# Utilizing crowdsourcing and machine learning in education: Literature review

Hadeel S. Alenezi[1] · Maha H. Faisal[1]

## Abstract

For many years, learning continues to be a vital developing field since it is the key measure of the world's civilization and evolution with its enormous effect on both individuals and societies. Enhancing existing learning activities in general will have a significant impact on literacy rates around the world. One of the crucial activities in education is the assessment method because it is the primary way used to evaluate the student during their studies. The main purpose of this review is to examine the existing learning and e-learning approaches that use either crowdsourcing, machine learning, or both crowdsourcing and machine learning in their proposed solutions. This review will also investigate the addressed applications to identify the existing researches related to the assessment. Identifying all existing applications will assist in finding the unexplored gaps and limitations. This study presents a systematic literature review investigating 30 papers from the following databases: IEEE and ACM Digital Library. After performing the analysis, we found that crowdsourcing is utilized in 47.8% of the investigated learning activities, while each of the machine learning and the hybrid solutions are utilized in 26% of the investigated learning activities. Furthermore, all the existing approaches regarding the exam assessment problem that are using machine learning or crowdsourcing were identified. Some of the existing assessment systems are using the crowdsourcing approach and other systems are using the machine learning, however, none of the approaches provide a hybrid assessment system that uses both crowdsourcing and machine learning. Finally, it is found that using either crowdsourcing or machine learning in the online courses will enhance the interactions between the students. It is concluded that the current learning activities need to be enhanced since it is directly affecting the student's performance. Moreover, merging both the machine learning to the crowd wisdom will increase the accuracy and the efficiency of education.

**Keywords** Machine learning · Crowdsourcing · Education · E-learning

✉ Hadeel S. Alenezi
hadeel.alenezi@grad.ku.edu.kw

Maha H. Faisal
maha.faisal@ku.edu.kw

[1] Department of Computer Engineering, Kuwait University, Kuwait City, Kuwait

Check for updates

🖄 Springer

## 1 Introduction

Education field is a vital field of interest that is keep enhancing over the years. Starting from learning and moving toward the e-learning, many methods and techniques were presented to produce higher quality experience that will enrich the whole educational sector. The road toward the improvement can utilize extra hands which will provide different perspective and changes and that can be done using the crowd.

Crowdsourcing is a way of using crowd wisdom to achieve a common goal. It can be applied in any field starting from the idea generation until performing real, small or significant tasks. Using the crowd is useful in terms of cost, time, and accuracy of the generated data. Utilizing the crowd can be inexpensive compared to hiring a professional (Schenk 2009). Moreover, the time used to provide feedback can be faster than the traditional way because in crowdsourcing the task is divided into micro-tasks where each person performs his/her small task and all the collected results are merged later. It is the seamless way to bring together a team of instructors, researchers, teachers, professionals and students wanting to contribute their knowledge and experience (Nkoana 2016). This contribution generated by the crowd can outperform groups of intelligent and expert people which were proven by real-world comparison between ideas generated from professionals and crowd users (Poetz and Schreier 2012).

The human factor in the crowd could provide imperfect results due to human nature. However, merging the machine with the human may provide more beneficial results (Kamar 2016; Vaughan 2018). Utilizing both crowdsourcing and machine learning in education will be a positive factor to the improvement process. As defined in the Oxford dictionary, "Machine Learning is the capacity of a computer to learn from experience" ("Oxford Dictionary," 2020). There are several machine learning models, but the most common two are supervised and unsupervised learning. In supervised learning, the data set is given to help the model learn while in unsupervised learning, the model will follow some procedures and learn by itself.

One of the fundamental practices in the education sector is the student/learner assessment. Educational assessment is the process of evaluating the learner's progress throughout the course to identify his/her level of understanding and improve his/her learning. This process affects students' motivations, developments, and learning strategies which described by Crooks (1988) as "one of the most potent forces influencing education." Likewise, Harlen et al. (2002) stated that "feedback on assessments has an important role in determining further learning" since it directly influences students' performance and their constant effort in future tasks.

In this work, we investigate the trends and identify limitations of utilizing crowdsourcing and machine learning in e-learning systems (Section 3) based on the research questions, the search process and the inclusion criteria we specified in Section 2. The results of the investigation revealed potential opportunities to improve the current e-Learning systems (Section 4).

## 2 Methodology

This paper will investigate and identify studies that discuss the current learning and e-learning approaches which use either crowdsourcing, machine learning, or

both crowdsourcing and machine learning in their proposed solutions. Furthermore, another purpose of this literature is to examine the assessment problem in the educational field in particular. Our goal is to find the unexplored gaps and the limitations of the existing assessment solutions for the learning process. In this study we investigate the following issues:

i.  How are learning and e-learning sectors utilizing machine learning and crowdsourcing approaches?
ii.  What are the existing approaches that address the exam assessment problem using machine learning or crowdsourcing?
iii.  Does utilizing crowdsourcing and machine learning enhance the interactions in online courses?

This review includes articles published in the period from 2013 to 2019. The search was conducted through the ACM digital library and IEEE Xplore using the keywords shown in Table 1. This table also provides an overview of the number of articles found and included per database. Figure 1 demonstrates the inclusion criteria and the article selection process.

## 3 Results

Based on our selection process and inclusion criteria, the obtained results can be separated into three different categories where the main approach of the articles can be either crowdsourcing approach, machine learning approach, or a hybrid of crowdsourcing and machine learning. To investigate and examine how those approaches are utilized in the education sector, the common learning activities were identified and summarized in Table 2. The most recurrent field of interest in education was enhancing the way exam questions and class material generated where many papers discuss various ways of creating a bank that collects efficient material and questions. Another typical learning activity was enhancing reviews

**Table 1** Online database search keywords and number of results

| Library Name | Search Keywords | Number of results | Total Results | Selected Papers |
|---|---|---|---|---|
| ACM | Crowdsourcing, machine, learning | 284 | 494 | 23 |
|  | Crowdsourcing, machine, learning, grading | 5 |  |  |
|  | Crowdsourcing, education | 184 |  |  |
|  | Crowdsourcing, machine, learning, education | 21 |  |  |
| IEEE | Crowdsourcing, machine learning | 200 | 313 | 7 |
|  | Crowdsourcing, machine learning, grading | 1 |  |  |
|  | Crowdsourcing, education | 108 |  |  |
|  | Crowdsourcing, machine learning, education | 4 |  |  |

Literature Search
    Database: IEEE Xplore, ACM.
    Included:
            Dates: ≥ 2013 – 2019 (Present)
            Language: English Only.
            Approach Type: Machine Learning,
            Crowdsourcing.
            Field of Problem: Learning and e-learning

Records identified through 2 databases'
searches after removing duplicates

(n = 807)

Records excluded due to
dates, language, approach
type
(n = 772)

Records screened
(n = 35)

Full-text articles excluded
because of discussing
problems for non-
educational field
(n = 5)

Full-text articles assessed
for eligibility
(n = 30)

Studies included in
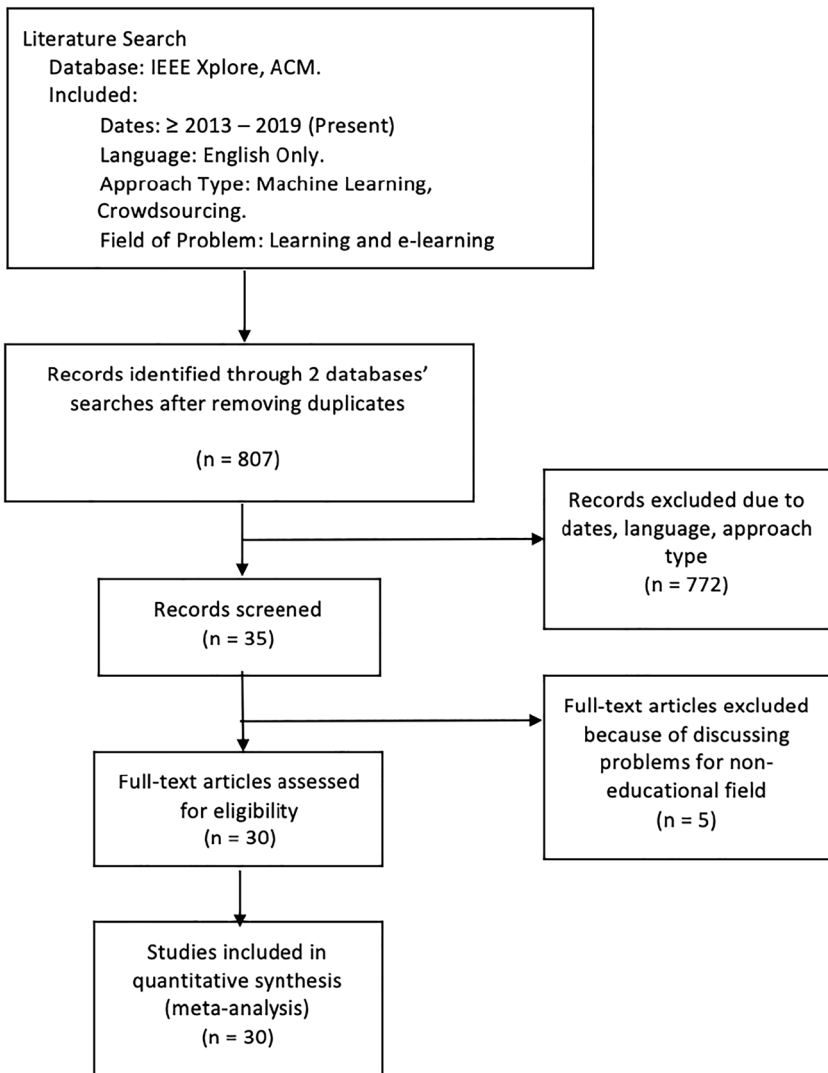quantitative synthesis
(meta-analysis)
(n = 30)

**Fig. 1** The article selection process and inclusion criteria

and explanations in online courses. The last popular activity was the assessment process (i.e., grading) of homework and exams.

The comparison based on the publication timeline will be useful for all the researchers to keep up with the most recent techniques and approaches in the educational sector. Moreover, the researchers can know what the fields that need more enhancement or focusing because they can see the evolution of these different solutions. Figure 2 and Table 3 show the publication timeline, and it is noticeable that most of the papers were published heavily after 2013 which explains the increase of interest in improving the learning and e-learning fields.

**Table 2** Results based on the learning activities

| Learning Activity | Crowdsourcing | Machine Learning | Hybrid |
|---|---|---|---|
| Research | (Vaish et al. 2017) | N. A | N. A |
| Students' Interactions in massive open online course (MOOC) | (Kulkarni et al. 2015a, b) | N. A | (Kamath et al. 2016) |
| Personalized Learning | (Faisal et al. 2015; Whitehill and Seltzer 2016) | (Li 2018) | N. A |
| Homework and Exams Assessment | (de Alfaro and Shavlovsky 2014; Pirttinen et al. 2018) | (Gutierrez et al. 2013; Kashi et al. 2016) | N. A |
| Exams and Material Generation | (Alghamdi et al. 2015; Farasat et al. 2017; Pirttinen et al. 2018; Taniguchi and Inoue 2015) | (Tang et al. 2016) | (Luger and Bowles 2013) |
| Cheating Detection | N. A | N. A | (Li et al. 2015) |
| Explanation/Reviews Enhancement | N. A | (Le et al. 2016; Watson et al. 2018) | (Huang et al. 2013; Krause et al. 2017; Williams et al. 2016) |

### 3.1 Crowdsourcing approach

Utilizing the crowd toward enhancing education is discussed heavily in the reviewed literature. Since crowdsourcing will help in reducing time and cost to achieve any goal, the education sector has adopted many approaches in this regard. Various learning activities are using the crowdsourcing techniques mentioned in Table 2, and more details are provided in this section.

### 3.1.1 Research

One of the essential aspects of education is the research filed. The lack of resources to write professional research is avoidable by creating a crowd of researchers who can share their resources and create a shared research. The proposed platform can help the students where each student can share his/her
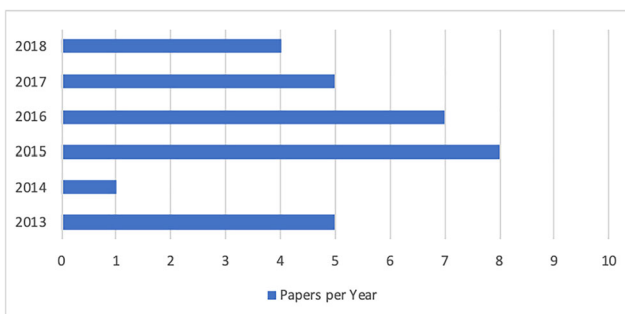


**Fig. 2** Paper publication timeline

**Table 3**  Timeline of paper publications

| Year of publication | Paper Reference | Total papers per year |
|---|---|---|
| 2013 | (Barbosa et al. 2013; Dow et al. 2013; Gutierrez et al. 2013; Huang et al. 2013; Luger and Bowles 2013) | 5 |
| 2014 | (de Alfaro and Shavlovsky 2014) | 1 |
| 2015 | (Alghamdi et al. 2015; Corrigan-Gibbs et al. 2015; Faisal et al. 2015; Kotturi et al. 2015; Kulkarni et al. 2015a, b; Li et al. 2015; Taniguchi and Inoue 2015) | 8 |
| 2016 | (Kamath et al. 2016; Kashi et al. 2016; Le et al. 2016; Nkoana 2016; Tang et al. 2016; Whitehill and Seltzer 2016; Williams et al. 2016) | 7 |
| 2017 | (Choi et al. 2017; Farasat et al. 2017; Kim et al. 2017; Krause et al. 2017; Vaish et al. 2017) | 5 |
| 2018 | (Li 2018; Pirttinen et al. 2018; Shongwe and Zuva 2018; Watson et al. 2018) | 4 |

ideas and contribute in being part of a significant research group that consists of many researchers each one of them performs a specific task (Vaish et al. 2017). Enriching the research field could move the education to a new era by gathering all the knowledge from different expertise which will elevate the education tools to a new level. Furthermore, it can help in providing the best learning experience since it will investigate all the imperfections in different areas and provide solutions to enhance them. However, utilizing the research crowd to perform joint research could lead to problems if certain aspects were neglected. Since the researchers (the crowd) are diverse group who are coming from different backgrounds, the gathered research data should be carefully reviewed to avoid any conflicts affecting the provided opinions or the focus of the research. Consequently, the quantity of the research papers may increase at the expense of their quality if the crowd group were large and diverse.

### 3.1.2 Students' interactions in MOOC

Another important topic in education is the minimal interactions between students in online courses. One of the proposed platforms is found to solve this issue by providing a peer discussion system that connects students from different countries to share their knowledge based on their regions and countries (Kulkarni et al. 2015a). Another platform was found to enhance the interaction in online courses by enabling the crowdsourcing technique in providing rapid feedback on online student's submission (Kulkarni et al. 2015b). This platform encourages students to create a draft for their assignment and seek feedback from the crowd for as many drafts as they need. All the proposed platforms could enhance the learning experience especially for students with lack of learning opportunities who are fully depending on online courses for their education. Therefore, focusing on students' interaction is essential to ensure the success of the learning process. Consequently, since the suggested solutions are fully dependent on non-expertise individuals, the process must be supervised to avoid delivering wrong information or giving imprecise feedback.

### 3.1.3 Personalized learning

Achieving a personalized learning goal is very crucial in the learning process. Whitehill and Seltzer (2016) suggested using crowdsourcing for online learning videos to achieve this goal where the crowdsourced videos are created and recorded by ordinary people not professional teachers. Furthermore, another platform used a framework which was built to help the learners choose the best learning plan based on a recommender system where the requested learning activity will be different based on each learner characteristics (Faisal et al. 2015). Crowdsourcing methods which are applied to enhance the personalized learning will open the door to many opportunities that will personalize the whole learning experience, including the material content, the way of delivering the material, and the examinations. Nevertheless, multiple challenges will arise if a remarkable rewarding system was not found for the crowd to motivate them achieve the highest quality standard for personalizing the learning.

### 3.1.4 Homework and exams assessment and material generation

One of the critical aspects to be discussed in education is crowdsourcing and enhancing assessment tools. An attempt to generate high-quality exam questions have been introduced (Alghamdi et al. 2015). This platform can involve as many instructors as possible in the process of creation and evaluation question items to develop large-scale exams in a university context. One of the suggested systems by de Alfaro and Shavlovsky (2014) let students submit and collaboratively review and grade homework. In this system, the students receive an overall crowd-grade that reflects both the quality of their homework, and the quality of their work as reviewers. In another system, users can create/edit exam questions and answer them where it will also check for the difficulty and the validity of the created questions based on a specific algorithm (Taniguchi and Inoue 2015). In the same vein, Pirttinen et al. (2018) introduced an online courses' embeddable tool that allows a student to create programming assignments where the students can also review each other's work and provide their evaluation. Farasat et al. (2017) introduced the concept of the crowd-Learning where students can have deeper learning by creating their learning materials. The proposed solution is an online platform that can be used as in-class practice or for online courses. Between all the previous methods, the assessment process specifically can become very delicate since it is the main way of evaluating any student. Total dependency on the crowd for the evaluation may lead to giving unfair or inaccurate grades. Researchers must focus on having a clear reward or punishment system that force the crowd to provide careful assessment.

All the previous studies have discussed and proposed ways of using crowdsourcing in e-learning focusing on three different aspects; online learning, instructors, and students. The reviewed literature can be categorized based on different criteria such as the crowd platform, the crowd type, the crowd size, and the crowd reward. These classifications can highlight the commonly used attributes for crowdsourcing researches in the education field. The Crowd platform is the tool used to obtain the crowd whether it is a stand-alone platform, workers from amazon mechanical Turk, contributors from MOOC platforms, or a real physical crowd (e.g., students in a classroom). On the other hand, the crowd type

can be either student, instructors, researchers, or online workers. The third criterion is the size of the crowd where it is estimated to be a small crowd if between (0–50) workers, a medium crowd if more than 50 and less than 100, and finally, a large crowd if they exceed 100 workers. The last criterion is the crowd reward whether it is a physical reward (e.g., a certain amount of money) or it is a logical reward (e.g., homework grade). All the previous categories were summarized in Table 4 and from the table, we can conclude that most of the research papers in our scope were focused on having a physical crowd where the crowd usually a small group of students. Therefore, this finding reflects the massive necessity for enhancing the educational tools in order to best satisfy the learners' needs.

### 3.2 Machine learning approach

Machine learning can be integrated with any field to enhance all its related tasks by providing real-time responses that reduce wasted time and eliminate manual interventions in a very efficient way. In the education sector specifically, diverse approaches are utilizing the different machine learning models like supervised or Natural Language Processing (NLP) to achieve their goal which are summarized in Table 5 and will be discussed in this section.

**Table 4**  Literature taxonomy

| Category | | Papers |
|---|---|---|
| Crowd Platform | Standalone Platform | (Alghamdi et al. 2015; de Alfaro and Shavlovsky 2014; Kulkarni et al. 2015b; Vaish et al. 2017) |
| | Amazon Mechanical Turk | (Dow et al. 2013; Kim et al. 2017; Whitehill and Seltzer 2016) |
| | MOOC platform | (Kotturi et al. 2015) |
| | Physical Crowd | (Faisal et al. 2015; Farasat et al. 2017; Nkoana 2016; Pirttinen et al. 2018; Taniguchi and Inoue 2015) |
| Crowd Type | Researchers / Authors | (Kim et al. 2017; Vaish et al. 2017) |
| | Students / Learners | (de Alfaro and Shavlovsky 2014; Faisal et al. 2015; Farasat et al. 2017; Kashi et al. 2016; Kotturi et al. 2015; Kulkarni et al. 2015a, b; Nkoana 2016; Pirttinen et al. 2018; Taniguchi and Inoue 2015) |
| | Instructors | (Alghamdi et al. 2015; Nkoana 2016) |
| | Online Workers | (Dow et al. 2013; Nkoana 2016; Whitehill and Seltzer 2016) |
| Crowd Size | Small | (de Alfaro and Shavlovsky 2014; Dow et al. 2013; Kulkarni et al. 2015a, b) |
| | Medium | (Vaish et al. 2017; Whitehill and Seltzer 2016) |
| | Large | (Kim et al. 2017; Pirttinen et al. 2018) |
| Crowd Reward | Physical | (Dow et al. 2013; Farasat et al. 2017; Kim et al. 2017; Whitehill and Seltzer 2016) |
| | Logical | (Kulkarni et al. 2015a, b; Vaish et al. 2017) |

**Table 5** Literature based on machine learning models

| Machine Learning Model | Papers |
| --- | --- |
| Supervised | (Le et al. 2016; Li 2018) |
| NLP | (Gutierrez et al. 2013; Kashi et al. 2016; Watson et al. 2018) |
| Other | (Choi et al. 2017; Tang et al. 2016) |

### 3.2.1 Supervised model

Many approaches utilized the supervised machine learning model for their proposed solution. One of the frameworks is proposed to automatically assess the quality of the community question answering where it uses a classification algorithm to assess the quality of any answer by following a group of different features (Le et al. 2016). The classification algorithm is one of the supervised learning approaches where the computer learns from input data and then produce classifications or categories from those data as the output. The proposed approach starts with feature extraction by collecting historical information about the community member which will help in building the classification model. Ultimately, the quality of all the new answers will be measured based on the trained models (Le et al. 2016). The proposed approach by Li (2018) used the supervised model also by following the regression algorithm where the output variable must be continues and real value. The author wanted to deeply understand and analyze the reasons behind each student's scores by applying the model to recognize the major causes affecting the student's learning performance. As a result, the approach will help the instructors track the learning effect and adjust their teaching strategy to suit the students' needs (Li 2018). The supervised model depends entirely on the learning process and past experiences. The past experience in a computer language is the provided set of data. Training the wrong set of data may provide inaccurate results which will affect the intended usage of the proposed approach. Hence, the challenge is choosing the right data with the right model for any new educational research because it is essential for its success.

### 3.2.2 Natural language processing (NLP) model

"NLP is a field in machine learning with the ability of a computer to understand, analyze, manipulate, and potentially generate human language" (Shetty 2018). One of the platforms that uses NLP in the reviewed literature is a web-based tool that visualizes the output from a machine-learning based model trained to guess Multiple Choice Questions (MCQs) where the primary goal is to identify and manipulate questions within a well-validated and high-quality bank of MCQs (Watson et al. 2018). Additionally, NLP is used in the exam assessment process where it compares the similarity between students' answers and the correct answer given that the ideal answers are available. As a result of this comparison, the proposed algorithm will calculate the student's recommended grade (Kashi et al. 2016). In addition to that, NLP is used to identify the correct and incorrect

statements in an assignment by combining extraction rules and machine learning based information extractors (Gutierrez et al. 2013). Natural language processing has created a wide set of opportunities that can be used to promote education and elevate education tools to give the best possible experience. However, utilizing NLP can be very tricky because interpreting the human's language to data that can be understood and analyzed by the computer is a complex procedure. The researchers who are considering this model must take into consideration the semantic, the syntactic, and the lexical meaning of any text to provide a correct analysis for the data and obtain proper results.

### 3.2.3 Other models

Different applications used other models of machine learning to reach the same goal of enhancing the education field. For example, creating high-quality test sets can be very crucial in any Automated Programming Assessment (APA) System, and it usually requires human experts. As a solution, a tool is created to generate test sets for APA systems where it uses the old submitted student's solution in order to generate test sets with better test coverage than the one created by experts (Tang et al. 2016). Another application of using a machine learning model is helping authors to create their books. The proposed adaptive e-book can be dynamically updating itself based on readers' responses where a feedback engine collects and aggregates their feedback to update the content (Choi et al. 2017).

### 3.3 Hybrid approach

Using crowdsourcing or machine learning approaches solely can be of great advantage to any educational field's dilemma. However, sometimes the lack of crowds' experience or the shortage of resources to build and train the machine learning model can reduce the quality of the work and provide undesirable and inaccurate results. Thus, some of the researchers provided their approaches using a hybrid of both crowdsourcing and machine learning to enhance and improve the quality of the solutions.

Some of the methods used to detect cheating in online exams without having many human resources included. For example, a platform was introduced to detect cheating using a hybrid approach by following three main steps; an automatic cheating detector, a peer cheating detector, and the final review committee (Li et al. 2015). Furthermore, some approaches use both machine learning and crowdsourcing to make the user reviews simpler and allow the reader to digest them (Huang et al. 2013). Moreover, a semi-automated way was presented to support feedback providers by analyzing feedback language using critique style guide (Krause et al. 2017). In the first part of the study, the crowd will rate the helpfulness of any feedback. From the generated data, machine learning is used to generate the critique style guide. Another automatic algorithm was proposed to measure the difficulty and power of MCQs exams and provide a way to build exams from crowdsourced data where the training data will be gathered from existing web resources to develop acceptable questions (Luger and Bowles 2013).

Moreover, some papers addressed the lack of engagement in the e-learning environment, unlike the classroom where the instructor can observe the

engagement level of the students. One platform proposes a system to auto-recognize students' engagement levels by observing their behavioral states (Kamath et al. 2016). Another concern in the e-learning environment is the lack of explanations in online problems where it costs a lot of money and time for MOOC to create a proper explanation for every problem. AXIS is a semi-automated system to create explanations for online problems where it engages the learners to generate and evaluate explanations and uses machine learning to analyze these data to identify the best explanation for future learners (Williams et al. 2016).

Merging multiple powerful tools is very common in the research area since some tools can help in covering the shortages of the other tool and provide enhanced results. However, sometimes using the hybrid approach for certain applications introduces new risks instead of being an advantage. For example, using crowdsourcing can compromise the confidentiality or it can increase the cost since the crowd must be rewarded. Likewise, using machine learning with the wrong trained data can provide inaccurate results. Hence, researchers must balance the pros and cons that each tool will provide to their study and assure that using the hybrid model is the best approach for their application.

## 4 Discussion

This study provides a systematic literature review where 30 research papers adopting crowdsourcing or machine learning in education were discussed. In order to utilize these techniques, the results should be highly effective and accurate due to the criticality of the education process. Due to this reason, this study aimed to identify all the learning activities in the education sector which are utilizing crowdsourcing and machine learning. Moreover, the study focused on addressing the proposed crowdsourcing and machine learning solutions for one of the important learning activities which is the exam assessment. Finally, the study investigated the interaction level in online courses after using crowdsourcing and machine learning approaches.

### 4.1 How is the learning and e-learning sector utilizing machine learning and crowdsourcing approaches?

These two techniques were utilized in various learning activities to enhance the learning process. Some of these activities, including, but not limited to; questions and material generation, reviews and explanations' enhancement, and exams' assessment. All the learning activities and their approaches were summarized in Table 2. In general, crowdsourcing is utilized in 47.8% of learning activities, while each one of the other two approaches utilized in around 26% of learning activities. This result was calculated based on Table 2 where the total number of discussed papers is 23, 11 of them were about crowdsourcing, 6 of them presented machine learning approaches, and the last 6 papers discussed the hybrid approaches. Similarly, for each learning activity, the percentage is calculated in terms of using either crowdsourcing, machine learning or the hybrid of both in Table 6.

**Table 6** Learning activity utilization

| Learning Activity | Crowdsourcing | Machine Learning | Hybrid |
|---|---|---|---|
| Research | 100% | 0% | 0% |
| Students' Interactions in MOOC | 67% | 0% | 33% |
| Personalized Learning | 67% | 33% | 0% |
| Homework and Exams Assessment | 50% | 50% | 0% |
| Exams and Material Generation | 67% | 16.5% | 16.5% |
| Cheating Detection | 0% | 0% | 100% |
| Explanation/Reviews Enhancement | 0% | 40% | 60% |

### 4.2 What are the existing approaches that address the exam assessment problem using machine learning or crowdsourcing?

Exam assessment is one of the most important and popular learning activities because it is the main way used to evaluate the students during their studies. For any assessment, providing prompt feedback will enhance the learning process and help the students to perform better in the final exams (Lemley et al. 2007). The assessment problem was discussed in the reviewed literature in four research papers. CrowdGrader by (de Alfaro and Shavlovsky 2014) and CrowdSorcerer by Pirttinen et al. (2018) are both utilizing the crowd to enhance the assessment but with a slightly different application. CrowdGrader is a public web-based system that lets students submit and collaboratively review and grade their homework. Students receive an overall crowd-grade that reflects both the quality of their homework, and the quality of their work as reviewers. On the other hand, CrowdSorcerer is an embeddable tool for online courses that allow the student to create programming assignments and review each other's work by providing their evaluation.

The assessment systems by Kashi et al. (2016) and Gutierrez et al. (2013) are relying on the machine learning to automate the procedure. The Score Recommendation System by Kashi et al. (2016) is using NLP to assign a score of the descriptive answers given that the ideal answer is available. While the system by Gutierrez et al. (2013) can identify the correct and incorrect statements by combining extraction rules and machine learning based information extractors. This method can provide detailed feedback for automatic summary grading.

Exams Assessment becomes more challenging with large classrooms because it will bring a dilemma between providing fast results or accurate results. This dilemma needs to be considered when proposing any new assessment method. How to grade essay questions within a short period of time efficiently? Relying on the crowd completely for the assessment process requires the researchers to find innovative ways which will motivate the crowd to provide precise assessment. Likewise, using only machine learning models without human intervention for this learning activity will raise concerns whether it is the ultimate robust solution.

### 4.3 Does utilizing crowdsourcing and machine learning enhance the interactions in online courses?

Interaction in the e-learning environment is essential for uplifting the students' enthusiasm during the course and eventually increase the gained value from it. Research papers by (Kamath et al. 2016; Kulkarni et al. 2015a, b) have covered this topic using different approaches. Talkabout (Kulkarni et al. 2015a) is a peer discussion system that connects students from different countries to share their knowledge based on their regions and countries. It connects students via guided video discussion to help them interact with peers from other parts of the world. Moreover, PeerStudio (Kulkarni et al. 2015b) enables the crowds to provide rapid feedback on the student's submission. The result of using this platform is improving course outcomes by giving the student feedback in just twenty minutes. The system described by Kamath et al. (2016) will auto-recognize students' engagement levels by observing their behavioral states. By using the front-facing camera, the system can detect the student's engagement and then the instructor can either contact the student for further clarification or adjust his/her learning content to fit the students' needs.

As a result of utilizing crowdsourcing or machine learning in the previous papers, the interaction in online courses has enhanced both "Participating in a video discussion with peers increases participation in quizzes and improves performance" (Kulkarni et al. 2015a) and "Students in the fast feedback condition did significantly better than those in no early feedback condition" (Kulkarni et al. 2015b).

## 5 Conclusion

The education field has great importance and a massive impact on different societies. Starting from learning and moving toward the e-learning, many methods and techniques were presented to produce a higher quality experience that will enrich the whole educational sector. The road toward improvement can be found by using tools like crowdsourcing and machine learning which were our main investigation in this literature.

After surveying the literature, a total of 30 papers discussed ways to enhance the education sector by using crowdsourcing and machine learning. Different learning activities were found including but not limited to; exam questions and class material generation, reviews and explanations' enhancement, and exams' assessment.

New opportunities can arise if a careful investigation and a thorough study are applied to find the best educational application tools using crowdsourcing and machine learning. These two methods can be extended to enhance the current found learning activities (see Table 6) and it can also be applied to other learning activities (e.g. critical thinking, problem solving, etc.). Furthermore, a hybrid method can be used to enhance the performance of the learning process. Lastly, crowdsourcing and machine learning can be widely used to keep enhancing the interaction on online courses and inside physical classes. However, the most important point all the researchers should always consider is the accuracy of collecting data; whether the collected data is from the crowd or used for training the machine learning model. Inaccurate input data will produce wrong results which will jeopardize the whole learning experience instead of benefiting it.

# References

Alghamdi, E. A., Aljohani, N. R., Alsaleh, A. N., Bedewi, W., & Basheri, M. (2015). CrowdyQ: A virtual crowdsourcing platform for question items development in higher education. *Proceedings of the 17th International Conference on Information Integration and Web-Based Applications &Services - IiWAS '15*, 1–4. https://doi.org/10.1145/2837185.2843852.

Barbosa, C. E., Epelbaum, V. J., Antelio, M., Oliveira, J., & Moreira de Souza, J. (2013). Crowdsourcing environments in E-learning scenario: A classification based on educational and collaboration criteria. *2013 IEEE International Conference on Systems, Man, and Cybernetics*, 687–692. https://doi.org/10.1109/SMC.2013.122.

Choi, J.-R., Kim, S., & Lim, S.-B. (2017). The feedback block model for an adaptive E-book. *Adjunct Publication of the 30th Annual ACM Symposium on User Interface Software and Technology - UIST '17*, 127–128. https://doi.org/10.1145/3131785.3131834.

Corrigan-Gibbs, H., Gupta, N., Northcutt, C., Cutrell, E., & Thies, W. (2015). Deterring cheating in online environments. *ACM Transactions on Computer-Human Interaction, 22*(6), 1–23. https://doi.org/10.1145/2810239.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*(4), 438–481. https://doi.org/10.3102/00346543058004438.

de Alfaro, L., & Shavlovsky, M. (2014). CrowdGrader: A tool for crowdsourcing the evaluation of homework assignments. *Proceedings of the 45th ACM Technical Symposium on Computer Science Education - SIGCSE '14*, 415–420. https://doi.org/10.1145/2538862.2538900.

Dow, S., Gerber, E., & Wong, A. (2013). A pilot study of using crowds in the classroom. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, 227. https://doi.org/10.1145/2470654.2470686.

Faisal, M. H., AlAmeeri, A. W., & Alsumait, A. A. (2015). An adaptive e-learning framework: Crowdsourcing approach. *Proceedings of the 17th International Conference on Information Integration and Web-Based Applications &Services - IiWAS '15*, 1–5. https://doi.org/10.1145/2837185.2837249.

Farasat, A., Nikolaev, A., Miller, S., & Gopalsamy, R. (2017). CROWDLEARNING: Towards collaborative problem-posing at scale. *Proceedings of the Fourth (2017) ACM Conference on Learning @ Scale - L@S '17*, 221–224. https://doi.org/10.1145/3051457.3053990.

Gutierrez, F., Dou, D., Martini, A., Fickas, S., & Zong, H. (2013). Hybrid ontology-based information extraction for automated text grading. 2013 12th International Conference on Machine Learning and Applications, 359–364. https://doi.org/10.1109/ICMLA.2013.73.

Harlen, W., Crick, R. D., Broadfoot, P., Daugherty, R., Gardner, J., James, M., & Stobart, G. (2002). A systematic review of the impact of summative assessment and tests on students' motivation for learning. *EPPI-Centre, University of London*, 151.

Huang, S.-W., Tu, P.-F., Fu, W.-T., & Amamzadeh, M. (2013). Leveraging the crowd to improve feature-sentiment analysis of user reviews. *Proceedings of the 2013 International Conference on Intelligent User Interfaces - IUI '13*, 3. https://doi.org/10.1145/2449396.2449400.

Kamar, E. (2016). *Directions in Hybrid Intelligence: Complementing AI Systems with Human Intelligence*. In IJCAI, 4070-4073.

Kamath, A., Biswas, A., & Balasubramanian, V. (2016). A crowdsourced approach to student engagement recognition in e-learning environments. 2016 *IEEE Winter Conference on Applications of Computer Vision (WACV)*, 1–9. https://doi.org/10.1109/WACV.2016.7477618.

Kashi, A., Shastri, S., Deshpande, A. R., Doreswamy, J., & Srinivasa, G. (2016). A score recommendation system towards automating assessment in professional courses. *2016 IEEE Eighth International Conference on Technology for Education (T4E)*, 140–143. https://doi.org/10.1109/T4E.2016.036.

Kim, J., Sterman, S., Cohen, A. A. B., & Bernstein, M. S. (2017). Mechanical novel: Crowdsourcing complex work through reflection and revision. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '17*, 233–245. https://doi.org/10.1145/2998181.2998196.

Kotturi, Y., Kulkarni, C. E., Bernstein, M. S., & Klemmer, S. (2015). Structure and messaging techniques for online peer learning systems that increase stickiness. *Proceedings of the Second (2015) ACM Conference on Learning @ Scale - L@S '15*, 31–38. https://doi.org/10.1145/2724660.2724676.

Krause, M., Garncarz, T., Song, J., Gerber, E. M., Bailey, B. P., & Dow, S. P. (2017). Critique style guide: Improving Crowdsourced design feedback with a natural language model. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, 4627–4639. https://doi.org/10.1145/3025453.3025883.

Kulkarni, C., Cambre, J., Kotturi, Y., Bernstein, M. S., & Klemmer, S. R. (2015a). Talkabout: Making Distance matter with small groups in massive classes. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15*, 1116–1128. https://doi.org/10.1145/2675133.2675166.

Kulkarni, C. E., Bernstein, M. S., & Klemmer, S. R. (2015b). PeerStudio: Rapid peer feedback emphasizes revision and improves performance. *Proceedings of the Second (2015) ACM Conference on Learning @ Scale - L@S '15*, 75–84. https://doi.org/10.1145/2724660.2724670.

Le, L. T., Shah, C., & Choi, E. (2016). Evaluating the quality of educational answers in community question-answering. *Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries - JCDL '16*, 129–138. https://doi.org/10.1145/2910896.2910900.

Lemley, D., Sudweeks, R., Howell, S., Laws, R. D., & Sawyer, O. (2007). The effects of immediate and delayed feedback on secondary distance learners. Quarterly Review of Distance Education, 8(3), 251.

Li, Y. (2018). An application of EDM: Design of a new Online System for correcting exam paper. 2018 13th International Conference on Computer Science & Education (ICCSE), 1–6. https://doi.org/10.1109/ICCSE.2018.8468821.

Li, X., Chang, K., Yuan, Y., & Hauptmann, A. (2015). Massive open online proctor: Protecting the credibility of MOOCs certificates. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15*, 1129–1137. https://doi.org/10.1145/2675133.2675245.

Luger, S. K. K., & Bowles, J. (2013). An analysis of question quality and user performance in crowdsourced exams. *Proceedings of the 2103 Workshop on Data-Driven User Behavioral Modelling and Mining from Social Media - DUBMOD '13*, 29–32. https://doi.org/10.1145/2513577.2513584.

Nkoana, T. H. (2016). E-learning: Crowdsourcing as an alternative model to traditional learning. 2016 International Conference on Advances in Computing and Communication Engineering (ICACCE), 423–428. https://doi.org/10.1109/ICACCE.2016.8073786.

Oxford Dictionary [Dictionary]. (2020). Retrieved January 1, 2020, from Lexico website: https://www.lexico.com/en/definition/machine_learning. Accessed 1 Jan 2020

Pirttinen, N., Kangas, V., Nikkarinen, I., Nygren, H., Leinonen, J., & Hellas, A. (2018). Crowdsourcing programming assignments with CrowdSorcerer. *Proceedings of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE 2018*, 326–331. https://doi.org/10.1145/3197091.3197117.

Poetz, M. K., & Schreier, M. (2012). The value of crowdsourcing: Can users really compete with professionals in generating new product ideas?: The value of crowdsourcing. *Journal of Product Innovation Management, 29*(2), 245–256. https://doi.org/10.1111/j.1540-5885.2011.00893.x.

Schenk, E., & Guittard, C. (2009). *Crowdsourcing: What can be Outsourced to the Crowd, and Why.* In Workshop on open source innovation, Strasbourg, France 72, 3.

Shetty, B. (2018). Natural Language Processing(NLP) for machine learning. Retrieved November 24, 2018, from towards data science website: towardsdatascience.com/natural-language-processing-nlp-for-machine-learning-d44498845d5b. Accessed 24 Sept 2019

Shongwe, T., & Zuva, T. (2018). Benefits of crowdsourcing and its growth in the education sector. *2018 International Conference on Advances in Big Data, Computing and Data Communication Systems (IcABCD)*, 1–6. https://doi.org/10.1109/ICABCD.2018.8465436.

Tang, T., Smith, R., Rixner, S., & Warren, J. (2016). Data-driven test case generation for automated programming assessment. *Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE '16*, 260–265. https://doi.org/10.1145/2899415.2899423.

Taniguchi, A., & Inoue, S. (2015). A method for automatic assessment of user-generated tests and its evaluation. *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers - UbiComp '15*, 225–228. https://doi.org/10.1145/2800835.2800931.

Vaish, R., Goel, S., Davis, J., Bernstein, M. S., Gaikwad, S. (Neil) S., Kovacs, G., … Belongie, S. (2017). Crowd research: Open and scalable university laboratories. *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology - UIST '17*, 829–843. https://doi.org/10.1145/3126594.3126648.

Vaughan, J. W. (2017). *Making better use of the crowd: How crowdsourcing can advance machine learning research.* Journal of Machine Learning Research, 18, 193-1.

Watson, P., Ma, T., Tejwani, R., Chang, M., Ahn, J., & Sundararajan, S. (2018). Human-level multiple choice question guessing without domain knowledge: Machine-learning of framing Effects. *Companion of the The Web Conference 2018 on The Web Conference 2018 - WWW '18*, 299–303. https://doi.org/10.1145/3184558.3186340.

Whitehill, J., & Seltzer, M. (2016). A crowdsourcing approach to collecting tutorial videos—Toward personalized learning-at-scale. *ArXiv:1606.09610 [Cs]*. Retrieved from http://arxiv.org/abs/1606.09610. Accessed 15 Aug 2019

Williams, J. J., Kim, J., Rafferty, A., Maldonado, S., Gajos, K. Z., Lasecki, W. S., & Heffernan, N. (2016). AXIS: Generating explanations at scale with Learnersourcing and machine learning. *Proceedings of the Third (2016) ACM Conference on Learning @ Scale - L@S '16*, 379–388. https://doi.org/10.1145/2876034.2876042.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.