MOOC Learner Behaviour: Attrition and Retention Analysis and Prediction Based on 11 Courses on the TELESCOPE Platform

Massimo Vitiello[1], Christian Gütl[1,2],
Hector R. Amado-Salvatierra[3](✉), and Rocael Hernández[3]

[1] IICM, Graz University of Technology, Graz, Austria
massimo.vitiello@student.tugraz.at, c.guetl@tugraz.at
[2] CLL, Curtin University, Perth, WA, Australia
[3] GES Department, Galileo University, Guatemala, Guatemala
{hr_amado, roc}@galileo.edu

Abstract. Massive Open Online Courses (MOOCs) have become an important online learning tool for educators and learners, but one of the major issues are the high drop-out rates. Recent research suggests not only to identify and support learners at-risk to drop-out but also to differentiate between the group of healthy attrition (intentionally leaving the MOOC) and unhealthy attrition (struggling to complete the MOOC). In this paper, we focus on two research questions: Firstly, can we already identify learners at-risk to drop-out a MOOC in an early stage? Secondly, can we differentiate between the group of healthy attrition and unhealthy attrition? Experimentation with Support Vector Machines based on learners logs from eleven MOOCs on the Telescope platform show first promising results.

Keywords: MOOC · Learning analytics · Attrition · Retention · Dropout prediction

1 Introduction

Information and communication technologies (ICT) have broadly changed the way how we learn and teach. Over the last decades, a movement has emerged which started with *open source* and provided a wide range of tools for the educational sector. This has been followed by *open content* and *open courseware*, the second important pillar of free and open education. Finally, *open online courses,* as the next logical step, have opened formal courses in educational institutions for virtually all potential learners over the globe without any restrictions. Due to the broad interest and need of open education, a big number of learners have enrolled such open online courses and the term Massive Open Online Courses (MOOCs) have been coined [3, 4, 17, 19].

Over the last years, MOOCs raised a lot of attention by learners, educators, educational institutions, and researchers. It was praised as a new form of education for the mass with a high potential. Advantages include free and open education without

L. Uden et al. (Eds.): LTEC 2017, CCIS 734, pp. 99–109, 2017.
DOI: 10.1007/978-3-319-62743-4_9

restrictions for everyone, even for the unprivileged and poor ones. It opens up courses from high renowned universities and enables self-guided learning. Very soon, also issues and challenges have been recognized and reported. This includes a big effort in coaching and technical support, and the feeling of isolation and the lack of communication and interaction with peers and tutors. Due to the lack of entry barriers, a great variety of pre-knowledge, computer literacy and expectation caused high drop-out rates, and attrition is reported as one of the major issues [3, 4, 17, 19].

Research in attrition has been an active research field for a very long time. Firstly, it focused on brick-and-mortar institutions and continued in computer-based and online learning. [19] Over the last years, research has increasingly focused on MOOC settings. Research varies from understanding user behavior and uncovering learner groups to models and predictions of learners at-risk to drop-out and leave the MOOCs [1, 2, 5–8].

Seeking a better understanding of the learning processes and learners' behavior, in particular to mitigate attrition issues, a collaborative research between University of Galileo in Guatemala, Curtin University of Technology in Australia, and Graz University of Technology in Austria has been established. Attrition research in MOOC settings and data analysis combined with users' questionnaires revealed three groups of risk for leaving a MOOC, which are (a) the personal learners' factors, (b) factors of the educational institutions and MOOC design, and (c) environmental factors influencing the learners. Narrowing down to the intention of enrolling and accessing a MOOC, questionnaire results indicate that not all enrolled users have as primary goal the completion of a MOOC. Rather, these users enrolled to have access to learning content or to just selectively participate in activities without formal completion. These situations can be described as *healthy attrition,* because the final completion of the course is not indented. In contrary, *unhealthy attrition* subsumes all other situations, in which users want to successfully pass the course but fail for various reasons. Based on these results, the *Attrition Model for Open Learning Environment Setting (AMOES)* has been proposed in order to better understand and differentiate reasons for learners at-risk to drop-out [3, 4].

As follow-up research, we have initiated the European project MOOC maker [18]. As part of this project, we are interested in modeling and predicting user behavior, identifying in an early stage learners at-risk to drop-out, uncovering features which indicate drop-out risks and developing support for learners accordingly. Related work aims at identifying and making early predictions on possible drop-out of learners [5–7]. Complementarily, there are different researches trying to develop strategies to increase learners' engagement by the use of game-based learning, social networks and effective communication within groups [8–11]. Our preliminary results on data of five MOOCs revealed that predicting learners at-risk to drop-out is feasible by taking into account the first half of activities of the MOOC with sufficient accuracy. Also, support vector machines classifier performs better than the K-means approach [15].

In this paper, we focus on two research questions: Firstly, can we already identify learners at-risk to drop-out a MOOC in an early stage? Secondly, can we differentiate between the group of healthy attrition and unhealthy attrition? To this end, logs from eleven MOOCs offered on the Telescope platform are considered [12]. Further related research and more detailed findings and best practices are covered in the MOOC maker deliverable [19]. The remainder of this paper is organized as follows: Sect. 2 gives an overview of the eleven MOOCs and the experimentation setup, followed by data analysis and findings in Sect. 3 and future perspectives in Sect. 4.

2 Experimental Setup

The experiments and studies are based on 11 MOOCs offered by Universidad Galileo on their Telescope platform and are briefly described in Table 1. Each of these MOOCs has a fix 8 weeks duration. In order to obtain a positive grade, learners have to successfully complete weekly activities (in form of quizzes or assignments) and, eventually, a final project and exam. Self-assessment activities are also planned for some MOOCs, but these do not influence learners' final grade. For each MOOCs a log file was created. These files reported each interaction taking place on the platform, and included information as the timestamp of the interaction, the ID of the learners and the particular tool the request referred to. The logs of the MOOCs in our dataset included a total of 21 different tools. However, out of these 21, more than 99% of the total interactions was accounted by 8 tools only, which are listed in Table 2.

Table 1. Description of analyzed MOOCs

MOOC	Target	Completers	Non-completers	Drop-out rate
Android	Students	77	516	87%
Cloud based learning	Professionals, Teachers	123	156	56%
E-learning	Professionals, Teachers	81	164	67%
Community manager	Professionals	320	501	61%
Medical urgencies	No specific target	49	69	58%
Client Attention	Professionals	60	31	34%
Cloud based learning (EDU)	Professionals, Teachers	99	83	45%
Cloud based learning (Tools)	Professionals, Teachers	131	186	59%
Digital interactive TV	Professionals	63	58	47%
E-learning (Tools)	Professionals, Teachers	101	156	60%
User experience	Students	62	127	67%
Combined	**–**	**1166**	**2047**	**58%**

We did the experiment based on the above listed MOOCs (see Table 1) and run 2 classification experiments. As all MOOCs have the same 8 weeks duration, we first analyze the logs on a weekly base and try to classify learners as either likely to complete a MOOC (completers) or at risk of dropping out (non-completers). For this

Table 2. Description of the tools available and used for the analysis

Tool	Description
Assessment	Mostly quizzes. Used to test user knowledge and satisfaction in the MOOC
Assignment	Link to the assignment page (e.g. tasks, projects, etc.)
Course board	Page including the weekly topics of the MOOC
Evaluation	Reports user's grade for each submitted task
File Storage	Link to the documents and resources of the course
Forum	Link to the discussion Forum
Learning content	Whole content of the MOOC. Includes links to files videos, audios, images and any resources used during the course
Peer evaluation	Tool used for the peer review of users

first experiment, we combine data from all MOOCs together and use this as our dataset. We refer to this setting as *completers experiment*.

As second experiment, we attempt to investigate the non-completers class and further categorize the samples within this one as either healthy or unhealthy attrition (see also Sect. 1 and [3, 4]). This second experiment is therefore a multi class classification aimed at finding the reasons that brought non-completers to eventually drop out of the course. In this case, the MOOCs are analyzed individually, without the weekly base split (we use the whole available interactions for each MOOC). We call this setting *attrition experiment*. In the following subsections, we first describe the general approach for features construction and then we specify each setting in details.

As second experiment, we attempt to investigate the non-completers class and further categorize the samples within this one as either healthy or unhealthy attrition (see also Sect. 1 and [3, 4]). This second experiment is therefore a multi class classification aimed at finding the reasons that brought non-completers to eventually drop out of the course. In this case, the MOOCs are analyzed individually, without the weekly base split (we use the whole available interactions for each MOOC). We call this setting *attrition experiment*. In the following subsections, we first describe the general approach for features construction and then we specify each setting in details.

2.1 Preprocessing and Features Creation

As first step, we clean the original log files removing any noisy record. Particularly, we remove interactions with inconsistent timestamp or wrongly tool entry. Interactions recorded with blank user id or tool id are also not considered. The considered learners for each MOOCs are those reported in Table 1. As first step, we sort the interactions of each user chronologically and calculate the sessions (i.e., a list of actions) of every users. Whenever the timespan between two actions is larger than 30 min, we create a new session for the corresponding user. From these sessions we compute the following features for our classification experiment: number of sessions, number of requests, average number of requests per session, total session length, average session length, total timespan within clicks, average timespan within clicks, active days, average requests per total day, average requests per active day. We refer to this first set of features as *session information*. We introduce a second set of features by considering the tool field of the log data. Particularly, from the tools listed in Table 2, we count the number of interactions that refers to each tool. This leads to the following 8 features: assessment, assignment, course board, evaluation, file storage, forum, learning content and peer evaluation. We refer to this second group of features as *request per tool information*.

We want to emphasize that the way we construct all the various set of features is not dependent on any particular domain of the MOOCs. All the features we have defined are, in fact, derived by the timestamp and the tool information.

After creating the features, we randomly split the labeled learner data into a train and a test set, using an 80/20 ratio and a stratified split approach, preserving the class distribution of the overall users' population. As an example, a random stratified split of a dataset consisting of 100 users where 90% are non-completers and 10% are completers (9:1 ratio) with a training-testing splitting of 80/20 will create a training set consisting of 72 non completers and eight completers and a testing set consisting of 18 non-completers and two completers.

After computing the train and test sets, we use for both experiments Support Vector Machines (SVM) for the classification task. SVM has been widely used for classification, with applications such as text classification [22], validation of cancer tissue [23] and user classification in online courses in our previous research [15]. With this classification algorithm, each users is represented as a point in a multidimensional space in the number of features. SVM finds the decision hyperplane that best split the examples (described in terms of different features) in the given number of classes. The best split is the one for which the distance of the closest point of each class to the decision hyperplane is maximized [13, 14]. We train SVM using the train set and predict the class for the examples in the test set. In the training phase, SVM computes the best fitting hyperplane for the examples in the train set. In the fitting phase SVM tries to classify the examples from the test set according to the hyperplane obtained in the training phase.

We evaluated the quality of the obtained classification using F1 score [16], a metric defined as the harmonic mean of precision and recall. Therefore, F1 score is naturally bounded between 0 and 1. A value of 0 indicates that every examples has been misclassified. The higher the F1 score, the more accurate the prediction, with a value of 1 indicating that all the examples have been correctly classified.

2.2 Completers and Attrition Experiments

Our goal in the *completers experiment* is to verify if it is possible to obtain a correct user classification using a limited amount of log data in the analysis. Particularly, we investigate how the accuracy varies by considering featured collected from the first week up to the 8 weeks total length of the MOOCs. Insights in the number of weeks learner behaviour data would indicate potential completers and not-completers is extremely valuable for MOOCs' administrator and tutors. The sooner they can identify at-risk users, the more time they have to reach out to them before the finally leave the course. To this end, for each week we run a separate classification experiment, considering all interactions that took place up to that week. For example, for week 2 we consider all interactions that happened during the first 2 weeks, while for week 5 we use all the interactions registered during the first 5 weeks. The *session information* and the *request per tool information* as described in Sect. 2.1 are therefore created in relation to the total amount of considered weeks.

In the attrition experiment we take a step forward and look at more detail at the non-completers. We attempt to identify subgroups of users within the non-completers using the AMOES model as reference [3, 4]. Particularly, we try to classify the non-completers as healthy or unhealthy attrition (see also Sect. 1). Such classification would help MOOCs' administrators and operators to concentrate on and support only those learners that are willing to complete a course but that, for some particular reasons, are facing problems to do so. In order to do so, we use non-completers answers to a set of surveys that was sent out at the end of each course, to label them as either healthy or unhealthy [17]. For non-completers who did not complete the surveys, no labeling was possible and, therefore, those users are excluded from this experiment. Table 3 summarizes the number of students per MOOC according to their labelling.

Table 3. Number of users according to their label for each MOOC being Completers (C), Healthy (H), Unhealthy (U), Unlabeled (N).

MOOC	Completer	Healthy	Unhealthy	Unlabeled
Android	77	46	38	432
Client attention	60	6	3	22
Cloud-based learning	121	39	15	104
CBL (Edu)	99	16	13	54
CBL (Tools)	131	74	22	90
Community manager	320	78	59	364
Digital interactive TV	63	16	4	38
E-learning	81	27	12	125
E-learning tools	101	20	18	178
Medical urgency	49	6	4	59
User experience	62	28	10	89
Combined	**1164**	**356**	**198**	**1555**

With this setting, the *session information* and the *request per tool information*, are calculated using the whole interactions for each MOOC. Thus, we consider the course's whole duration but run experiments for individual MOOCs and also for the combined data set.

3 Experiment Results and Lessons Learned

The aim of the completers experiment is to verify how the accuracy of the prediction varies when different number of week are considered. Correctly detect users at-risk as soon as possible would allow course instructors to intervene and develop strategies to keep the at-risk users engaged and eventually prevent them from dropping out. The result of this experiment are presented in Table 4. Not surprisingly, we can see that the more weeks are considered the higher is the accuracy to correctly identify completers from the non-completers. From week 6 on, it is possible to obtain an average F1 Score that is higher than 0.8. However, when considering fewer weeks, the predictions exhibit a low accuracy. We believe that the reason for this is the features not providing sufficient information to the SVM, which is unable to correctly discern between completers and not-completers. This is due to too less interactions being available, the fewer weeks are considered. Furthermore, some of the analyzed MOOCs are characterized by the first two/three weeks with low amount of interactions, and not sufficient

Table 4. Results for the eight weeks analyzed in the experience, presenting the three metrics.

Analyzed weeks	Label	Result		
		Precision	Recall	F1 score
Week 1	Completer	0.75	0.60	0.67
	Non completer	0.00	0.00	0.00
	Average	0.62	0.50	0.56
Week 2	Completer	0.70	0.88	0.78
	Non completer	0.00	0.00	0.00
	Average	0.51	0.64	0.57
Week 3	Completer	0.60	0.55	0.57
	Non completer	0.29	0.33	0.31
	Average	0.49	0.47	0.48
Week 4	Completer	0.67	0.90	0.77
	Non completer	0.60	0.25	0.35
	Average	0.64	0.66	0.61
Week 5	Completer	0.76	0.57	0.65
	Non completer	0.44	0.67	0.53
	Average	0.65	0.60	0.61
Week 6	Completer	0.82	1.00	0.90
	Non completer	1.00	0.58	0.74
	Average	0.88	0.86	0.85
Week 7	Completer	0.88	0.91	0.89
	Non completer	0.83	0.77	0.80
	Average	0.86	0.86	0.86
Week 8	Completer	0.92	1.00	0.96
	Non completer	1.00	0.85	0.92
	Average	0.95	0.94	0.94

Table 5. Results from the SVM form five of the MOOCs and a General average for the eleven courses evaluated in the experience.

MOOC	Attrition classes	Precision	Recall	F1 score
Digital interactive TV	Completer	0.92	0.92	**0.92**
	Healthy	0.50	0.67	**0.57**
	Unhealthy	0	0	**0**
	Average	0.79	0.82	**0.81**
E-learning	Completer	1	0.94	**0.97**
	Healthy	0.75	0.60	**0.67**
	Unhealthy	0.25	0.50	**0.33**
	Average	0.88	0.83	**0.85**
E-learning tools	Completer	0.94	0.85	**0.89**
	Healthy	0.43	0.75	**0.55**
	Unhealthy	0.33	0.25	**0.29**
	Average	0.78	0.75	**0.76**
Medical urgency	Completer	1	0.90	**0.95**
	Healthy	0.33	1	**0.50**
	Unhealthy	0	0	**0**
	Average	0.86	0.86	**0.83**
User experience	Completer	1	1	**1**
	Healthy	0.60	0.50	**0.55**
	Unhealthy	0	0	**0**
	Average	0.78	0.75	**0.76**
Global	Completer	0.96	1	**0.98**
	Healthy	0.77	0.67	**0.71**
	Unhealthy	0.20	0.25	**0.22**
	Average	0.83	0.83	**0.83**

for correctly predicting at-risk users. Thus, these MOOCs introduce a certain noise, which subsequently cause a worsening of the overall predictions score. Weekly combined analyses on MOOCs with different structure of the course and organization, can therefore led to low accuracy, especially when only few initial weeks are considered. Increasing the number of analyzed weeks and, therefore, with more available interactions, the accuracy of the predictions increases accordingly.

In the *attrition experiment*, we aim at further classify the non-completers as either healthy or unhealthy. Healthy attrition indicates users enrolled that are however not interested into complete the course. On the other hand, unhealthy attrition is typical of those users that are motivated to complete the course but that are facing problems in doing so. If not supported and scaffolded in time, these users would eventually abandon the course and drop out. Because of this distinction, it is clear that instructors and administrators of MOOCs should focus and be interested on the unhealthy attrition group only. Table 5 reports the results of this experiment for some of the MOOCs. Due to the small number of considered users of most of the MOOCs, some of the results exhibit a low F1 score, which indicates that SVM misclassified most of the learners.

This is the case for the MOOC “Client Attention”, “Digital Interactive TV” and “Medical Urgency”.

Small datasets could introduce difficulties for the classification task performed by SVM. However, the F1 score grows for bigger datasets. Generally, the scores for the completer’s class are always higher than those of both healthy and unhealthy class. We can therefore conclude that users within the class of completers are more similar with each other than users of the class healthy and unhealthy attrition group. In this scenario, it is easier for SVM to distinguish completers from non-completers than to distinguish users within the two subgroups of non-completers. Additionally, it could be possible that the considered features are not good enough to further split the non- completers into the groups of healthy and unhealthy attrition.

4 Conclusions and Future Works

In this paper, we focused based on previous research on two research questions: Firstly, can we already identify learners at-risk to drop-out a MOOC in an early stage? Secondly, can we differentiate between the group of healthy attrition and unhealthy attrition? Experimentation with Support Vector Machine (SVM) based on learner logs from eleven MOOCs on the Telescope platform show first promising results but also leave much room for further improvements.

Correct classification of users into completers and non-completers or in relation to healthy and unhealthy attrition, is harder the fewer amount of data and weeks are analyzed. On the other hand, the earlier it is possible to identify non-completers or users exhibiting unhealthy attrition, the more time tutors or professors have in order to take actions to motivate such users to continue and complete the course. Correct identification of healthy and unhealthy attrition users is a crucial point, because in order to mitigate the overall numbers of dropouts, the focus should be only on the latter group. Healthy attrition users should not be accounted as dropouts, as their final goal is not to successfully complete a course, and therefore to describe them as at-risk, is incorrect.

Identify and construct more valuable features becomes necessary when the amount of information available is generally low or when the amount of analyzed time only include the first few weeks (or any short amount of time). Entry tests or graded assignments could offer more insights and indications of user behaviors already within the first weeks of a MOOC. Analyzing users’ scores, time needed to complete the assignments together with the numbers of wrong or right answers are all factors that would lead to more accurate predictions.

Encouraging social interactions through forum activities and peer evaluations also represent potentially valid ways to increase user engagement and thus mitigate dropouts. Users who constantly engage in the MOOC forum, create discussions, reply to other users’ questions, show high interest in the MOOC and therefore have higher motivations to succeed. Trying to engage a larger number of users to participate in a forum is advisable, but it is not a trivial task and most of the time it is not easily achievable. However, a highly active forum, even if only animated by a few users, could encourage other users as well, who do not actively participate, to at least spending time reading existing discussions and maybe finding answers to their concerns. Although in this way the forum participation would not be improved, it may however bring some improvements and more knowledge to all users. It is important to notice that the MOOC platform and especially the content should be accessible with a special benefit for students with disabilities as a way to increase user engagement through usability and accessibility [20, 21].

As future work, we want to investigate further what are the most important features to predict at an early stage learners at-risk to drop out but also to separate the group of healthy from the group of unhealthy attrition. We also want to research ways to apply generalize models or models for particular types of MOOCs in order to predict learner groups, in particular learners at-risk, for individually designed MOOCs. Final, we want to provide MOOCs' administrators and experts with a dashboard interface to help them individuate at risk users, and provide appropriate follow-up exploiting the results from the SVM classification.

Acknowledgment. This work is partially supported by European Union through the project MOOC-Maker www.moocmaker.org. Reference: 561533-EPP-1-2015-1-ES-EPPKA2-CBHE-JP.

References

1. Berge, Z.L., Huang, Y.-P.: A model for sustainable student retention: a holistic perspective on the student dropout problem with special attention to e-learning. Deosnews **13**(5), 26 (2004)
2. Angelino, L.M., Williams, F.K., Natvig, D.: Strategies to engage online students and reduce attrition rates. J. Educ. Online **4**(2), 1–14 (2007)
3. Gütl, C., Rizzardini, R.H., Chang, V., Morales M.: Attrition in MOOC: lessons learned from drop-out students. In: Uden, L., Sinclair, J., Tao, Y.H., Liberona, D. (eds.) Learning Technology for Education in Cloud. MOOC and Big Data. LTEC 2014. Communications in Computer and Information Science, vol. 446, pp. 37–48. Springer, Cham (2014)
4. Hernández, R., Morales, M. Gütl, C.: An attrition model for MOOCs. In: Formative Assessment, Learning Data Analytics and Gamification, pp. 295–311 (2016)
5. Xing, W., Chen, X., Stein, J., Marcinkowski, M.: Temporal predication of dropouts in MOOCs: reaching the low hanging fruit through stacking generalization. Comput. Hum. Behav. **58**, 119–129 (2016)
6. Li, W., Gao, M., Li, H., Xiong, Q., Wen, J., Wu, Z.: Dropout prediction in MOOCs using behavior features and multi-view semi-supervised learning. In: 2016 International Joint Conference on Neural Networks (IJCNN), pp. 3130–3137. IEEE (2016)
7. Yang, D., Sinha, T., Adamson, D., Rosé, C.P.: Turn on, tune in, drop out: anticipating student dropouts in massive open online courses. In: Proceedings of the 2013 NIPS Data-Driven Education Workshop, vol. 11, p. 14 (2013)
8. Lubis, F.F., Rosmansyah, Y., Supangkat, S.H.: Experience in learners review to determine attribute relation for course completion. In: 2016 International Conference on ICT For Smart Society (ICISS), pp. 32–36. IEEE (2016)
9. Pirker, J., Riffnaller-Schiefer, M., Tomes, L.M., Gütl, C.: Motivational active learning in blended and virtual learning scenarios: engaging students in digital learning. In: Handbook of Research on Engaging Digital Natives in Higher Education Settings, p. 416 (2016)

10. Hernández, R., Guetl, C., Amado-Salvatierra, H.R.: Facebook for e-moderation: a Latin-American experience. In: Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies, p. 37. ACM (2011)
11. Zhang, Q., Peck, K.L., Hristova, A., Jablokow, K.W., Hoffman, V., Park, E., Bayeck, R.Y.: Exploring the communication preferences of MOOC learners and the value of preference-based groups: is grouping enough? Educ. Technol. Res. Dev. **64**(4), 809–837 (2016)
12. Rizzardini, R.H., Gütl, C., Chang, V., Morales, M.: MOOC in Latin America: implementation and lessons learned. In: Uden, L., Tao, Y.H., Yang, H.C., Ting, I.H. (eds.) The 2nd International Workshop on Learning Technology for Education in Cloud. Springer Proceedings in Complexity, pp. 147–158. Springer, Dordrecht (2014)
13. Suykens, J.A., Vandewalle, J.: Least squares support vector machine classifiers. Neural Process. Lett. **9**(3), 293–300 (1999)
14. Joachims, T.: Text categorization with support vector machines: learning with many relevant features. In: Nédellec, C., Rouveirol, C. (eds.) ECML 1998. LNCS, vol. 1398, pp. 137–142. Springer, Heidelberg (1998). doi:10.1007/BFb0026683
15. Vitiello, M., et al.: Classifying students to improve MOOC dropout rates. In: EMOOCs 2016 Conference, Research Track, S. 501 (2016)
16. Sammut, C., Webb, G.I.: Encyclopedia of Machine Learning, p. 1335. Springer, Heidelberg (2011)
17. Guetl, C., et al.: Must we be concerned with the massive drop-outs in MOOC? An attrition analysis of open courses. In: Proceedings of the International Conference Interactive Collaborative Learning (ICL 2014) (2014)
18. MOOC Maker. Official Website (2017). http://www.mooc-maker.org/. Accessed 7 May 2017
19. MOOC Maker. Attrition and Retention Aspects in MOOC Environments. Version 5. Deliverable 1.6 (2017b). http://www.mooc-maker.org/wp-content/files/WPD1.6_INGLES.pdf. Accessed 7 May 2017
20. Amado-Salvatierra, H.R., Hilera, J.R., Tortosa, S.O., Rizzardini, R.H., Piedra, N.: Towards a semantic definition of a framework to implement accessible e-learning projects. J. Univ. Comput. Sci. **22**(7), 921–942 (2016)
21. Martín, J.L., Amado-Salvatierra, H.R., Hilera, J.R.: MOOCs for all: evaluating the accessibility of top MOOC platforms. Int. J. Eng. Educ. **32**, **5**(B), 2374–2383 (2016)
22. Tong, S., Koller, D.: Support vector machine active learning with applications to text classification. J. Mach. Learn. Res. **2**, 45–66 (2001)
23. Furey, T.S., et al.: Support vector machine classification and validation of cancer tissue samples using microarray expression data. Bioinformatics **16**(10), 906–914 (2000)