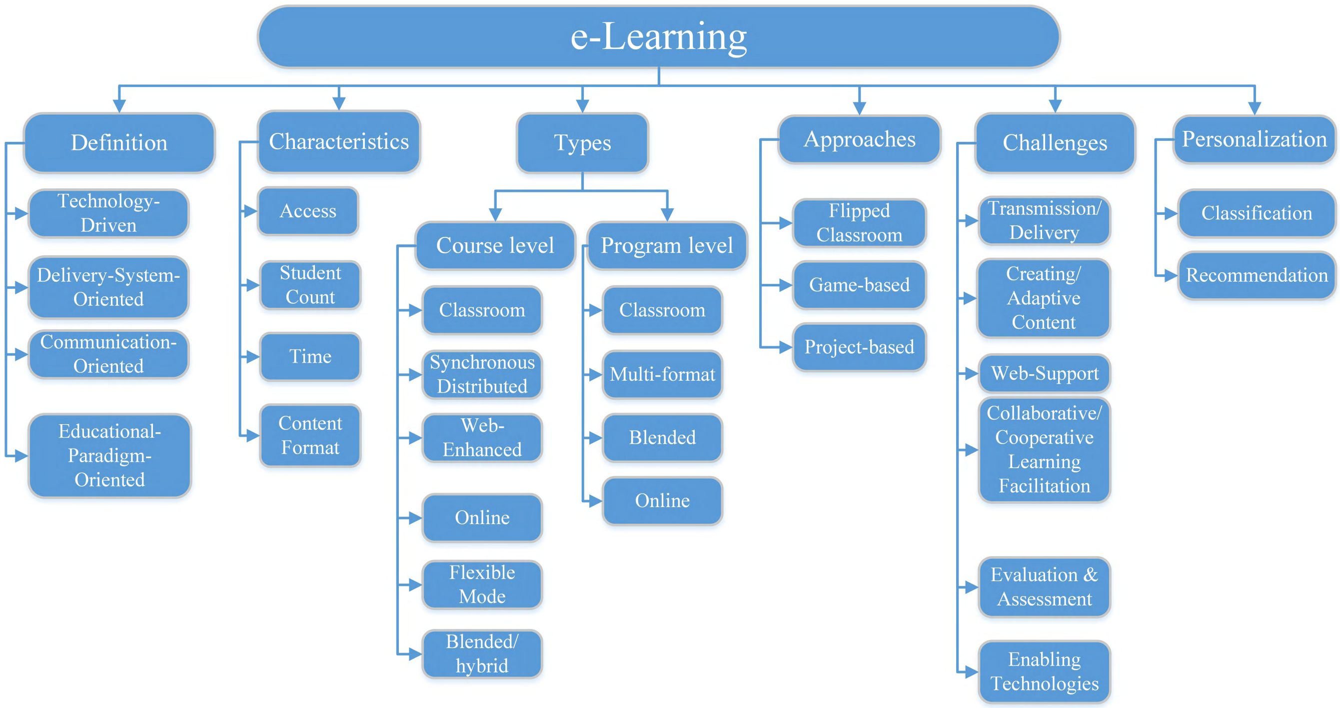

FIGURE 1. **Overview of e-learning.**

students in different e-learning environments. Additionally, ML classification techniques can be used to classify students based on their learning style. Therefore, ML and DA can play a crucial role in improving the learning experience.

This paper aims to provide a comprehensive view of the field of e-learning and the challenges it is facing. Furthermore, it discusses some of the literature that has been proposed to tackle these challenges. Moreover, the paper presents several research opportunities that investigate how to use ML and DA techniques as potential solutions to overcome the discussed challenges.

The remainder of this paper is organized as follows: Section II introduces the general topic of e-learning along with its main characteristics. Section III presents the different challenges in e-learning as well as some of the previous works that tackled these challenges. Section IV gives a brief background about the different machine learning and data analytics techniques. Section V shows why ML and DA techniques offer potential solutions to the various challenges facing the field of e-learning and discusses some of the research opportunities that can be explored using such techniques. Finally, Section VI concludes this paper.

II. INTRODUCTION TO E-LEARNING

A. DEFINITIONS

With the proliferation of technology throughout the world and with the boom in information access, distance learning has become more popular as it allows individuals to learn new skills without having a mentor physically present teaching them. Distance learning forms part of the e-learning process as it allows people to share knowledge despite geographical boundaries and limitations. As mentioned earlier, e-learning can be defined to be the access to educational curriculum outside of a traditional classroom by utilizing electronic technologies [1]. Fig. 1 gives an overview of the different topics encompassed by e-learning.

More complex and technical definitions of e-learning are given throughout the literature. For example, the authors in [7] categorize the definitions into four main categories as follows:

- **Technology-Driven Definitions:** This category focuses on the technological aspects of e-learning. For example, one definition under this category defines e-learning as "the use of technology to deliver learning and training programs" [7].
- **Delivery-System-Oriented Definitions:** This category focus on the accessibility of resources rather than the results of any achievement. For example, one definition under this category defines e-learning as "the delivery of education (all activities relevant to instructing, teaching, and learning) through various electronic media" [7].
- **Communication-Oriented Definitions:** This category focuses on the the tools of communication and interaction between the involved parties. For example, one definition under this category defines e-learning as "E-learning is learning based on information and communication technologies with pedagogical interaction between students and the content, students and the instructors or among students through the web" [7].

- **Educational-Paradigm-Oriented Definitions:** This category focuses on the educational side of e-learning by considering it a new way of learning or an improvement on an existing educational paradigm. For example, one definition under this category defines e-learning as "E-learning is the use of new multimedia technologies and the Internet to improve the quality of learning by facilitating access to resources and services, as well as remote exchange and collaboration" [7].

It is worth noting that one common point among the different definitions and categories is the use of technology and technological devices such as computers and handheld devices as a means to access and share information. Based on the aforementioned categories, a more inclusive definition was given by the authors as follows:

"E-learning is an approach to teaching and learning, representing all or part of the educational model applied, that is based on the use of electronic media and devices as tools for improving access to training, communication and interaction and that facilitates the adoption of new ways of understanding and developing learning" [7].

B. CHARACTERISTICS

There are some characteristics that can describe an e-learning course/module/program, which are listed below [8]:

- **Access:** This characteristic describes how students access the content of the course/module/program.
  1) Online: Students/learners access the content through the Internet. This means that students can learn on their computers, laptops, phones, or any other device that is connected to the Internet. This makes it easier for them to access the content whenever and wherever they want.
  2) Offline: Students/learners access the content offline through CDs, DVDs, etc. This restricts the students to use computers and laptops to access the information. However, it frees them from the distractions of accessing it through the Internet where they might end up going to other irrelevant sites.
- **Student Count:** This characteristic describes whether students can interact among themselves or not.
  1) Individual: Students/learners do not interact with each other. Each student only interacts with the course instructor and does all of the work on an individual basis. This can help the student move at his/her own pace. However, this will limit the peer knowledge sharing process that can help students learn faster.
  2) Group: Students can interact with each other via emails, discussion boards, forums, and chat rooms among other means. This encourages peer knowledge sharing between students. However, it might discourage some students when they feel that the course/module/program pace does not match their own.
- **Time:** This characteristic describes the timing of when students can access the content.
  1) Synchronous: Content is given in real time in which students "attend" classes at the same time via conference calls, teleconferencing... This style mimics a real life class that is being held on a virtual platform. However, this might become a constraint especially if the students are geographically dispersed around the world as this would cause some students to "attend" the class late at night or in the early hours of the day.
  2) Asynchronous: Content is accessed at any time by students. Forums and discussion boards are available for students to interact with instructors or with each other. This overcomes the obstacle of having to learn at inconvenient times of the day since it allows the students/learners to learn at a time that suits them best.
- **Content Format:** This characteristic describes the format of the content.
  1) Static: The format is the same and the content remains the same throughout the course/module/program. This type ensures consistency and easily maintained. However, this might alienate some students/learners as it might not cater to their wants and needs.
  2) Dynamic: The format changes and adapts based on student behavior or knowledge acquisition level. This type caters more to the wants and needs of the students as it tries to modify the content according to the feedback and data collected from the students. However, this poses a challenge of how to provide new content and how to maintain it in the long run.

C. TYPES

There are different ways to categorize e-learning courses/modules/program. One way is the following that looks at e-learning at the course level [9]:

- **Classroom Course:** Course is given in a traditional classroom. However, it involves the use of computers for simulation or design purposes.
- **Synchronous Distributed Course:** Course is given in a traditional classroom in addition to being streamed using web conferencing for students at off-campus locations.
- **Web-enhanced Course:** Course is mostly face-to-face with some requirements being online (typically, 20% or less).
- **Blended/Hybrid Course:** Course is a mixture of face-to-face instruction and online instruction. These are traditionally used for courses whose students are within commuting distance of campus. There are two main types of blended/hybrid courses listed below.
  1) Blended Classroom Course: A slightly bigger portion of the course is given in a traditional classroom.
  2) Blended Online Course: A slightly bigger portion of the course is given online.
- **Online Course:** All requirements are done online. There is no face-to-face sessions required for the course. These are effective for students who do not have effective access to campus or for people that would like to take

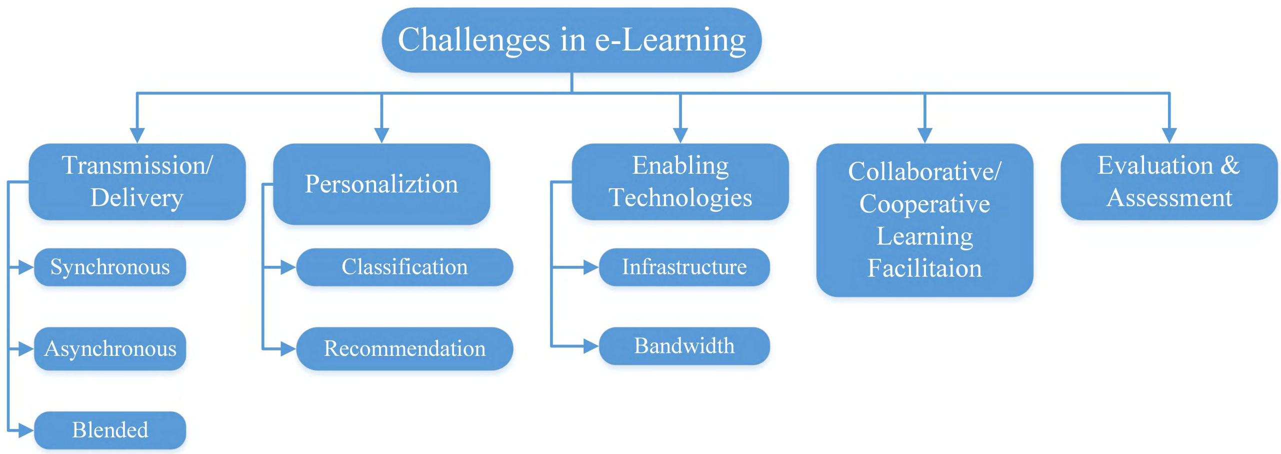

FIGURE 2. Challenges in e-learning.

a course to improve their knowledge in a specific area or in a particular topic.

- **Flexible Mode Course:** The course is offered in multiple delivery modes, allowing the students to choose the mode that suits them best (can be purely online or blended or classroom course).

A similar style of categorization can be given for e-learning at the program level [9]:

- **Classroom Program:** This can include a mix of web-enhanced, traditional, or hybrid courses. However, all courses require a portion of face-to-face sessions.
- **Multi-format Program:** This mixes classroom courses with other formats such as web-enhanced, blended, fully online courses. . .
- **Blended program:** Typically up to 30 % of the course requirements are given as face-to-face or blended course while the rest is fully online.
- **Online Program:** The whole program is given online.

D. APPROACHES

The teaching approaches adopted for instruction can also differ. In what follows, a brief discussion about three such approaches is given.

- **Flipped Classroom:** This approach is an innovative teaching approach in which the traditional instruction approach is flipped. When adopting it, class time is used for assignments while learning material is pre-recorded and made available for students to watch before class [10]. Jonathan Bergmann, a founder of this approach, states that the videos are not just for students to watch, but rather they are given to instigate discussion in-class. He says that this approach can also help struggling students as it allows them to complete assignments in class and hence promote engagement between them and himself [10]. This approach can be useful in blended classrooms that have an online portion and a face-to-face portion.
- **Game-based:** As the name suggests, this teaching approach is based on using games as a means for students to learn [11]. In this case, students work towards a specific goal by taking actions within a game. This allows them to experiment and gain achievement points along the way. Research showed that adopting such an approach can help improve cognitive and psychomotor processes [11]. This approach can be beneficial online courses since learners do not need to interact heavily with instructors.
- **Project-based:** This approach mainly focuses on the students working on a project by exploring real-world problems and challenges [12]. This allows students to identify and explore problems of interest to them as well as develop innovative solutions for them. This in turn can promote their creativity and provides them with more hands-on experiences when compared to traditional teaching approaches [12]. This approach again can be helpful in both blended courses as well as online courses. This is because instructors mainly act as facilitators that help guide the students in their projects in such an approach.

It is worth noting that there are other teaching approaches that can be adopted such as inquiry-based learning and kinesthetic learning [13]. However, focus was put more on approaches that may be suitable for e-learning environments that promote increased engagement and discussion as well as continuous access to content [14].

III. CHALLENGES IN E-LEARNING

There are many challenges within e-learning that are related to the different aspects of it. Some of these challenges are observed from the educational point of view and some are observed from the technological point of view. Fig. 2 summarizes these some of these challenges.

In what follows, a more detailed discussion about these challenges that were identified by leading experts within the field as well as some implemented case studies is presented. Also, some of the recent and representative literature that has been done on them is given [15]–[19].

A. CHALLENGE 1 - TRANSMISSION/DELIVERY

1) DESCRIPTION

Delivery and transmission of content can be done using different methods. Each method poses different challenges to both the learners as well as the teachers.

- Synchronous: In this delivery method, learners and teachers interact in real-time using tools such as video-conferencing and chat. However, this introduces several challenges. For example, since the class is being given in real time, teachers need to pace themselves in a manner to accommodate students with slower connections that may be suffering from some sort of lag. Also, it is important for teachers to be able to identify whenever any student poses a question. Thus, he/she must provide a clear structure for questions and answers so as to mimic in class participation. Another concern is for students and teachers to be tech savvy. This is important because a lot of students and teachers know only the basics of computers and computer networks. This often appears in courses such as literature and social studies [16]. Therefore, providing training for both learners and teachers about the tools and software that might be used is imperative. Last but not least, the issue of simulating hands on experience is also a concern within such courses. Some courses require some hands on experimentation. Hence, it is essential to develop ways to simulate and mimic these experiments to ensure learners fully grasp the content.
- Asynchronous: In this method of delivery, learners and teachers communicate via email and discussion board as not everyone is online simultaneously (as is the case with synchronous learning). Despite the flexibility it offers to learners, it still introduces some challenges. One challenge is keeping content stable and inclusive in such a way that it considers learners' diversity and addresses their knowledge gaps. Another challenge is keeping students motivated and not allow them to become isolated. This is essential since some students might become discouraged if they feel their learning pace is slower than others. Therefore, teachers need to find a way to keep students engaged and motivated. Moreover, learners do not get immediate feedback on their questions or concerns in the same manner that a synchronous style would provide. This might affect the learner especially if he/she do not know their current position within the course. Thus, it is important for teachers to set up a clear mechanism to provide learners with prompt feedback so that all learners feel engaged and appreciated.
- Blended: This delivery method combines both the online and classroom learning. This allows the teachers to design and modify the course so that it suits the students, as well as provide the needed learning objectives. This poses some challenges as it requires teachers to carefully select the material to be given online and the material to be given in class. Furthermore, efficient communication between the teachers and the IT department is imperative to ensure the content is posted on time so that the learning process be optimized.

2) PREVIOUS WORK

Tabak and Rampal discussed how to design and develop synchronous conferencing tool to compliment face-to-face content delivery [20]. The paper emphasizes the merits of incorporating new technologies into traditional courses to improve the students' experience. The technologies can also help overcome spatial limitations as it allows students to interact despite not being in the same geographical location.

Obasa *et al.* compare two e-learning platforms: an asynchronous platform called Modular Object Oriented Dynamic Learning Environment (MOODLE) and a synchronous platform called Elluminate [21]. The two platforms were compared in terms of resource requirements needed to implement them as well features they offer. Also, the benefits and challenges of each platform were discussed. The study showed that the decision of which platform is dependent on the resources available (both software and hardware) as well as the technological skills of its intended users (instructors and students).

Akram *et al.* also presented a case study based on the SCHOLAT learning system to try get insights into learners' behaviours [22]. Analysis showed that students were mostly interested in activities that they thought would improve their grade like submitting their assignments on time. However, other activities like asking questions on the discussion forum were not as important and influential as some students never asked any questions there. The analysis is said to be the basis for more complicated analysis to determine how to get students more involved and encourage them to be more active on the SCHOLAT system.

B. CHALLENGE 2 - PERSONALIZATION

1) DESCRIPTION

It is well known that different people learn in different ways. Some prefer instruction while other prefer doing it themselves. Therefore, being able to create and adapt the content in such a manner that caters to the different learners' styles/preferences and needs is extremely important since it can help maximize and speed up the learning process. This is called ''Personalizing'' e-learning.

Personalization in e-learning is a topic that is garnering significant attention within this field [23]–[25]. Since students deal with different technological devices, it becomes imperative to adapt the content to maximize and speed up the learning process. Two main tasks are needed in e-learning personalization, namely classification and recommendation. Classification tries to categorize the dataset into different classes based on some metrics. Recommendation uses the aforementioned classification to recommend

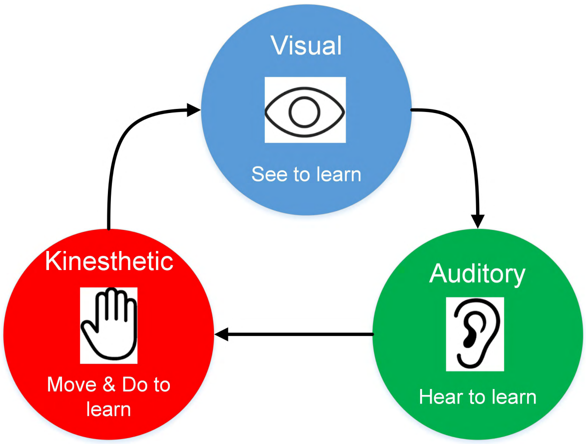

FIGURE 3. VAK learning styles model.

a course of action that can maximize or improve a performance metric.

a: CLASSIFICATION

Classification is an essential step in the e-learning personalization process. It forms the basis of the process as it groups similar points within the dataset available under one umbrella. This helps the system make better recommendations. There are several aspects that need classification such as question/task classification [26] and student status classification [27]. However, one important classification process for e-learning is the student learning style/preference classification.

It is well known from educational research that different people learn in different ways [28]–[30]. Some people learn through verbal instruction. Other people learn by doing things themselves. There are different models adopted to classify the learning style/preference of learners. Three of the most popular models are discussed below.

(i) *VAK Model:* The VAK model is the simplest learning style model that can be used for classification as shown in Fig. 3. VAK stands for visual, auditory, and kinesthetic (the three main sensory receivers). This style assumes that a person either prefers to learn by seeing the subject matter, hear about the subject matter, or do a task related to the subject matter [28], [31].
Each style/preference has some characteristics that differentiates it from the other styles. In what follows, we summarize some of the characteristics of each style/preference.
- **Visual Learners:** This type of learners often prefer visual stimulants. For example, they like to read or write something repeatedly for them to learn. They also often favor figures and diagrams in order to better understand concepts. Students with this learning style prefer to use visual aids during lessons as well as have the handouts included concept maps and outlines. They also often highlight important points in the text with colors [28], [31].
  - Use visual aids such as charts and graphs during lessons.
  - Have the content in handout form with a lot of white space for them to take notes.
  - Have the handouts include concept maps, outlines, agendas, etc. for them to read at later times.
  - Highlight important points in the text with colors.
- **Auditory Learners:** This type of learners prefer auditory stimulants. They like to discuss things and talk about them with their colleagues or instructors. The analyze the tone, pitch, and speed of the speech to help them better interpret the meaning of the content they are encountering. Students with this learning style prefer to work in groups and discuss ideas with someone else. They also prefer having a short summary at the beginning/end of the session about the material to be learned/learnt [28], [31].
  - Work in groups.
  - Discuss their ideas with someone else.
  - Recite the information out loud to emphasize it.
  - Receive a short summary at the beginning of the session about what they will learn and at the end of of the session about what they learned.
- **Kinesthetic Learners:** This type of learners prefer tactile stimulants. They learn better when they move and do the tasks with their own hands. Students with this learning style prefer to keep moving and tend to take frequent breaks to stretch and refocus [28], [31].
  - Keep moving.
  - Take frequent study breaks to stretch and refocus.
  - Glance through the reading material to get the big picture first before delving into the details.
  - Work at a standing position as it allows them to move more freely.

(ii) *Kolb's Model:* Another model used for learning styles classification is Kolb's model [29]. David Kolb developed a model containing four distinct learning styles that are based on a four-stage learning cycle. "Cycle of learning" is a central principle in his experiential learning theory. The four-stage cycle includes: Concrete Experience (CE), Reflective Observation (RO), Abstract Conceptualization (AC), and Active Experimentation (AE). Based on his learning cycle, **immediate or concrete experiences** provide a basis for **observations and reflections**, which are assimilated and distilled into **abstract concepts** producing new implications for action which can be **actively tested** in turn creating new experiences. Kolb explains that different people naturally prefer a certain single different learning style. The strength of this model is that it takes into consideration two continuums: Processing continuum (how we process information) and Perception continuum (how we receive information). Perception continuum $\Rightarrow$ CE (feeling) vs AC (thinking) while Processing continuum $\Rightarrow$ RO (watching) vs AE (doing). The four different learning styles are the result of the person's preferred

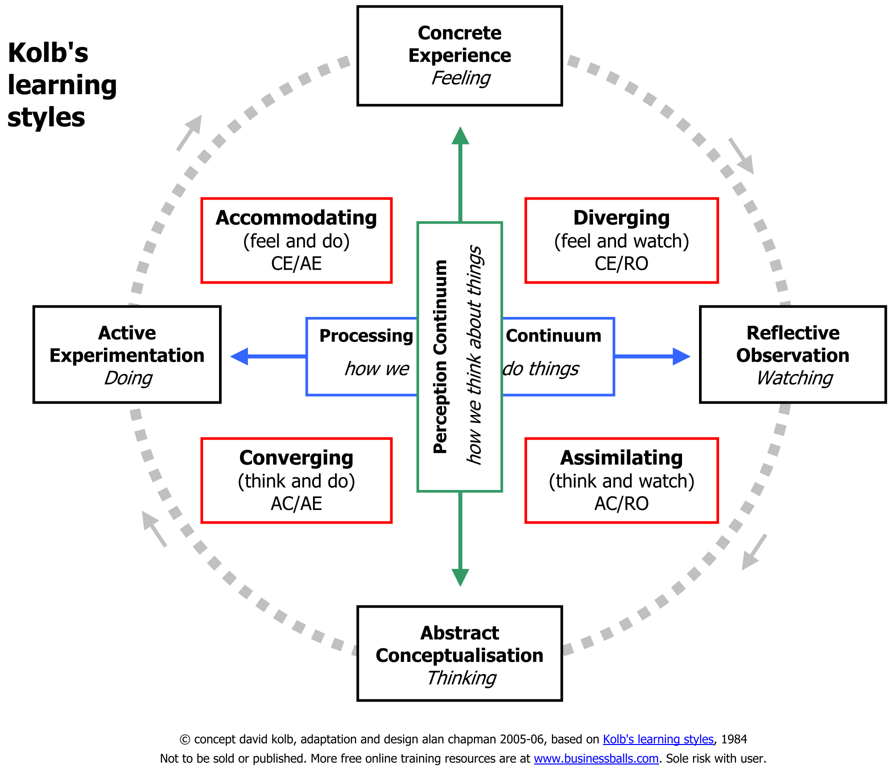

FIGURE 4. Kolb's learning styles model [29].

choice of which type of perception and processing styles to learn. These styles are described below:

– **Diverging (CE/RO)** ⇒ Feeling and watching. Such people perform best in situations that require ideas-generation such as brainstorming. They prefer to work in groups.
– **Assimilating (AC/RO)** ⇒ Watching and thinking. Such people prefer abstract ideas and concepts and tend to excel in information and science careers.
– **Converging (AC/AE)** ⇒ Thinking and doing. Such people are good at finding practical uses for ideas and theories. They prefer experimenting and working with practical applications.
– **Accommodating (CE/AE)** ⇒ Doing and feeling. Such people rely on other people's analysis. They rely on their "gut" instincts. The also prefer to work in groups to complete a task.

Kolb's model is correlated to other models. For example, this model is correlated to Carl Jung's model (learning styles result from people's preferred ways of adapting in the world). Kolb's model also resembles the VAK model, but is a bit more descriptive. Fig. 4 illustrates the cycle and the different learning styles proposed by Kolb.

(iii) *Felder-Soloman Model:* Another learning style classification model is the Felder-Soloman learning style model (FSLSM) shown in Fig. 5 below [30]. The model proposes 16 possible classifications with each classification being based on how we process, perceive, receive, and understand information (for example, a person's

Information Processing	
Active	Reflective
Doing something active with information.	Thinking and reflecting about information.
Information Perception	
Sensing	Intuitive
Learning the facts.	Discovering relationships
Information Reception	
Visual	Verbal
See pictures, flowcharts, timelines….	Read and hear written and spoken explanations
Information Understanding	
Sequential	Global
Understand in linear steps (Gradual)	Understanding in large jumps

FIGURE 5. Felder-Soloman learning style model.

learning style can be: active, sensing, verbal, and global). This model is more extensive than Kolb's model as it provides a classification at a more granular level. This model tries to categorize students/learners based on the different levels of the learning process. Knowing to which category a student/learner belongs to can help the instructor deliver the content as it gives him the guidelines of what information to give and how to give it. This in turn can speed up and improve learning since the whole process would be designed based on the students/learners preferred learning styles. For example, active learners prefer working in groups as it allows them to discuss the information they received. Therefore, instructors can encourage active learners to form groups and discuss the presented concepts. On the other hand, reflective learners prefer to work individually because it allows them to think and reflect about the received information. Hence, instructors can give such students a few minutes to think about the presented concepts before trying to engage with them and make sure that they fully absorbed the material.

It is worth noting that "learning styles" theories and models are largely unproven using research, but they do offer benefit when used by teachers and educators as guidelines for teaching [32].

b: RECOMMENDATION

Recommendation is the second step of e-learning personalization. This step involves providing an informed suggestion on an appropriate course of action based on the results of the aforementioned classification step. The proposed action is thought to maximize a certain satisfaction/utility metric of the user [33]. Such systems have been used in several fields such as e-commerce and route planning using public transportation [34], [35].

In the context of e-learning, recommender systems can play a vital role. For example, e-learning platforms such as "Coursera" and "edX" use the previous searches that a user performed to recommend new courses to take. This aims to direct the users of the platform towards courses that are thought to be of interest to them. Recommender systems can also play a role in adapting the content of a course by suggesting content format based on any previous classification of the content or of the users of the content [36], [37].

2) PREVIOUS WORK

Hogo proposed the use of different fuzzy clustering techniques such as Fuzzy C-Means (FCM) and Kernelized Fuzzy C-Means (KFCM) to classify the learners based on their profile [23]. The goal of the work was to try to determine how to help bad students improve their performance. Experimental results showed that the KFCM algorithm outperformed FCM as it achieved an accuracy of around 78%.

Joseph proposed a machine learning-based framework for feedback and monitoring in an e-learning environment [24]. The goal of the study was to see how machine learning algorithms can be used to adapt the content presentation rather than the content itself. This would be done based on the feedback from the students. The challenge was to determine what to monitor within the e-learning environment and how to interpret it as feedback. Decision trees and neural networks were used to map the interactions and forecast the interest/disinterest of learners with the material. A small control study consisting of 22 students was used as means to evaluate the effectiveness of the proposed framework. The experimental results showed high agreement between the interest/disinterest forecasting of the framework and the student views on the content, providing an important evidence to the potential of the framework.

Virvou *et al.* proposed a data analytics tool that collects student interaction information in order to help teachers make better decisions [25]. The tool collects data about the student interaction by storing their activities on the e-learning platform. Then a text analysis is performed with results being reported in a single document that is sent to the course instructor. The hope was that this document can help the instructors adapt their content based on the student feedback.

Klašnja-Milićević *et al.* propose a recommendation module of a programming tutoring system (PROTUS) that can automatically adapt to interests and knowledge levels of students [38]. They were inspired by recommender systems (RSs) implemented in e-commerce and wanted to develop a similar system for tutoring programming. The recommender module they proposed has three main tasks: recognize the different learning styles, cluster students based on styles, and recommend a series of activities for each group. The Felder-Soloman learning style model was adopted for student classification.

Another work that aimed to offer a "personalized" e-learning environment was presented in [39]. The work was a Master's thesis that aimed to prove that e-Learning usage data can be used to derive general metrics that can be easily and economically applied to improve e-learning processes. The author believed that usage data plays an important role to reflect the activities and systems. He used a six layer model to create a dynamic learner model and implemented a testbed to evaluate the model. He defined the learner model as "a set of information structures that represent learners' behaviours and preferences".

The author adopted a dynamic method that aimed to learn information about the students through real-time collected data. Hence, a six layer model (with each layer depending on the lower layer) was proposed:

- **Raw Data Layer:** Set of data collected directly from the system.
- **Fact Data Layer:** Set of data that is filtered, sorted and combined to correctly describe a student's behavior.
- **Data Mining Layer:** Set of data mining techniques used on the fact data.
- **Measurement Layer:** Set of results of the patterns produced by the machine learning and data mining techniques.
- **Metric Layer:** Represents an aspect of the student's characteristics, experiences and preferences in their learning model (examples: learning style, activities, social connections).
- **Application Layer:** Set of educational applications that can be done using the information and collected results from lower layers.

C. CHALLENGE 3 - COLLABORATIVE/COOPERATIVE LEARNING FACILITATION

1) DESCRIPTION

One important aspect of many courses is group projects/tasks. Instructors often ask students to work collectively on projects/tasks in an attempt to foster learners' teamwork skills as well as allow the peer-to-peer knowledge sharing. However, this may become an obstacle especially in online courses as learners often only communicate via email or discussion boards. The question becomes, how can collaboration be facilitated and improved in e-learning? Can intelligent educational systems improve collaboration and knowledge-sharing? How can tasks be intelligently assigned? All these questions need to be considered and addressed in an e-learning environment [40].

It is worth noting that there is a slight difference between collaborative learning and cooperative learning. Cooperative learning focuses on team building by allowing learners within a group to complete individual tasks that aim to achieve success for the whole group (for example, a group project). On the other hand, collaborative learning focuses on collective problem-solving by encouraging students to effectively communicate with each other to collectively analyze and discuss possible solutions for the tackled problem (for example, working on a new research).

2) PREVIOUS WORK

Suciu *et al.* developed an online collaboration and e-learning platform in the cloud [41]. The platform allows users to chat, share documents, have voice and video calls, and has enabled calendar scheduling. Such a platform can encourage and facilitate collaboration and cooperation between students as it allows both real-time and non real-time interaction between them. The fact that the platform is cloud-based further facilitates access and makes use of the computational and storage resources available to satisfy the needs of the different users.

Rajam *et al.* also proposed a cloud-based collaboration environment for e-learning [42]. The environment offers students a software development platform to encourage task management and distribution. The environment allows students to tackle different tasks within a group project. This facilitates and encourages collaboration among students because it allows each one to focus on a specific task within the project.

Franceschi *et al.* proposed using virtual worlds as a means to foster group collaboration in e-learning environments [43]. The idea is that using avatars, both verbal and non-verbal communications can be better exchanged between students and promote more creative thinking. This in turn will promote effective group learning and collaboration as it would mimic a real-life situation in which students are working together in the same geographical location.

Mahadevan *et al.* on the other hand proposed a framework that is based on forming online group studies to encourage collaboration in an e-learning environment [44]. Similar to the typical physical group study sessions that students often organize to prepare for exams, the authors suggest a framework that integrates seamlessly with e-learning platforms and allows students to create an online study group in which they can share ideas and documents.

D. CHALLENGE 4- EVALUATION & ASSESSMENT

1) DESCRIPTION

This challenge is two folds. The first is from the educational point of view while the second is from the technological point of view. The first challenge is how to evaluate the student performance. This task becomes harder especially when little to no face-to-face interaction occurs. Therefore, instructors need to know how to evaluate students and track their knowledge acquisition efficiently. This is of significant importance in asynchronous online courses specifically where is there is no real-time interaction between instructors and students at all.

The second challenge is how to evaluate the efficiency of the e-learning tool or environment used. This is important because the tool or environment can have a big impact on the overall learning process. Therefore, it is essential that instructors/institutions properly evaluate the tool being used to be able to determine whether students are meeting the required learning outcomes or not when using this tool/environment.

2) PREVIOUS WORK

Chang focused in her work on the importance of examining and evaluating the e-learning environment [45]. To that end, a frame work that scores the e-learning environment based on four metrics: access, interaction, response, and results was proposed as means to evaluate the appropriateness of the environment. These four metrics can give instructors/institutions a quantitative measure of how well the e-learning tool/environment is and can guide them towards areas in which they can improve it.

Balogh *et al.* proposed the to perform usage analysis to better evaluate the efficiency of the e-learning tool/environment [46]. Data is collected using questionnaires, as well as the log files of the online course to determine the behaviour rules of the students. This can measure the engagement level of the students with the course and give instructors/institutions insights into the effectiveness of the adopted tool/environment.

Tahereh *et al.* presented a multi-dimensional framework to evaluate the different components of the e-learning environment [47]. The framework takes into consideration the standards used in the environment, the evaluation methods, the context, and the indicators to evaluate the quality of the environment. Each dimension is evaluated qualitatively to give the stakeholders a better understanding of efficiency of the e-learning environment. This allows them to determine which dimension needs improvement and which is appropriate for their goals.

E. CHALLENGE 5 - ENABLING TECHNOLOGIES

1) DESCRIPTION

Another important choice in e-learning is that of the enabling technologies. This is important because these technologies (such as cloud environment [48]) will be the tools that learners use to acquire the knowledge and access the content.

Hence, it is important to carefully choose the supporting technologies so that it maximizes the learning process in different scenarios. The proper infrastructure and bandwidth requirements needed should also be considered as they might act as the system's bottleneck.

2) PREVIOUS WORK

Caminero *et al.* proposed load forecasting algorithm for an e-learning system [49]. The goal of the algorithm is to efficiently provision the needed computational and storage resources to maintain a desired quality of service (QoS) as perceived by the users of the system. With this algorithm, cloud technologies are used as the main infrastructure to provide the users with an efficient e-learning system.

Dong *et al.* also presented a cloud computing-based e-learning system [50]. The system adopts the cloud architecture consisting of three main layers: infrastructure layer, content layer, and the application layer. This architecture allows the universities to track the situation of their resources and allocate them on demand, allows workloads to recover from any faults, and promotes the evolution of the learning process in terms of content and engagement.

Mondal *et al.* posited a low cost low bandwidth virtual classroom system for distance learning [51]. The system is based on using UDP protocol for voice streaming while a TCP connection is used for visual content transfer. Small control messages are sent from the teacher's end to all students. These control messages invoke specific events (such as changing a slide in the presentation). Hence, most of the traffic within the proposed system consists of voice packets and some small control messages which typically only requires low bandwidth.

IV. MACHINE LEARNING & DATA ANALYTICS

One can see that the field of e-learning can significantly contribute to the notion of big data as increasing number of students access educational material online and generate more data flows. It has been reported that around 58 million students have registered for online courses worldwide with nearly 7000 courses available [3]. With typical courses generating thousands of event log records per student per course, the amount of data generated is growing exponentially [3]. Thus, machine learning and data analytics also become crucial in order to make use of the growing amount of collected data generated in the field of e-learning.

Machine learning and data analytics have been proposed as possible solutions to process the increased amounts of data collected. These algorithms are the tools that can make use of the data by "learning" the behavior and finding interesting patterns within it.

In what follows, the different machine learning and data analytics types and algorithms are discussed briefly.

A. MACHINE LEARNING

As shown in [52], data has become more abundant and easy to obtain. However, extracting knowledge from this collected data is often expensive. With the help of computers which can perform calculations at tremendous speeds, more complex data analysis has become available. Moreover, having computers that can "learn" without being told what to do is essential as this would give it a greater capacity to adapt based on new inputs. To this end, the field of machine learning has been developed. Machine learning, as per Arthur Samuel, is defined to be the "field of study that gives computers the ability to learn without being explicitly programmed" [5]. The learning is performed using some data or observations such as examples, direct experience, or instruction [5]. This is crucial in many cases such as when the solution changes over time or when it needs to be adapted to specific cases. Thus, "learning" based on experience is imperative in many future applications.

Machine learning algorithms have garnered significant attention in recent years and have been used in several applications such as pricing prediction, optical character recognition, image recognition, spam filtering, fraud detection, healthcare, transportation, and many others [5], [53]–[56]. Fig. 6 shows some of the different machine learning types and algorithms that will be discussed in this work.

1) TYPES

Machine learning algorithms are divided into four main categories: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. Supervised learning includes having a dataset with the correct output that is used to "train" the system. On the other hand, unsupervised learning includes trying to find relations among the points in the dataset without having the correct results during training [57]. This means that the algorithm tries to "cluster" points that it believes to be highly correlated under one label based on their statistical properties only. Semi-supervised learning combines the previous two types by training the system using a dataset containing labeled and unlabeled data points. The goal is to improve the performance of the model by making use of both types of data points [58]. Last but not least, reinforcement learning in contrast uses trial-and-error to discover the set of actions that maximize some cumulative reward metric [59].

a: SUPERVISED LEARNING

Supervised learning is a branch of machine learning algorithms in which a function is inferred based on labeled training data [60]. The training data is formed of a group of training examples, each of which is a pair (x, y) where x is an input vector and y is the output value. The algorithm produces a function that can be used for mapping future unknown inputs.

Supervised learning algorithms often fall into one of two main categories: regression algorithms (output is continuous) or classification algorithms (output is discrete) [61]. Within each category, several algorithms exist which will be presented below.

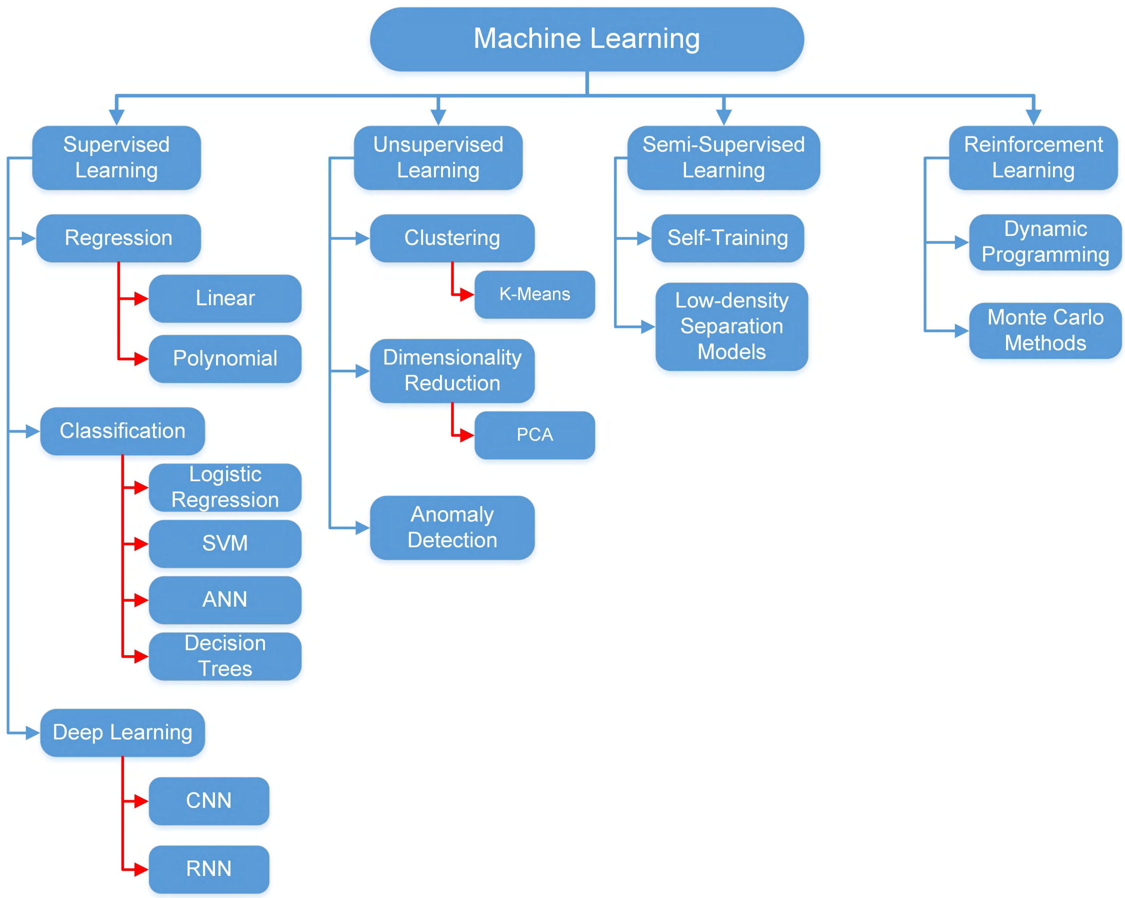

FIGURE 6. **Different machine learning categories and algorithms.**

1) **Regression:** Regression algorithms try to find the best fit function for the training data available. Two main algorithms are discussed below: linear regression and polynomial regression.
   - **Linear Regression:** One of the most common regression algorithms used in machine learning is the linear regression algorithm. This algorithm tries to find the best fit line/hyperplane for the available training data. The goal of the algorithm is to find the value of the optimal coefficient vector $\theta_{opt} = [\theta_0, \theta_1, ..., \theta_N]$ such that the predictive function has a linear form. This is done by using the minimum mean squared error function.
   - **Polynomial Regression:** Another common regression algorithm used is the polynomial regression algorithm. This algorithm tries to find the best fit polynomial for the available training data. Similar to its linear counterpart, the goal of the algorithm is to find the value of the coefficient vector $\theta_{opt} = [\theta_0, \theta_1, ..., \theta_{kN}]$ such that the predictive function is a polynomial of order k.
2) **Classification:** In contrast to regression algorithms that try to find the best fit function for the training data, classification algorithms try to find the best fit class for the data by putting each input in its correct class. In such cases, the output of the predictive function is discrete with the possible values being one of the different classes available as part of the training data. Four important classification algorithms are discussed below, namely logistic regression, artificial neural networks, support vector machines, and decision trees.
   - **Logistic Regression:** Logistic regression is an extremely popular classification algorithm used in the literature [62]–[65]. Despite its name, this algorithm is used for classification (i.e. its output is discrete) rather than being a regression algorithm. It is typically used a binary classifier where the output belongs to one of two categories only. The predictive function $z_{\theta_{opt}}(x)$, also known as the hypothesis function, calculates the probability of the output being equal to 1 given a specific input. In other words, $z_{\theta_{opt}}(x) = P(y = 1/x; \theta)$. If this probability is greater than 0.5, the output is defined to be 1. Otherwise, the output is defined to be 0.

     To determine the coefficient vector θ_{opt}, a cost function needs to be used. However, the squared error function will not work because the sigmoid function used to determine the hypothesis function will make the output wavy and thus have many local optima. Instead, a different cost function based on the log function is used that ensures the output is convex [66]. This algorithm can also be extended for multi-class classification. In that case, instead of having just one hypothesis function $z_{\theta_{opt}}(x)$, we would have multiple functions with each calculating the probability of the output being class i given the input available.

- **Support Vector Machines:** Support vector machines (SVM) is another supervised classification algorithm. It tries to find the optimal hyperplane that separates the labeled data with the maximum margin from the closest point. It is a more powerful and restrictive classifier than the logistic regression algorithm. This algorithm replaces the sigmoid function used in logistic regression with a new function called the hinge loss function [67].
  Note that the hypothesis function given by the SVM algorithm is not interpreted as the probability of the output being 1 or 0, but rather is a discrimination function that outputs either 1 or 0.
- **Artificial Neural Networks:** Artificial neural networks (ANN) is a popular supervised classification algorithm. It is often used whenever we have abundant labeled training data with many features and a non-linear hypothesis function is desired [68]. ANN tries to mimic the way our brain works as it has been proven that the brain uses one ''learning algorithm'' for all its different functions [69].
  Similar to neurons that act as computational units that take electrical inputs (through dendrites) and channel them towards an output (axon), the ANN algorithm adopts a model in which the features act as dendrites (nerve cell) and outputs the value of the hypothesis function. Often, one ''hidden'' layer is used that acts as an intermediate layer. This layer helps extract more information from the set of features available as part of the training data and is called the activation layer.
  The sigmoid function used in logistic regression is used here at each layer of the network. Since the sigmoid function is used in ANN, then the cost function used to determine the values of the coefficient vector $\theta_{opt}^{(f)}$ at layer f is the same one used for logistic regression.
- **Decision Trees:** Decision trees are another popular choice of supervised learning classification algorithms. These algorithms are often referred to as statistical classifiers since they use statistical metrics to determine the branching of the nodes [70]. To classify an instance, decision trees sort the instance down the tree from the root node to a specific leaf node. Each node within the tree represents a test of a particular feature of the instance while each branch represents a possible value of the tested feature.
  There are several decision tree-based algorithms such as ID3, ASSISTANT, and random forests [70]. One of the most well-known decision trees algorithms among them is the C4.5 algorithm. The algorithm was first proposed by Quinlan [71] and is based on the notion of information entropy. Effectively, the C4.5 chooses the feature that best divides its set of samples into smaller subsets rich in one class or the other. To determine the division criterion, the normalized information gain metric (difference in entropy) is used. The feature chosen is the one with the highest information gain.

3) **Deep Learning:**
One special class of supervised machine learning algorithms is deep learning. In essence, deep learning can be thought of as a large scale neural network [72]. However, due to the fact that deep learning is also able to perform automatic unsupervised feature extraction, also commonly referred to as feature learning [73], it cannot be classified as a traditional neural network. Hence, deep learning is then considered a special case of supervised machine learning. In general, deep learning tries to model abstractions found in data using a graph with multiple processing layers [74]–[76]. These processing layers contain units that apply linear and non-linear transformations on the data to extract as much useful information as possible.
Deep learning algorithms are very similar to artificial neural networks. In fact, ANN can be classified to be one of the deep neural networks learning algorithms. However, deep learning algorithms are more broad as they can be applied to both labeled and unlabeled data. Moreover, they can be applied to a much larger scale of neural networks. Andrew Ng, co-founder of Coursera and the Chief Scientist at Baidu Research, said that deep learning is just applying ANN on a large scale that can be trained with more data and have better performance because of that [74].
There are many different deep learning algorithms other than ANN. In what follows, two popular algorithms are briefly described and discussed.
- **Convolutional Neural Networks:** Convolutional Neural Networks (CNN) is a version of artificial neural networks that was inspired by the connectivity patterns found in the visual cortex of animals. These connectivity patterns have been proven to be mathematically described by a convolution operation [77]. This operator replaces the sigmoid function that is typically used in artificial neural networks.
- **Recursive Neural Networks:** Recursive Neural Networks (RNN) is another version of deep neural networks. RNNs are created by using the same set of weights in a recursive manner over a tree like structure [78], [79]. The tree is traversed in topological order. This algorithm is often used to process sequential data. Similar to the idea that CNNs are used

to process data grouped in the form of a grid, RNNs are specialized in processing a sequence of values.

b: UNSUPERVISED LEARNING

In contrast to supervised learning, unsupervised learning is the branch of machine learning in which a function/pattern is inferred based on unlabeled training data [80]. The training data consists of only inputs $x^1, x^2, ..., x^M$ and no known outputs. Therefore, unsupervised learning algorithms aim to make sense of the training data by finding relations and patterns within it.

Unsupervised learning algorithms can be divided into 3 main categories: clustering, dimensionality reduction, and anomaly detection [80]. Several algorithms exist within each category. In what follows, some of the most popular algorithms in these categories are presented.

1) **Clustering:** One of the easiest way to make sense of a set of data points is to group/cluster them. This makes the data more understandable as it gives more structure to it by forming a finite set of groups rather than having a multitude of random data points. This is especially important in applications such as market segmentation and social network analysis. Since we do not know whether this grouping is correct or not, the term "clustering" is used rather than "classification" because the data points are not labeled to be belonging to specific classes. In what follows, a well-known clustering algorithm is discussed.
   - **K-Means Algorithm:** K-means is one of the most popular unsupervised clustering algorithms for automatic data grouping into coherent clusters. This algorithm tries to group the data into K clusters by finding the cluster centroid (also known as cluster mean) and group with it the data points closest to it.

     To be able to properly cluster the data points into the K clusters, a cost function needs to be minimized. This cost function is dependent on 3 variables: $c^{(j)}$ which is the index of the cluster to which example $x^{(j)}$ is currently assigned to, μ_k centroid of cluster k, and $\mu_{c^{(j)}}$ the centroid of the cluster to which example $x^{(j)}$ is currently assigned to.

     The algorithm aims to find the indices and centroids that will minimize the average distance of every example to its corresponding cluster centroid.
2) **Dimensionality Reduction:** Dimensionality reduction is another essential topic in the field of machine learning. The motivation behind dimensionality reduction can be summarized as follows:

   i- Remove redundant data

   ii- Reduce the storage and computational needs

   iii- Simplify the visualization of data by only considering a few features.

   To this end, one well-known algorithms is discussed which is the principal component analysis (PCA) algorithm.
   - **Principal Component Analysis:** Principal component analysis is one of the most popular dimensionality reduction algorithms in unsupervised learning. Its aim is to find the subset of features that best represents the data. For example, assuming two features x_1 and x_2 are given, PCA tries to find a single line that can describe both these features effectively at the same time. This is different from linear regression since the goal of PCA is to reduce the average projection error (orthogonal distance from the feature to the projection line) while linear regression tries to reduce the average error (vertical distance) to the line. Performing PCA analysis involves four steps: data preprocessing, covariance matrix computation, eigen value decomposition, and choosing first k eigen vectors.
3) **Anomaly Detection:**

   Another important unsupervised learning algorithm is the anomaly detection algorithm. From its name, this algorithm tries to determine whether the given new example $x^{(new)}$ is anomalous or not. To do so, a probability function/model $p(x)$ is calculated which gives the probability of an example not being anomalous. A threshold value, denoted by ϵ, is used as the dividing value between identifying the example as normal or anomalous.

   To be able to calculate the probability function $p(x)$, the set of features are assumed to be independent and hence the probability function $p(x)$ becomes the product of the probabilities of the features $p(x_i)$ $\forall i$. Another assumption that is made is that the features are normally distributed, i.e. $p(x_i)$ $\forall i$ follows the gaussian distribution. Hence, determining whether $x^{(new)}$ is anomalous or not is done by comparing $p(x^{(new)})$ with ϵ: if $p(x^{(new)}) < \epsilon$ then the example is anomalous, otherwise it is normal.

c: SEMI-SUPERVISED LEARNING

Semi-supervised learning is a branch of machine learning techniques in which a function/patter is inferred based on partially labeled training data [58], [81]. It combines elements from supervised and unsupervised learning. The aim of semi-supervised learning is to try to make use of both the labeled and unlabeled data to get better learning models [58], [81]. Typically, these algorithms are implemented whenever the dataset has a small amount of labeled data points and an abundance of unlabeled data points.

To be able to make use of the unlabeled training data examples, some assumptions need to be made to have some structure to the data distribution. All semi-supervised learning algorithms make at least one of the following assumptions [82]:

– **Smoothness Assumption:** It is assumed that points that are close to each other are more probable to share the same label.

– **Cluster Assumption:** It is assumed that the data tend to form discrete clusters with points within the same cluster being more probable to share the same label.
– **Manifold Assumption:** It is assumed that the data lies approximately on a manifold of lower dimension than the input space. This is a practical assumption in cases where high-dimensional data is generated such as in voice recognition. In that case, it is know that the voice is controlled by a small number of vocal folds and thus only the space surrounding these folds is considered rather than the space of all possible acoustic waves.

Semi-supervised learning differs from transductive learning. That is because the aim of transductive learning is to predict the correct label of the set L using the labeled training set M. On the other hand, semi-supervised learning, which is inductive, aims to use both the labeled and unlabeled training sets to obtain a function that can predict the label for any "future" example [58], [81]. There are several algorithms that fall under the umbrella of semi-supervised learning. In what follows, some of these algorithms are discussed.

- **Self-training:** Self-training is the simplest semi-supervised learning algorithm. This algorithm first obtains a prediction function z using the labeled data points. It then applies this function to a subset of the unlabeled dataset and adds the outcome $(x^{M+l}, z(x^{M+l}))$ to the labeled data. It then repeats the process to obtain a better prediction function z^* until all the unlabeled points are given a label and are used to learn a final model z^{final}.
- **Low-density Separation Models:** Another group of semi-supervised learning algorithms is the low-density separation models. These algorithms try to place the dividing boundaries in regions with a few data points (labeled or unlabeled). Among this class of algorithms is the transductive support vector machine algorithm (TSVM), also known as semi-supervised support vector machine algorithm (S3VM). Similar to the supervised learning SVM, TSVM aims to find the decision boundary that maximizes the margin over all the data. To do that, the unlabeled data needs to be labeled. Therefore, the TSVM algorithm enumerates all possible labeling of the unlabeled data points (2^L possible labels) and uses SVM to find a decision boundary. After finding all the possible decision boundaries, the one with the largest margin is chosen.

d: REINFORCEMENT LEARNING

Reinforcement learning (RL) is a branch of machine learning in which the action is taken in such a manner so as to maximize a cumulative reward metric [59]. This is often done using a trial-and-error sort of way in an attempt to discover the actions with the highest rewards. The decision taken often not only affects the immediate reward, but also subsequent ones. These two features, which are trial-and-error and delayed reward, are the two most distinguishing characteristics of reinforcement learning [59].

RL problems are often modelled as a stochastic finite state machine problem with inputs being the actions taken by the agent and the output being the observations and rewards obtained by the agent [83]. The environment consists of:
– A set of states S describing the environment.
– A set of possible actions A.
– Transitioning rules between states.
– Scalar reward rules of a transition and the observation rules of an action.

RL agents take actions in discrete time steps. Hence, at time instance t, the agent makes observation O_t which includes a reward r_t. The agent then chooses an action a_t that makes the environment move from state S_t to state S_{t+1}. This transition is then associated with a reward r_{t+1}. The goal of the agent is to maximize the collected cumulative reward [84]. The RL model depicting the interaction between the agent and the environment is shown in Fig. 7.

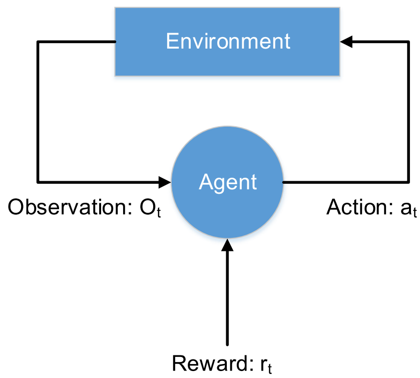

FIGURE 7. RL model.

There are a variety of algorithms that fall under the RL category. These algorithms have proved to be popular in a variety of applications including control theory, telecommunications, gaming (backgammon and checkers), and scheduling [59]. In what follows, some of the most popular RL algorithms are discussed [85], [86].

- **Dynamic Programming:** Dynamic programming is one of the most well known RL algorithms. This algorithm is considered a value function-based approach. Value function-based approaches try to determine a policy π, a mapping that assigns a probability distribution of the actions to all possible histories, that maximizes the reward by maintaining a set of estimates of the expected rewards for different policies. Using the recursive relationships that all value functions follow, the optimal policy π^* can be determined using the transition probabilities between states and the expected rewards gained [86].
- **Monte Carlo Methods:** Another group of algorithms that are used for RL is Monte Carlo Methods. The benefit of these algorithms is that they do not assume complete knowledge of the environment, but rather learn about it by experience (sample sequence of stations, actions, and rewards) [85].

TABLE 1. Summary of ML algorithms.

Algorithm	Data Type		Problem Type	
	Labeled Data	Unlabeled Data	Prediction & Classification	Pattern & Structure Discovery
Linear Regression	x		x	
Polynomial Regression	x		x	
Logistic Regression	x		x	
SVM	x	x	x	
ANN	x	x	x	
Decision Trees	x		x	x
K-means		x		x
PCA		x		x
Anomaly Detection	x	x	x	
Self-training	x	x		
Low-density Separation models	x	x	x	x
Graph-based Algorithms	x	x		x
Dynamic Programming				
Monte Carlo Methods				
Heuristic Methods				
CNN	x	x	x	
RNN	x	x	x	

This group of algorithms depend on running several iterations to evaluate the policy for the different values of the state-action pairs (S_i, a_i). The different iterations start at random from different states. The computed values are then averaged out to get a good estimate of the action value function of the considered policy.

Table 1 summarizes the different machine learning algorithms, the data types they can work on, and the problem types they can be applied for.

B. DATA ANALYTICS

Data analytics has become an increasingly important tool for users to extract knowledge and inferences from the abundant amount of collected data [52]. This analysis has become easier with the emergence of sophisticated computers and software that are able to perform exhaustive computations in lightning quick speeds. In contrast to machine learning which tries to "teach" computers to perform certain tasks without explicitly programming them, data analytics try to analyze and make inferences based on the raw collected data in order to take better decisions [6].

To this end, different techniques and algorithms have been developed. These algorithms have been applied in various fields such as business, marketing, health care, transportation, and education [6], [87]–[89]. This has helped individuals make more informed decisions about the data available at their hands by making use of the results of such techniques.

Fig. 8 shows the different data analytics types and algorithms that will be discussed.

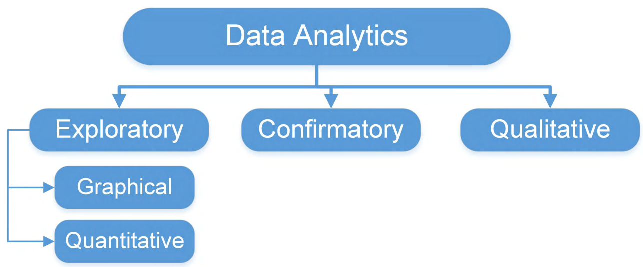

FIGURE 8. Different data analytics categories and algorithms.

1) TYPES

Data analytics algorithms and techniques can be divided into three main categories: exploratory, confirmatory, and qualitative. Each category represents a different goal and approach when it comes to dealing with the data.

(i) **Exploratory Data Analytics (EDA):** Exploratory Data Analytics (EDA) is a group of techniques that tries to determine/discover new features within the data [90]. These techniques often give insights into the available data and uncover the underlying structure found within the data. The difference between EDA techniques and machine learning techniques is that EDA techniques analyze the data and try to infer the adequate model. On the other hand, machine learning techniques first try to find the model and then analyze its parameters [90]. EDA techniques can be divided into two main groups: graphical and quantitative. Some of these techniques are discussed below [90]:

- Graphical:
  - Histograms: Histograms give a graphical summary of the distribution of a univariate dataset. This can be done to one feature of the collected data.
  - Probability Plot: Probability plots are another graphical technique that helps analysts assess whether the data follows a specific distribution or not. The data is plotted against a theoretical distribution. If the plot has a linear form, then this implies that theoretical distribution fits the data well. If the plot does not have a linear form, then the considered distribution is not representative of the data.
- Quantitative:
  - Confidence intervals for the mean: These intervals provide a range of where the true mean may lie given the dataset available. Since the data contains only a sample of the actual possible instances of a specific feature, then the calculated mean of the data is not the true mean of the feature. Hence, confidence intervals are used to indicate the lower and upper limits of the true mean.
  - Variance: Variance is a measure of how the data varies around the mean. It gives an idea of how wide or narrow the distribution is. This can help the analyst determine how sensitive the dataset is to a variation in a specific variable.

(ii) **Confirmatory Data Analytics (CDA):** Confirmatory Data Analytics (CDA) is a group of techniques that try to confirm whether a hypothesis is true or false. Such techniques assume that there is a specific hypothesis and perform tests that either confirm the validity or invalidity of the hypothesis. Two well known tests are discussed below [90].

- Analysis of Variance (ANOVA): ANOVA techniques are often used to compare data collected from two or more processes. The idea is to test whether the processes have equal variances. Bartlett's test performs ANOVA by testing for what is called homogeneity of variance. The hypothesis is assumed to be that all the variances are equal. Performing this analysis answers the question of whether the variances are truly equal or not with some specific confidence level. This type of tests is often used when the dataset seems to be normally distributed and is one of a set of statistical tests called "parametric tests" which also include t-tests [91], [92].
- Chi-Square Test for variance: This test is used to see if the variance of a dataset is equal to a specific value. This is often very important in manufacturing processes because manufacturers desire that their production process's variability is greater or smaller than some threshold.

These tests output a value that allows the analyst to determine whether the original hypothesis is valid or not.

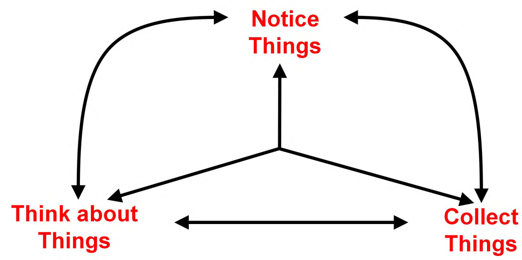

FIGURE 9. **Qualitative data analysis process.**

Based on this, he/she can take action and change some aspects of the considered process. This type of tests is often used when the dataset does not seem to be normally distributed, but rather follows other distributions. It is one of a set of statistical tests called "non-parametric tests" that include tests like Wilcoxon rank sum test and Spearman's rank correlation coefficient [91], [92].

(iii) **Qualitative Data Analytics (QDA):** Another type of data analytics is Qualitative Data Analytics (QDA). This type encompasses the set of processes and techniques that try to explain/interpret qualitative (non-numeric) data such as videos, interview transcripts, images, and documents [93]. These techniques are very prevalent in social sciences in which a lot of the data collected is non-numeric in nature. They are also prevalent in text analysis by searching for recurring words or topics.

There are some key terms that often pop up in QDA techniques. Words such as "Theme" that provides a general idea of the topics apparent in the data and "Characteristic" which represents a single item in a text often appear when performing QDA. The process of performing QDA involves having to iteratively and progressively notice, collect, and think about things [94]. This process is not linear since the cycle keeps on repeating and you do not go through it step by step, but rather do these steps simultaneously. This process is shown in Fig. 9.

To perform QDA, techniques like sorting and categorizing the collected data into groups based on theme can help simplify the analysis. After sorting the data, further analysis is performed by trying to interpret the significance of the sorting or by trying to explain the findings based on the sorted data.

These techniques have been used in different applications like text mining, marketing research, business, criminology, and sociology [95], [96].

It is worth mentioning that machine learning can be used as one method for data analysis. This is because in essence, machine learning can be thought of as a data analysis method that builds the analytical model automatically [97]. Hence, it can help identify and uncover patterns found within the data.

V. RESEARCH OPPORTUNITIES USING MACHINE LEARNING & DATA ANALYTICS FOR E-LEARNING CHALLENGES

As discussed before, one of the sources of data growth is the data being generated through online learning websites and learning management systems (LMSs) as part of e-Learning environments. Statistics show that online course websites such as Coursera, edX, and Udacity have more than 78 million students combined with around 10 thousand courses being offered by more than 600 universities [3]. Moreover, it has been estimated that there are almost one thousand event log entries per student every month and around 60,000 course visits every month for online courses [4]. This gives an idea of how big the data streams are expected to be with online courses specifically and the field of e-learning in general being a main contributor to the "Big Data" concept. Thus, a need to analyze and extract useful information from the collected data has risen in order to take more informed decisions and have more efficient systems. These systems can become more intelligent and responsive to users as they better cater to their needs. This can be done using ML and DA techniques such as quantitative data analysis to study students' performance and supervised classification to perform task such as sentiment analysis and students' style classification.

Based on the statistics above and the previously provided discussion on the challenges facing the field of e-learning, ML and DA techniques become promising tools to improve the e-learning platforms and processes. In what follows, a set of possible research opportunities to tackle each challenge discussed earlier is presented.

- **Transmission/Delivery - Research Opportunities:**
  Data analytics can play an important role in tackling this challenge. For example, a case study can be conducted that studies the performance of two groups of students. The first group would take the course in a synchronous manner and the second group would take the course in an asynchronous manner. The performance of both groups can be compared based on the collected data of both groups using both quantitative exploratory and qualitative data analytics. This can help instructors determine the optimal mode of delivery for the specific course based on metrics such as students' location distribution, their interaction level on forums, and their grades throughout the course. This is important because previous work have only compared the two possible delivery mechanisms in terms of the resources needed rather than their impact on student learning. Therefore, studying the merits of each delivery mode for students should be studied.

  Another possible area where qualitative data analytics can play a role is through text mining for sentiment analysis. This can be done to identify discouraged or disgruntled students based on their posts in the discussion forums. Moreover, supervised machine learning techniques such as SVM and ANN can be used to classify such posts. This will help instructors gain more insights into how the students are feeling about the material and the course. Based on this, instructor can try different ways to keep the students engaged and motivated. It can also help them make informed decisions on what to change and what to keep in the course in terms of material, assignments, topics, and activities.
- **Personalization - Research Opportunities:**
  As can be seen, several works have already started using machine learning and data analytics algorithms such as student clustering and content classification to tackle the challenge of personalizing the e-learning experience [23]–[25], [38], [39]. However, more opportunities are available. For instance, students can be classified based on any of the previously mentioned learning styles/preferences models. This classification can be used in several ways. One way is to recommend new courses to take based on the students' learning style/preference. This is important for online programs and courses where there is an abundance of courses available to be taken. Another possible recommendation is to recommend a course format for the student taking a course. Although this might be a demanding task for instructors, it can be offset by applying this to supplemental material rather than the main material of the course. This would provide the students with extra material that caters to their preference and can help them to better understand the course they are taking.

  A different opportunity is to predict students' final grade and classify them into different groups based on their performance in the first part of a course using techniques such as logistic regression and ANN. This can help instructors identify weak students that may need help in the course. This analysis can also pinpoint which concepts or learning outcomes the students seem to be struggling with. This can also help instructors adapt the content a bit or provide extra information about confusing concepts to ensure students better grasp it. Additionally, this can be used for peer student recommendation. The instructor might recommend that a weak student collaborate with a better performing student so that he/she can help him/her with the course.

  Another research opportunity is to determine relationships between certain activities and the performance of students in courses. For example, case studies that investigate the statistical significance of the number of logins to an e-learning platform on the students' performance at the end of a course portion should be performed to test the validity of any assumptions or hypotheses made beforehand. This would give instructors insights into what material was accessed more and whether it helped improve students' understanding of the material or not. This analysis can be done using confirmatory data analytic techniques. Accordingly, the instructor might adapt the content throughout the course to improve the

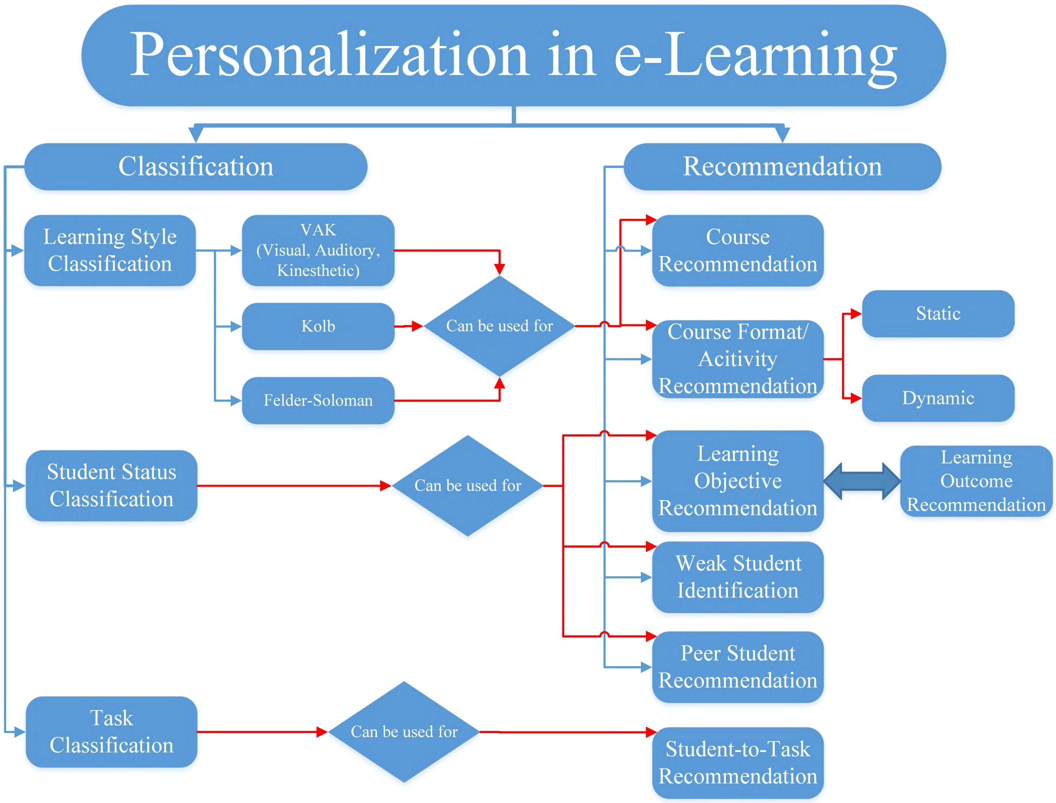

FIGURE 10. Personalization in e-learning.

students' e-learning experience. Fig. 10 gives an overview of some of the different classifications that can be done and recommendations that can be given within the field of e-learning.

- **Collaborative/Cooperative Learning Facilitation - Research Opportunities:**
One opportunity to use machine learning and data analytics algorithms is to analyze the level of collaboration between students on group projects based on their posts and engagement in group forums. The statistical significance of the engagement level on the students' performance in group projects can be studied to get better insights on how effective was the collaboration between members. This can be done using unsupervised clustering algorithms such as k-means or using qualitative data analytic algorithms.
Another opportunity to enhance and facilitate collaboration and cooperation between students is by assigning students to groups based on their learning style. Students with similar learning styles can cooperate better and share the information among them in a better way because they tend to have the same learning preferences. To that end, students' learning styles/preferences can be determined and groups can be formed based on this classification. The classification can be done using unsupervised learning techniques such as k-means algorithm.
- **Evaluation & Assessment - Research Opportunities:**
To tackle this challenge, several possible researches can be conducted that are based on machine learning and data analytics. One possible idea is to give a course in two different formats: one being in a traditional format with the course being given in a traditional classroom, while the second being given online using the e-learning tool/environment. The performance of the two groups of students can then be compared to see the efficiency of giving the course online using quantitative exploratory analytics algorithms such as means and variances. Also, the feedback from the group which take the course online is essential. The combination of the student performance and the feedback collected can give insights into the efficiency of the e-learning tool/environment.
Another possible research is to perform sentiment analysis to the feedback of students using an e-learning tool/environment. This can be done using qualitative data analytics. This analysis can be combined with a more quantitative analysis of the activity log of the students in a given course to get a comprehensive measure of how effective the e-learning tool/environment is and whether it is indeed helping to improve the learning experience of students.
- **Enabling Technologies - Research Opportunities:**
Machine learning and data analytics algorithms can be beneficial in tackling this challenge as well.

TABLE 2. Summary of e-learning challenges, previous works, & research opportunities.

Challenge	Reference	Used Techniques	Research Opportunity Using Machine Learning & Data Analytics
Transmission/Delivery	Ref. [20], Ref. [21], Ref. [22]	Qualitative Data Analytics Quantitative (Statistical) Data Analytics	a- Study & compare performance of students taking course synchronously and asynchronously using EDA. b- Sentiment analysis to identify discouraged students using EDA and supervised machine learning classification algorithms such as SVM and ANN.
Personalization	Ref. [23], Ref. [24], Ref. [25], Ref. [38], Ref. [39]	K-Means Decision Trees and Neural networks Quantitative (Statistical) Data Analytics	a- Classify students based on learning style and recommend courses/course format using classification algorithms such as logistic regression, SVM, and ANN. b- Classify students based on performance and recommend extra material/recommend peers for help using classification algorithms such as logistic regression, SVM, and ANN. c- Study statistical significance of activity to determine the relationship between activity level and performance using EDA.
Collaborative/Cooperative Learning Facilitation	Ref. [41], Ref. [42], Ref.[43], Ref. [44]	N/A	a- Study the collaboration level between students based on engagement level in group discussion forums. b- Recommend forming groups of students based on learning style/preference.
Evaluation & Assessment	Ref. [45], Ref. [46], Ref. [47]	Quantitative (Statistical) Data Analytics	a- Study & compare performance of students taking course in traditional classroom and those taking it online using EDA. b- Perform sentiment analysis to determine the level of satisfaction of students about the e-learning tool/environment using EDA and supervised machine learning classification algorithms.
Enabling Technologies	Ref. [49], Ref. [50], Ref. [51]	Exponential Smoothing	a- Perform load forecasting of e-learning traffic using supervised regression algorithms such as linear or polynomial regression. b- Modify content based on available resources.

For example, regression algorithms can be used to train the system for load forecasting. These algorithms can use labeled data to determine the model that best describes the traffic generation process. This can help the infrastructure providers better provision resources for the e-learning system as they can predict the time and size of the traffic load they will need to satisfy.

Another possible benefit of utilizing machine learning algorithms is by detecting and identifying high bandwidth applications. Based on usage data collected, infrastructure providers can make informed decisions on how to handle the data generated by the e-learning system. For example, the course format can be modified from synchronous to asynchronous to distribute the traffic load. The content can also be modified taking into consideration the bandwidth resources available. For instance, rather than having video be streamed, a different format of the content can be provided that is less bandwidth demanding in order to reduce the traffic load. These decisions can be made based on the patterns and models developed using machine learning and data analytics algorithms.

Table 2 summarizes the different challenges, previous work, and research opportunities within the field of e-learning that can benefit from using machine learning and data analytics algorithms.

VI. CONCLUSION

The growth of data available from various devices has led to the need to analyze and extract useful information from it. Machine learning and data analytics techniques have been proposed as means to satisfy this need. From supervised learning and unsupervised learning to exploratory and confirmatory data analytics, these algorithms are proving to be beneficial and essential in various applications and fields such as healthcare, business, and energy. Among these applications is the field of e-learning. E-learning is a field that has gather interest in recent times due to the proliferation of technology throughout the world. Despite its ability to provide access to information, e-learning is facing many challenges such as how to deliver content and how to personalize the e-learning experience. In this survey, we discussed the various machine learning and data analytics algorithms and glanced over some of the works in which they were used in different fields. Moreover, the field of e-learning was explored thoroughly by providing definitions and the characteristics describing it. Furthermore, some of the challenges faced within this field have been investigated. Also, some of the works that implemented machine learning and data analytics in e-learning were discussed. Finally, a few research opportunities have been proposed to give insights into the areas that need further exploration.

REFERENCES

[1] E-Learning Inc. (2016). *About E-Learning*. [Online]. Available: http://www.elearningnc.gov/about_elearning/

[2] J. Shailaja and R. Sridaran, "Taxonomy of E-learning challenges and an insight to blended learning," in *Proc. Int. Conf. Intell. Comput. Appl. (ICICA)*, Coimbatore, India, Mar. 2014, pp. 310–314.

[3] D. Shah. (2016). *By The Numbers: MOOCS in 2016*. [Online]. Available: https://www.class-central.com/report/mooc-stats-2016/

[4] S. Eichhorn and G. W. Matkin, "Massive open online courses, big data, and education research," *New Directions Inst. Res.*, vol. 2015, no. 167, pp. 27–40, 2016.

[5] M. J. Kearns and U. V. Vazirani, *An Introduction to Computational Learning Theory*. Cambridge, MA, USA: MIT Press, 1994.

[6] A. Cooper, "What is analytics? Definition and essential characteristics," *CETIS Anal. Ser.*, vol. 1, no. 5, pp. 1–10, 2012.

[7] A. Sangrà, D. Vlachopoulos, and N. Cabrera, "Building an inclusive definition of E-learning: An approach to the conceptual framework," *Int. Rev. Res. Open Distrib. Learn.*, vol. 13, no. 2, pp. 145–159, 2012. [Online]. Available: http://www.irrodl.org/index.php/irrodl/article/view/1161

[8] T. Ayodele, C. A. Shoniregun, and G. Akmayeva, "Towards E-learning security: A machine learning approach," in *Proc. Int. Conf. Inf. Soc. (i-Society)*, London, U.K., Jun. 2011, pp. 490–492.

[9] F. Mayadas, G. Miller, and J. Sener, "Definitions of E-learning courses and programs, developed for discussion within the online learning community," Online Learning Consortium, Newburyport, MA, USA, Tech. Rep. Version 2.0, Apr. 2015.

[10] B. Tucker, "The flipped classroom," *Educ. Next*, vol. 12, no. 1, pp. 1–6, 2012.

[11] S. Tobias, J. D. Fletcher, and A. P. Wind, "Game-based learning," in *Handbook of Research on Educational Communications and Technology*. Springer, 2014, pp. 485–503.

[12] J. S. Krajcik and P. C. Blumenfeld, *Project-Based Learning*. Cambridge, U.K.: Cambridge Univ. Press, 2006.

[13] Teach.Com. *Teaching Methods*. Accessed: Mar. 1, 2018. [Online]. Available: https://teach.com/what/teachers-know/teaching-methods/

[14] *Best Practices in Online Teaching Strategies*, Hanover Res. Council, Washington, DC, USA, Jul. 2009.

[15] A. Zafar and N. Ahmad, "Towards adaptive E-learning: Technological challenges and enabling technologies," in *Proc. Annu. IEEE India Conf.*, New Delhi, India, Sep. 2006, pp. 1–6.

[16] C. Xin, "E-learning applications and challenges," in *Proc. 2nd Int. Conf. Future Inf. Technol. Manage. Eng. (FITME)*, Sanya, China, Dec. 2009, pp. 580–583.

[17] A. Ravi, J. Ravi, H. Jani, M. S. Krishnan, and T. Patel, "Perceived challenge in E-learning and its causes," in *Proc. IEEE 7th Int. Conf. Technol. Edu. (T4E)*, Warangal, India, Dec. 2015, pp. 95–100.

[18] B. Kerr, "Challenges in global E-learning: What does the literature tell us," in *Int. Conf. Interactive Collaborative Learn. (ICL'14)*, Dubai, UAE, Dec. 2014, pp. 156–160.

[19] Witdono, "Utilization of e-learning at Universitas Siswa Bangsa International: Issues and challenges," in *Proc. IEEE 63rd Annu. Conf. Int. Council Educ. Media (ICEM)*, Oct. 2013, pp. 1–6, doi; 10.1109/CICEM.2013.6820205.

[20] F. Tabak and R. Rampal, "Synchronous E-learning: Reflections and design considerations," *Int. J. Educ. Develop. Unsing Inf. Commun. Technol.*, vol. 10, no. 4, pp. 80–92, 2014.

[21] A. I. Obasa, A. A. Eludire, and T. A. Ajao, "A comparative study of synchronous and asynchronous E-learning resources," *Int. J. Innov. Res. Sci., Eng. Technol.*, vol. 2, no. 11, pp. 5938–5946, 2013.

[22] A. Akram, C. Fu, Y. Tang, Y. Jiang, and X. Lin, "Exposing the hidden to the eyes: Analysis of SCHOLAT E-Learning data," in *Proc. IEEE 20th Int. Conf. Comput. Supported Cooper. Work Design (CSCWD)*, Nanchang, China, May 2016, pp. 693–698.

[23] M. A. Hogo, "Evaluation of e-learning systems based on fuzzy clustering models and statistical tools," *Expert Syst. Appl.*, vol. 37, no. 10, pp. 6891–6903, Oct. 2010.

[24] H. R. Joseph, "Promoting education: A state of the art machine learning framework for feedback and monitoring E-Learning impact," in *Proc. IEEE Global Humanitarian Technol. Conf.-South Asia Satell. (GHTC-SAS)*, Kerala, India, Sep. 2014, pp. 251–254.

[25] M. Virvou, E. Alepis, and S.-C. Sidiropoulos, "A learning analytics tool for supporting teacher decision," in *Proc. 6th Int. Conf. Inf., Intell., Syst. Appl. (IISA)*, Corfu, Greece, Jul. 2015, pp. 1–3.

[26] T. Fei, W. J. Heng, K. C. Toh, and T. Qi, "Question classification for e-learning by artificial neural network," in *Proc. Joint Conf. 4th Int. Conf. Inf., Commun. Signal Process., 4th Pacific Rim Conf. Multimedia*, Singapore, vol. 3, Dec. 2003, pp. 1757–1761.

[27] S. Manne, S. Yelisetti, M. Kakarla, and S. Fatima, "Mining VRSEC student learning behaviour in moodle system using datamining techniques," in *Proc. Int. Conf. Comput. Commun. Technol. (ICCCT)*, Allahabad, India, Dec. 2014, pp. 1–7.

[28] D. Clark. (2011). *Visual, Auditory, and Kinesthetic Learning Styles (VAK)*. [Online]. Available: http://www.nwlink.com/~donclark/hrd/styles/vakt.html

[29] S. McLeod. (2013). *Kolb-Learning Styles*. [Online]. Available: http://www.simplypsychology.org/learning-kolb.html

[30] R. Felder and B. Soloman. *Learning Styles and Strategies*. Accessed: Mar. 1, 2018. [Online]. Available: http://www4.ncsu.edu/unity/lockers/users/f/felder/public/ILSdir/styles.htm

[31] Southwestern Community College. *The VAK Learning System*. Accessed: Mar. 1, 2018. [Online]. Available: http://www.southwesterncc.edu/sites/default/files/VAK_Learning_Styles.pdf

[32] F. Coffield, D. Moseley, E. Hall, and K. Ecclestone. (2004). *Learning Styles and Pedagogy in Post-16 Learning: A Systematic and Critical Review*. [Online]. Available: http://www.lsda.org.uk/files/PDF/1543.pdf

[33] P. Melville and V. Sindhwani, *Recommender Systems*. New York, NY, USA: Springer, 2017.

[34] Z. Gao, Z. Li, and K. Niu, "Solutions for problems of existing E-commerce recommendation system," in *Proc. IEEE Int. Conf. Consum. Electron.-Taiwan (ICCE-TW)*, Taipei, Taiwan, May 2016, pp. 1–2.
[35] M. M.-U. Hussain, B. Avci, G. Trajcevski, and P. Scheuermann, "Incorporating weather updates for public transportation users of recommendation systems," in *Proc. 17th IEEE Int. Conf. Mobile Data Manage. (MDM)*, Porto, Portugal, vol. 1, Jun. 2016, pp. 333–336.
[36] K. Verbert *et al.*, "Context-aware recommender systems for learning: A survey and future challenges," *IEEE Trans. Learn. Technol.*, vol. 5, no. 4, pp. 318–335, Oct./Dec. 2012.
[37] M. Erdt, A. Fernández, and C. Rensing, "Evaluating recommender systems for technology enhanced learning: A quantitative survey," *IEEE Trans. Learn. Technol.*, vol. 8, no. 4, pp. 326–344, Oct./Dec. 2015.
[38] A. Klašnja-Milićević, B. Vesin, M. Ivanović, and Z. Budimac, "E-Learning personalization based on hybrid recommendation strategy and learning style identification," *Comput. Educ.*, vol. 56, no. 3, pp. 885–899, Apr. 2011.
[39] W. Liu, "Using data mining to dynamically build up just in time learner models," M.S. thesis, Dept. Comput. Sci., Univ. Saskatchewan, Saskatoon, SK, Canada, Dec. 2009.
[40] I. Magnisalis, S. Demetriadis, and A. Karakostas, "Adaptive and intelligent systems for collaborative learning support: A review of the field," *IEEE Trans. Learn. Technol.*, vol. 4, no. 1, pp. 5–20, Jan./Mar. 2011.
[41] G. Suciu, T.-L. Militaru, C.-G. Cernat, G. Todoran, and V.-A. Poenaru, "Platform for online collaboration and E-learning in open source Distributed Cloud systems," in *Proc. 54th Int. Symp. ELMAR*, Zadar, Croatia, Sep. 2012, pp. 337–340.
[42] S. Rajam, R. Cortez, A. Vazhenin, and S. Bhalla, "E-learning computational cloud (eLC2): Web services platform to enhance task collaboration," in *Proc. IEEE/WIC/ACM Int. Conf. Web Intell. Intell. Agent Technol. (WI-IAT)*, Toronto, Canada, vol. 3, Aug. 2010, pp. 350–355.
[43] K. G. Franceschi, R. M. Lee, and D. Hinds, "Engaging E-learning in virtual worlds: Supporting group collaboration," in *Proc. 41st Annu. Hawaii Int. Conf. Syst. Sci.*, Waikoloa, HI, USA, Jan. 2008, p. 7.
[44] D. Mahadevan, A. Kumar, and K. Bijlani, "Virtual group study rooms for MOOCS," in *Proc. Int. Conf. Adv. Comput., Commun. Inform. (ICACCI)*, Kerala, India, Aug. 2015, pp. 1711–1714.
[45] V. Chang, "An evaluation instrument for e-learning ecosystem," in *Proc. 6th IEEE Int. Conf. Ind. Inform.*, Daejeon, South Korea, Jul. 2008, pp. 1244–1249.
[46] Z. Balogh, M. Munk, and M. Turcani, "Assessment tools and techniques for e-learning evaluation: Usage analysis and survey sampling," in *Proc. Sci. Inf. Conf. (SAI)*, London, U.K., Oct. 2013, pp. 38–44.
[47] M. Tahereh, T. M. Maryam, M. Mahdiyeh, and K. Mahmood, "Multi dimensional framework for qualitative evaluation in e-learning," in *Proc. 4th Int. Conf. E-Learn. E-Teach. (ICELET)*, Shiraz, Iran, Feb. 2013, pp. 69–75.
[48] K. Salah, M. Hammoud, and S. Zeadally, "Teaching cybersecurity using the cloud," *IEEE Trans. Learn. Technol.*, vol. 8, no. 4, pp. 383–392, Oct. 2015.
[49] A. C. Caminero, S. Ros, R. Hernández, A. Robles-Gomez, and R. Pastor, "Load forecasting mechanism for E-learning infrastructures using exponential smoothing," in *Proc. 11th IEEE Int. Conf. Adv. Learn. Technol. (ICALT)*, Athens, GA, USA, Jul. 2011, pp. 364–365.
[50] B. Dong, Q. Zheng, J. Yang, H. Li, and M. Qiao, "An E-learning ecosystem based on cloud computing infrastructure," in *Proc. 9th IEEE Int. Conf. Adv. Learn. Technol. (ICALT)*, Riga, Latvia, Jul. 2009, pp. 125–127.
[51] P. Mondal, S. Misra, and I. S. Misra, "A low cost low bandwidth real-time virtual classroom system for distance learning," in *Proc. IEEE Global Humanitarian Technol. Conf., South Asia Satell. (GHTC-SAS)*, Kerala, India, Aug. 2013, pp. 74–79.
[52] P. Pietzuch, R. C. Fernandez, M. Weidlich, and A. Gal, "Grand challenge: Scalable stateful stream processing for smart grids," in *Proc. 8th ACM Int. Conf. Distrib. Event-Based Syst. (DEBS)*, Mumbai, India, 2014, pp. 1–6. [Online]. Available: http://lsds.doc.ic.ac.uk/pubs
[53] A. Mukherjee, A. Pal, and P. Misra, "Data analytics in ubiquitous sensor-based health information systems," in *Proc. 6th Int. Conf. Next Gener. Mobile Appl., Services Technol.*, Paris, France, Sep. 2012, pp. 193–198.
[54] A. Rodríguez-González, J. Torres-Niño, G. Hernández-Chan, E. Jiménez-Domingo, and J. M. Alvarez-Rodríguez, "Using agents to parallelize a medical reasoning system based on ontologies and description logics as an application case," *Expert Syst. Appl.*, vol. 39, no. 18, pp. 13085–13092, Dec. 2012.
[55] C. L. Mak and H. S. L. Fan, "Heavy flow-based incident detection algorithm using information from two adjacent detector stations," *J. Intell. Transp. Syst.*, vol. 10, no. 1, pp. 23–31, 2006.
[56] T. Liebig and C. Bockermann, "Predictive trip planning—Smart routing in smart cities," in *Proc. Workshop EDBT/ICDT Joint Conf.*, Athens, Greece, 2014, pp. 1–8.
[57] C. Donalek. Supervised and unsupervised learning. Caltech Center for Data-Driven Discovery. Accessed: Feb. 23, 2018. [Online]. Available: http://www.astro.caltech.edu/ donalek/Teaching.html
[58] X. Zhu. (2007). *Semi-Supervised Learning Tutorial*. [Online]. Available: http://pages.cs.wisc.edu/*textasciitilde*jerryzhu/pub/sslicml07.pdf
[59] R. Sutton and A. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 1998.
[60] A. Rostamizadeh and A. Talwalkar, *Foundations of Machine Learning*. Cambridge, MA, USA:MIT Press, 2012.
[61] W.-L. Chao and J. J. Ding, "Integrated machine learning algorithms for human age estimation," M.S. thesis, Dept. Elect. Eng., National Taiwan Univ., Taipei, Taiwan, Jun. 2011.
[62] S. Ahmad, N. M. Ramli, and H. Midi, "Outlier detection in logistic regression and its application in medical data analysis," in *Proc. IEEE Colloq. Humanities, Sci. Eng. (CHUSER)*, Kota Kinabalu, Malaysia, Dec. 2012, pp. 503–507.
[63] C. Zhang *et al.*, "A visual analytics approach to high-dimensional logistic regression modeling and its application to an environmental health study," in *Proc. IEEE Pacific Vis. Symp. (PacificVis)*, Kobe, Japan, Apr. 2016, pp. 136–143.
[64] S. Yadav and S. Selvakumar, "Detection of application layer DDoS attack by modeling user behavior using logistic regression," in *Proc. 4th Int. Conf. Rel., Infocom Technol. Optim. (ICRITO)*, Noida, India, Sep. 2015, pp. 1–6.
[65] D. Han, L. Ma, and C. Yu, "Financial prediction: Application of logistic regression with factor analysis," in *Proc. 4th Int. Conf. Wireless Commun., Netw. Mobile Comput. (WiCom)*, Dalian, China, Oct. 2008, pp. 1–4.
[66] D. W. Hosmer, Jr., S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, vol. 398. Hoboken, NJ, USA: Wiley, 2013.
[67] A. Andrew, "An introduction to support vector machines and other kernel-based learning methods by Nello Christianini and John Shawe-Taylor, Cambridge University Press, Cambridge, 2000, xiii+189 pp., ISBN 0-521-78019-5 (Hbk, $27.50)," *Robotica*, vol. 18, no. 6, pp. 687–689, 2000.
[68] R. J. Schalkoff, *Artificial Neural Networks*, vol. 1. New York, NY, USA: McGraw-Hill, 1997.
[69] J. Cannell. (2015). *The Brain as a Universal Learning Machine*. [Online]. Available: http://lesswrong.com/lw/md2/the_brain_as_a_universal_learning_machine/
[70] T. Mitchell, *Decision Tree Learning*. New York, NY, USA: McGraw-Hill, 1997.
[71] J. R. Quinlan, *C4.5: Programs for Machine Learning*. San Francisco, CA, USA: Morgan Kaufmann, 1993.
[72] A. Ng, "Deep learning, self-taught learning and unsupervised feature learning," IPAM Graduate Summer School, UCLA, Los Angeles, CA, USA, 2012.
[73] Y. Bengio, "Deep learning of representations for unsupervised and transfer learning," in *Proc. ICML Workshop Unsupervised Transf. Learn.*, Edinburgh, Scotland, 2012, pp. 17–36.
[74] J. Brownlee. (2016). *What is Deep Learning?* [Online]. Available: http://machinelearningmastery.com/what-is-deep-learning/
[75] L. Lab and U. O. Montreal. (2015). *Deep Learning Tutorial: Release 0.1*. [Online]. Available: http://deeplearning.net/tutorial/deeplearning.pdf
[76] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org
[77] LeNet. (2016). *Convolutional Neural Networks (LeNet)—DeepLearning 0.1 Documentation*. [Online]. Available: http://deeplearning.net/tutorial/lenet.html
[78] C. Goller and A. Kuchler, "Learning task-dependent distributed representations by backpropagation through structure," in *Proc. IEEE Int. Conf. Neural Netw.*, Washington, DC, USA, vol. 1, Jun. 1996, pp. 347–352.
[79] R. Socher, C. C.-Y. Lin, A. Ng, and C. Manning, "Parsing natural scenes and natural language with recursive neural networks," in *Proc. 26th Int. Conf. Mach. Learn. (ICML)*, Washington, DC, USA, 2011, pp. 129–136.
[80] Z. Ghahramani. (Sep. 2004). *Unsupervised Learning*. [Online]. Available: http://mlg.eng.cam.ac.uk/zoubin/papers/ul.pdf
[81] P. Rai. (2011). *Semi-Supervised Learning*. [Online]. Available: https://www.cs.utah.edu/~piyush/teaching/8-11-slides.pdf
[82] O. Chapelle, B. Scholkopf, and A. Zien, *Semi-Supervised Learning*. Cambridge, MA, USA: MIT Press, 2006.
[83] K. Murphy. (1998). *A Brief Introduction to Reinforcement Learning*. [Online]. Available: http://www.cs.ubc.ca/~murphyk/Bayes/pomdp.html

[84] D. Silver. (2015). *Introduction to Reinforcement Learning*. [Online]. Available: http://www0.cs.ucl.ac.uk/staff/d.silver/web/Teaching_files/intro_RL.pdf
[85] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *J. Artif. Intell. Res.*, vol. 4, pp. 237–285, 1996.
[86] C. Szepesvari. (2009). *Algorithms for Reinforcement Learning*. [Online]. Available: https://sites.ualberta.ca/~szepesva/papers/RLAlgsInMDPs.pdf
[87] N. Mohamed and J. Al-Jaroodi, "Real-time big data analytics: Applications and challenges," in *Proc. Int. Conf. High Perform. Comput. Simulation (HPCS)*, Bologna, Italy, Jul. 2014, pp. 305–310.
[88] S. H. Koppad and A. Kumar, "Application of big data analytics in healthcare system to predict COPD," in *Proc. Int. Conf. Circuit, Power Comput. Technol. (ICCPCT)*, Nagercoil, India, Mar. 2016, pp. 1–5.
[89] R. P. Biuk-Aghai, W. T. Kou, and S. Fong, "Big data analytics for transportation: Problems and prospects for its application in China," in *Proc. IEEE Region 10 Symp. (TENSYMP)*, Bali, Indonesia, May 2016, pp. 173–178.
[90] NIST/SEMATECH. (2013). *E-Handbook of Statistical Methods*. [Online]. Available: http://www.itl.nist.gov/div898/handbook/
[91] T. Hoskin, "Parametric and nonparametric: Demystifying the terms," in *Proc. Mayo Clinic*, 2012, pp. 1–5.
[92] T. Neideen and K. Brasel, "Understanding statistical tests," *J. Surgical Educ.*, vol. 64, no. 2, pp. 93–96, 2007.
[93] M. B. Miles, A. M. Huberman, M. A. Huberman, and M. Huberman, *Qualitative Data Analysis: An Expanded Sourcebook*. Newbury Park, CA, USA: SAGE, 1994.
[94] J. Seidel, "Qualitative data analysis," in *Originally Published as Qualitative Data Analysis, in the Ethnograph v5.0: A Users Guide, Appendix E*, Colorado Springs, CO, USA: Qualis Research, 1998.
[95] M. U. Çakir and S. Güldamlasioğlu, "Text mining analysis in Turkish language using big data tools," in *Proc. IEEE 40th Annu. Comput. Softw. Appl. Conf. (COMPSAC)*, Atlanta, GA, USA, vol. 1, Jun. 2016, pp. 614–618.
[96] N. Denzin and Y. Lincoln, *The Sage Handbook of Qualitative Research*, 3rd ed. Newbury Park, CA, USA: SAGE, 2005.
[97] *Machine Learning: What it is and Why it Matters*, SAS, Cary, NC, USA, 2017.

ABDALLAH MOUBAYED received the B.E. degree in electrical engineering from Lebanese American University, Beirut, Lebanon, in 2012, and the M.Sc. degree in electrical engineering from the King Abdullah University of Science and Technology, Thuwal, Saudi Arabia, in 2014. He is currently pursuing the Ph.D. degree in electrical and computer engineering with Western University, London, ON, Canada. His research interests include wireless communication, resource allocation, wireless network virtualization, performance and optimization modeling, machine learning and data analytics, computer network security, cloud computing, and e-learning.

MOHAMMADNOOR INJADAT received the B.Sc. degree in computer science from Al Al-Bayt University, Jordan, in 2000, and the M.Sc. degree in computer science from the University Putra Malaysia, Malaysia, in 2002, and the M.E. degree in electrical and computer engineering from Western University in 2015, where he is currently pursuing the Ph.D. degree in software engineering with the Department of Electrical and Computer Engineering. His research interests include data mining, machine learning, social network analysis, data analytics, and cloud computing.

ALI BOU NASSIF received the Ph.D. degree in electrical and computer engineering from Western University, London, ON, Canada, in 2012. He is currently an Assistant Professor with the University of Sharjah, United Arab Emirates. He has published several papers in international journals. His interests are machine learning and soft computing, software engineering, cloud computing and service oriented architecture, and mobile computing. He is a member of the IEEE Computer Society and the Association for Computing Machinery. He is a registered professional engineer in Ontario.

HANAN LUTFIYYA received the B.Sc. degree from Yarmouk University, Jordan, the M.Sc. degree from the University of Iowa, Ames, and the Ph.D. degree from the University of Missouri, Rolla, in 1992, all in computer science. She has been with the Computer Science Department, Western University, London, ON, Canada, since 1992, where she is currently a Professor and the Chair of the Computer Science Department. Her research interests include distributed systems management and software engineering.

ABDALLAH SHAMI received the B.E. degree in electrical and computer engineering from Lebanese University in 1997 and the Ph.D. degree in electrical engineering from the Graduate School and University Center, City University of New York, in 2002. In 2002, he joined the Department of Electrical Engineering, Lakehead University, Thunder Bay, ON, Canada, as an Assistant Professor. Since 2004, he has been with Western University, where he is currently a Professor with the Department of Electrical and Computer Engineering. His current research interests are in the areas of network optimization, cloud computing, and wireless networks.