Pre-Evaluation with a Personalized Feedback Conversational Agent Integrated in Moodle

https://doi.org/10.3991/ijet.v18i06.36783

Wijdane Kaiss[1,2(✉)], Khalifa Mansouri[1], Franck Poirier[2]
[1]Laboratory Signals, Distributed Systems and Artifical Intelligence, ENSET Mohammedia, University Hassan II of Casablanca, Casablanca, Morocco
[2]Lab-STICC, University Bretagne Sud, Vannes, France
wijdane.kaiss@univ-ubs.fr

Abstract—Pre-evaluation of the learner's level is a common learning strategy designed to determine the prior knowledge and skills of learners. A pre-evaluation is carried out at the beginning of the course and based on the results obtained, personalized resources will be provided that respond to individual learner needs. This paper presents a pre-evaluation for a C programming language course by providing, at the end of the quiz, a personalized formative feedback and recommendation to the learners. We have developed our conversational chatbot named QuizCbot, which allows learners to go directly to the parts where they need the most help through the personalized feedback provided to them, including their final scores, the questions they answered correctly and the questions they answered incorrectly with the correct answer and explanation. Hence, the chatbot makes a recommendation on the concepts in which the learner did not obtain the average, identifying the concepts not mastered where the learner needs more (or less) support. Determining what learners know and don't know can help to improve the learning experience. We have integrated our QuizCbot chatbot, which is based on Natural Language Understanding (NLU), into the Moodle learning environment.

Keywords—evaluation, conversational agent, chatbot, personalized feedback, LMS, moodle

1 Introduction

Evaluation plays a critical role in providing meaningful information to guide teaching, help learners achieve next steps, and verify accomplishments and progress. Evaluation should be planned with these aims: evaluation for learning, evaluation as learning, and evaluation of learning, each has a role to play in supporting and enhancing student learning. Evaluation for learning informs teachers and the online learning environment about what learners understand and allows them to know what learning objects to adjust and to plan and guide teaching while providing useful feedback to learners. Evaluation as learning allows learners to become aware of how they are learning, to adjust and progress their learning with increased responsibility. Evaluation of learning allows learners, teachers and the online learning environment to be informed of the learning outcomes achieved at a specific time to highlight successes, plan interventions and continue to promote success.

Three of the most popular types of evaluation are: Pre-evaluation, Formative evaluation, Summative evaluation. Before starting learning, it is essential to know what type of students it is for, by situating their level of knowledge through the Pre-evaluation. Tracking learners' progress through Formative evaluations. A Summative evaluation aims to assess if the most important knowledge has been acquired at the end of the training, in order to adjust the continuation of his/her teaching according to the level reached.

Formative feedback is considered an essential approach to facilitating the development of students as independent learners to monitor, evaluate, and regulate their own learning [1]. Formative feedback involves the feedback given to learners during evaluations to improve their learning [2]. Thus, helping learners who need more support, and recommending them the contents, can help to improve the learning experience [3].

In this article, we will address the first type of evaluation to determine the level of knowledge of learners and rank them in order to identify individual difficulties with the diversity of learners to provide individualized formative feedback and recommendations.

Determining the level of knowledge of each learner and providing individualized feedback and recommendations would be difficult to do by teachers who do not have the time or availability for this task.

To solve this problem, the new form of evaluation is made, which is the use of conversational chatbot. In most of the previous researches, the simple chatbot solutions are widely used in many fields such as medicine, product and service industries. However, the use of chatbots in the educational context is still limited [4], [5]. With the integration of our conversational chatbot, multiple choice questions become amusing and interactive.

The objective of this contribution is to improve and personalize the learner's learning experience based on a pre-evaluation to determine what learners know and don't know and situate them in the appropriate level and recommend them to the appropriate content and provide them with personalized feedback. The solution developed is a conversational chatbot that uses Natural Language Understanding (NLU) to make the learning process engaging and motivating for learners. To facilitate the use of our solution we have integrated our chatbot, named QuizCbot, into the Moodle platform.

The rest of this paper is organized as follows: section 2 describes the state of the art, section 3 describes the architecture and design of our solution, section 4 explains the details of the experimental design for the C language programming quiz and population and we will discuss the results obtained and section 5 is reserved for the conclusion.

2 Related works

2.1 Quizzes and feedback

Due to the intense critique of evaluation and its inability to reflect the real state of learners' competence level in a particular subject or skill due to problems such as test validity, the theory of evaluation was developed [6]. In this context, the learner will receive formative feedback before, during, or after an evaluation. The feedback should then be adapted to the user, the task and the environment in order to facilitate the learner's improvement in a subject or a skill.

In general, the feedback has one of the most persuasive influences on performance and learning, but this impact can be negative or positive [7]. Educational research has revealed dependencies between variables such as the learner's skill level, the learner's motivation, the properties of the task and the effectiveness of feedback [6]. Additionally, it has been determined that adaptive feedback is more beneficial to the learner than generic feedback. When a learner uses the information he/she receives in the form of feedback to enhance his/her performance, this is called the formative feedback [8]. There are mainly two types of feedback, called formative and summative feedback.

Formative feedback includes feedback given to learners during evaluations to improve their learning [2]. To be effective, formative feedback must be constructive, timely, personal and motivating. There are two types of formative feedback, namely directive and facilitative. To let students know about their mistakes and learning improvements, directive feedback is used. The facilitative feedback guides learners in their revision by providing suggestions, advice and comments.

Summative feedback involves feedback given in the form of a mark or grade after an evaluation has been completed. It assesses what students have learned at the end of a subject or semester [2]. A study indicates that summative evaluations allow learners to study more, learn more, and feel more inspired. Along with the grade, summative evaluation also ensures that learners have certain knowledge, abilities, and skills where strengths and weaknesses are identified [9].

In [10], study the effect of correct/incorrect feedback in generic quizzes and they demonstrated that simple quizzes combined with simple feedback are effective. This gives an idea about the ideal placement of this type of activity.

The validity of the results of the tests is an important prerequisite for meaningful feedback. Evaluation based on multiple-choice questions allows learners to easily guess the correct answer. For example, the study [11] demonstrates how multiple-choice questions can be passed when learners simply guess the answers. Therefore, the success of such a form of evaluation depends on the honesty of the learner and is therefore suitable as a self-feedback tool without any consequences for official test scores or course results. In such cases, an opportunity to reflect on their self-confidence can create an environment for learners in which they develop a deeper level of self-regulation.

The study presented in [12] illustrates that quizzing helps learners grasp more information than rereading. This is also known as the "test effect" or "retrieval practice". The authors considered the concepts of dynamic testing and evaluation to enhance learning.

2.2 Chatbots in education

Conversational AI provides new possibilities for alternative and innovative information and communication technologies tools, such as AI chatbots. The explosive development of information and communication technology will inevitably have a profound impact on every sector, including education [13], [14]. Today, integrating information and communication technology to facilitate online learning is crucial [15].

Artificial intelligence (AI) chatbots have gained popularity over time and have been widely used in e-commerce, online banking, and healthcare, among others. In general, the use of chatbots is gaining popularity in several sectors.

However, in the field of education, chatbot as a learning tool to improve learning is still in its infancy [16]. Most studies on chatbot in education are based on online foreign language learning [17], [18]. It has been identified that a chatbot can help learning with the same benefits as those obtained from a "real" interview [19]. In addition to making the learning process enjoyable, the conversational chatbot could make learners more likely to self-evaluate and improve their level, because they do not feel judged.

Our approach is to evaluate students through the chatbot in order to track each learner individually, to guide them based on their deficiencies to the appropriate content, and to provide them a personalized feedback based on the results of the evaluation. The following section describes the architecture and design of our solution.

3 Proposed design and architectural design

Figure 1 describes the schematic representation of our chatbot that allows to have a conversation with learners by detecting their intentions based on natural language understanding (NLU). We developed our conversational chatbot named QuizCbot that performs a pre-evaluation, online ranking test, for a C programming language course integrated into the Moodle platform to evaluate learners' skills.

Our chatbot is based on the open source version of Google's Dialogflow Machine Learning framework that allows users to develop human-computer interaction technologies that can handle natural language understanding (NLU). Basically, this allows us to create digital programs that interact with the final users via natural languages.

Fig. 1. Schematic representation of our chatbot

When the learner sends a message to QuizCbot in Moodle to start the quiz, the chatbot transmits it to Dialogflow which tries to understand the received text by associating it to an intention using Natural Language Understanding. Once the intention is detected by Dialogflow, an action to be executed is chosen. The Webhook sends a formatted response corresponding to the intention. All actions are coded in Node.js and hosted by the fulfillment Webhook. The fulfillment webhook is a service that allows a dynamic response by searching for response elements in an external database. At the Webhook stage, questions and answers are processed, the learner's level is detected (learner's classification as shown in Figure 2), a recommendation and a personalized feedback are provided to the learner.

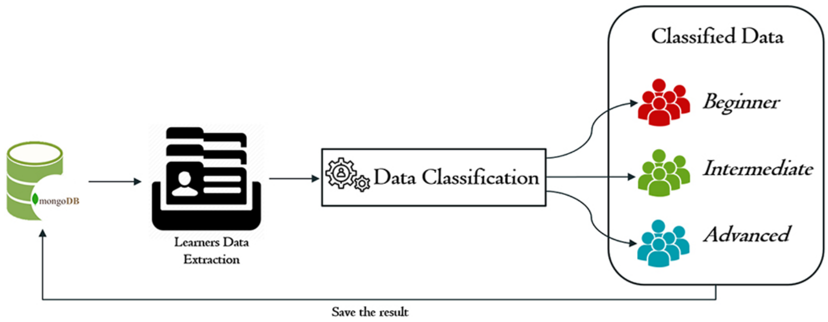

Fig. 2. Learner classification process

The webhook records all this information and data in the Mongodb database. Figure 3 shows the interaction between the different components.

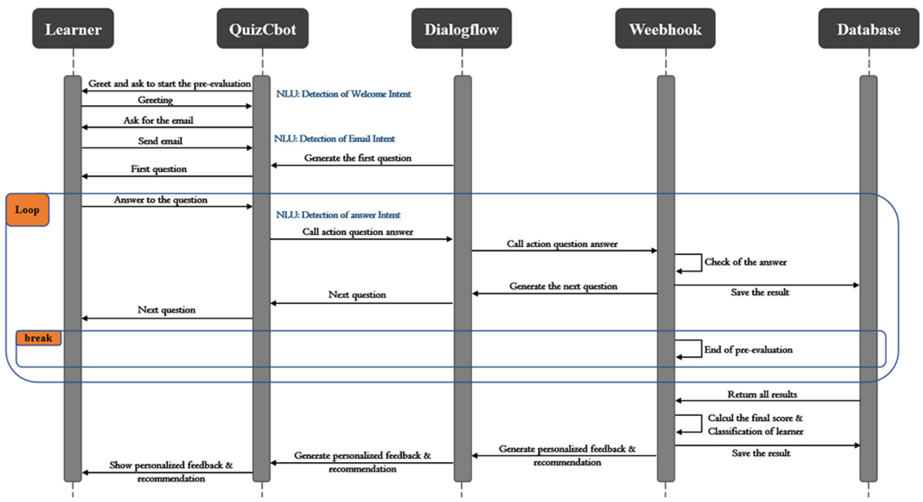

Fig. 3. Interaction between the different components

4 QuizCbot evaluation

This section describes the evaluation of the QuizCbot in terms of the participants' perception of the usefulness of the personalized feedback provided to them and the recommendation that was given. Each learner was asked to participate in a survey at the end of the quiz. Section 4.1 explains the details of the experimental design for the C language programming quiz and the population. In Section 4.2 we discuss the results obtained.

4.1 Design of the experiment

We have chosen the Ecole Normale Supérieure de l'Enseignement Technique de Mohammedia (ENSET), an engineering school in Morocco to conduct our experiment. The participants are first-year engineering students in the Software Engineering and Distributed Computing Systems (GLSID) and Computer Engineering Big Data and Cloud Computing (BDCC) fields. The experiment was conducted on the C programming techniques module.

The pre-evaluation consists of 40 questions of different levels of difficulty and in different concepts of the course (see Table 1), validated by the professor experienced in the C programming language, to correctly classify the level of knowledge of the learner and then assign him/her to the appropriate group.

Table 1. The different course concepts addressed in QuizCbot

20 questions classified as Beginner Level	– Introduction to programming in C – Primitive types and variable declarations – Affectation instruction – Arithmetic operators – Increment/decrement operators – Relational operators – Logical operators – Standard Input/Output functions – The Control Structures
10 questions classified as Intermediate Level	– The arrays – The Pointers – The structures – The functions
10 questions classified as Advanced Level	– Pointers and dynamic memory allocation (Function pointer, Array of pointers, Dynamic memory management, …) – Unions – File management in C

We have integrated our QuizCbot chatbot into Moodle and the students of GLSID and BDCC access in, before starting the course, to pass the pre-evaluation on the C programming language via our chatbot to determine their level of knowledge and at the end of the test, it provides them recommendations on the concepts not mastered where the learners need more (or less) support and provides them a personalized formative feedback, i.e., giving the learners the opportunity to use the feedback information to regulate their learning process to improve their skills.

The sample included 71 students (52 males and 19 females), who participated in this study and passed the test on our chatbot QuizCbot. As mentioned above, the participants are first-year engineering students in the Software Engineering and Distributed Computing Systems (GLSID) and Computer Engineering Big Data and Cloud Computing (BDCC) fields, and are between the ages of 20 and 21 years old. Participants were allowed to use the QuizCbot on any device of their choice. Although the chatbot offers the option to restart the quiz, only the first fully completed attempt is considered for this study. An important observation is that none of the students needed a demonstration or explanation on how to use the chatbot, adoption of the technology was on the fly due to familiarity with messengers and quizzes.

4.2 Results and discussion

As mentioned above, learners were allowed to use QuizCbot on any device of their choice. As shown in Figure 4, an example of a learner connected through their mobile phone (Figure 4a), and another learner connected through their computer (Figure 4b).

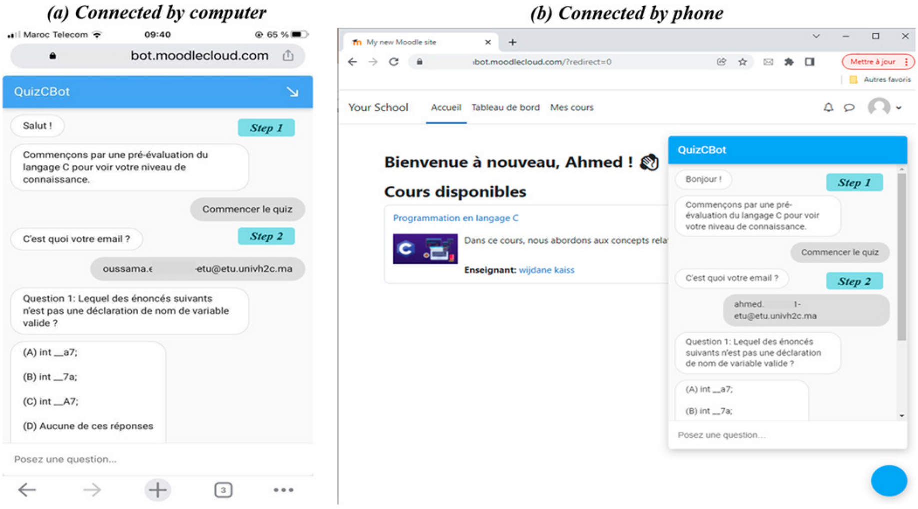

Fig. 4. Student identification

In both cases, the chatbot greets the learner and asks him/her to start the quiz in step 1 and then, to identify the learner, the chatbot asks the learner for his/her email in step 2. When the learner provides it, the chatbot starts the conversational quiz permitting the pre-evaluation in the C programming language, as shown in Figure 5, with the associated choices. The learner answers by making sentences or by giving the options (A, B…).

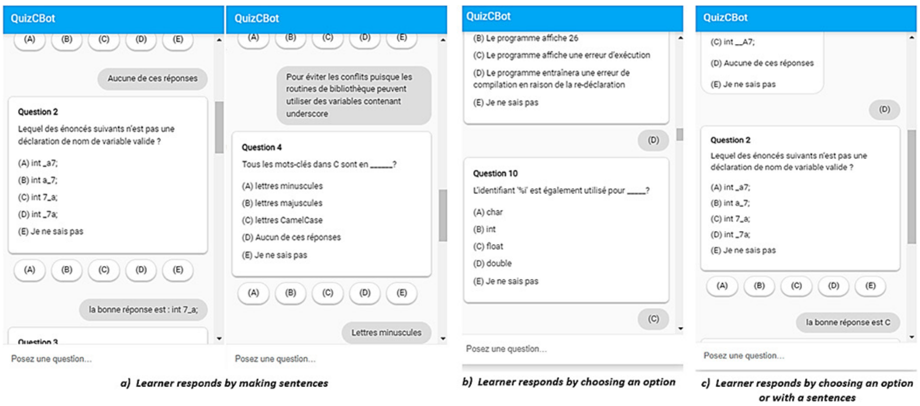

Fig. 5. Pre-evaluation in C programming language

The chatbot successfully retrieved each student's score for each question and stored it in the database (Mongodb database), and then it provided learners personalized feedback at the end of the test, as shown in Figure 6, including his/her final score, the questions he/she answered correctly, and the questions he/she answered incorrectly, and it also provided a QuizC-Solution.pdf document containing the answers to the questions with explanation.

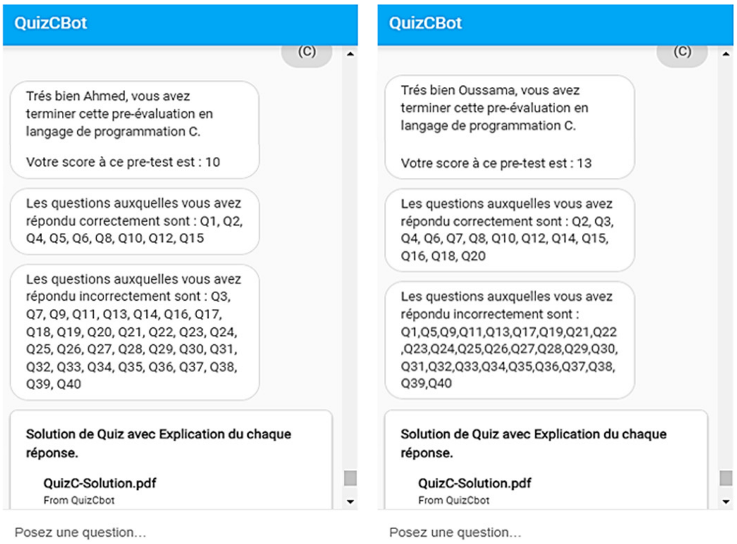

Fig. 6. An example of personalized feedback

We then categorized each student into the appropriate level (beginner, intermediate, advanced) to provide personalized resources that respond to individual learners' needs. Figure 7 illustrates three histograms of learners' scores for questions answered correctly at different levels (left histogram for beginner level, middle histogram for intermediate level, and right histogram for advanced level).

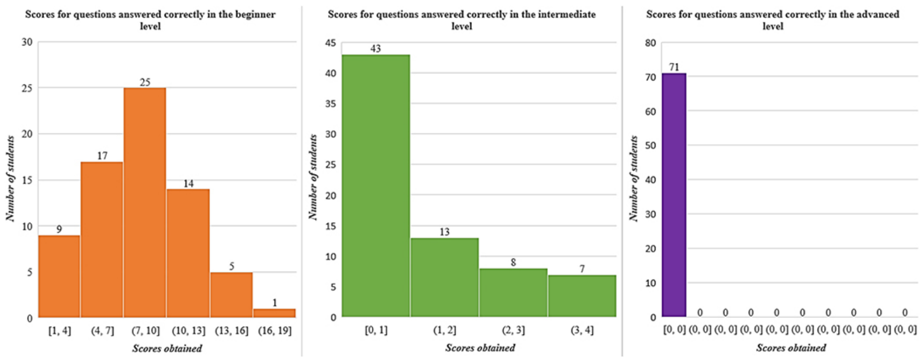

Fig. 7. Learners' scores for correctly answered questions at different levels

Analyzing these results, we can see that for the advanced level questions, no students answered correctly. For intermediate level questions, 43 of the learners scored 0 or 1 out of 10, 13 learners scored 2 out of 10, 8 learners scored 3 out of 10, and only 7 out of 71 learners scored 4 out of 10. And for the beginner level questions, 51 of the learners scored between 1 and 10 out of 20, 19 learners scored between 11 and 15 out of 20, and only one learner scored 16 out of 20, so all of these learners still lack mastery of some concepts at this beginner level in C programming language. Hence the conclusion of this part is that all learners are clearly classified at the beginner level.

But we are not stopping there because, for example, there is a learner of beginner level who has a score of 16 out of 20 (he/she still needs support for some notions), and another learner in the same level who has a score of 7 out of 20, the concepts that these two learners master are completely different even if they are classified in the same level. Therefore, since our approach aims to personalize learning subsequently by providing each learner with concepts not mastered, we decided to divide this beginning level into three supplementary levels (low, medium, high) in order to provide each learner with recommendations on the concepts that need more support according to the level he/she belongs.

We wanted to do this method to perfectly personalize the recommendation. Figure 8 shows a pie chart on the learners' final classification. There are 51 learners classified as low beginner level, 18 learners as medium beginner level and only two learners as high beginner level.

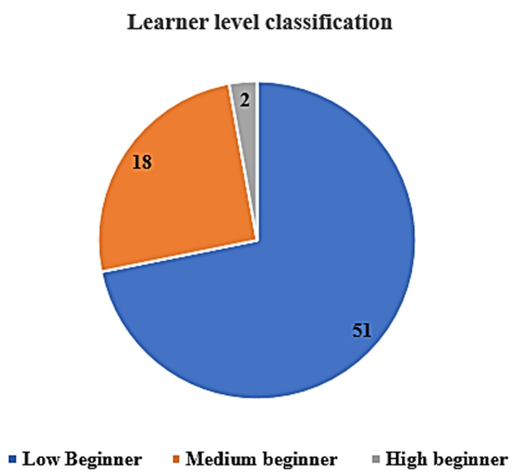

Fig. 8. Learner level classification

At the end of this pre-evaluation, after our chatbot provides personalized feedback to each learner and ranks them in the level they belong to, the chatbot recommends the concepts in which the learner did not get the average (see Figure 9), identifying the non-mastered concepts where learners need more (or less) support only in the level in which he/she is classified. This means that, for example, if a learner is classified in the beginner level, our chatbot will first recommend the concepts in which he/she did not get the average in the level he/she belongs to, no need to recommend him/her, for example, to the intermediate level concepts and he/she has not mastered the beginner level concepts yet.

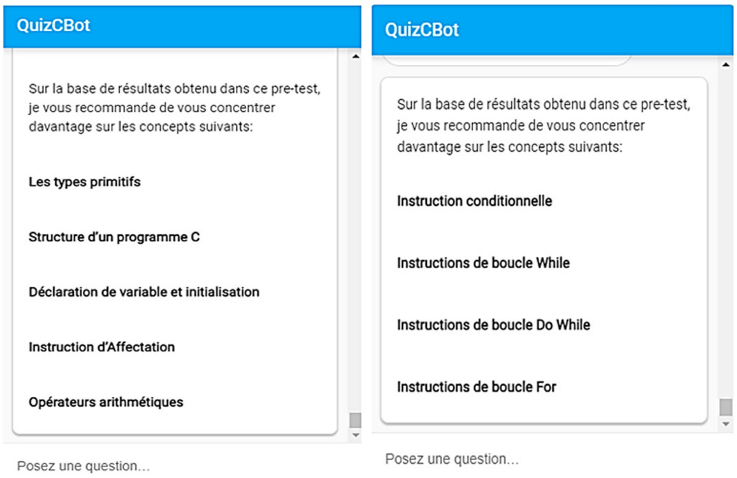

Fig. 9. Examples of personalized recommendations

All students then answered the questions in Table 2 at the end of the test. The results show a high degree of acceptance and a positive attitude towards a chatbot-based evaluation from all students.

Table 2. QuizCbot evaluation questions

Questions	Very Interesting	Interesting	Not at all Interesting
How would you rate the feedback provided by QuizCbot?	57%	43%	0%
How would you rate the recommendation provided by QuizCbot?	97%	3%	0%
How would you rate the explanation of wrong answers provided by the chatbot?	98%	2%	0%
How would you rate the quiz in chat format?	86%	14%	0%
How would you rate the overall experience with QuizCbot?	95%	5%	0%

In all survey questions, all students selected the very interesting or interesting responses. Table 2 shows the distribution of responses to each survey question. Based on these results, we conclude that our chatbot QuizCbot was perceived as helpful and interesting, which shows a positive attitude of learners towards artificial learning companions.

5 Conclusion

In this paper, we have presented a research study involving the design and implementation of a pre-evaluation chatbot, integrated into Moodle, that allows learners to pass a test at the beginning of the course to determine and classify their knowledge level. Based on the results obtained of the evaluation, our conversational chatbot named QuizCbot provides personalized feedback to learners including their final score, the questions they answered correctly and the questions they answered incorrectly with the correct answer and explanation. It also provides a recommendation for the concepts in which the learner did not achieve the average, identifying the concepts not mastered where learners need more (or less) support. However, on the basis of the results obtained from the survey conducted to evaluate QuizCbot, we concluded that our chatbot QuizCbot was perceived as interesting and helpful and shows a positive attitude from learners towards the chatbot-based evaluation. The learners also mentioned that the feedback, explanations of wrong answers, and recommendations provided by QuizCbot are interesting or very interesting to them. It allows learners to self-evaluate and can create an environment in which they improve their learning experience, which allows them to become aware of how they learn, adjust and advance their learning by assuming more responsibility. For our future work, we will address the design and implementation of a chatbot to personalize the formative evaluation questions for each learner according to the level in which he/she belongs, in order to guide him/her towards the resources most adapted to his/her needs.

6 Acknowledgment

This work is supported by the University Hassan II of Casablanca-Morocco and Professor Mohammed Qbadou.

7 References

[1] Md. Mamoon-Al-Bashir, Md. R. Kabir, and I. Rahman, "The Value and Effectiveness of Feedback in Improving Students' Learning and Professionalizing Teaching in Higher Education," *Journal of Education and Practice*, vol. 7, pp. 38–41, 2016.

[2] K. Dekker, "Effective Learning Methods: The Importance of Formative Feedback," *2018*. https://www.studiosity.com/blog/effective-learning-methods-the-importance-of-formative-feedback (accessed Oct. 20, 2022).

[3] N. Manouselis, H. Drachsler, K. Verbert, and E. Duval, "Recommender Systems for Learning," 2013, https://doi.org/10.1007/978-1-4614-4361-2

[4] R. Winkler and M. Soellner, "Unleashing the Potential of Chatbots in Education: A State-Of-The-Art Analysis," *Academy of Management Proceedings*, vol. 2018, no. 1, p. 15903, Aug. 2018, https://doi.org/10.5465/AMBPP.2018.15903abstract

[5] S. Yang and C. Evans, "Opportunities and Challenges in Using AI Chatbots in Higher Education," ICEEL 2019: 2019 3rd International Conference on Education and E-Learning, pp. 79–83, Nov. 2019, https://doi.org/10.1145/3371647.3371659

[6] P. Black and D. Wiliam, "Developing the Theory of Formative Assessment," *Educ Assess Eval Account*, vol. 21, no. 1, pp. 5–31, Feb. 2009, https://doi.org/10.1007/s11092-008-9068-5

[7] J. Hattie and H. Timperley, "The Power of Feedback," *Rev Educ Res*, vol. 77, no. 1, pp. 81–112, 2007, https://doi.org/10.3102/003465430298487
[8] D. Wiliam, "Embedded Formative Assessment," p. 225, Accessed: Oct. 20, 2022. [Online]. Available: https://www.solutiontree.com/embedded-formative-assessment-second-ed.html
[9] J. States, R. Detrich, and R. Keyworth, "Summative Assessment (Wing Institute Original Paper)," Project: Wing Institute Original Papers, 2018, https://doi.org/10.13140/RG.2.2.16788.19844
[10] O. Bälter, E. Enström, and B. Klingenberg, "The Effect of Short Formative Diagnostic Web Quizzes with Minimal Feedback," *Computers & Education*, vol. 60, no. 1, pp. 234–242, Jan. 2013, https://doi.org/10.1016/j.compedu.2012.08.014
[11] P. Novacek, "Confidence-Based Assessments within an Adult Learning Environment.," *International Association for Development of the Information Society*, Oct. 2013.
[12] H. Roediger and J. D. Karpicke, "The Power of Testing Memory: Basic Research and Implications for Educational Practice," *Perspectives on Psychological Science*, vol. 1, no. 3, pp. 181–210, 2006, https://doi.org/10.1111/j.1745-6916.2006.00012.x
[13] L. L. Khoroshko, P. Ukhov, and A. L. Khoroshko, "The Use CAD/CAE Systems to Create E-Learning Courses on Technical Subjects at University," *International Journal of Engineering Pedagogy (iJEP)*, vol. 8, no. 2, pp. 64–71, 2018. https://doi.org/10.3991/ijep.v8i2.8134
[14] S. Svetsky, O. Moravcik, P. Tanuska, and I. Markechová, "The Personalized Computer Support of Teaching," *International Journal of Engineering Pedagogy (iJEP)*, vol. 8, no. 4, 2018. https://doi.org/10.3991/ijep.v8i4.8149
[15] L. N. Fewella, L. M. Khodeir, and A. Suidan, "Impact of Integrated E-learning: Traditional Approach to Teaching Engineering Perspective Courses," *International Journal of Engineering Pedagogy (iJEP)*, vol. 11, no. 2, pp. 82–101, 2021. https://doi.org/10.3991/ijep.v11i2.17777
[16] D. E. Gonda, J. Luo, Y. L. Wong, and C. U. Lei, "Evaluation of Developing Educational Chatbots Based on the Seven Principles for Good Teaching," Conference: 2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE), pp. 446–453, Jan. 2018, https://doi.org/10.1109/TALE.2018.8615175
[17] A. M. Rahman, A. al Mamun, and A. Islam, "Programming Challenges of Chatbot: Current and Future Prospective," 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), vol. 2018, pp. 75–78, Feb. 2017, https://doi.org/10.1109/R10-HTC.2017.8288910
[18] S. Ruan *et al.*, "BookBuddy: Turning Digital Materials Into Interactive Foreign Language Lessons Through a Voice Chatbot," the Sixth (2019) ACM Conference, Jun. 2019, https://doi.org/10.1145/3330430.3333643
[19] J. Q. Pérez, T. Daradoumis, and J. M. M. Puig, "Rediscovering the Use of Chatbots in Education: A Systematic Literature Review," *Computer Applications in Engineering Education*, vol. 28, no. 6, pp. 1549–1565, Nov. 2020, https://doi.org/10.1002/cae.22326

8 Authors

Wijdane Kaiss is a PhD student in Computer Science at the University of Bretagne Sud in France, and at the University Hassan II in Morocco. Her research aims to enhance learning experience in an e-learning system. (email: wijdane.kaiss@univ-ubs.fr).

Khalifa Mansouri is a teacher of computer science and researcher at the University Hassan II Casablanca, ENSET Institute. His research is focused on Real Time Systems,

Information Systems, e-Learning Systems, and Industrial Systems (Modeling, Optimization, and Numerical Computing). He earned a Diploma ENSET Mohammedia in 1991, CEA in 1992 and PhD (Calculation and Optimization of Structures) in 1994 from Mohammed V University in Rabat, HDR in 2010, and Ph.D (Computer Science) in 2016 from Hassan II University in Casablanca. (email: khmansouri@hotmail.com).

Franck Poirier is Professor of Computer Science at Université Bretagne Sud (UBS). He is co-head of CAPE (Collaboration Assessment Personalisation for Education) research team and head of the Mathematics, Computer Science and Statistics department of the Faculty of Science at UBS. He has worked in Human-Computer Interaction (HCI) for over 30 years and more recently in e-Education and Augmentative and Alternative Communication (AAC). Currently his research concerns recommander systems, student satisfaction, self-regulated learning, e-learning success model, structural equation modeling, and learner experience (email: franck.poirier@univ-ubs.fr).

Article submitted 2022-11-21. Resubmitted 2023-01-24. Final acceptance 2023-01-25. Final version published as submitted by the authors.