An Initial Analysis of Prediction Techniques as a Support for the Flipped Classroom

Aarón Rubio-Fernández[1], Pedro Manuel Moreno-Marcos[1], Pedro J. Muñoz-Merino[1], Carlos Delgado Kloos[1]

[1]Universidad Carlos III de Madrid
{aarubiof, pemoreno, pedmume, cdk}@it.uc3m.es

Abstract. With the increasing use of active learning methodologies such as the Flipped Classroom (FC), many approaches have been taken to enhance the students' learning in such contexts. Prediction techniques can be used in combination with Learning Analytics (LA) dashboards for the improvement of the FC model. In this direction, we analyze some theoretical cases in which this approach can provide academic benefits (e.g. providing additional resources or re-designing the class). Furthermore, we present several initial ideas on how to combine two existing software tools, one which provides LA dashboards and the other that implements prediction techniques, that can be used successfully in such scenarios for the FC. This is a preliminary work for the joint use of prediction techniques and LA dashboards in FC contexts.

Keywords: Learning Analytics, Prediction, Flipped Classroom, Visualizations, Course Design, Orchestration.

1 Introduction

Nowadays, active learning methodologies such as the Flipped Classroom (FC) are widely used. In a FC, students access to academic resources before the face-to-face lesson, and the lesson is devoted to do active learning methodologies [1]. However, the FC has some disadvantages such as the need of preparing the face-to-face lesson, or the time investment related to the creation of the academic resources [1]. In order to overcome some of these limitations, learning analytics (LA) can be used. In particular, LA may allow to prevent several problems which appear in a FC environment such as the lack of preparation of the face-to-face lesson or the student dropout rate.

Particularly, the combination of predictive models and LA dashboards may be useful in FC contexts, since they may allow to solve some of the issues such as e.g., the lack of preparation of the face-to-face lesson or the student dropout rate. Nonetheless, despite of the fact that those systems have been used in educational scenarios, and they are sometimes combined to provide visualizations about the predictions (e.g. [2]), the potential of this combination to support FC contexts has not been deeply analyzed. For

1 Copyright © 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0)

CEUR-WS.org/Vol-2671/paper02.pdf

this reason, this paper aims to provide some ideas on how LA dashboards and predictive models can be combined to support FC contexts. In this sense, the objectives of this paper are: (1) to explore some cases where visualizations and predictions might be used to improve FC, and (2) to present how two existing LA tools can be used to implement together the analytical and the predictive approaches in order to enhance FC.

2 Related Work

Due to the increasingly popularity of the FC model, it is not surprising that many researchers have tried to improve this model in order to maximize the educational benefits that it can bring. Among many other approaches, we focus on the use of LA in FC contexts due to the advantages that it provides. For example, LA can allow students to receive information to self-reflect about their learning, and teachers to re-design and improve classroom orchestration [3]. In this sense, teachers can encourage students to review some materials or they can provide additional support based on videos' visualization patterns [4]. Moreover, researchers have developed predictive models to forecast dropout (e.g. [5]), learning outcomes (e.g. [6]) and students' behaviours (e.g. [7]). These models can serve to detect which students may have difficulties in the course, in order to provide them with additional support or in order to adapt the teaching methodology. In addition, other researchers have developed visualization systems [8], and frameworks to use them to support FC (e.g., [9]).

The FC is suitable to take advantage of LA techniques in order to improve the students' learning model. For instance, LA can provide teachers with information to adapt the face-to-face lessons to the students' needs [9], or information to better understand the learning process in e-learning contexts [10]. In the context of this work, we focus on researches from the perspective of LA dashboards, and works focused on prediction techniques.

For example, regarding the analytical approach, we can remark the LA tool used in [11], which provides teachers with information about the students' interactions with the videos, and the tool shown in [12], which allows teachers to know how students use the #C compiler used in the practices. However, as far as we are concerned, none of this type of works makes use of prediction techniques.

As for the predictive approach, there are also works focused on FC environments. The main objective is to predict the students' performance using the interactions between the students' and the academic resources. For instance, in [13], the students' performance is predicted using the students' interaction with the course material, the discussion forums, the digital assessments, and the period when students interacted (e.g., whether they interacted before, during, of after the class). A similar approach was used by [14] to predict performance. In the latter case, they separated variables considering all the period or only before classes (e.g., videos watched on time or on the whole period). Nonetheless, as similar to the previous case, these works are focused only on the predictive approach and none of them consider using LA dashboards in combination with predictions.

Therefore, as far as we know, there are no works which specifically propose the joint use of both analytical and predictive techniques for the improvement of the FC. For this reason, we present some initial ideas following this direction.

3 Discussing the Combination of LA Dashboards and Prediction Techniques for the FC

The FC can be significantly improved if the appropriate information is provided to teachers. In this section, we present some typical cases in LA, and we discuss how the predictive approach along with LA dashboards can enhance the FC.

Providing Personalized Additional Resources and Alerts Before the Face-to-face Sessions. When students prepare the face-to-face sessions for the FC, the provided resources may not be tailored to their needs. For example, high-achievers can take advantage of resources focused on complex concepts while low-achievers can need resources focused on the basic knowledge. In this sense, prediction techniques can help in several ways, e.g.: (1) predicting if students will interact a lot with the resources, and determining students at risk of dropout; (2) predicting if students will receive very easy or very difficult resources and estimating their grades in each case; (3) predicting students' performance in each possible exercise.

These types of predictions should be redone for each face-to-face session. Therefore, the prediction should be done several times during the time for giving information about different points of time. Moreover, LA dashboards allow teachers to know the information related to the most used resources for each student, so that we can provide students with their most "appealing" types of resources (e.g., videos, lectures) and with resources tailored to their needs. In addition, we can alert students who are at risk and have a prediction of not preparing the face-to-face sessions.

Re-designing the Face-to-face Sessions. Using the students' performance predictions, we can infer whether they are ready to study complex concepts, or simpler ones, and which concepts they should review. For example, we can analyze the students' visualization patterns in order to detect whether they are struggling with some parts of the video (e.g., if they are watching one part over and over again), so that we can check that the face-to-face lesson's exercises are suitable for them. If the students have not had any issues with the preparatory resources, and they are predicted to obtain high scores, then the face-to-face session can be designed to explain the most difficult concepts. On the other hand, if students have difficulties with some concepts, and they are expected to obtain medium or low scores, then the exercises done during the class can be focused on that difficult concepts. Moreover, performance predictions can be used as a metric for grouping students for collaborative activities. Therefore, instead of using other type of data (e.g. in [15] students' interactions with videos are used), we could use these predictions.

Evaluating the Student Learning Behavior Trend. It would be also interesting to analyze how the students have used the academic resources in the different weeks of the course, as well as if they are going to use those resources in the future weeks in the same way. In this direction, we can infer if students prefer one type of resource over the others (e.g., videos over online exercises), so that the most appealing for them can be used. Furthermore, we can also infer if students will ask for help if they struggle with the exercises. Prediction techniques can give us this type of information which can be used to enhance the learning process.

4 Learning Analytics Tools and Proposal of Use

In this section, two LA tools are presented (one providing LA dashboards, and the other implementing prediction functionalities), and an initial proposal of joint use for the improvement of the FC.

4.1 Course Context and Data Collection

These tools have been designed to support university courses at Universidad Carlos III of Madrid which offer SPOCs (Small Private Online Course) [16] to support the face-to-face classes. These SPOCs are hosted in a local instance of Open edX, and they typically contain videos and exercises. The use of the SPOCs allows retrieving student data about activity, and interactions with videos and problems. Despite designing the tools to be used in the abovementioned SPOCs, it is worth mentioning that the tools could be used in other Open edX contexts, as the data would be the same.

For the development of the tools, data from the events (tracking logs) was considered. These events are: problem_check, play_video, pause_video, and seek_video. With these events, some variables were retrieved for the pair "student-item", where an item can be either a video or exercise. These variables are used in the first tool.

- Variables retrieved for each video: 1) percentage viewed and 2) number of times each second of the video has been watched by the student.
- Variables retrieved for each exercise: 1) grade, number of attempts and number of times the exercise is solved correctly by the student.

Similarly, for the predictive models, several variables have been obtained. However, these variables are obtained at student level (and not the pair "student-item") so that they can be used as predictors:

- Variables related with videos: 1) percentage of opened videos, 2) completed videos, 3) percentage of viewed time, 4) average number of repetitions and 5) average number of pauses.
- Variables related with exercises: 1) percentage of attempted exercises, 2) average number of attempts, 3) average grade of attempted exercises with all attempts and 4) with only the first attempt, 5) percentage of correct exercises with all attempts and

6) only considering first attempts, and 7) maximum number of consecutive correct exercises.

- Variables related to activity: 1) percentage of days the student accesses to the SPOC, 2) maximum number of consecutive days the student accesses, and 3) average number of consecutive days the student accesses to the SPOC.

4.2 Tool for LA Dashboards

This tool provides teachers with information about the past and present students' learning process. With this aim, the tool offers many visualizations such as the most watched parts of a video, or the students' performance when solving exercises [17]. In particular, the tool shows information related to the students' interaction with: videos, exercises, lectures (PDFs or HTML pages), and online discussion forums. Moreover, it also includes a functionality to group students using their learning data [15].

Regarding the technologies, PHP and JavaScript (web interface), MySQL (database), and Python (analysis of the data) are used. Moreover, the tool is integrated within the GEL platform, which was developed by the "Servicio de Informática y Comunicaciones" of the Universidad Carlos III of Madrid.

Finally, the tool allows teachers to know: 1) whether students are preparing the face-to-face lesson, and the specific students who are doing so; 2) the use of the academic resources; and 3) the most difficult concepts for students.

4.3 Prediction tool

The second tool aims to provide the predictions of two variables about what students will have done by the end of the course. The first one is dropout, which is defined as completing 75% of the exercises at least (a typical threshold, used in other platforms, such as MiríadaX, [18]). The second one is success, which is defined as achieving an average grade of 50% at least, considering all the exercises of the SPOC (non-attempted exercises count as 0). In order to develop the predictive models, variables mentioned in the previous section were used. In addition, four of the most common algorithms, according to [7], were tested: Logistic Regression, Random Forest, Support Vector Machines and Decision Trees. From those models, Random Forest was selected as it provided better results in this context. Taking this into account, one model was trained each of the 16 weeks of each SPOC (all university courses lasts 16 weeks), considering all the data available until that moment. This way, when a student is in week X (e.g., week 2), their predictions are obtained with the model trained in week X (e.g., week 2). Using this approach, it can be possible to compute the interactions of the student each week and update predictions based on models trained for each period.

With these models, it is important to note that first predictions are less accurate, but they provide more anticipation. Particularly, predictions of dropout achieve an Area Under the Curve (AUC) of 0.8 from week 3 and 0.9 from week 6, and predictions of success achieve and AUC of 0.8 from week 2 and 0.9 from week 6. This means that it is possible to obtain early predictions that could anticipate possible problems.

When predictions are computed, a probability is obtained for each variable (dropout and success). The results of the prediction tool are not currently connected to any dashboard, although there are some ideas about how these visualizations could be. For example, some traffic lights with the probabilities could be used (as they were used by [19]). In addition, some aggregated values for each class could be given using a similar format (e.g., number of students with high, medium, low risk of dropout/failure). This way, instructors could get an idea of the situation of the whole class and then delve into specific students. That could be particularly relevant for big groups, where the instructor may be more focused on aggregated data.

4.4 Combination of LA Tools to Support Flipped Classroom

In this section we present different ways to use the two presented tools within some of the cases described in section 3.

Personalized Resources Before the Face-to-face Lessons. Using the prediction tool, teachers can determine which students are expected to be high-achievers and which ones low-achievers for the specific lesson. After that, teachers can use the tool that provides LA dashboards to analyze how much the students use the preparatory videos, exercises, etc. With this information, teachers would provide tailored resources to the different types of students (e.g., basic materials for low-achievers, or videos for the high-achievers which prefer videos over exercises). For instance, teachers can analyze the students' number of attempts related to one exercise and if they are high-achievers or low-achievers, as we can see in the following figure (Fig. 1).

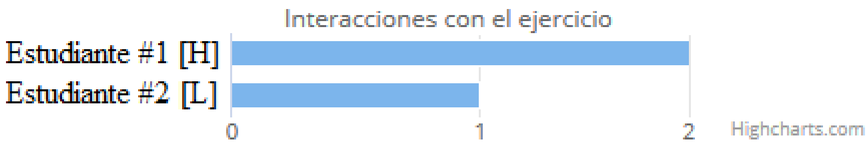

Fig. 1. Students' number of attempts and if they are high-achievers [H] or low-achievers [L]

Re-designing the class. In this case, teachers start analyzing the most watched parts of the preparatory videos in order to know whether their students are struggling with the explained concepts. Next, teachers analyze the predicted dropout rate, in order to know if the face-to-face lesson has to be focused on engaging activities devoted to encourage students not to dropout, and/or on activities focused on the most difficult concepts for students. An example of the integrated visualization of these two types of information is shown below (Fig. 2).

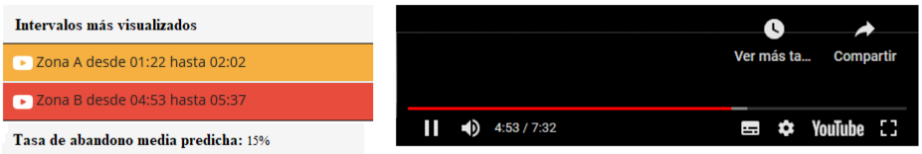

Fig. 2. Most watched part of the video (red parts are the most watched ones and oranges are the parts more watched than the average), and predicted dropout rate.

Furthermore, teachers can also analyze the number of times that the preparatory exercises have been attempted and have been solved correctly, along with the performance prediction. If students are predicted not to pass, basic activities can be done related to the concepts of the most difficult exercises for the students. However, if students are predicted to pass, the class can be devoted to more complex activities.

5 Conclusions and Future Work

In this paper, we propose some initial ideas about the combination of LA dashboards with LA prediction techniques to enhance different aspects of the FC model. Although the joint use of LA an FC has been analyzed in some previous works (e.g., in [11] LA is used to find out the students' use of the resources), as far as we know there are no research focused on the proposed approach.

Along with the definition of some scenarios where the approach would enhance the FC, we present two LA tools which can be combined to obtain the advantages of LA dashboards and prediction techniques. Furthermore, we provide some specific examples that show how these tools would be used to obtain the benefits of the approach.

This paper is a very preliminary work in which some initial ideas are presented. There are other scenarios and possibilities that can be explored in the future, e.g., through a specific framework focused on combining LA dashboards and prediction techniques with the FC. Moreover, the tools are separated tools, so that an integration process is needed to ease the implementation of the approach. Finally, the learning outcomes have to be validated in real contexts.

Acknowledgements

This work has been partially funded by: FEDER/Ministerio de Ciencia, Innovación y Universidades – Agencia Estatal de Investigación/Smartlet project (TIN2017-85179-C3-1-R). In addition, this work has been partially funded by the e-Madrid-CM project with grant no. S2018/TCS-4307, which is funded by the Madrid Regional Government (Comunidad de Madrid), by the Fondo Social Europeo (FSE) and by the Fondo Europeo de Desarrollo Regional (FEDER). This work also received partial support by Ministerio de Ciencia, Innovación y Universidades, under an FPU fellowship (FPU016/00526).

References

1. Bergmann, J., Sams, A.: Flip your classroom: Reach every student in every class every day. International society for technology in education (2012).
2. Chen, Y., Chen, Q., Zhao, M., Boyer, S., Veeramachaneni, K., Qu, H.: DropoutSeer: Visualizing learning patterns in Massive Open Online Courses for dropout reasoning and prediction. In: 2016 IEEE Conference on Visual Analytics Science and Technology (VAST), pp. 111-120. IEEE, (2016, October).

3. Rodríguez-Triana, M. J., Martínez-Monés, A., Asensio-Pérez, J. I., Dimitriadis, Y.: Scripting and monitoring meet each other: Aligning learning analytics and learning design to support teachers in orchestrating CSCL situations. British Journal of Educ Technology 46(2), 330-343 (2015).
4. Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., Miller, R. C.: Understanding in-video dropouts and interaction peaks inonline lecture videos. In: Proceedings of the first ACM conference on Learning@ scale conference, pp. 31-40. (2014, March).
5. Xing, W., Du, D.: Dropout prediction in MOOCs: Using deep learning for personalized intervention. Journal of Educational Computing Research 57(3), 547-570 (2019).
6. Moreno-Marcos, P. M., Pong, T. C., Muñoz-Merino, P. J., Delgado Kloos, C.: Analysis of the Factors Influencing Learners' Performance Prediction With Learning Analytics. IEEE Access 8, 5264-5282 (2020).
7. Moreno-Marcos, P. M., Alario-Hoyos, C., Muñoz-Merino, P. J., Delgado Kloos, C.: Prediction in MOOCs: A review and future research directions. IEEE Transactions on Learning Technologies 12(3), 384-401 (2018).
8. Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., Gillet, D., Dillenbourg, P.: Perceiving learning at a glance: A systematic literature review of learning dashboard research. IEEE Transactions on Learning Technologies 10(1), 30-41 (2016).
9. Rubio-Fernández, A., Muñoz-Merino, P. J., Delgado Kloos, C.: Scenarios for the application of learning analytics and the flipped classroom. In: 2018 IEEE Global Engineering Education Conference (EDUCON), pp. 1619-1628. IEEE, (2018, April).
10. Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., Leony, D., Delgado Kloos, C.: ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform. Computers in Human Behavior 47, 139-148 (2015).
11. Redondo-Martínez, D., Muñoz-Merino, P. J., Ruipérez-Valiente, J. A., Delgado Kloos, C., Pijeira Díaz, H. J., Santofimia Ruiz, J.: Combining Learning Analytics and the Flipped Classroom in a MOOC of maths. In: CHANGEE/WAPLA/HybridEd@ EC-TEL, pp. 71-79. (2015).
12. Estévez-Ayres, I. M., Arias-Fisteus, J., Uguina, L., Alario-Hoyos, C., Delgado Kloos, C.: Uncovering flipped-classroom problems at an engineering course on systems architecture through data-driven learning design. Int. Journal Engineering. Education. 34(3), 865-878 (2017).
13. Nouri, J., Saqr, M., Fors, U.: Predicting performance of students in a flipped classroom using machine learning: towards automated data-driven formative feedback. In: 10th International Conference on Education, Training and Informatics (ICETI 2019), pp. 17-21 (2019).
14. Martínez-Carrascal, J. A., Cebrián, D. M., Sancho-Vinuesa, T., Valderrama, E.: Impact of early activity on flipped classroom performance prediction: A case study for a first-year Engineering course. Computer Applications in Engineering Education (2020).
15. Rubio-Fernández, A., Muñoz-Merino, P. J., Kloos, C. D.: Analyzing the group formation process in intelligent tutoring systems. In: International Conference on Intelligent Tutoring Systems, pp. 34-39. Springer, Cham, (2019, June).
16. Fox, A.: From moocs to spocs. Communications of the ACM 56(12), 38-40 (2013)
17. Rubio-Fernández, A., Muñoz-Merino, P. J., Delgado Kloos, C.: A learning analytics tool for the support of the flipped classroom. Computer Applications in Engineering Education 27(5), 1168-1185. (2019).
18. MiríadaX Web page, https://miriadax.net/en/faq?faqid=8635212, last accessed 31/05/2020.
19. Gutiérrez, F., Seipp, K., Ochoa, X., Chiluiza, K., De Laet, T., Verbert, K.: LADA: A learning analytics dashboard for academic advising. Computers in Human Behavior 107, (2018).