

## LectureBuddy: Towards Anonymous, Continuous, Real-time, and Automated Course Evaluation System

Muhammad Abdullah Shah Bukhari<sup>1</sup>, Faisal Bukhari<sup>2</sup>, Muhammad Idrees<sup>2</sup>, Syed Ameer Hamza Bokhari<sup>3</sup>, Ashfaq Ahmad<sup>4</sup>

<sup>1</sup>Department of Computer Science, University of the Punjab, Lahore, Pakistan.

<sup>2</sup>Department of Data Science, University of the Punjab, Lahore, Pakistan.

<sup>3</sup>Ghulam Ishaq Khan Institute of Technology, Swabi, Pakistan.

<sup>4</sup>Department of Computer Science & IT, The University of Lahore, Lahore, Pakistan.

\* **Correspondence:** Muhammad Idrees ([idrees@pucit.edu.pk](mailto:idrees@pucit.edu.pk))

**Citation |** Bukhari. M. A. S, Bukhari. F, Idrees. M, Bokhari. S. A. H, Ahmad. A, “Lecture Buddy: Towards Anonymous, Continuous, Real-time, and Automated Course Evaluation System”, *IJIST*, Vol. 5, No. 3, pp. 284-297, Sep. 2023.

**Received |** Sep 01, 2023; **Revised |** Sep 18, 2023; **Accepted |** Sep 24, 2023; **Published |** Sep 26, 2023.

Student’s course evaluations are a primary tool for measuring teaching effectiveness. The traditional practice in course evaluation at most institutes is carried out once, at the end of each semester. The effectiveness of this system requires candid participation from the students, followed up by the administration, and the faculty. While corrective action took place behind the scenes over a long period, students never observed any immediate change(s) based on the feedback they were provided through the existing course evaluation systems. This discourages students from considering the evaluation seriously. In this paper, we investigate the need for an innovative system to replace the existing course evaluation systems. We conducted two separate surveys from 210 students and 67 teachers to gain insight into the existing course evaluation systems. The survey participants answered questions based on the tendency of feedback provided by students, method of teacher’s evaluations, frequency of evaluations conducted by institutes, and steps to make classrooms more interactive. We also conducted a comprehensive statistical analysis of the data collected from the surveys, both qualitative and quantitative. Our study showed a need for an innovative course evaluation system to continuously gather student feedback throughout the semester anonymously. These findings led us to develop the prototype of an innovative course evaluation system, “Lecture Buddy”, which is anonymous, continuous, real-time, and automated and which alleviates the shortcomings of the traditional course evaluation systems.

**Keywords:** improving classroom teaching, evaluation methodologies, interactive learning environments, and computer-mediated communication.

### Acknowledgment.

The authors acknowledge the student group for the survey they conducted and for writing the initial draft of the

course term project report.

### Conflict of interest.

The authors have no conflicts of interest relevant to this article.

### Author’s Contribution.

Both authors completely participated in the research work at all times.

**Project Detail.** Nil



## Introduction

Student's course evaluations have always been the most common method for measuring teaching effectiveness [1][2][3]. Most of the higher education institutes conduct these evaluations using an online system [4][5]. The primary purpose of this process is to collect and analyze student's feedback to measure the effectiveness of courses and instructors [6][7][8]. However, there is a wide range of research surrounding the question of the reliability, validity, and usefulness of student evaluations [9][10][11].

Student evaluation of teaching (SET) is one of the primary tools for gauging teaching effectiveness [12]. However, the results of teacher's evaluations are only known at the end of each semester, mostly after the declaration of grades [9]. Moreover, the current evaluation systems do not reflect an accurate picture of students learning [13].

The information and feedback provided by students as part of course evaluation are essential for the self-improvement of the teachers and the overall quality of the institute [14], [15]. The effectiveness of this system requires candid participation from the students, followed by the administration and the faculty [16]. While corrective action took place behind the scenes over a long period, students never observed any immediate changes based on the feedback they were provided through the existing course evaluation systems. This discourages students from considering the evaluation seriously. Therefore, the existing course evaluation systems are not sufficient and do not help improve the quality of education.

Typically, course evaluations can be used to collect formal and informal feedback from students through Likert-scale-based questionnaires [17]. The formal way includes collecting feedback about course policies, grading schemes, assignments, curriculum, and syllabus. The informal way includes collecting information about course experience and opinions about the teacher's characteristics, personality, and communication effectiveness. Mostly, the purpose of the evaluation questions is typical and old-fashioned; they might not help enough to improve the quality of a course. When students are asked to evaluate a teacher and a course at the end of the semester, the questions are the same for both theoretical and applied subjects, so there is no right way to rate the capability of a teacher regarding the nature of a discipline. Students get bored of answering the same, repeated questions.

Student-teacher class interactions play a pivotal role in improving course quality [18]. Most lecture-oriented classes are effective in delivering course content in a limited amount of time. However, a lack of student-instructor interactions often makes most of the students uninterested in course content during the class and, as a result, makes them less motivated to learn. In contrast, classes that emphasize student learning provide adequate time for students to think about concepts, give feedback, and actively participate in class. Some of the teachers need to restructure their class so that it contains activities where they can have student's feedback to decide if there is a need to adopt better methods. Moreover, it helps students generate different solutions and come up with new ideas in the context of the lesson.

In this paper, we conducted two separate surveys of 210 students and 67 teachers to understand the effectiveness of the current evaluation system and the need for an innovative course evaluation system. Both questionnaires are centered around the tendency of feedback provided by students, the method of teacher's evaluations, the frequency of evaluations conducted by institutes, and steps to make classrooms more interactive. Survey results indicate that in most of the institutes, semester evaluations are done at the end of the semester. In the rest of this paper, we present related work, discuss the material and methods used to collect the surveys, discuss the results of the student's survey and the teacher's survey, discuss the proposed solution named "Lecture Buddy", draw some conclusions, and discuss future work.

**Objectives:**

The aim of this study is a mobile app or web-based system to access the progress of course conduction. This assessment may be used to improve the quality of student-teacher interaction in the future and also provide an effective measure in the ranking of education in institutes. This research aims to understand the limitations of traditional course evaluation systems, gather data from students and teachers, perform a statistical analysis of this data, and propose an innovative course evaluation system that overcomes these limitations and provides continuous, real-time feedback for improvement.

**Novelty Statement:**

This research and the development of a related app or web-based is the first of its kind. Its effective use will help different stakeholders in making decisions. The novelty of "Lecture Buddy" is its holistic approach to course evaluation, which addresses the limitations of traditional methods by providing continuous, real-time, anonymous, and automated feedback to improve teaching effectiveness. This innovative system aims to enhance the overall learning experience for students and the teaching experience for instructors.

**Related Work:**

There has been some work to study the effectiveness of the course evaluation systems. For example, [15] study the satisfaction of students and teachers with the typical course evaluation system. The results reveal that optional questions are more relevant to understanding course satisfaction. [16] claims that the end-of-course evaluations contribute significantly to faculty development, but usually, faculty members are not satisfied with the quality of the feedback. The author supports the idea of student's feedback in the middle of the execution of a course. A survey conducted by [17] reported that 41% of faculty members are not satisfied with the end-of-course evaluations. A study by [18] reveals biases in student evaluations of teaching against female faculty members. The authors report that favoritism varies in evaluations in different disciplines. Another study by [19] discusses a lenient grading policy that can have a significant bias in student's evaluations of teaching. This work shows that a positive correlation between student's grades and evaluations is mainly due to bias instead of a valid teaching methodology. A recent study [20] supports collecting student feedback regularly to improve teaching and learning practices. Some of the authors study the impact of web-based course evaluation systems. For example, [21] shows low response rates using web-based evaluation methods, but that does not affect the overall mean evaluation score. [22] show a simple start-stop-continue evaluation method could enhance the standard of student's evaluations. They show faculty evaluations are not significantly affected by switching from paper-based to web-based evaluations. Researcher [23] claims that online evaluations have the drawback of low student participation. Some of the recent studies identified the need for enabling SET during the class using mobile devices [24].

Some of the recent work proposed automated tools to evaluate student performance. For example, [25] built automated tools to estimate the performance of students based on the neural network classification method. To estimate the performance of any student, their proposed solution uses prior knowledge about the student and also utilizes the knowledge of other students having similar characteristics. Another recent study by [26] identifies the impact of artificial intelligence technology in the education learning process. In this paper, we present our study conducted to understand the merits and demerits of the current evaluation systems. The study is based on two separate surveys conducted by students and teachers. Based on our findings, we develop, "Lecture Buddy" an anonymous real-time course evaluation system for student's evaluation of teaching. Our proposed system overcomes the limitations of existing state-of-the-art methods, including end-of-course, start-stop-continue, and typical web-based evaluations. A recent work by [27] also proposed an online Teacher Evaluation System which

uses a web-based approach for student feedback and then it uses data analysis to classify the feedback.

Fuzzy logic was employed in this paper [28] for teacher evaluation in the context of developing academic institutions, particularly in India, to create a more effective and equitable education system. The paper proposed a fuzzy-based educator feedback system that aimed to assess faculty performance and provide essential feedback to enhance teaching and learning processes. This fuzzy system used a linguistic model, the multiple input and single output (MISO) Mamdani model, to categorize educators based on student feedback collected from various parameters. The system sought to identify areas where educators needed improvement, thereby facilitating better engagement with students and enabling confidential appraisal reports for institution administrators. Fuzzy logic was chosen due to its ability to handle imprecision and uncertainty inherent in qualitative data, making it a valuable tool for evaluating educator performance more flexibly and efficiently compared to traditional, manual methods.

This article [29] emphasizes the significance of feedback in online learning environments, especially when instructors and students are physically separated. It highlights the challenges instructors face in delivering timely and useful feedback, particularly in large online cohorts. To address this, automatic feedback systems were proposed. The paper presents a systematic literature review on automatic feedback generation in learning management systems, summarizing findings from 63 selected studies published between 2009 and 2018. The review aims to identify trends, goals, and outcomes of these systems to enhance feedback practices in online education.

This case study [30] explores how five Chinese learners of English engaged with automated, peer, and teacher feedback in an online EFL writing course over a 17-week semester. It aims to understand the dynamic and interactive nature of learners' engagement with different feedback sources and the factors influencing their feedback uptake decisions over time. The study addresses a gap in research on learner engagement with multiple feedback sources in online EFL writing contexts and emphasizes the need for a naturalistic approach to capture longitudinal developments in feedback-revision processes.

This study [31] aimed to improve teacher feedback by providing detailed and actionable automated feedback, overcoming the limitations of infrequent and performance-focused human classroom observations. To achieve this, they developed a method for teachers to easily record high-quality audio from their classes, resulting in 89% usable recordings out of 142 sessions. Using speech recognition and machine learning, they created computer-scored estimates of key aspects of teacher discourse, finding that these automated models were moderately accurate compared to human coders, with speech recognition errors having minimal impact. The next step is to integrate these automatic models into an interactive visualization tool to offer teachers objective feedback on the quality of their teaching discourse.

This paper uses Teaching Analytics [32] which is a novel approach that combines teaching expertise, visual analytics, and design-based research to support teachers in using data and evidence to enhance the quality of teaching. TA is gaining significance, offering opportunities to improve teaching performance and engage teachers in reflective dialogue. Teachers need to develop data literacy and understand the connection between TA, Learning Analytics (LA), and Learning Design (LD). This research reviews TA literature, aims to provide a comprehensive framework, and introduces the concept of the Teaching Outcome Model (TOM) to guide teachers in using data for better teaching. The study systematically analyzed articles from 2012 to 2019, revealing a need for further development of TA concepts.

This study [33] introduces a supervised aspect-based opinion mining system that utilizes a two-layered LSTM model to handle student's qualitative feedback for evaluating faculty teaching performance. The first layer predicts aspects mentioned in the feedback, while the

second layer determines the sentiment orientation (positive, negative, neutral) of these aspects. The model achieves high accuracy in both aspect extraction (91%) and sentiment polarity detection (93%). It addresses the challenge of processing qualitative opinions efficiently in academic feedback and aims to automate the analysis of student's comments. The research contributes by preparing an academic domain dataset, proposing a two-stage LSTM model for aspect and sentiment analysis, and providing advancements in sentiment analysis through deep learning techniques. The study suggests the potential applicability of this model in various domains with minor parameter adjustments.

This paper [34] addresses the limitation of question-score-based student evaluations of teaching (SET) by proposing two methods, knowledge-based and machine learning-based, to automatically extract opinions from student's short reviews. These methods aim to capture additional facets of the teaching process that may not be covered by predefined questionnaires. The study also highlights the diversity in the themes and styles of reviews with the same sentiment polarity, demonstrating that reviews with similar sentiments share common language patterns. The experimental results indicate that these methods achieve high accuracy in sentiment classification of student reviews (78.13% and 84.78%). The paper concludes by presenting a real-world application scenario for using these methods in the SET process.

### **Material and Methods:**

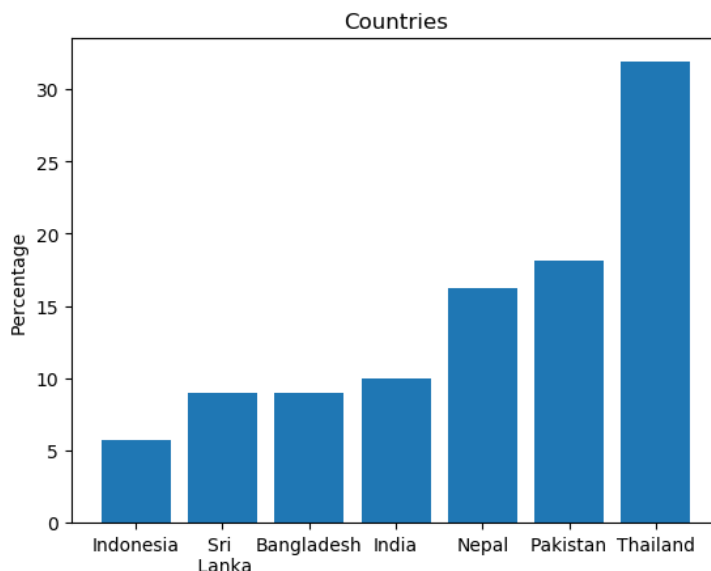
We conducted two separate online surveys of students and teachers. These surveys were based on diversified questions answered by students and teachers from different countries and universities. We asked students how frequently they provided feedback to their teachers. Our questionnaire designed for teachers inquired about how often they carried out informal evaluations during their classes apart from the formal evaluations conducted by their respective institutes. We asked students whether they would prefer an automated evaluation system that disguised their identity during teacher's evaluations. The data obtained from the student's survey was named the student's dataset, and the data obtained from the teacher's survey was named the teacher's dataset. We performed extensive statistical analysis based on estimation theory and hypothesis testing using an appropriate z-test. Some of our survey questions were based on the ordinal scale, so we applied the nonparametric Kolmogorov-Smirnov test. We plotted the results using pie charts for both datasets.

### **Student's Dataset:**

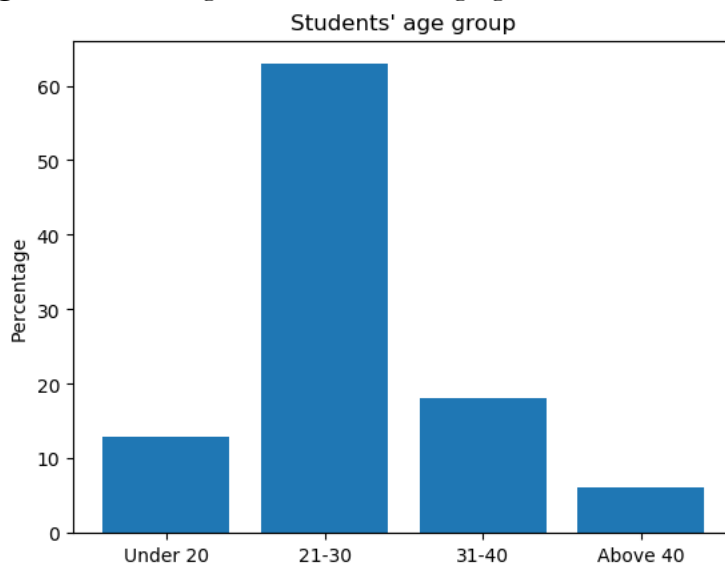
In this section, we explain the results of the student survey. Our sample size consisted of 210 students from seven different countries, i.e., Indonesia, Sri Lanka, Bangladesh, India, Nepal, Pakistan, and Thailand. Our student dataset consisted of 32% of graduate students, 35% of master's students, and 31% of Ph.D. students. The results of the survey showed that the average age and standard deviation of students were 27.45% and 1.36%, respectively. Figure 1 shows the country-wise participation of the students. The majority of the students in our survey belonged to Thailand (31.9%) and Pakistan (18.1%). Figure 2 shows age-wise participation in our student's survey. The majority of the students belonged to the 21–30 (62.9%) age group. The questionnaire was based on five questions. All the questions were ordinal. In the first question, students were asked to choose the Likert-type scale ranging from never, frequently, and occasionally, while in the remaining four questions, students were asked to respond to the Likert-type scales, i.e., strongly disagree (SD), disagree (D), neutral (N), agree (A), and strongly agree (SA).

Table 1 shows the percentage of student's responses across the total sample size of students. The response to question P1 showed that the majority of students (60.41%) occasionally or never (23.35%) provide feedback with their identity revealed to teachers. However, a majority of the students (82.89%) are willing to provide anonymous feedback to teachers. This is evident from the responses of P2. Responses based on question P3 suggest that

most of the students (59.41%) were comfortable raising questions during the class. However, responses from P4 were mixed: 43.37% of students supported the idea, and 29.08% did not support the idea of raising questions anonymously during the class. Responses to P5 revealed that the majority of the students (67.02%) were in favor of using an electronic system to provide anonymous feedback to teachers. We concluded that the majority of students were not afraid to ask questions during the class about their identities. However, most of the students supported the idea of a real-time anonymous feedback system that provides feedback to teachers during class. See Table 1 for details of P1 to P5.



**Figure 1.** Percentage of students belonging to different countries.



**Figure 2.** Percentage of students belonging to different age groups.

Table 2 shows the percentage of teacher’s responses across the total sample size of teachers. The response to question P6 showed that the majority of teachers (72.24%) performed informal evaluations. However, a majority of the teachers (53.97%) conduct informal evaluations once at the end of the semester. 12.7% of teachers conduct informal evaluations monthly. This is evident from the responses of P7. The responses of teachers based on question P8 suggested that most of them (66.67%) preferred paper-based formal evaluations. P9 suggested that most of the institutes (80.00%) conducted formal evaluations.

**Table 1.** Percentage of student’s responses across the total sample size.

Questions based on informal evaluations	Never	Occasionally	Frequently	Neutral	Agree	Strongly Agree	Disagree	Strongly Disagree
$P_1$ : How often do you provide feedback to your teacher about their content delivery through any channel that identifies you?	23.35	60.41	16.24	-	-	-	-	-
$P_2$ : I want to provide feedback to my teachers about their content delivery and lecture pace without disclosing my identity.	-	-	-	18.23	49.89	33.00	4.93	2.96
$P_3$ : I am usually comfortable to raise questions during class.	-	-	-	29.21	42.08	17.33	9.41	1.98
$P_4$ : I want to be anonymous when raising questions during class.	-	-	-	27.55	18.88	10.20	30.61	12.76
$P_5$ : I want to use an electronic system to provide anonymous feedback to my teachers.	-	-	-	23.4	35.11	31.91	4.79	4.79

Note:  $P_1 - P_5$  are tags assigned to each question asked in a students’ questionnaire.

**Table 2.** Percentage of teacher’s responses across the total sample size

Questions	Yes	No	After Every Lecture	Weekly	Monthly	After Mid Term	Once After the Semester	Twice or Thrice a Semester	Not Applicable
Questions based on informal evaluations									
$P_6$ : Do you ever have the students in your classes informally evaluate you and the course to improve the course?	74.24	25.76	-	-	-	-	-	-	-
$P_7$ : If you have conducted such evaluations, how often do you do it?	-	-	4.76	4.76	12.7	1.59	53.97	1.59	20.63
Questions based on formal evaluations									
$P_8$ : When you have conducted evaluations, what is your preferred method?	Paper Based		Raising Hands		Computer Based		Oral	Not Applicable	
	66.67		8.33		1.67		3.33	20	
$P_9$ : Does your institute conduct formal course evaluations?	80.00	20.00	-	-	-	-	-	-	-
$P_{10}$ : How frequently does your institute conduct course evaluations?	-	-	1.69	-	-	-	74.58	1.69	22.03
Questions based on formal evaluations									
$P_{11}$ : If your institute does conduct formal course evaluations, when do you receive the evaluations?	Within a Week	Within a Month	After 2 Months	Within a semester	Within a Year	Does not share evaluations with faculty		Not Applicable	
	21.54	23.08	2.00	6.08	20.00	4.62		23.08	
$P_{12}$ : Would you like to use an automated real-time system during your lecture to get students’ feedback immediately without knowing their identity?	72.73	27.27	-	-	-	-	-	-	-

Note:  $P_6 - P_{12}$  are tags assigned to separate question asked in a teachers’ questionnaire.

However, the frequency of formal evaluations conducted by institutes based on P10 is limited to once after the semester (74.58%). P11 revealed 23.08%, 21.54%, and 20.00% of teachers received formal evaluations within a month, within a week, and within a year, respectively. Teacher’s responses based on question P12 showed most of them (72.73%) recommended a real-time anonymous evaluation system. See 2 for details of P6 to P12.

**Experimental Results and Analysis:**

Some hypotheses were proposed targeting the most important questions from our questionnaire. We used the level of significance of  $\alpha = 0.05$ . See Table 3 for details. We applied the Kolmogorov-Smirnov test to all the questions in Table 3, as all the questions are based on the ordinal scale. We accepted the alternative hypothesis, i.e., H1B, and concluded that the results showed a significant preference for occasional It means students occasionally provide informal feedback to teachers through any means that identify them. The Kolmogorov-Smirnov test result suggested accepting the alternative hypothesis, i.e., H2B, and concluded that the results showed a significant preference for agreeing. It suggests that students agree to provide informal evaluations to teachers without disclosing their identities.

Our survey results in Table 3 negated hypothesis H1A that the majority of students would not like to hide their identity while doing a formal evaluation. The null hypothesis H1A

was rejected and suggested that the majority of students were curious about their anonymity during evaluation. We failed to reject the null hypothesis H2A that the majority of students were comfortable raising questions during the class. The next hypothesis, H3A, was rejected because it suggested that the majority of students did not like to be anonymous while asking questions in class. Finally, the result of hypothesis H4A concluded that the majority of students preferred an automated electronic system for anonymous feedback.

**Table 3.** Hypotheses testing of student’s sample population based on Kolmogorov-Smirnov Test

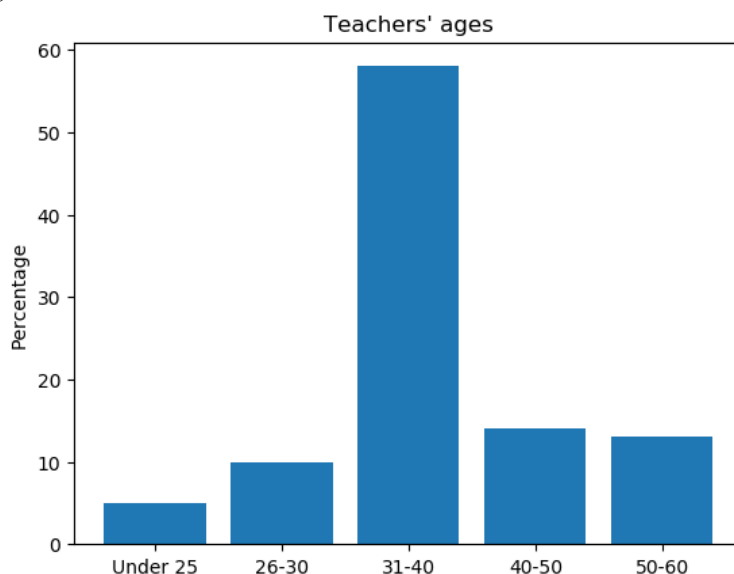
Hypotheses	Computed D	Critical D*	Conclusion
$H_{1A} : P_1 =$ All options (corresponding to $P_1$ ) are equally preferred. $H_{1B} : P_1 =$ All options are not equally preferred.	0.0162	0.171	Accept the alternative hypothesis i.e., $H_{1B}$ .
$H_{2A} : P_2 =$ All options (corresponding to $P_2$ ) are equally preferred. $H_{2B} : P_2 =$ All options are not equally preferred.	0.3212	0.0955	Accept the alternative hypothesis i.e., $H_{2B}$ .
$H_{3A} : P_3 =$ All options (corresponding to $P_3$ ) are equally preferred. $H_{3B} : P_3 =$ All options are not equally preferred.	0.3129	0.0957	Accept the alternative hypothesis i.e., $H_{3B}$ .
$H_{4A} : P_4 =$ All options (corresponding to $P_4$ ) are equally preferred. $H_{4B} : P_4 =$ All options are not equally preferred.	0.0755	0.0971	Accept the null hypothesis i.e., $H_{4A}$ .
$H_{5A} : P_5 =$ All options (corresponding to $P_5$ ) are equally preferred. $H_{5B} : P_5 =$ All options are not equally preferred.	0.3017	0.1741	Accept the alternative hypothesis i.e., $H_{5B}$ .

Note:  $P_1 - P_5$  are explained in Table 1.

\* Indicate critical value of Kolmogorov-Smirnov test i.e D at  $\alpha = 0.05$  level significance.

**Teacher’s Dataset:**

In this section, we explained the results obtained from the teacher’s survey. Our sample size consisted of 67 teachers from seven different countries, i.e., Indonesia, Sri Lanka, Bangladesh, India, Nepal, Pakistan, and Thailand. Teacher’s average age was 38.20 years, with an SD = 3.12 years. The average experience of teachers was 9.5 years, with an SD = 1.40 years. Teacher’s questionnaires consisted of seven questions, each based on multiple options depending on the nature of the question. Figure 3 shows that the majority of teachers (58%) fall in the 31 – 40 age bracket.



**Figure 3.** Percentage of teachers of different age groups.

Figure 4 shows that 36% of the teachers have 6 to 10 and 25% have 11 to 15 years of teaching experience.



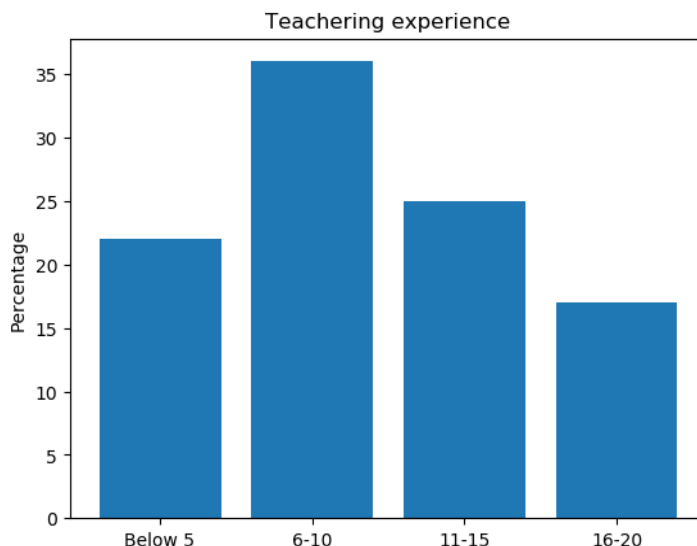


Figure 4. Percentage of teachers with different teaching experience.

**Hypotheses Testing and Statistical Analysis:**

We formulate hypotheses based on questions defined in Table 3, and rejection and acceptance of hypotheses are based on p-values using the z-test. We set the level of significance at  $\alpha = 0.05$ . We designed separate hypotheses for formal (H6A and H7A) and informal (H9A, H10A, H11A, and H12A) evaluations.

**Hypotheses Based on Informal Evaluations:**

We constructed hypotheses H6A and H7A based on the questions defined for informal evaluations in Table 2. Table 4 results showed that an alternative hypothesis, i.e., H6B, was accepted, and we concluded that more than 50% of teachers did informal evaluations to seek improvement in their respective courses. We also accepted the alternative hypothesis H7B and concluded that the results showed a significant preference for informal evaluations at the end of the semester.

Table 4. Hypotheses testing of teacher’s sample population.

Hypotheses	p – value <sup>1</sup>	Computed D	Critical D <sup>2</sup>	Conclusion
Hypotheses based on informal evaluation				
H <sub>6A</sub> : P <sub>06</sub> = 0.50	4.1 × 10 <sup>-5</sup>	-	-	-
H <sub>6B</sub> : P <sub>06</sub> > 0.50	-	-	-	Accept the alternative hypothesis i.e., H <sub>6B</sub> .
H <sub>7A</sub> : P <sub>07</sub> = All options (corresponding to P <sub>7</sub> ) are equally preferred	-	0.4128	0.1713	-
H <sub>7B</sub> : P <sub>07</sub> = All options are not equally preferred.	-	-	-	Accept the alternative hypothesis i.e., H <sub>7B</sub> .
Hypotheses based on formal evaluation				
H <sub>9A</sub> : P <sub>09</sub> = 0.50	1 × 10 <sup>-4</sup>	-	-	-
H <sub>9B</sub> : P <sub>09</sub> > 0.50	-	-	-	Accept the alternative hypothesis i.e., H <sub>9B</sub> .
H <sub>10A</sub> : P <sub>10</sub> = All options (corresponding to P <sub>10</sub> ) are equally preferred	-	0.2627	0.1771	-
H <sub>10B</sub> : P <sub>10</sub> = All options are not equally preferred.	-	-	-	Accept the alternative hypothesis i.e., H <sub>10B</sub> .
H <sub>11A</sub> : P <sub>11</sub> = All options (corresponding to P <sub>11</sub> ) are equally preferred	-	0.1677	0.1886	Fail to reject the null hypothesis i.e., H <sub>11A</sub> .
H <sub>11B</sub> : P <sub>11</sub> = All options are not equally preferred.	-	-	-	-
H <sub>12A</sub> : P <sub>12</sub> = 0.50	1 × 10 <sup>-4</sup>	-	-	-
H <sub>12B</sub> : P <sub>12</sub> > 0.50	-	-	-	Accept the alternative hypothesis i.e., H <sub>12B</sub> .

Note: P<sub>6</sub> – P<sub>12</sub> are explained in Table 2.

1 The level of significance is set at  $\alpha = 0.05$  for p-value based critical region.

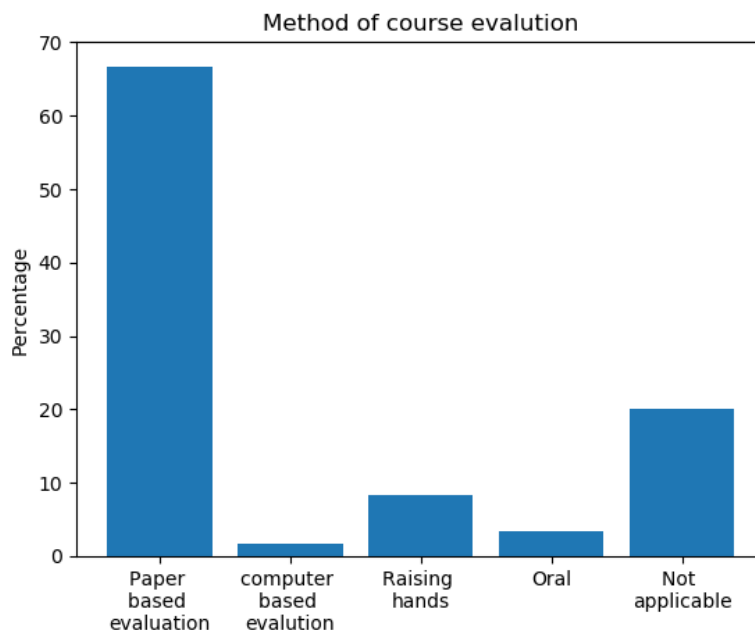
2 Indicate critical value of Kolmogorov-Smirnov test i.e D at  $\alpha = 0.05$  level significance.

**Hypotheses Based on Informal Evaluations:**

We designed hypotheses (H9A, H10A, H11A, and H12A) based on questions based on formal evaluations in Table 2. Table 4 indicated the acceptance of the alternative hypothesis H9B and concluded that more than 50% of institutes conduct formal evaluations. We accepted the alternative hypothesis H10B and concluded that there was a significant preference for once at the end of the semester. This means most of the institutes conduct formal evaluations once

at the end of the semester. We failed to reject the null hypothesis, i.e., H11A, and concluded that all levels are equally preferred.

Lastly, the acceptance of the alternative hypothesis H12B very strongly advocated teacher's (more than 50%) preference for a real-time evaluation system. Figure 5 explains the response of P8, which is defined in Table 2. The majority of teachers (66.7%) preferred paper-based informal evaluations. Some teachers (8.33%) got informal evaluations from students by asking them to raise their hands. A slim majority (3.3%) of teachers conducted such evaluations based on oral communication with students.



**Figure 5.** Percentage of teacher's responses based on question P8 as explained in Table 2.  
**Lecture Buddy Real-Time Course Evaluation System:**

To overcome the limitations of existing course evaluation systems, we developed a prototype course evaluation system, namely "Lecture Buddy". It is a simple web-based, easy-to-use, anonymous, continuous, and real-time student evaluation system. The proposed system consists of two main modules.

The Student's module shown in Figure 6 is used to collect student feedback on the content's understandability and the pace of the teacher. The student can optionally give comments to the teachers. This view is automatically enabled for the students during the lecture, and students can anonymously provide feedback to the teacher during the lecture.

The teacher's module shown in Figure 7 aggregates the student feedback in real time, and the teacher can enable alerts on a specific number of new responses obtained to see the student's feedback in real-time and adjust the lecture accordingly. One can criticize the proposed system for making it annoying for the teachers to check the comments. However, the teachers have the convenience to check the student feedback at regular intervals and ignore the alerts. Currently, we are using this system to collect the data and intend to perform a comparative study of the "Lecture Buddy" with the end-of-semester course evaluation system.

### Discussion:

Web-based systems like "Lecture Buddy" are designed to be easily accessible for both students and faculty, featuring user-friendly interfaces that require minimal technical expertise. This accessibility ensures that students and teachers can use the system without significant barriers. Web-based systems offer a high degree of anonymity, which encourages students to provide more candid feedback compared to traditional methods. This anonymity allows students

to share their thoughts without fear of repercussions. The anonymity provided by web-based systems positively influences the quality and candidness of evaluations, as students feel secure in sharing their genuine thoughts and concerns. This leads to more valuable feedback for instructors.

Web-based systems can be customized to meet the specific needs and requirements of different educational institutions or departments, ensuring flexibility and alignment with unique goals. This customization allows institutions to tailor the system to their specific evaluation criteria. Web-based systems are highly adaptable to evolving evaluation criteria or methods, allowing for easy updates and modifications as educational standards change. This adaptability ensures that the system remains relevant over time. Integration of web-based systems into regular classroom activities promotes ongoing feedback, as instructors can use real-time data to adjust teaching methods during the semester. This integration encourages a continuous feedback loop between students and teachers.

Web-based systems enhance the interactive learning environment by fostering communication between students and instructors, resulting in a more responsive and engaging classroom atmosphere. This interaction can lead to improved learning outcomes. Web-based systems prioritize security and confidentiality through encryption protocols, secure data storage, and access controls to safeguard evaluation data. This ensures that sensitive information remains protected. Sensitive information about both students and teachers is protected in web-based systems through anonymization of responses and robust security measures to prevent unauthorized access to user data. This protection is essential for maintaining privacy and data security.

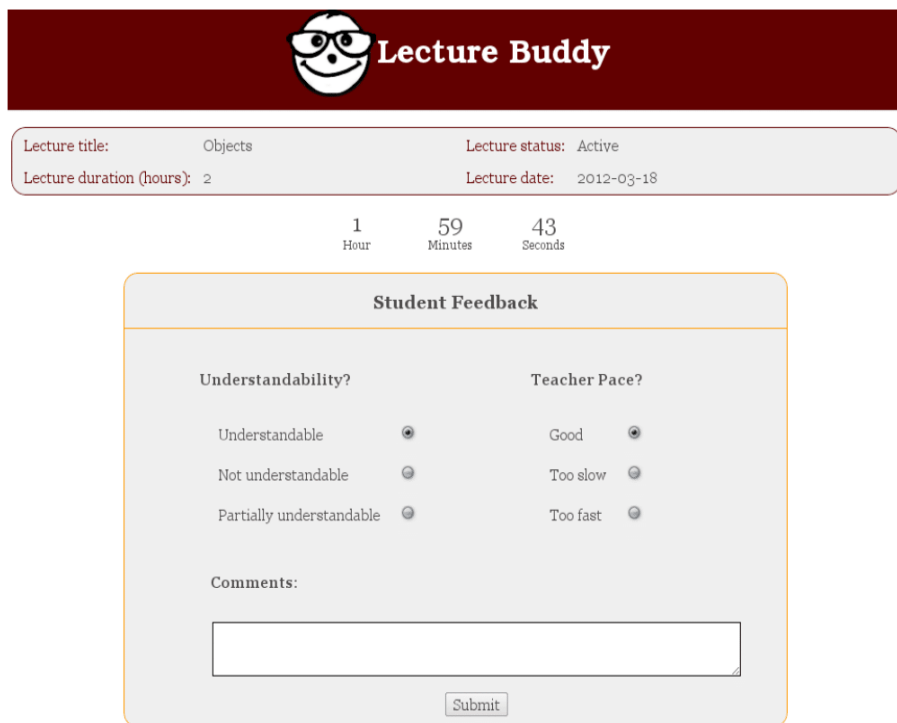


Figure 6. Lecture Buddy interface for student's view.

**Conclusion and Future Work:**

In this paper, we investigate the need for an innovative system to replace the existing course evaluation systems. We conducted two separate surveys of students and teachers. Based on the survey results, we concluded that instructors usually do not carry student's evaluations by themselves, and students prefer to provide more frequent feedback to teachers anonymously. Therefore, the majority of the student participants support the idea of a real-time anonymous

course evaluation system. We proposed and developed a prototype system named “Lecture Buddy,” which is an anonymous real-time course evaluation system for the students to give prompt feedback to their teachers during their classes.

Currently, we are using “Lecture Buddy” to obtain data and intend to perform another study to compare it with the traditional course evaluation system. There should be frequent evaluations with the help of a real-time automated system so the teacher can review student’s opinions at the end of every lecture because teachers need to know student’s learning during the semester rather than at the end. It will help instructors access their data electronically over time; they will be able to track their results across different courses.

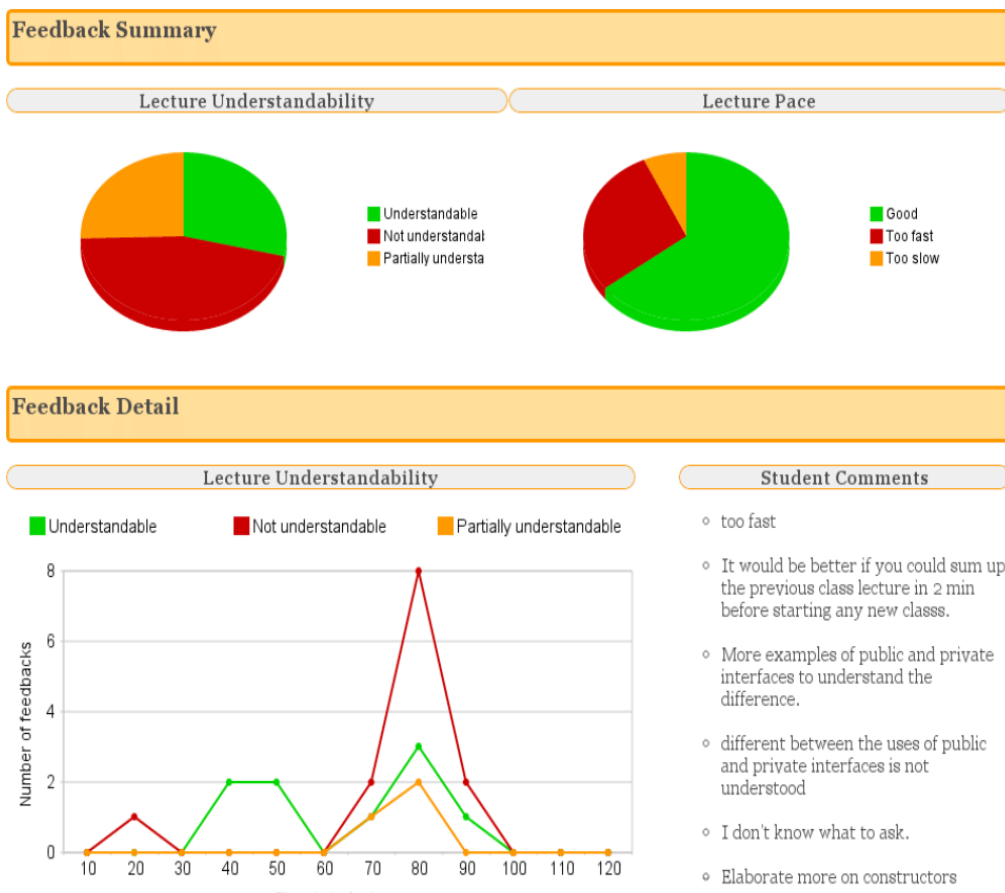


Figure 7. Lecture Buddy teacher’s view.

**References:**

- [1] C. Steyn, C. Davies, and A. Sambo, “Eliciting student feedback for course development: the application of a qualitative course evaluation tool among business research students,” *Assess. Eval. High. Educ.*, vol. 44, no. 1, pp. 11–24, Jan. 2019, doi: 10.1080/02602938.2018.1466266.
- [2] H. W. Marsh and D. Hocevar, “Student’s evaluations of teaching effectiveness: The stability of mean ratings of the same teachers over a 13-year period,” *Teach. Teach. Educ.*, vol. 7, no. 4, pp. 303–314, Jan. 1991, doi: 10.1016/0742-051X(91)90001-6.
- [3] O. Mitchell and M. Morales, “The effect of switching to mandatory online course assessments on response rates and course ratings,” *Assess. Eval. High. Educ.*, vol. 43, no. 4, pp. 629–639, May 2018, doi: 10.1080/02602938.2017.1390062.
- [4] A. S. Rosen, “Correlations, trends and potential biases among publicly accessible web-based student evaluations of teaching: a large-scale study of RateMyProfessors.com

- data,” *Assess. Eval. High. Educ.*, vol. 43, no. 1, pp. 31–44, Jan. 2018, doi: 10.1080/02602938.2016.1276155.
- [5] J. V. Adams, “Student evaluations: The ratings game.” 1997. Accessed: Sep. 18, 2023. [Online]. Available: <https://philpapers.org/rec/ADASET>
- [6] D. E. Clayson, “Student evaluation of teaching and matters of reliability,” *Assess. Eval. High. Educ.*, vol. 43, no. 4, pp. 666–681, May 2018, doi: 10.1080/02602938.2017.1393495.
- [7] A. VANACORE and M. S. Pellegrino, “An agreement-based approach for reliability assessment of Student’s Evaluations of Teaching,” *Third Int. Conf. High. Educ. Adv.*, p. 2017, Jun. 2017, doi: 10.4995/HEAd17.2017.5583.
- [8] H. W. Marsh, “Student’s Evaluations of University Teaching: Dimensionality, Reliability, Validity, Potential Biases and Usefulness,” *Scholarsh. Teach. Learn. High. Educ. An Evidence-Based Perspect.*, pp. 319–383, Jun. 2007, doi: 10.1007/1-4020-5742-3\_9.
- [9] B. Uttl, C. A. White, and D. W. Gonzalez, “Meta-analysis of faculty’s teaching effectiveness: Student evaluation of teaching ratings and student learning are not related,” *Stud. Educ. Eval.*, vol. 54, pp. 22–42, Sep. 2017, doi: 10.1016/J.STUEDUC.2016.08.007.
- [10] P. B. Stark and R. Freishtat, “An Evaluation of Course Evaluations,” *Sci. Res.*, vol. 0, no. 0, Sep. 2014, doi: 10.14293/S2199-1006.1.SOR-EDU.AOFRQA.V1.
- [11] L. McClain, A. Gulbis, and D. Hays, “Honesty on student evaluations of teaching: effectiveness, purpose, and timing matter!,” *Assess. Eval. High. Educ.*, vol. 43, no. 3, pp. 369–385, Jul. 2018, doi: 10.1080/02602938.2017.1350828.
- [12] K. Young, J. Joines, T. Standish, and V. Gallagher, “Student evaluations of teaching: the impact of faculty procedures on response rates,” *Assess. Eval. High. Educ.*, vol. 44, no. 1, pp. 37–49, Jan. 2019, doi: 10.1080/02602938.2018.1467878.
- [13] M. A. Bush, S. Rushton, J. L. Conklin, and M. H. Oermann, “Considerations for Developing a Student Evaluation of Teaching Form,” *Teach. Learn. Nurs.*, vol. 13, no. 2, pp. 125–128, Apr. 2018, doi: 10.1016/J.TELN.2017.10.002.
- [14] K. Sedova, M. Sedlacek, and R. Svaricek, “Teacher professional development as a means of transforming student classroom talk,” *Teach. Teach. Educ.*, vol. 57, pp. 14–25, Jul. 2016, doi: 10.1016/J.TATE.2016.03.005.
- [15] N. Denson, T. Loveday, and H. Dalton, “Student evaluation of courses: what predicts satisfaction?,” *High. Educ. Res. Dev.*, vol. 29, no. 4, pp. 339–356, Aug. 2010, doi: 10.1080/07294360903394466.
- [16] “What Can We Learn from End-of-Course Evaluations?” <https://www.facultyfocus.com/articles/faculty-development/can-learn-end-course-evaluations/> (accessed Sep. 18, 2023).
- [17] P. Brickman, C. Gormally, and A. M. Martella, “Making the grade: Using instructional feedback and evaluation to inspire evidence-based teaching,” *CBE Life Sci. Educ.*, vol. 15, no. 4, Dec. 2016, doi: 10.1187/CBE.15-12-0249/ASSET/IMAGES/LARGE/AR75FIG5.JPEG.
- [18] A. Boring, K. Ottoboni, P. B. Stark, and G. Steinem, “Student Evaluations of Teaching (Mostly) Do Not Measure Teaching Effectiveness,” *Sci. Res.*, vol. 0, no. 0, Jan. 2016, doi: 10.14293/S2199-1006.1.SOR-EDU.AETBZC.V1.
- [19] W. Stroebe, “Why Good Teaching Evaluations May Reward Bad Teaching,” <https://doi.org/10.1177/1745691616650284>, vol. 11, no. 6, pp. 800–816, Nov. 2016, doi: 10.1177/1745691616650284.
- [20] L. Mandouit, “Using student feedback to improve teaching,” *Educ. Action Res.*, vol. 26, no. 5, pp. 755–769, Oct. 2018, doi: 10.1080/09650792.2018.1426470.

- [21] R. J. Avery, W. K. Bryant, A. Mathios, H. Kang, and D. Bell, "Electronic Course Evaluations: Does an Online Delivery System Influence Student Evaluations?," *J. Econ. Educ.*, vol. 37, no. 1, pp. 21–37, Dec. 2006, doi: 10.3200/JECE.37.1.21-37.
- [22] A. Hoon, E. Oliver, K. Szpakowska, and P. Newton, "Use of the 'Stop, Start, Continue' method is associated with the production of constructive qualitative feedback by students in higher education," *Assess. Eval. High. Educ.*, vol. 40, no. 5, pp. 755–767, Jul. 2015, doi: 10.1080/02602938.2014.956282.
- [23] T. H. Reisenwitz, "Student evaluation of teaching: An investigation of nonresponse bias in an online context," *J. Mark. Educ.*, vol. 38, no. 1, pp. 7–17, 2016.
- [24] T. Standish, J. A. Joines, K. R. Young, and V. J. Gallagher, "Improving SET Response Rates: Synchronous Online Administration as a Tool to Improve Evaluation Quality," *Res. High. Educ.*, vol. 59, no. 6, pp. 812–823, Sep. 2018, doi: 10.1007/S11162-017-9488-5/METRICS.
- [25] F. Yang and F. W. B. Li, "Study on student performance estimation, student progress analysis, and student potential prediction based on data mining," *Comput. Educ.*, vol. 123, pp. 97–108, Aug. 2018, doi: 10.1016/J.COMPEDU.2018.04.006.
- [26] M. Chassignol, A. Khoroshavin, A. Klimova, and A. Bilyatdinova, "Artificial Intelligence trends in education: a narrative overview," *Procedia Comput. Sci.*, vol. 136, pp. 16–24, Jan. 2018, doi: 10.1016/J.PROCS.2018.08.233.
- [27] M. Amjad and N. Jahan Linda, "A Web Based Automated Tool for Course Teacher Evaluation System (TTE)," *Int. J. Educ. Manag. Eng.*, vol. 10, no. 2, pp. 11–19, Apr. 2020, doi: 10.5815/IJEME.2020.02.02.
- [28] R. Lalit, K. Handa, and N. Sharma, "FUZZY BASED AUTOMATED FEEDBACK COLLECTION AND ANALYSIS SYSTEM REEMA LALIT, KARUN HANDA and NITIN SHARMA," *Adv. Appl. Math. Sci.*, vol. 18, no. 8, 2019.
- [29] A. P. Cavalcanti et al., "Automatic feedback in online learning environments: A systematic literature review," *Comput. Educ. Artif. Intell.*, vol. 2, p. 100027, Jan. 2021, doi: 10.1016/J.CAEAI.2021.100027.
- [30] L. Tian and Y. Zhou, "Learner engagement with automated feedback, peer feedback and teacher feedback in an online EFL writing context," *System*, vol. 91, p. 102247, Jul. 2020, doi: 10.1016/J.SYSTEM.2020.102247.
- [31] E. Jensen et al., "Toward Automated Feedback on Teacher Discourse to Enhance Teacher Learning," *Conf. Hum. Factors Comput. Syst. - Proc.*, Apr. 2020, doi: 10.1145/3313831.3376418.
- [32] I. G. Ndukwe and B. K. Daniel, "Teaching analytics, value and tools for teacher data literacy: a systematic and tripartite approach," *Int. J. Educ. Technol. High. Educ.*, vol. 17, no. 1, pp. 1–31, Dec. 2020, doi: 10.1186/S41239-020-00201-6/FIGURES/6.
- [33] I. Sindhu, S. Muhammad Daudpota, K. Badar, M. Bakhtyar, J. Baber, and M. Nurunnabi, "Aspect-Based Opinion Mining on Student's Feedback for Faculty Teaching Performance Evaluation," *IEEE Access*, vol. 7, pp. 108729–108741, 2019, doi: 10.1109/ACCESS.2019.2928872.
- [34] Q. Lin, Y. Zhu, S. Zhang, P. Shi, Q. Guo, and Z. Niu, "Lexical based automated teaching evaluation via student's short reviews," *Comput. Appl. Eng. Educ.*, vol. 27, no. 1, pp. 194–205, Jan. 2019, doi: 10.1002/CAE.22068.



Copyright © by authors and 50Sea. This work is licensed under Creative Commons Attribution 4.0 International License.