R&E SOURCE

Vol. 12 (2025), Special Issue 1
To the Roots: Engineering Pedagogy

Application of Automatic Evaluation Systems for Flipped Classroom

József Udvaros[1], *Ildikó Pšenáková*[2]

DOI: https://doi.org/10.53349/re-source.2025.is1.a1411

Abstract

Automatic evaluation systems play a significant role in the methodology of flipped classroom, as they provide the opportunity for continuous, immediate feedback to students, thus increasing learning efficiency. In this article, we present how these systems can be used in the environment of flipped classroom, with particular attention to the teaching of programming and other technological fields. Automated assessment tools, such as code review and testing platforms, allow students to learn at their own pace while allowing classroom time to be spent on interactive activities. The study explores the benefits and challenges of the systems and their impact on student performance and motivation. The results show that automatic assessment systems can increase student engagement, improve learning outcomes and promote the effectiveness of flipped classroom.

Keywords: Automatic Evaluation Systems, Flipped Classroom, Teaching Methods, Educational Technologies, Innovation

1 Introduction

The development of automated evaluation systems is one of the fastest changing areas of educational technology, especially in programming and technology education. The first generation of systems started to spread in the 1960s and 1970s and focused mainly on detecting syntactic errors and simple code validation. Such systems provided limited feedback, merely indicating when there was an obvious error in syntax or logical structure.
Since the 1980s, however, evaluation systems have evolved to include semantic analysis of programs, i.e. the logical correctness of programs as a function of their inputs. At the same

[1] University of Trnava in Trnava, Faculty of Education, Department of Mathematics and Computer Science, Priemyselná 4, 918 43 Trnava, Slovakia.
E-Mail: *jozsef.udvaros@truni.sk*
[2] University of Trnava in Trnava, Faculty of Education, Department of Mathematics and Computer Science, Priemyselná 4, 918 43 Trnava, Slovakia.

time, 'online judge' systems for programming competitions appeared, which automatically evaluated solutions based on predefined test cases.

From the 2000s onwards, the systems also considered the runtime and memory usage of programs, i.e. their efficiency, which became particularly important for larger data processing systems and programming competitions. Newer platforms also compared algorithms with different efficiencies, highlighting solutions that were correct but less efficient – so students could now aim for optimised solutions.

The emergence of systems that integrate formative feedback has added a new dimension to evaluation systems, providing not only summative but also developmental feedback. These tools, such as code editors that provide syntactic highlighting and auto-completion, visualisation of the program's operation and a partial scoring system, support learners in correcting errors and improving solutions (Pacheco-Venegas et al., 2015).

In recent years, systems have supported multiple programming languages and provide the possibility to manage more complex projects where the code consists of multiple files and modules. Cloud computing allows students to access them from anywhere, which is a significant facilitator for distance learning and continuous practice (de la Peña et al., 2012).

Recent developments integrate artificial intelligence and machine learning tools, and systems now provide personalised feedback, recognising typical errors and learning paces to provide more targeted support. These systems motivate students with challenges and practical tasks that support individual progress (Flament et al., 2021).

Modern systems provide teachers with detailed statistical analyses and learning reports, giving them a comprehensive picture of students' progress, failures and strengths. And they provide students with continuous feedback on their performance and areas where further practice is needed, making the learning process more effective and targeted.

The development of automated evaluation systems provides comprehensive support for both students and teachers, enabling personalised, immediate and continuous feedback that motivates students and improves learning effectiveness. These systems are increasingly contributing to the renewal of digital education, particularly in programming and technology training.

The flipped classroom methodology is becoming increasingly widespread in education, as it fundamentally changes the traditional structure of learning and teaching. The methodology involves taking some parts of the learning process outside the classroom, while focusing on collaborative work and interactive, experiential activities in the classroom. Prior acquisition and processing of knowledge takes place at home, where students usually learn the basics through digital resources such as pre-recorded videos, online materials or other learning aids. This approach allows classroom time to be used more efficiently, as students arrive with prior knowledge and can work through the material through more in-depth, practice-oriented exercises, discussions, projects and other interactive activities, guided by the teacher. The flipped classroom methodology is particularly suitable in areas such as programming,

mathematics or other technology subjects where practical application and problem solving are central (Pšenáková et al., 2024).
Another advantage of the method is that it gives students the opportunity to learn the basics at their own pace, so they can spend more time working together on more difficult, complex problems in the classroom. It also supports the development of independent learning skills and students' ownership of their own learning process, which has a positive long-term impact on learning outcomes and motivation. New challenges in this kind of unconventional education should also consider students' rights (Bódi, 2023).
The aim behind the popularity of the flipped classroom methodology is to better adapt the educational process to the individual needs of students, facilitating active participation and in-depth knowledge acquisition. The integration of online learning technologies, automated evaluation systems and interactive learning materials further enhances the effectiveness of the methodology and allows for personalisation of the learning process, while maximising classroom time for collaborative activities and individual feedback (Szabó & Pšenáková, 2017a; Szabó & Pšenáková, 2017b; Pšenáková, 2019).
In this article, we address the following research question: How can automated evaluation systems support flipped classroom methodology in programming and technology education?

2 Theoretical Background and Literature Review

Modern tools for learning computer programming include learning environments and automatic evaluating systems. These learning environments include basic programming environments, example-based interactive environments, visualisation/animation systems and intelligent tutoring systems (Gabaľová et al., 2023; Stoffova et al., 2023; Gabaľová, 2022). However, most of these systems lack evaluation tools that can be useful for formative evaluation of students during practical tasks. Automatic evaluation systems, on the other hand, often focus only on the evaluation of programs (online judges) and usually adapt the evaluation methods of programming competitions. Although these systems are effective in supporting the development of practical skills, the spirit of competition is motivating for students, but they are not suitable for formative assessment. This is due to limited feedback and lack of integration with the content and tools of the learning environment.
Educational environments provide comprehensive and developmental feedback to evaluate the functionality of computer programs. For a given programming task, students are asked to produce a solution, which is checked using the automatic evaluation tool of the learning environment. Summative feedback is provided through the analysis of the program, giving feedback on the correctness of syntax, semantics and efficiency, as well as possible errors, and a grade is assigned to the submitted solutions based on this feedback. In addition, the environment also provides formative feedback through a variety of tools to support students in improving their solutions before the assessment. These tools include a code editor for syntactic highlighting and auto-completion, automatic checking of correct programming

practices, interactive visualisation of programs from source code, a partial scoring system that gives partial points for incomplete solutions, a test module with customisable inputs, multiple file and project submission options, and support for multiple programming languages (Java, Python, JavaScript, C/C++/C#), as well as a manual evaluation module and interactive statistical reports for students and teachers (Restrepo-Calle et al., 2018).

The wider introduction of programming courses in educational institutions poses significant challenges in evaluating the diversity of students. Traditional, manual grading methods are increasingly difficult to maintain due to their demanding nature and the potential for human error and bias. In response, we have developed an automated evaluation system adapted to the complexity of modern programming education, which draws heavily on previous work in developing similar systems (Pietrikova et al., 2015). This system is designed to provide continuous and regular feedback, like the iterative processes in professional software development environments where code is regularly tested and improved. This approach not only supports the incremental development of programming skills but also responds to ongoing changes in industry standards.

The security and privacy of automated systems is key, as they handle sensitive student data and are exposed to potential breaches, including threats from malicious code submissions (Horváth, 2024).

Automated evaluation is a challenging area for the academic community, which plays an important role in reducing the evaluation workload. The authors of this paper present a new feature of the Easy Java Simulations (EJS) tool, the Automatic Evaluation Element (AEE), which provides automatic evaluation for virtual and remote labs created using EJS through the Goodle Grading Management System (GMS) server application. The integration of the two tools allows instructors to create interactive virtual and remote labs and automatically grade student work (Farias et al., 2016).

Automatic evaluation systems are playing an increasingly important role in education by making the evaluation process more efficient and objective, but they also face a few challenges. One of the main advantages of automated systems is that they significantly reduce the workload of teachers. By automating the evaluation process, teachers can save time, which they can then spend on supporting students and personalising their teaching. In addition, the systems offer the possibility to manage large numbers of assignments at once, which is particularly useful for mass online courses (MOOCs) and large class sizes where traditional manual evaluation would be extremely time-consuming (Dadi et al., 2020).

Another advantage of automated systems is that they ensure consistency and objectivity in classification. In contrast to human evaluation, where bias and error can occur, automated evaluation systems follow a consistent set of criteria, making the evaluation objective and reliable. In addition, automated evaluation allows students to receive immediate feedback on their work, which facilitates continuous learning and improvement by allowing them to react quickly to errors and suggestions for improvement. As an added benefit, these systems can

also be more cost-effective, as they reduce costs associated with traditional paper-based evaluation, including production and administrative expenses (Pushak et al., 2011).
However, automated evaluation systems face several challenges. One of the biggest obstacles is the management of open-ended evaluations. While these systems perform well for multiple-choice or other objective questions, they are less reliable for assessing essays and open-ended responses. Current systems often use superficial measures such as grammar and coverage that do not always reflect the quality of content, making them unsuitable for objectively evaluating tasks that require deeper thought and reasoning (Ramesh, & Sanampudi, 2022).
The implementation and operation of these systems has additional technical and infrastructural requirements that are not within the reach of all institutions. The need for a significant technical infrastructure and expertise may limit the uptake of the system, especially in educational institutions that do not have the appropriate technical background. In addition, the attitudes of students may also affect the success of the system. Some students may be averse to automatic evaluation, especially if they are not familiar with the technology or are distrustful of the feedback provided by the system (Bello & Abdullah, 2021).
Furthermore, it is challenging to create flexible systems that can evaluate different subjects. Different disciplines – especially those requiring complex problem solving, creativity or critical thinking – expect students to have a wide range of skills that automated evaluation systems are not always able to measure adequately (Maaruf et al., 2023).
The flipped classroom is an innovative blended learning approach that reverses traditional teaching methods: it makes lectures and processing of course material part of out-of-class learning, often through video lessons, while in-class time is devoted to interactive, hands-on activities such as problem solving and group discussions (González-Gómez et al, 2016). This model encourages active participation and student engagement as classroom time focuses on direct practical applications and deeper understanding (Bosch Farré et al., 2024; Schmitt, & Cequea, 2020).
The flipped classroom offers several advantages. For one, it allows for better learning outcomes, as students often perform better compared to traditional teaching methods, having mastered the basics at their own pace. This model allows for personalised learning: students can work through material at their own pace, rewind and re-watch videos, repeating what they have learned as necessary. In addition, the model increases student engagement and motivation by encouraging active participation through interactive classroom activities (Lepkova et al., 2024).
The flipped classroom also helps to develop higher-level thinking skills and practical competences, as students not only passively absorb knowledge, but also apply and deepen it. Consequently, student´s satisfaction tends to be higher in this type of course, as the model is interactive and flexible, better adapted to the needs of the students.
The flipped classroom can be successfully applied in a wide range of educational contexts. For example, in engineering education, where basic language and technical skills are taught, it can

increase student engagement and improve student performance. In health sciences, particularly in community mental health courses, it can enhance learning and increase student engagement. It is also popular in primary education, for example in science courses, where it leads to improved student performance and more positive student feedback. It can also be used in logistics management courses, where task-driven teaching methods help to develop practical skills and teaching effectiveness. In the field of physical education, for example in basketball courses, the use of flipped classrooms through digital platforms improves students' initiative and learning efficiency (Koltsova & Boyko, 2022).

Alongside the benefits, the flipped classroom also has its challenges. One of the main obstacles is teacher preparation time, as creating and maintaining digital content requires significant time and effort. In addition, student preparedness is a key factor, as success depends on their willingness to engage with the material before class. If students are not prepared with the pre-learning material, the interactive part of the classroom will be less effective (Akçayır & Akçayır,2018).

3 Results

Automated evaluation systems can be effectively used in flipped classroom environments, especially in teaching programming and other technology areas, as they allow students to practice and receive feedback at their own pace outside of class time. In this model, students can engage with new material, such as video tutorials and online learning resources, at home or outside the classroom, and then spend classroom time primarily on interactive, problem-solving activities. Automated evaluation systems ensure that students receive feedback on their work in the pre-preparation phase, so that they can focus on deeper questions, in-depth exercises and more complex problems in the classroom.

Automatic evaluation systems are especially useful in teaching programming, as they give students the opportunity to test and improve their code independently. These systems provide immediate feedback on syntax, semantics and code effectiveness, so that students can react quickly to any errors. A partial scoring system built into the evaluation process encourages students to keep trying to improve their partially correct solutions. Code editors are often available in the system to assist coding with auto-completion and syntactic highlighting, and interactive testing facilities are also available to simulate the running of the program.

In other technology areas, such as data processing or web development, automated evaluation systems are also well suited for use in a flipped classroom environment. Students can practise writing code or software tasks at home and then use the classroom time for direct support and solving more complex problems. In addition, automated systems support multiple programming languages (e.g. Python, Java, JavaScript), so teachers can introduce them in different technological contexts.

The benefits of automatic evaluation systems linked to flipped classrooms include instant feedback, which helps students make continuous progress and reduces the evaluation workload for teachers. With automated feedback, students receive feedback more frequently than in a traditional classroom environment, allowing them to continuously improve their skills. By freeing up classroom time, teachers can spend more time explaining more complex topics and providing individual support to students, making the learning process more effective and personalised.

Take an example of the flipped classroom method being used in a programming course in a university introductory Python course. In this course, the teachers use an automated evaluation system to support the flipped classroom model, which gives students the opportunity to practice at home and then use the classroom time to solve more complex problems together.

The course design and use of the automatic evaluation system using the flipped classroom method:

1. Home preparation with video lessons and self-practice: students watch the video lessons and read a short introduction note on basic Python syntax, such as if-else structures and loops, before the class. At the end of the lessons, each student is given a short coding exercise that builds on the newly learned material, such as writing a simple calculator that can perform basic operations on user-specified numbers.

2. The system provides immediate feedback on the correctness of their code, including any syntax or logic errors. The feedback includes whether the code runs correctly, performs the calculations correctly and displays the expected result. If there is an error in the code, the system offers specific suggestions for improvements in syntax, variable handling or error handling.

3. Classroom lesson – Interactive, higher-level activities: at the beginning of the classroom lesson, students will already have a basic working code that they have created at home using the automated evaluation system. The teacher now gives more complex tasks that require, for example, the development of a calculator with several functions (e.g. adding additional operations, angle functions, memory management). During the lesson, students work in teams and the teacher helps them to understand more complex aspects of the program and to solve more difficult coding problems.

4. This means that they arrive at class with a basic understanding of the basics, where classroom time can be spent on solving more difficult programming problems rather than correcting basic mistakes. The interactive part of the class allows the teacher to focus more on the parts where students get stuck the most, which provides a more personalised and deeper learning experience.

Based on the results published in the literature and our own experience, we can conclude that this approach improves learning outcomes, as students receive quick and direct feedback on their work at home and can then focus on higher level, more complex problem-solving tasks in the classroom. The use of an automated evaluation system allows students to progress at their own pace, while teachers can save valuable time that can be spent on interactive tasks that require deeper understanding.

4 Conclusion

Automatic evaluation systems offer a few advantages, such as increasing efficiency, ensuring scalability and maintaining objectivity. However, they also face significant challenges, particularly in evaluating complex, open-ended tasks, and in being widely acceptable and adaptable to different educational contexts.

The flipped classroom methodology offers several benefits, including improved learning outcomes, personalised learning opportunities and increased student engagement. It has been widely and successfully applied in different educational settings, demonstrating its versatility and effectiveness.

Automatic evaluation systems in the flipped classroom environment can make a significant contribution to the effectiveness of the learning process in technological fields. They allow students to practise and develop independently, while teachers can use classroom time for activities that require deeper critical thinking and problem-oriented tasks.

Future research should move in several directions to develop the use of flipped classrooms and automated evaluation systems in technology education. One important area is the development of adaptive automated systems that can consider the individual progress, learning style and needs of students. These adaptive systems could provide personalised feedback based on students' prior knowledge levels to facilitate effective learning, particularly in programming and other technology courses.

Another important research direction is the development of improved error handling and deeper feedback, especially for open-ended tasks. While automatic evaluation systems are well suited for closed questions, they still face challenges in the evaluation of open-ended and complex tasks. Therefore, there is a need to develop algorithms that can assess solutions that require creative and critical thinking, which would contribute to the development of students' thinking skills.

Supporting real-time, collaborative feedback can also be important in the context of the flipped classroom model. The development of collaborative project work and real-time feedback systems could encourage collaboration and increase student engagement in technology training. In addition, measuring the effectiveness of the flipped classroom across different subjects and levels could be a promising area for research. This would allow understanding in which subjects and groups of students the flipped classroom is effective, and where additional support or different methodological approaches are needed.

The integration of artificial intelligence (AI) and machine learning into automated evaluation systems could also be an important future direction. AI-based systems could be able to detect student errors and support progress by providing personalised improvement suggestions, which is particularly useful in programming and other technology areas. In addition, studying student and teacher satisfaction could be key to the widespread adoption of flipped classroom and automated evaluation systems. Based on student and teacher feedback, systems could be refined to better meet user needs (Forman et al., 2023; Udvaros & Forman, 2023a; Udvaros & Forman, 2023b).

The issue of security and privacy is of particular importance for automated systems used in flipped classroom models. Protecting sensitive student data and defending against malicious code is essential for reliable operation, and future research should also address this area in depth.

These proposals can contribute to the further development of flipped classroom and automated evaluation systems and support the effectiveness and accessibility of technology education. Through research, personalised and active learning can become more widely available, improving learning experiences and outcomes in different educational environments.

Acknowledgments

Financial support for this study was provided from the project KEGA 015TTU-4/2024 "The Implementation of the Flipped Classroom Method in Higher Education".

References

Akçayır, G., & Akçayır, M. (2018). The flipped classroom: A review of its advantages and challenges. Computers & Education, 126, pp. 334–345.

Bello, H., & Abdullah, N. A. (2021). Investigating the influence of quality factors on user satisfaction with summative computer-based assessment. Electronic Journal of e-Learning, 19(6), pp. 490–503.

Bódi, S. (2023). Presentation of the fundamental law of Hungary. The Lawyer Quarterly, *13*(3).

Bosch Farré, C., Cicres i Bosch, J., Patiño Masó, J., Morera Basuldo, P., Toran Monserrat, P., Lladó Martínez, A., & Malagón Aguilera, M. C. (2024). Effectiveness of the flipped classroom methodology in higher education. A systematic review. Revista Educación XX1, *27*(1), pp. 19–56.

Dadi, R., Pasha, S. N., Sallauddin, M., Sidhardha, C., & Harshavardhan, A. (2020, December). An overview of an automated essay grading systems on content and non content based. *IOP Conference Series: Materials Science and Engineering 981*(2), p. 022016). IOP Publishing.

de la Peña, D. M., Gómez-Estern, F., & Dormido, S. (2012). A new Internet tool for automatic evaluation in Control Systems and Programming. Computers & Education, *59*(2), pp. 535–550.

Farias, G., Muñoz de la Peña, D., Gómez-Estern, F., De la Torre, L., Sánchez, C., & Dormido, S. (2016). Adding automatic evaluation to interactive virtual labs. Interactive Learning Environments, *24*(7), pp. 1456–1476. *Technologies and Applications (ICETA),* pp. 1–6. IEEE.

Flament, F., Lee, Y. W., Lee, D. H., Passeron, T., Zhang, Y., Jiang, R., & Aarabi, P. (2021). The continuous development of a complete and objective automatic grading system of facial signs from selfie pictures: Asian validation study and application to women of three ethnic origins, differently aged. *Skin Research and Technology, 27*(2), pp. 183–190.

Forman, N., Udvaros, J., & Avornicului, M. S. (2023). ChatGPT: A new study tool shaping the future for high school students. Future, *5*(6), 7.

Gabal’ová, V., Karpielová, M., & Stoffová, V. (2023). Beginners Online Programming Course for Making Games in Construct 2. *Proceedings of International Conference on Recent Innovations in Computing: ICRIC 2022*, Vol. 1, pp. 547–558. Springer Nature Singapore, Singapore.

Gabaľová, V. (2022). Detské programovacie jazyky a ich využitie v pedagogickej praxi.– 1.vyd.– Trnavská univerzita v Trnave. Pedagogická fakulta. Katedra matematiky a informatiky. 79 p., ISBN 978-80-568-0510-7.

González-Gómez, D., Jeong, J. S., Airado Rodríguez, D., & Cañada-Cañada, F. (2016). Performance and perception in the flipped learning model: an initial approach to evaluate the effectiveness of a new teaching methodology in a general science classroom. *Journal of Science Education and Technology, 25,* pp. 450–459.

Horváth, M., Kormaník, T., & Porubän, J. (2024). Adaptation of Automated Assessment System for Large Programming Courses. *5th International Computer Programming Education Conference (ICPEC 2024).* Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

Koltsova, E. A., & Boyko, S. A. (2022). Flipped Classroom Approach in Language Classes for Oil and Gas Engineering Master Students. In Proceedings of the Conference “Integrating Engineering Education and Humanities for Global Intercultural Perspectives”, pp. 207–215. Cham: Springer International Publishing.

Lepkova, N., Gülseçen, S., & Talan, T. (2024). Flipped classroom method application case study analysis. *Baltic journal of modern computing, 12*(2).

Maarouf, O., El Ayachi, R., & Biniz, M. (2023). Realization of an intelligent evaluation system. Int J Inf & Commun Technol ISSN, 2252(8776), 8776.

Pacheco-Venegas, N. D., López, G., & Andrade-Aréchiga, M. (2015). Conceptualization, development and implementation of a web-based system for automatic evaluation of mathematical expressions. *Computers & Education, 88*, pp. 15–28.

Pietrikova, E., Juhar, J., & Šťastná, J. (2015). Towards automated assessment in game-creative programming courses. *13th International Conference on Emerging eLearning.*

Pšenáková, I. (2019). Interatív tesztek a tudásszint mérésére. *OxIPO, Debrecen.* K+F Stúdió. ISSN 2676-8771, pp. 85–93. DOI 10.35405/OXIPO.2019.1.85

Pšenáková, I., Pšenák, P., & Szőköl, I. (2024). Flipped Classroom in Pedagogical Practice. In: Auer, M.E., Cukierman, U.R., Vendrell Vidal, E., Tovar Caro, E. (eds) Towards a Hybrid, Flexible and Socially Engaged Higher Education. ICL 2023. *Lecture Notes in Networks and Systems, vol. 911*. Springer, Cham. https://doi.org/10.1007/978-3-031-53382-2_26

Pushak, A., Carter, D., Wrzesniewski, T., & Lawrence, R. (2011). Experiences using an automated testing and learning system. *Proceedings of the 14th IASTED International Conference on Computers and Advanced Technology in Education*, pp. 7-11.

Ramesh, D., & Sanampudi, S. K. (2022). An automated essay scoring system: a systematic literature review. *Artificial Intelligence Review, 55*(3), pp. 2495–2527.

Restrepo-Calle, F., Ramírez-Echeverry, J. J., & Gonzalez, F. A. (2018). Uncode: interactive system for learning and automatic evaluation of computer programming skills. *EDULEARN18 Proceedings*, pp. 6888–6898, IATED.

Schmitt, V. G. H., & Cequea, M. M. (2020). Flipped classroom: A change in paradigms in higher education. [Aula invertida: Uma mudança nos paradigmas no ensino superior]. *Interciencia, 45*(11), pp. 501–507.

Stoffova, V., Gabal'ová, V., & Katyetova, A. (2023). How to Teach Programming to Beginners in a Playful Way. *Science and Information Conference,* pp. 801–811. Cham: Springer Nature Switzerland.

Szabó, T., & Pšenáková, I. (2017a). Computer-aided education and evaluation of student exam tasks from MS Excel. *15th International Conference on Emerging eLearning Technologies and Applications (ICETA)*, pp. 1–6, IEEE.

Szabó, T., & Pšenáková, I. (2017b). Automatické hodnotenie testov z MS Excelu. *30th DIDMATTECH, New methods and technologies in education and practice*. Trnavská univerzita v Trnave. Pedagogická fakulta, pp. 242–247. ISBN 978-80-568-0029-4.

Udvaros, J., & Forman, N. (2023a). Artificial intelligence and Education 4.0. *INTED2023 proceedings,* pp. 6309–6317, IATED.

Udvaros, J., & Forman, N. (2023b). Using Artificial Intelligence *Education 4.0. Journal of Technology & Information Education, 15*(1).