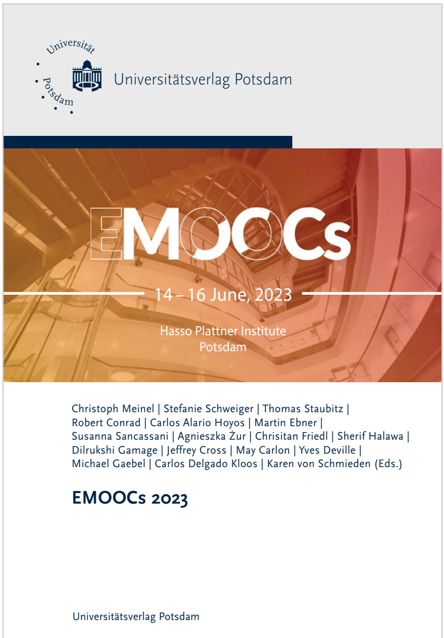
Universitätsverlag Potsdam

Article published in:

Christoph Meinel, Stefanie Schweiger, Thomas Staubitz, Robert Conrad, Carlos Alario Hoyos, Martin Ebner, Susanna Sancassani, Agnieszka Żur, Christian Friedl, Sherif Halawa, Dilrukshi Gamage, Jeffrey Cross, May Kristine Jonson Carlon, Yves Deville, Michael Gaebel, Carlos Delgado Kloos, Karen von Schmieden (Eds.)

EMOOCs 2023

2023 – vii, 350 p.
DOI https://doi.org/10.25932/publishup-57645



Suggested citation:
Julie Wasilewski; Mehdi Khaneboubi; Éric Bruillard: How to detect At-Risk learners in professional finance MOOCs, In: Christoph Meinel, Stefanie Schweiger, Thomas Staubitz, Robert Conrad, Carlos Alario Hoyos, Martin Ebner, Susanna Sancassani, Agnieszka Żur, Christian Friedl, Sherif Halawa, Dilrukshi Gamage, Jeffrey Cross, May Kristine Jonson Carlon, Yves Deville, Michael Gaebel, Carlos Delgado Kloos, Karen von Schmieden (Eds.): EMOOCs 2023 : Post-Covid Prospects for Massive Open Online Courses - Boost or Backlash?, Potsdam, Universitätsverlag Potsdam, 2023, S. 305–316.
DOI https://doi.org/10.25932/publishup-62481

This work is licensed under a Creative Commons License: Attribution 4.0
This does not apply to quoted content from other authors. To view a copy of this license visit:
https://creativecommons.org/licenses/by/4.0/

How to Detect At-Risk Learners in Professional Finance MOOCs

Step One

Julie Wasilewski, Mehdi Khaneboubi, and Éric Bruillard

Université Paris Cité
EDA, F-75006 Paris, France
julie.wasilewski@free.fr | mehdi.khaneboubi@u-cergy.fr | eric.bruillard@u-paris.fr

"Financial Analysis" is an online course designed for professionals consisting of three MOOCs, offering a professionally and institutionally recognized certificate in finance. The course is open but not free of charge and attracts mostly professionals from the banking industry. The primary objective of this study is to identify indicators that can predict learners at high risk of failure. To achieve this, we analyzed data from a previous course that had 875 enrolled learners and involve in the course during Fall 2021. We utilized correspondence analysis to examine demographic and behavioral variables.

The initial results indicate that demographic factors have a minor impact on the risk of failure in comparison to learners' behaviors on the course platform. Two primary profiles were identified: (1) successful learners who utilized all the documents offered and spent between one to two hours per week, and (2) unsuccessful learners who used less than half of the proposed documents and spent less than one hour per week. Between these groups, at-risk students were identified as those who used more than half of the proposed documents and spent more than two hours per week. The goal is to identify those in group 1 who may be at risk of failing and those in group 2 who may succeed in the current MOOC, and to implement strategies to assist all learners in achieving success.

1 Introduction

Despite being offered by reputable universities and validated with certificates, assessing MOOC participants' skills and their recognition can be challenging. Unlike traditional degrees, MOOCs do not confer formal academic credentials, making it difficult to evaluate the acquired skills even with certificates and badges. To address this, badges and micro-credentials have emerged as opportunities to recognize knowledge acquisition and certify skills, both professionally and institutionally, including the potential for academic credit [5, 12].

Micro-credentials, as defined by [16], encompass formation that spans multiple courses but falls short of a full degree. They are increasingly prevalent across various platforms and are referred to by different names such as certifications, specializations, or nanodegrees. These paid certificates consist of a series of independent or non-independent MOOCs, where successful completion of each course leads to the award of a certificate.

However, limitations persist for MOOCs and e-learning, besides the issue of certification, a recurrent issue in the MOOC is dropout. [13] found that approximately 10 % of registered participants actually complete the course. Dropout rates are prevalent across all online courses not just MOOC, [11] showed that they tend to be higher in online courses compared to face-to-face courses.

Numerous studies have examined dropout and failure in MOOCs, by trying to determine the factors that determine them. [10] identified three main categories of factors: student factors, course factors, and environmental factors. Student factors encompass demographics, motivation, and goals. Course factors refer to the design, implementation, and support of the course, as well as learner interactions on the platforms. Environmental factors pertain to external influences such as work constraints and supportive study environments. Course designers and researchers can directly address course factors to decrease dropout risks.

Learners in MOOCs typically juggle professional, personal, and family responsibilities, which necessitates available time to complete the courses. The interests and motivations of learners are closely tied to their own objectives [17]. Indeed, several studies have highlighted the link between the motivation of participants and the completion of the MOOC [3, 2]. Participant's motivation is an important aspect of learners.

Research on course factors has highlighted the increasing use of learning analytics to understand student behavior and predict at-risk students. Techniques such as logistic regression, decision trees, factor analysis, and neural networks have been employed for this purpose [21, 14]. However, despite the multitude of studies, no singular model has emerged as significantly superior [14].

Regarding forum interactions, participants' engagement can enhance their performance. Research on dropout and failure based on forum interactions has shown that individuals who actively use forums have a higher likelihood of success [22]. Moreover, forums might not provide a meaningful context for studying dropouts, as few at-risk students consistently participate. This emphasizes the importance of employing alternative methods to identify diverse dropout patterns [22]. [1] demonstrated that, in their context, factors observed in the first week of the course could predict dropout with 82 % accuracy.

Analyzing learning analytics allows for the classification of learners based on their platform behavior. These classifications reveal distinct groups of student trajectories. For instance, [9] identified four sub-groups: “Completing” learners who successfully completed most assessments, “Auditing” learners who infrequently participated in assessments, “Disengaging” learners who initially engaged but then decreased their participation, and “Sampling” learners who only watched video lectures for a few assessment periods. While these classifications primarily rely on behavioral data from the platform and remain relatively stable on a weekly basis, [20] found that demographic characteristics are linked to changes in student behavior and group dynamics within the first half of the course.

The aim of this article is to investigate the online behavior of learners in a French MOOC called “Financial Analysis” offered in 2021 as part of a program named International Certificate in Corporate Finance (ICCF).

We seek to understand how to support at-risk students in achieving success.

By examining the factors influencing dropout rates and exploring effective strategies to engage learners, this study aims to contribute to the improvement the MOOC and enhance the overall learning experience for participants. Understanding the dynamics of online learning and identifying successful approaches can pave the way for more effective educational offerings in the digital age.

2 Context

2.1 Micro-accreditation ICCF

The International Certificate in Corporate Finance (ICCF) is a certification offered through a partnership between *HEC Paris* business school, which provides the pedagogical content, and *First Finance*, a company specialized in professional learning and certification in finance, which implements the program. In 2013, *First Finance* launched the first MOOC in finance [6], which was followed by two other MOOCs in 2015, resulting in the creation of the ICCF program. The program offers two training sessions each year in March and October and costs 2950 €. French workers can use various public funding options for continuing education. To obtain the ICCF certificate, learners must complete all three courses independently and pass a final exam at an assessment center.

ICCF is composed of three MOOCs: Financial Analysis, Valuation, and Financial Decisions. These three courses have a similar structure: a prerequisite phase, four weeks of lessons released on a weekly basis, and a case study. Each week is composed of a new chapter divided into sections. A section is comprised of videos, quizzes, documents, and live sessions. Each course is estimated to require three to five hours of personal work per week. In addition to the MOOCs, learners can attend synchronous moments:

- a masterclass (once per course) takes place in presence and online, it's composed by a presentation by a reputed financial professional, and by a revisit of the concepts covered in the week.
- a meet-up (once per course) is for remote case study preparation sessions (organized two weeks before the submission of the case study), supervised by didactic tutors.
- an online live session (once a week) to revisit the concepts covered in the week's chapter, to correct the difficult questions of the week's quiz, and to answer questions from the learners.

These moments are recorded, and learners have access to the replays. The final exam is a quiz of 120 questions randomly selected from a database made of 2000 questions. If a learner fails at the case study a re-take may be offered. Successful completion of the three courses and the final exam allows students to obtain ICCF certificate.

2.2 Financial Analysis course

The MOOC *Financial Analysis* is the first part of the ICCF certification. A new chapter is delivered each week during five weeks and one assessment week. Chapters are organized in two parts: content with videos and documents and content with training and evaluative quizzes. The pedagogical content of the course aims to:

- Analyze the financial documents of a company (income statement, balance sheet, cash flow statement) in different strategic contexts.
- Construct and interpret the key indicators of a company's financial equilibrium.
- Carry out complete financial analyses independently.
- Make a relevant diagnosis of the financial situation of companies.

The validation of this course is achieved through the completion of a case study evaluated by peers. To prepare for the case study, participants can attend an online meetup with tutors. The case study and peer review are both mandatory.

Problematics: A dataset from the *MOOC Financial Analysis* has been extracted during a complete course done from September 2021 to December 2021. Based on this data we focused on understanding the behavior of learners in the course so that we could determine the common characteristics of failed students to be able to identify those at risk in future sessions. This article is a preliminary attempt to describe online behavior of learners in micro-credentials.

3 Method

3.1 Data

We have collected data from 875 learners in two forms:

- Variables produced by the Edx platform: time spent watching videos, number of videos viewed, quizzes 'scores and results.
- Declarative demographic variables: gender, age, nationality, level of study.

To explain and describe failure or success in the *MOOC Financial Analysis* we have used R software [18] to proceed to a Multiple Correspondence Analysis (MCA). MCA is an unsupervised algorithm permitting to explore the relationships between categorical variables. Similar to the Principal Component Analysis (PCA), MCA provides an overall view of the links between variables [7, 19].

Ignoring missing data can decrease the precision and bring strong biases in the analysis models. We used the *missData* R package to handle missing value by imputing missing data by proportions with the method of [8]. In addition, we performed a Hierarchical Agglomerative Clustering (HAC) to visualize clusters of learners based on their proximity.

3.2 Participants profile

The sample was composed by 277 females (mean = 36.2 years old, s.d = 8.3), 539 males (mean = 36.7 years old, s.d = 8.5) and 59 who did not respond to the gender question. They mostly came from Europe (63.3 %), Africa (27.5 %). 74 % of the learners have at least a master's degrees. All participants were registered at the beginning of the course and were French speakers.

4 Results

4.1 Influence of socio-demographic data on achievement

The influence of the socio-demographic variable on achievement has been made by searching for correlation between gender, age, levels of diploma and localization. A Pearson's chi-squared test does not permit to find influence of gender on success or failure ($X^2 = 1.08$, $p = 0.299$). Learners with a master's degree succeed more than the others but not significantly ($X^2 = 3.6$, $p = 0.05766$). However, learners aged between 35 and 45 years old (Figure 1), have a greater risk of failure than other age

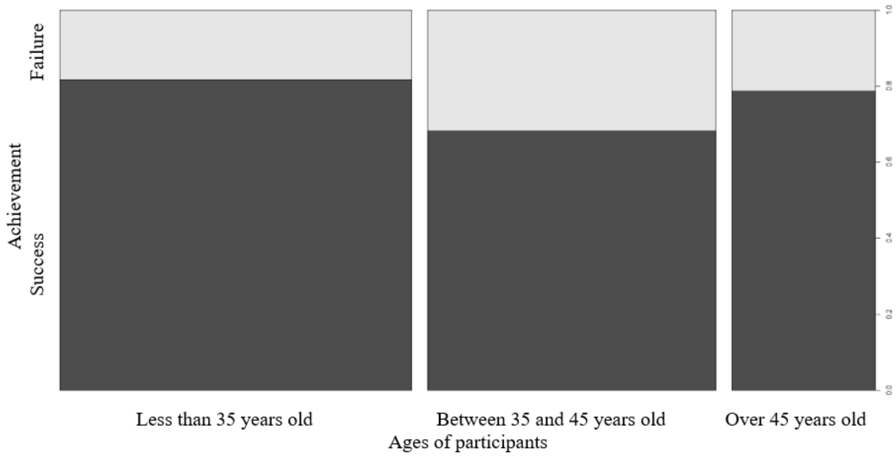


Figure 1: Frequency of results by learners' age

categories ($X^2 = 16.35$, $p < 0.001$). People living in Europe have a higher chance of success ($X^2 = 10.7$, $p = 0.005$).

4.2 Factors influencing success according to correspondence analysis

Data produced by the Edx platform have been recoded into qualitative variables:

- Completion of the quizzes (yes/no)
- Time spent on the platform (less than 1 hour, between 1 and 2 hours, more than 2 hours)
- Number of videos viewed (less than half, more than half, all videos)

Achievement's variable is produced at the end of the course and have two modalities: Success and failure.

The Figure 2 is a representation of learners according to their achievement, we observe a parabolic representation named Gutmann's effect occurring if there are multiple links between the variables [4]. This representation shows that individuals who failed the courses are grouped on the right whereas those who succeed are grouped on the left of the scatter plot. In between these groups, there are "at-risk" students who use more than half of the proposed documents and spend more than two hours per week.

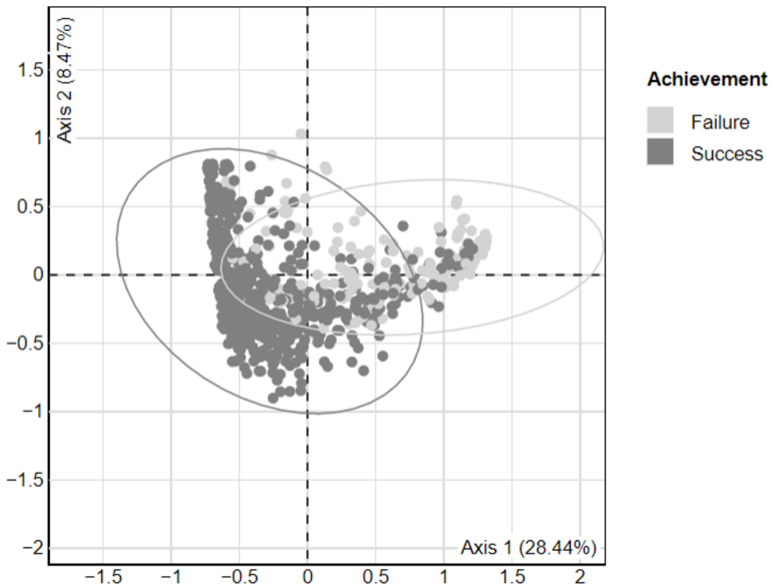


Figure 2: Representation of Individuals factor map (MCA)

The variable linked to the failure of the course are closed to the one not submitting the quizzes, not watching all the videos, and spending less than an hour on the platform during the first week of the course, while the variable identifying success is close to the learners who complete the quizzes, view all the videos and spend between 1 and 2 hours on the platform during the first week of the course.

To have a better view of the groups, we performed a Hierarchical Agglomerative Clustering (HAC). It allows us to have a more precise view of the clusters. The HAC produced a dendrogram, we decided to focus on a 3-cluster solution.

Figure 3 shows two main groups, the first one concern learners with a high tendency to fail. They share common characteristics:

- Not completing the weekly quizzes
- Watching less than half of the videos offered each week during the course.
- Spending less than two hours on watching videos on the platform during prerequisite courses
- Spending less than an hour on watching videos on the platform each week of the course.

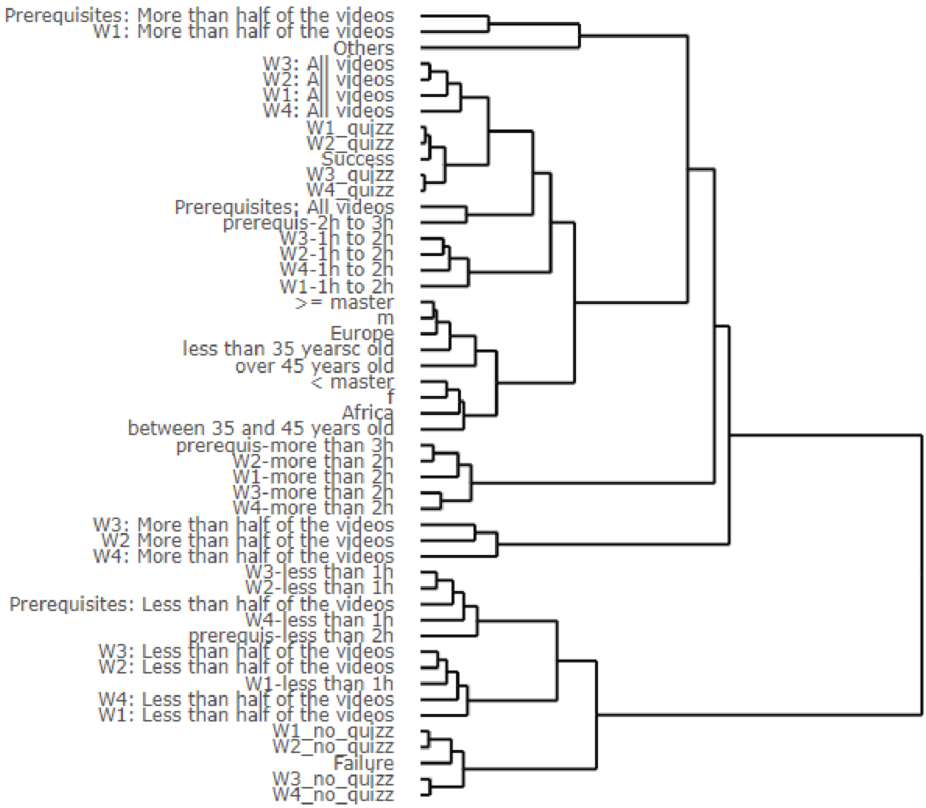


Figure 3: Classification of learners' achievement

The second one is people with a high tendency to succeed share common characteristics:

- Completing the weekly quizzes
- Watching all the videos offered each week during the course.
- Spending one hour to two hours on watching videos on the platform during prerequisite courses
- Spending one hour to two hours on watching videos on the platform each week of the course.

Between them, there are characteristics belonging to the same cluster of successful learners, but which are close to characteristic with high risk to fail. They share common characteristics:

- Spending more than three hours on watching videos at prerequisite course
- Spending more than two hours on watching videos each week of the course
- Not watching all of videos each week during the course

These results highlight that learners' early behaviors on the platform can indicate whether they will succeed or fail in the course.

5 Conclusion

We conducted an analysis of learners enrolled in the initial course of the ICCF program. Our analysis focused on various learner behaviors on the platform, including the time spent watching videos, the number of videos viewed, completion of quizzes, and course completion. The results revealed three distinct behavioral clusters. Cluster 1 consisted of individuals who successfully completed the course, demonstrated by high numbers of video views, completed quizzes, and spending one to two hours on the platform each week. Cluster 2 comprised individuals who failed the course, characterized by low levels of video views, completed quizzes, and time spent on the platform each week. Cluster 3 consisted of at-risk students who managed to succeed but exhibited similarities to Cluster 2, as they did not watch all the videos and spent more than two hours on the platform each week.

Interestingly, we observed that learner demographics had minimal influence on course success, contrary to previous studies [22, 15]. This discrepancy can be attributed to the fact that ICCF learners possess more homogeneous profiles compared to the general MOOC population which can be explained by the cost of this micro-certificate (2950 €) and its specific thematic on corporate finance.

We have been able to determine factors that can be identified at the beginning of the course, which allows us to detect learners at risk. We found that learner behavior on the platform during the early weeks of the course can serve as predictive indicators of their success or failure. These indicators enable us to determine risk profiles for learners. However, there are certain limitations to our research, as we were unable to ascertain the learners' specific objectives upon enrollment and whether their perceived failure aligned with their primary objectives. Nevertheless, previous studies [6] on the ICCF program have identified several objectives directly linked to their engagement and motivation, such as obtaining certification or expanding one's professional network. The success of these objectives requires the completion and success of the three MOOCs composing the ICCF.

Our next step involves analyzing data from other courses within the ICCF program to determine if similar patterns emerge and to develop strategies that promote success for all learners. We will thoroughly investigate the impact of success or failure in the initial course on subsequent micro-accreditation courses and explore whether the characteristics of at-risk learners in the first course remain consistent throughout subsequent courses.

Acknowledgment

We are grateful to First Finance and its staff for supporting this research by providing the data.

References

[1] A. Alamri, M. Alshehri, A. Cristea, F. D. Pereira, E. Oliveira, L. Shi, and C. Stewart. "Predicting MOOCs Dropout Using Only Two Easily Obtainable Features from the First Week's Activities". In: *Intelligent Tutoring Systems*. 2019, pages 163–173. DOI: 10.1007/978-3-030-22244-4_20.

[2] M. Badali, J. Hatami, S. K. Banihashem, E. Rahimi, O. Noroozi, and Z. Eslami. "The role of motivation in MOOCs' retention rates: A systematic literature review". In: *Research and Practice in Technology Enhanced Learning* 17.1 (2022), page 5. DOI: 10.1186/s41039-022-00181-3.

[3] L. Y. Chaw and C. M. Tang. "Driving High Inclination to Complete Massive Open Online Courses (MOOCs): Motivation and Engagement Factors for Learners". In: *Electronic Journal of E-Learning* 17.2 (2019), Article 2. DOI: 10.34190/JEL.17.2.05.

[4] P. Cibois. *Les figures de l'analyse des correspondances*. 2014.

[5] D. Gibson, N. Ostashewski, K. Flintoff, S. Grant, and E. Knight. "Digital badges in education". In: *Education and Information Technologies* 20.2 (2015), pages 403–410. DOI: 10.1007/s10639-013-9291-7.

[6] S. Huguenin. "Étude du projet de formation individuel des participants à la formation ICCF @ HEC Paris et de sa mise en oeuvre: Entrée, suivi et valorisation". 2021.

[7] F. Husson, S. Lê, and S. Pagès. *Analyse de données avec R (Pu De Rennes)*. 2016.

[8] J. Josse and F. Husson. "Handling missing values in exploratory multivariate data analysis methods". In: *Journal de la société française de statistique* 153.2 (2012), pages 79–99.

[9] R. F. Kizilcec, C. Piech, and E. Schneider. "Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses". In: *Proceedings of the Third International Conference on Learning Analytics and Knowledge*. 2013, pages 170–179. DOI: 10.1145/2460296.2460330.

[10] Y. Lee and J. Choi. "A review of online course dropout research: Implications for practice and future research". In: *Educational Technology Research and Development* 59.5 (2011), pages 593–618. DOI: 10.1007/s11423-010-9177-y.

[11] Y. Levy. "Comparing dropouts and persistence in e-learning courses". In: *Computers & Education* 48.2 (2007), pages 185–204. DOI: 10.1016/j.compedu.2004.12.004.

[12] M. Lewis and J. Lodge. "Keep Calm and Credential on: Linking Learning, Life and Work Practices in a Complex World". In: *Foundation of Digital Badges and Micro-Credentials: Demonstrating and Recognizing Knowledge and Competencies*. 2016, pages 41–54. DOI: 10.1007/978-3-319-15425-1_3.

[13] T. R. Liyanagunawardena, A. A. Adams, and S. A. Williams. "MOOCs: A systematic study of the published literature 2008-2012". In: *The International Review of Research in Open and Distributed Learning* 14.3 (2013), page 202. DOI: 10.19173/irrodl.v14i3.1455.

[14] K. S. Na and Z. Tasir. "Identifying at-risk students in online learning by analysing learning behaviour. A systematic review". In: *2017 IEEE Conference on Big Data and Analytics (ICBDA)*. 2017, pages 118–123. DOI: 10.1109/ICBDAA.2017.8284117.

[15] J.-H. Park and H. J. Choi. "Factors Influencing Adult Learners' Decision to Drop Out or Persist in Online Learning". In: *Educational Technology & Society* 12 (2009), pages 207–217.

[16] L. Pickard, D. Shah, and J. J. De Simone. "Mapping Microcredentials Across MOOC Platforms". In: *2018 Learning With MOOCS (LWMOOCS)*. 2018, pages 17–21. DOI: 10.1109/LWMOOCS.2018.8534617.

[17] I. Quentin. *Les intentions des apprenants de MOOC*. 2014.

[18] *R: A language and environment for statistical computing*. 2022.

[19] É. Vrillon. "Une typologie de MOOC de France Université Numérique (FUN): Méthode et enjeux". In: *Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation* 24.2 (2017), pages 87–111. DOI: 10.3406/stice.2017.1741.

[20] B. Yang, L. Shi, and A. Toda. “Demographical Changes of Student Subgroups in MOOCs: Towards Predicting At-Risk Students”. In: *International Conference on Information Systems Development (ISD)*. 2019.

[21] E. Yukselturk, S. Ozekes, and Y. K. Türel. “Predicting Dropout Student: An Application of Data Mining Methods in an Online Education Program”. In: *European Journal of Open, Distance and E-Learning* 17.1 (2014), pages 118–133. DOI: 10.2478/eurodl-2014-0008.

[22] J. Zhang, M. Gao, and J. Zhang. “The learning behaviors of dropouts in MOOCs. A collective attention network perspective”. In: *Computers & Education* 167 (2021), page 104189. DOI: 10.1016/j.compedu.2021.104189.