Learning Analytics and Recommender Systems toward Remote Experimentation

Alexandre L. Gonçalves[1], Gustavo R. Alves[2], Lucas M. Carlos[1], Juarez B. da Silva[1], João B. da M. Alves[1]

[1] Federal University of Santa Catarina, Araranguá, Brazil
[2] Polytechnic of Porto, Porto, Portugal
a.l.goncalves@ufsc.br; gca@isep.ipp.pt; lucas.mellos@posgrad.ufsc.br; juarez.silva@ufsc.br; jbosco@inf.ufsc.br

Abstract. This paper presents a process based on learning analytics and recommender systems to provide suggestions to students about remote laboratories activities in order to scaffold their performance. For this purpose, the records of remote experiments from the VISIR project were analyzed taking into account one of its installations. Each record is composed of requests containing the assembled circuits and the configurations of the measuring equipment, as well as the response provided by the measurement server that evaluates whether a particular request can be performed or not. With the log analysis, it was possible to obtain information in order to determine some initial statistics and provide clues about the student's behavior during the experiments. Using the concept of recommendation, a service is proposed through request analysis and returns to the students more precise information about possible mistakes in the assembly of circuits or configurations. The process as a whole proves consistent in what regards its ability to provide suggestions to the students as they conduct the experiments. Furthermore, with the log, relevant information can be offered to teachers, thus assisting them in developing strategies to positively impact student's learning.

Keywords: Remote Experimentation, Learning Analytics, Recommender Systems.

1 Introduction

Nowadays, taking into account the stage of science and technology, new approaches to education are required in order to positively impact student's performance. In the context of engineering education, solid knowledge is required not only from theoretical classes, but also from experimentation in laboratories [1]. In this sense, calculus classes, hands-on laboratories, simulations and remote laboratories are important resources in the training of students. As stated by [2], students have to become *fluent in the language of nature and a successful designer, and for that (...) must perform numerous experiments, practice, laboratory work*.

Thus, the skills developed by the students throughout the course will impact on their professional careers. In general, experimental work has traditionally been developed in laboratories. However, the increase in the number of higher education students in the last decades has put pressure on the physical structures and resources of laboratories. To overcome this, researchers have developed computational simulations and remote laboratories, enabling the expansion of educational boundaries.

This scenario provides new opportunities to enhance the student's learning process. With the advent of online systems, the data generated by student interaction in remote laboratories and simulations can be collected and analyzed. From this, some areas have been promoting support, among them, Learning Analytics (LA) and Recommender Systems (RS).

Learning analytics (LA) appears as an important tool that can leverage students' learning experiences as well as provide insights to teachers so they can learn and improve their classes. LA as a knowledge discovery paradigm can help stakeholders involved with the learning process to better understand its potential and interconnections [3]. Additionally, taking into account the collected and analyzed data, it is an opportunity to offer stakeholders recommendations about the educational context. In this way, Recommender Systems (RS) can supply suggestions to increase student's performance in learning activities. Generically, RS intends to recommend items that may be of interest for a user [4][5]. Originally coming from e-commerce, RS has evolved to compose solutions in a couple of areas, including e-learning. RS toward e-learning usually aim to help students in choosing courses, subjects and learning materials or activities [6]. Also, this kind of system can scaffold students by providing them with means to improve their performance in remote laboratory activities.

This paper proposes a process based on LA and RS in order to assist students in their remote lab activities. The process has two main goals, as follows: a) to analyze the data generated from student interaction through remote experimentation environments aiming to offer insights to stakeholders in the educational context; and b) to generate recommendations that can increase students' performance in learning activities. Section 2 introduces the background of the study. Section 3 presents the proposed process. Section 4 shows the experimental design. The results, the scenario analysis, and a general discussion about the process are presented in Section 5. Lastly, Section 6 draws conclusions.

2 Background

2.1 Remote Experimentation

There are several educational resources able to scaffold the students' learning process. Calculus classes pose abstract and methodic aspects, dealing with mathematics and knowledge about certain topics [7]. Hands-on lab activities allow achieving more complex competences by strengthening the connection between theory and practice and enabling students to achieve haptic skills and instrumentation awareness [7] [8] [9]. Simulation represents another important resource, although students should understand that they are dealing with a simulated reality as this may lead to some disconnection between the real and the virtual world [10]. Nevertheless, studies such as those published by [11] [12] point out that simulations can complement calculus classes and hands-on lab exercises.

Just like hands-on laboratories, remote laboratories require space and devices to compose the infrastructure. However, this approach goes beyond the traditional one and enables students to carry out real experiments controlled by computers through the internet. It increases the frequency and places in which experiments can be executed [9]. Additionally, by using remote laboratories students can access real equipment, which can leverage their out-of-classroom experiences [13]. Experimental devices can be shared by enhancing the infrastructure of traditional laboratories. In this way, remote laboratories, regarding the student's learning process, are seen as additional tools with some of the benefits of hands-on laboratories and computer simulations. However, there may be some difficulties in terms of availability of use since remote laboratories are connected to real equipment. On the other hand, students have access to simulators available on the internet, being a resource that does not require any kind of physical mechanism. Thus, remote and simulation labs have a further role in the educational context for providing teachers with complementary tools [9][14].

2.2 Learning Analytics

Learning has several impacts on student's lives. It is increasingly distributed across space, time and media, generating a large volume of data about students and the learning process [15]. All students' interactions through online educational environments leave traces about their experiences, making it possible to carry out a wide variety of analyses. Taking into account this behavior, Learning Analytics (LA) is more and more becoming a relevant tool that can positively impact student's performance.

Among the LA definitions, the following is the most cited one: "the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs" [16]. As stated by [17], LA has been incorporated in the context of educational institutions and has its origins or basis from the business intelligence field. Other fields include web analysis, educational data mining, and recommendation systems [18]. Focusing initially on the capture, analysis, and report of data by educational stakeholders as well as on the provision of information to enhance the performance of educational institutions, learning analytics has currently a mostly operational perspective. It intends to supply tools toward students and teachers for the achievement of higher performance and a broader understanding of the learning process.

2.3 Recommender Systems

Since the mid-1990s, Recommender Systems (RS) have evolved and become an important research field [4][5][19]. The objective is to provide suggestions generally by analyzing a great amount of options in situations where users may find some difficulties in making their choices [20]. This kind of system is suitable for the user and the service provider due to its capacity to help during the selection of items, making it a more enjoyable process in addition to leading to the achievement of better results. As stated by [21], "the purpose of RS is to generate valid recommendations for items that may be of interest to a set of users". An "item" is a piece of information representing, for instance, a product, a paper or a service, which is suggested to users when they interact with RS via the web, email or text message [22]. According to [23], an "item" is the general designation to denote what is going to be recommended to users by the RS.

There are some approaches presented in the literature, among them, content-based filtering (CBF), collaborative filtering (CF), and hybrid filtering [24][25]. More recently, the semantic web technology has empowered RS to deal with the overload of information and heterogeneous data sources [26]. Approaches based on formal structures of knowledge, such as ontologies, have been developed [27]. Also, in the educational context, there are e-learning recommender systems, an evolution from traditional e-learning systems [6]. These systems provide suggestions about what students should take into account, such as courses, subjects or learning activities, aiming to scaffold their performance.

3 Process Proposition

This section shows the proposed process considering the context of learning analytics and recommender systems. It aims to analyze the data generated from the interaction of students with a specific remote experimentation environment and generate recommendations that can help them to carry out the simulations. It intends to provide ways to scaffold students' performance on remote experimentation. Fig. 1 demonstrates the process in which a student performs experiments and receives more detailed information of possible problems found from the established configurations. Each experiment is composed of some elements that will be described in Section 4.

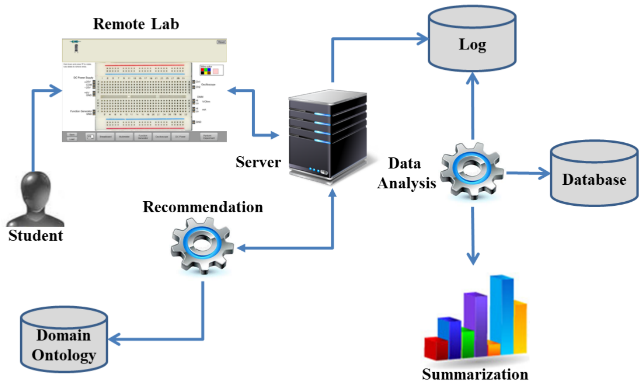

Fig. 1. The process toward students' performance analysis and suggestions of possible errors.

During the process, there are three phases consisting of logging, recommendations, and data analysis.

All student requests are sent to the server, which executes two main tasks. The first one consists of sending the request containing the settings made by the student to the recommendation service. The service then, considering a domain ontology, makes an inference to determine whether the request is correct or not. If it is not, the service recommends one or more types of errors. These errors are then sent to the remote lab interface so students can analyze the settings and carry out the necessary modifications. Also, both the request and the response, correct or not, provided by the server are logged.

Data analysis also has two essential functions. The first one focuses on monitoring the log acquiring all requests and responses from the experiments. A request is composed of a set of configurations which will be detailed in Section 4. New log entries are analyzed and stored in a database to provide means to easier analyze the result of the experiments achieved by the students. Besides, the data is summarized to provide information that may help teachers to better understand the students' performance during remote experimentation activities. In general, the summarization can provide interesting inputs for teachers to have information about the difficulties faced by the students. It allows an analysis of the causes of poor performance in specific subjects and can therefore guide teachers in actions of revision or improvements in their theoretical and hands-on classes.

3.1 Support Structures

Whenever a given experiment is configured, the student can send those settings to be evaluated by the server and receive a response. This information is characterized as the log of the remote experimentation process. From this, the analysis and persistence of each log entry in the database are carried out to evaluate student's performance. To support that, a database model was devised as illustrated in Fig. 2.

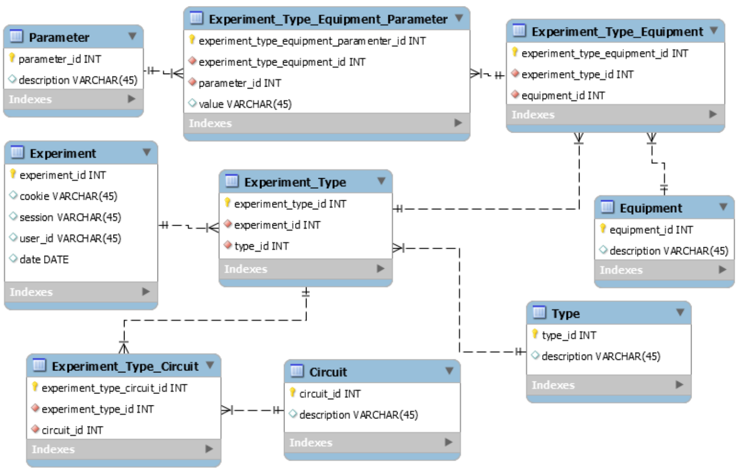

Fig. 2. Database model devised to enable log persistence and analysis toward teachers' understanding about students' performance.

The main table represents the experiment, which is named **Experiment**. Each experiment is composed of configurations performed by students taking into account circuits assembly and Multimeter, Function Generator, Oscilloscope and DC Power settings. In addition, there are two basic types registered into the **Type** table as request and response. After those experiment configurations, the student can send the data representing a request. From that point, the remote experimentation server analyzes the request to determine if all parameters were correctly selected. In affirmative case, all the measurements carried out are returned thus enabling the results to be presented through the interface.

After each request or response, the data are recorded in the **Experiment_Type** relationship table, hence allowing to store the information of which circuits were used and configured, which equipment was configured for the experiment and which parameters were defined. The **Equipment** table stores the available equipment in the remote experiment environment, while the **Parameter** table keeps the possible parameters for each equipment that will be used in a particular remote experiment.

To support the process as a whole, a domain ontology is also used. The ontology represents the knowledge base with the rules that make it possible to determine whether a given experiment has an error and, if so, what type of error. The ontology is presented in Fig. 3 and represents an overview of a multimeter.

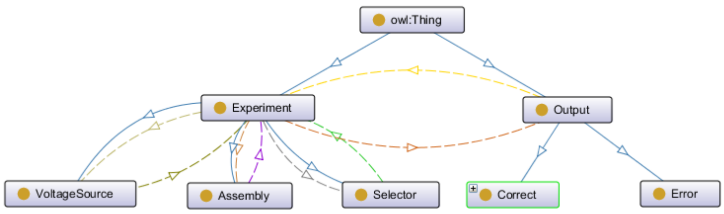

Fig. 3. Domain ontology used in the analysis of the experiments and recommendation of possible errors.

The ontology is composed of a set of classes, and the two main classes are **Experiment** and **Output**. The **Experiment** class makes it possible to define an instance through a set of properties. The instance represents a request made by the student and related with instances already defined in the **VoltageSource**, **Assembly**, and **Selector** classes. Using this information and through a reasoning process, it determines whether the output represents an error or not. In case of error, a more detailed message is provided.

4 Experimental Design

The process proposed here was implemented in the VISIR project. In order to detail the experimental design, both the VISIR project and the log are described as follows.

4.1 Remote Experimentation

The Virtual Instruments Systems In Reality (VISIR) project aims to provide support to the area of Electrical and Electronics Engineering focusing on the subject of circuit theory and practice.

Thus, by means of remote experimentation as an additional approach to other educational resources, such as calculus classems, hands-on lab activities, and simulations, the student has the opportunity to leverage their skills. Fig. 4 presents a demonstration board with components donated by Toyota®.

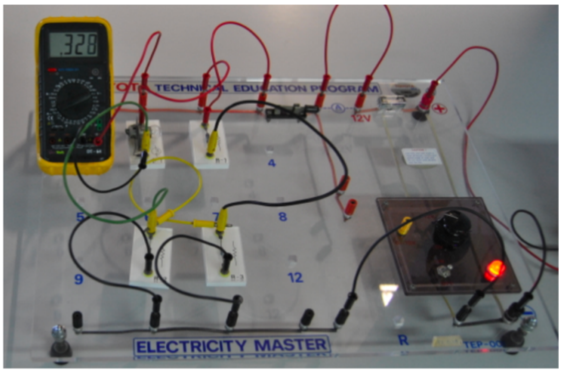

Fig. 4. Example of a circuit that students had to assemble in the Toyota® demo board.

A VISIR remote lab installation is used to interact with the physical boards and components. Through the environment of remote experimentation, the student can carry out the assembly of the circuits as well as define all the measurement parameters. Fig. 5 shows an example of configuration and measurement.

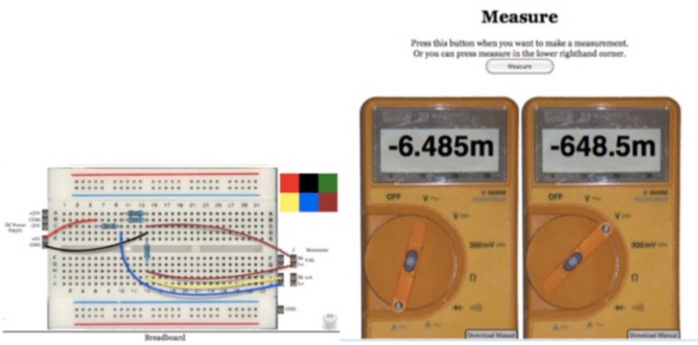

Fig. 5. Using VISIR to measure the voltage/current in the 100Ω resistor, in one of the circuits used in the individual lab assessment.

4.2 Data

After assembling the circuits and defining the measurement parameters for a particular experiment, the student then executes it. When doing so, a request is sent to the server, which performs all the checks and calculations, returning a response with the measurements or the information that the experiment was not successful, without however informing the specific type of error. Both the request and the response generated by the server are then logged.

For the present paper, a copy of the VISIR logs installed in the Polytechnic of Porto - School of Engineering (ISEP) was used. The log has a total of 545.152 records (requests, responses or errors) ranging from 2010-07 to 2018-03.

As already explained, a record in the log consists of a request and a response. The request has all the settings made by the student through the remote lab interface and the response has all the measurements performed by the server. If the settings are misconfigured or put the equipment from the physical laboratory at risk, a general error is sent. Fig. 6 shows an example of a partial log considering a request.


```
<protocol version="1.3">
 <request sessionkey="a05c194678883d9f55ee5ae129a8b518">
  <circuit>
   <circuitlist>
      W_X A25 DMM_VHI
      W_X A26 DMM_VLO
      POT_X A25 A26 A27 100k 64
  </circuitlist>
  </circuit>
  <multimeter>
   <dmm_function value="resistance"/>
   <dmm_resolution value="3.5"/>
   <dmm_range value="10"/>
  </multimeter>
  .... Other configurations ...
 </request>
</protocol>
```


Fig. 6. Partial log taking into account a request message.

Basically, the request log stores all the components used, indicating the positions where they are arranged on the breadboard, indicated by the <circuitlist> element. In addition, if the student has selected and configured a multimeter, the values used for that are indicated by the <multimeter> element. Other types of equipment are available at the remote lab interface, including a function generator, an oscilloscope, and a DC power, being these resources available to be used simultaneously.

5 Results and Analysis

In this section, the main results achieved so far are summarized taking into account the data analysis and the recommendation phases, as shown in the process described in Section 3.

5.1 Data Analysis

In this session, some analyses obtained from the data of the experiments registered in the log are discussed. The log is composed of 272,576 requests made by students from the interface of the remote laboratory considered in this paper. Of these requests, 238,949 (87.66%) had an adequate response, that is, after the evaluation, the server sent a response with the result of the measurements. Of the remaining responses provided by the server, 33,627 (12.34%) represent measurement errors. In the current VISIR version, the answer is generic and does not detail the type of error committed in the assembly of the components or in the configuration of the measuring equipment.

Each request belongs to the context of a remote lab session in which the student sets up a given experiment and sends it to the server. During the session, the parameters can be modified and resubmitted. Thus, multiple experiments can be performed. A total of 37,645 distinct sessions were identified, averaging 7.24 requests.

Finally, a comparison between the types of instruments used in the remote experiments is presented (see Fig. 7). As can be seen, the multimeter is the most used instrument with 79.46%, followed by DC Power, Function Generator and Oscilloscope with 78.64%, 48.83, and 47.52%, respectively.

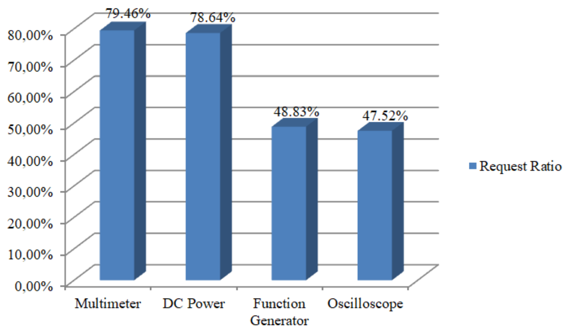

Fig. 7. Comparison between the types of instruments used in the requests made by the students.

5.2 Recommendation Approach

In this phase of the process, the requests made by students when using a specific remote laboratory are analyzed. As already mentioned, the request is sent to the server that accesses the recommendation service.

The recommendation service receives the request parameters involving the configuration of the circuits and the measurement equipment and fulfills an instance of the **Experiment** class in the domain ontology using object properties. Fig. 8 shows an instance for an experiment named **Experiment_1**.

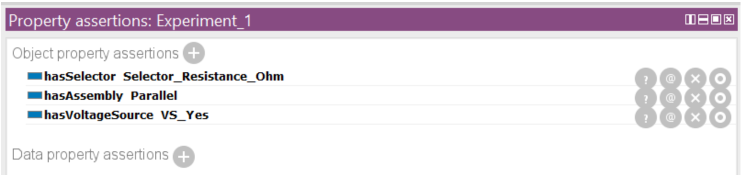

Fig. 8. Instance of an experiment named Experiment_1.

An experiment instance relates to the **VoltageSource**, **Assembly**, and **Selector** classes through the **hasVoltageSource**, **hasAssembly**, and **hasSelector** properties, respectively. In the previous example, instances related to the experiment are VS_Yes (Yes or No), Parallel (Series or Parallel) and Selector_Resistance_Ohm (V-, V ~, A-, A ~, Ω or OFF).

Based on the relationship of the experiment with instances of **VoltageSource**, **Assembly**, and **Selector** classes, it is possible to make the inference to determine whether there is an error or not in the configuration. Considering the relationships between instances of classes, there are 24 output possibilities. Fig. 9 presents two rules based on first-order logic.

Rules:
hasVoltageSource(?x, VS_Yes), hasAssembly(?x, Parallel), hasSelector(?x, Selector_Resistance_Ohm) -> hasOutput(?x, Type_AD)
hasVoltageSource(?x, VS_Yes), hasAssembly(?x, Parallel), hasSelector(?x, Selector_Voltage_V-) -> hasOutput(?x, Type_AB)

Fig. 9. Examples of rules that can be analyzed during the inference process.

The first rule, after the evaluation, will return a Type_AD output. This output (instance) represents an error and has an associated message, namely "Resistance reading with the circuit in tension". On the other hand, the second rule returns a Type_AB output instance that represents a possible and correct configuration.

Finally, after receiving the return from the recommendation service, the server composes the error in the form of response and returns it to the remote lab interface. It also records the request and the response in the log for further analysis.

6 Conclusion

The current scenario of education presents new challenges that require the combination of strategies and tools with a more sustainable vision. In this sense, remote experimentation allows overcoming some obstacles and limitations faced by hands-on laboratories. The present study focused on the application of concepts of learning analytics and recommender systems in the context of remote experimentation. For this purpose, the student interaction records made available by the VISIR project were used from one of its installations.

Experiment log analyses can reveal relevant information that help understand difficulties faced by students and provide subsidies for teachers to improve their classes and increase students' learning performance. In the present paper, the total of 12.34% of measurement errors seems to indicate acceptable figures since, at first, in addition to the theoretical and practical classes, there is a learning curve about the remote experimentation environment. Relating the students' errors to the duration of the course could provide additional information to better understand the learning process.

Currently, taking into account the response to a given experiment that was evaluated with error, the server only logs a general message without describing a specific type. In this sense, this paper uses an ontology to provide a knowledge base that can be used to clearly typify the error. The ontology presents only a part of the knowledge necessary to map all the possible errors, but it allows an initial visualization of how the errors can be made available to the students and stored for future analysis.

The results obtained are incipient but consistent in the scope of the proposed process. Knowing the main errors made during the experiments and allowing them to be returned to students is fundamental for improving the student's learning process.

With the development of this study, a better understanding is sought about the difficulties faced by students in an environment of remote experimentation. In addition, it is intended for the identification of the main errors produced, as well as their correlation with the executed experiments in order to provide teachers with information for class improvement.

Acknowledgment

We gratefully acknowledge the financial support from UFSC-Brazil and ISEP-P.PORTO-Portugal.

References

[1] C. A. Jara, F.A. Candelas, S.T. Puente, and F. Torres. "Hands-on experiences of undergraduate students in Automatics and Robotics," Computer and Education, vol. 57, pp. 2451-2461, 2011.

[2] I. Gustavsson, G. Alves, R. Costa, K. Nilsson, J. Zackrisson, U. Hernandez-Jayo, and J. Garcia-Zubia, "The VISIR Open Lab Platform 5.0-an architecture for a federation of remote laboratories," In: REV 2011 - 8th International Conference on Remote Engineering and Virtual Instrumentation. Brasov, Romania, July 2011.

[3] A. Peña-Ayala, "Learning analytics: A glance of evolution, status, and trends according to a proposed taxonomy," WIREs Data Mining Knowledge Discovery, 2017.

[4] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry, "Using collaborative filtering to weave an information tapestry," Communications of the Association of Computing Machinery, vol. 35, n. 12, pp. 61-70, 1992.

[5] P. Resnick, N. Iacovou, M. Suchak, P. Bergstorm, and J. Riedl, "GroupLens: An open architecture for collaborative filtering of netnews," In: Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work, 1994, Chapel Hill, North Carolina. Anais. ACM. pp. 175–186.

[6] J. Lu, D. Wu, M. Mao, W. Wang, and G. Zhang, "Recommender system application developments: A survey," Decision Support Systems, vol. 74, pp. 12-32, 2015.

[7] C. Viegas, N. Lima, G. Alves, and I. Gustavsson, "Improving students experimental competences using simultaneous methods in class and in assessments," In: TEEM '14 - Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality, pp.125-132, 2014.

[8] J. B. Lopes, C. Viegas, and J.P. Cravino, "Improving the learning of physics and development of competences in engineering students," International Journal of Engineering Education (IJEE), vol. 26, n. 3, pp. 612-627, 2010.

[9] M. Abdulwahed and Z. Nagy, "The TriLab, a novel ICT based triple access mode laboratory education model," Computers & Education, 56, pp. 262-274, 2011.

[10] J. Ma and J. Nickerson, "Hands-on, simulated, and remote laboratories: A comparative literature review," ACM Computing Surveys, vol. 38, n. 3, 2006.

[11] K-E. Chang, Y-L. Chen, H-Y. Lin, and Y-T Sung, "Effects of learning support in simulation-based physics learning," Computers & Education, vol. 51, pp. 1486-1498, 2008.

[12] N. Rutten, W.R. Joolingen, and J.T. Veen, "The learning effects of computer simulations in science education," Computers & Education, vol. 58, pp. 136-153, 2012.

[13] J. A. Rossiter, B. Pasik-Duncan, S. Dormido, L. Vlacic, B. Jones, and R. Murray, "A survey of good practice in control education," European Journal of Engineering Education, 2018.

[14] C. Gravier, J. Fayolle, B. Bayard, M. Ates, and J. Lardon, "State of the art about remote laboratories paradigms - Foundations of ongoing mutations," International Journal of Online Engineering (iJOE), vol. 4, n. 1, 2008.

[15] M. A. Chatti, V. Lukarov, H. Thüs, A. Muslim, A. M. F. Yousef, U. Wahid, C. Greven, A. Chakrabarti, and U. Schroeder, "Learning Analytics: Challenges and future research directions," E-learning and education (eleed), 2016.

[16] G. Siemens and P. Long, "Penetrating the fog: Analytics in learning and education," EDUCAUSE Review, vol. 46, n. 5, 2011.

[17] R. Ferguson and S.B. Shum. Social Learning Analytics: Five Approaches, In: LAK '12 - Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, pp. 23-33, 2012.

[18] R. Ferguson, "Learning analytics: Drivers, developments and challenges," International Journal of Technology Enhanced Learning, vol. 4, n. 5/6, pp. 304–317, 2012.

[19] U. Shardanand and P. Maes, "Social information filtering: algorithms for automating "Word of Mouth"," In: CHI '95 - Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, vol. 1, pp. 210–217, 1995.

[20] S. Senecal and J, Nantel, "The influence of online product recommendations on consumers' online choices," Journal of Retailing, vol. 80, n. 2, pp. 159-169, 2004.

[21] P. Melville and V. Sindhwani, "Recommender Systems", In: SAMMUT, Claude; WEBB, Geoffrey I.. Encyclopedia of Machine Learning. S. I.: Springer, pp. 829-838, 2011.

[22] D. Kotkov, J. Veijalainen, and S. Wang, "Challenges of serendipity in recommender systems," In: Proceedings of the 12th International Conference on Web Information Systems and Technologies., vol. 2, SCITEPRESS, pp. 251–256, 2016.

[23] F. Ricci, L. Rokach, and B. Shapira, "Introduction to Recommender Systems Handbook," In: RICCI, Francesco et al (Ed.). Recommender Systems Handbook. S. I.: Springer, 2011. pp. 1-35.

[24] X. Su and T.M. Khoshgoftaar, "A survey of collaborative filtering techniques," Advances in Artificial Intelligence. pp. 1–20, 2009.

[25] D. Kotkov, S. Wang, and J. Veijalainen, "A survey of serendipity in recommender systems," Knowledge-Based Systems, 111, pp. 180-192, 2016.

[26] L. O. Colombo-Mendoza, R. Valencia-García, A. Rodríguez-González, G. Alor-Hernández, J. J. Samper-Zapaterd, "RecomMetz: A context-aware knowledge-based mobile recommender system for movie showtimes," Expert Systems With Applications, S. I., vol. 42, n. 3, pp. 1202-1222, 2015.

[27] R-C. Chen, Y-H. Huang, C-T. Bau, and S-M. Chenb, "A recommendation system based on domain ontology and SWRL for anti-diabetic drugs selection," Expert Systems With Applications, S. I., vol. 39, n. 5, pp. 3995-4006, 2012.