Does failing the first assessment affect withdrawal decisions at course level? An empirical evidence

Juan Antonio Martínez-Carrascal [1] [0000-0002-7696-6050] and Teresa Sancho-Vinuesa[1] [0000-0002-0642-2912]

[1] Universitat Oberta de Catalunya, Rambla del Poblenou, 156, 08018 Barcelona, Spain

[jmartinezcarra, tsancho]@uoc.edu

Abstract. Classical evaluation models were based on measuring the grades of a test or set of tests performing along a specific course. The method had drawbacks, and continuous evaluation is nowadays a preferred method.

Continuous evaluation performs a continuous measure of the progress the student fulfils to achieve learning outcomes. However, and in practice, continuous evaluation systems usually consist – or at least, include – a set of assessments along the course which are graded and computed.

In this article, we demonstrate the relevance of the first test grade in terms of withdrawal. We make use of a quite uncommon method in learning analytics, such as survival curves. The method is commonly used in other fields, and in particular in medicine, to detect differences among populations under study.

Practical implementation will be carried out by analyzing a dataset containing higher education online university courses. Results show that the students failing the first grade show not only higher withdrawal rates, but also disengage earlier from the course. This evidence should reinforce the need to design actions targeted to this group, but also to use an initial test as a measure of engagement.

Keywords: withdrawal, assessment failure, student grades, survival analysis, learning analytics.

1 Introduction

Dropout in general, and withdrawal in particular, is one of the core problems of higher education institutions. Dropout means inefficient use of resources, and at the same time, implies frustration for students who leave either the system as a whole or a specific subject in particular.

This fact was one of the topics the Bologna process would aim to resolve, in a general framework of trans-European coordination [1]. As a specific factor, continuous evaluation was encouraged. Instead of the classical approach, where the students had a reduced set of tests to evaluate performance, competencies were considered the key, and continuous evaluation was encouraged.

From a practical point of view, this continuous evaluation would require "the evaluation of a subject through daily classwork, course-related projects, and/or practical work(s) instead of a final examination system." [2]. In practice, however, tests are still being performed and competence acquisition is often measured through grades gathered in a set of tests.

In this scenario, we hypothesize that continuous evaluation is conditioned in some way by the first of the evaluative assessments. Although its relative weight can be minor compared to the whole set of activities, a low grade can have a discouraging effect on the student. To validate our hypothesis, we raise the following research question:

- RQ: To what extent does failing the first evaluative assessment condition withdrawal decision of online students?

As it can be seen we look for an answer that goes beyond a pure affirmative question, indicating that there can be an influence. We aim to compare withdrawal depending on this first test, quantifying the specific impact.

To provide additional significance, we will analyze an assorted set of courses from an open database provided by the Open University. It includes data about a set of courses that have not been specifically designed for our research, including 22 editions of 7 different courses with over 30.000 total enrollments. Results will show that the students failing this first test show not only a higher withdrawal ratio but also tend to abandon the course earlier.

2 Theoretical framework

The theoretical models behind dropout were established around 1975. Works by Tinto [3] establish the first model on the topic. Tinto's model was known as the student integration model. It included both academic factors related to the student herself and factors related to the institution. As a whole, the model considered a set of interactions that conditioned the decision to drop out.

After this initial model, we can find different works that rely on this theory. [4] introduced the 'student attrition model' which relies on the concept of behavioural intention, where dropout is conditioned by a mixture of factors, which include academic, social-psychological, environmental, and socialization factors.

None of these theories considers specifically online studies. [5] performs an analysis, based on the previous theories, and conforms the 'composite persistence model'. In this model, academic performance and dropout are finally a combination of student characteristics, student skills, external factors, and internal factors. The three models abovementioned are the most cited references for the study of dropout but are not unique [6]. We can also cite models by Kember [7] and Lee and Choi [8].

The core of these models is centred on university dropout, which could be defined as 'leaving the university study in which they have enrolled before they have obtained a formal degree'[9]. However, this phenomenon can also be analyzed at a micro-level. In particular, considering the fact of a student leaving a course she is enrolled in. In this case, the term withdrawal is preferred, although compilation works show that the formal definition is unclear. 78% of the recent studies do not provide a clear definition of the term [6]. There are also no specific theoretical models for withdrawal.

For the sake of our research, we will consider it as “voluntary or involuntary removal from a course before completion”, a consistent definition with references in the literature [10], [11]. It is noticeable that the concept includes not only the decision to abandon the course but also considers the time, as withdrawal is carried out before the end of the course.

Withdrawal analyses are normally set up on specific course analyses. Besides, we can find a mixture of quantitative and qualitative analysis. Early works related to withdrawal in online environments detected that the pressure of work, technical problems, and lack of time were withdrawal determinants[12]. More recently, focuses on family and organizational support, and course satisfaction and relevance.

Despite the relevance of time in withdrawal, time analysis is uncommon. Most studies are limited to a classification problem, aimed to determine variables the influence whether a student withdraws or not. Among those references to studies considering the relevance of time, we can cite [13], [14] which are focused on university dropout. Focused on a specific course, we can find a MOOC case example[15]. It must be pointed out that most studies focus on the institutional level, and not specifically on withdrawal at the course level.

Among those techniques to approach the problem, we can find correlation analysis, classifiers – both Bayesian and different decision trees -, variance analysis, logistic regression, support vector machines, neural networks, or machine learning techniques. These techniques are found both at university and course level and also in traditional university courses and MOOCs [16], [17]. MOOCs are one of the fields where withdrawal has been more analysed due to its higher rates[16].

Despite survival analysis is commoly used in other disciplines[18], references to survival analysis techniques in e-Learning problems are not so common. The basics behind the technique are described in [19]. The interest on it is more than justified, due to its focus on time – which is particularly relevant when analysing withdrawal – but also for providing better results that classical approaches in terms of prediction[20].

[20] suggests that more research should be performed using this approach. Among those works focusing on time, we can cite [13], [21]. Results in [21] indicate that the beginning of the course is a critical moment that concentrates a high number of withdrawals. [13] performs a survival analysis over time from a university-level perspective, with results showing that grade point average at first semester, gender, and location are relevant for determining university dropout.

Whichever method, the relevance of early activity is considered in different studies. Early activity in general is considered a predictor of final course performance[22]–[24]. Assignment grades in particular constitute a strong predictor of the final performance in MOOC courses [25].

In this scenario, we analyze the specific impact of early grades in evaluative assessments on the withdrawal decision of the student.

3 Methodology

From a methodological perspective, two critical factors arise. First, the technique to be used. Second, a specific database to work with. Regarding the method, and due to the relevance of time, we will map our study as a survival analysis problem, as described in Section 3.1. Regarding the data, we will make use of a publically available database created by the Open University. Details of this database are included in Section 3.2..

3.1 Mapping withdrawal as a survival analysis problem

Survival analysis is 'a collection of statistical procedures for data analysis where the outcome variable of interest is time until an event occurs' [19]. The method is commonly used in other disciplines such as medicine, where survival time or time to relapse is under consideration. A really interesting view of the technique with a practical approach can be seen in a series of articles[19], [26]–[28].

References to survival analysis are scarce in the field of education in general and withdrawal in particular. As indicated in Section 2, we can cite a couple of analyses of university dropout [13], [14] and another focused on MOOC courses[15].

Two specific aspects are needed to perform survival analysis, which are the event under consideration, and the time to event. The event under consideration will be the fact of withdrawing, while the time to event will be the number of days the student remains enrolled in the course.

As different courses will be analyzed we will consider as t=0 the initial day of the course. Times above t=0 will be interpreted as the number of days after the course starts. Negative values reflect a withdrawal after enrolling but before the course effectively starts.

As specific tools we will make use of Kaplan-Meier curves to visualize and analyze the relevance of the variable under analysis, looking for statistical significance and clear interpretation[19]. Statistical validation will be performed considering the null hypothesis that different groups generated based on the grade of the first assessment share the same hazard functions. Log-rank test (in particular, Peto's) will be used for being more robust, and also as it provides more weight to earlier events [29]. As a limitation, Kaplan-Meier does not allow quantifying hazard. Hazard can be computed by using a simple non-parametric method, such as Nelson-Aalen[19].

To quantify the impact, and considering we are not segregating populations based on multiple parameters, we can use a non-parametric method. In particular, we will use the Nelson-Aalen method to estimate cumulative hazard. Although non-cumulative hazard at a specific time can also be computed, cumulative estimation is preferred for being more stable.

3.2 Working dataset

The search for a dataset that allows for analysis linked to our RQ has lead us to consider the public dataset offered by the Open University[30]. At the highest level, this dataset provides information about 22 editions – namely presentations in the OU nomenclature – of 7 different courses. All courses present at least two editions. A total of 32,593 students are enrolled in these courses – modules in the OU nomenclature -.

This database includes both personal and academic data of the students under consideration. For the sake of our purpose, specific information to determine withdrawal – and in particular, withdrawal date - is included. Regarding assessments, the dataset also includes the whole set of evaluative activities linked to every course, with its weight and grades obtained by the students in the different editions of the course.

Table 1 includes information about the first evaluative assessment of the different courses under analysis, as well as its weight. We also provide the total course duration.

Table 1. Data regarding course characteristics

Module	Presentation	Presentation length (days)	Date of 1st assessment (days since presentation start)	Weight of 1st assessment (%)
AAA	2013J	268	19.0	10.0
AAA	2014J	269	19.0	10.0
BBB	2013J	268	19.0	5.0
BBB	2014J	262	19.0	0.0
BBB	2013B	240	19.0	5.0
BBB	2014B	234	12.0	5.0
CCC	2014J	269	32.0	9.0
CCC	2014B	241	32.0	9.0
DDD	2013J	261	25.0	10.0
DDD	2014J	262	20.0	5.0
DDD	2013B	240	25.0	7.5
DDD	2014B	241	25.0	10.0
EEE	2013J	268	33.0	16.0
EEE	2014J	269	33.0	16.0
EEE	2014B	241	33.0	16.0
FFF	2013J	268	19.0	12.5
FFF	2014J	269	24.0	12.5
FFF	2013B	240	19.0	12.5
FFF	2014B	241	24.0	12.5
GGG	2013J	261	61.0	0.0
GGG	2014J	269	61.0	0.0
GGG	2014B	241	61.0	0.0

Table 2 shows an overview of the course enrolment and dropout, including percentage of withdrawal, failure and pass. The pass group includes also those students qualified with distinction:

Table 2. Enrolment and ratios linked to dropout analysis for courses in the OU dataset

Course	#Students	Withdrawals	Fail	Pass
AAA	747	16.73%	12.18%	71.08%
BBB	7903	30.18%	22.32%	47.50%
CCC	4434	44.54%	17.61%	37.84%
DDD	6266	35.86%	22.49%	41.65%
EEE	2934	24.61%	19.15%	56.23%
FFF	7758	30.96%	22.02%	47.03%
GGG	2534	11.52%	28.73%	59.75%

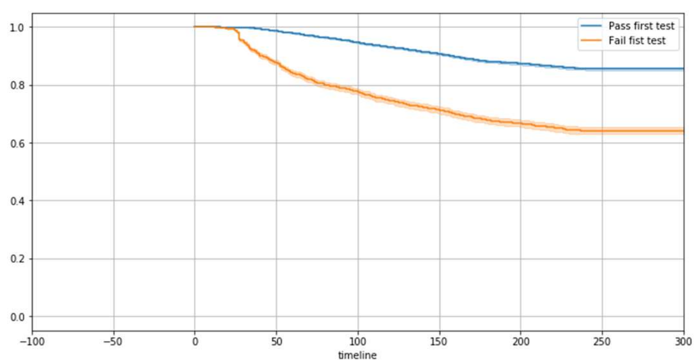
With this information, we can map the problem as a survival analysis problem as indicated in Section 3.1. It is also noticeable that the data used is suitable for performing Kaplan-Meier analysis. The OU has an active policy to manage dropout. Withdrawal time is always recorded. Due to this fact, independence of censoring and survival is guaranteed.

4 Results

4.1 Relevant difference in withdrawal depending on first assessment failure

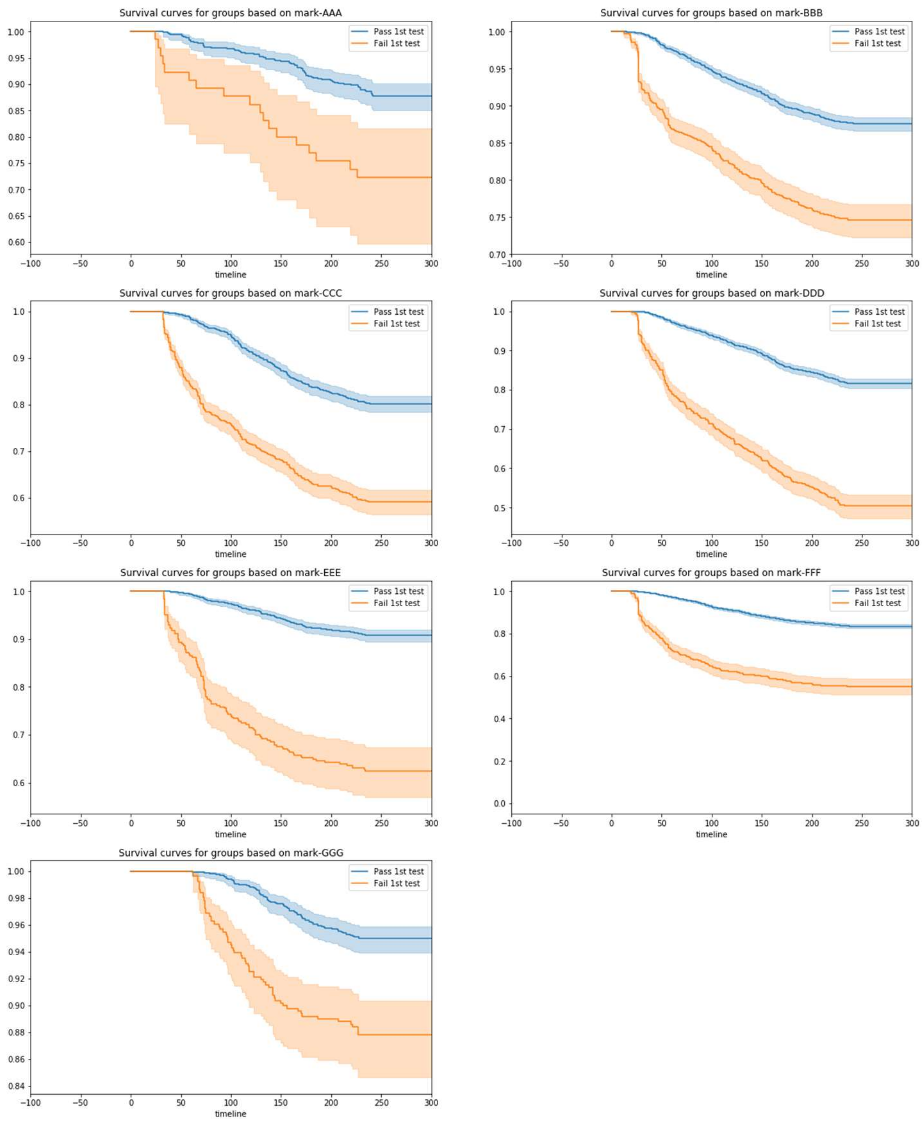
Before quantifying the potential impact, a detailed analysis must be performed to determine whether passing students and failing students of the first test show statistically significant differences in withdrawal patterns. Figure 1 shows the results of the Kaplan-Meier estimates comparing both groups:

Fig. 1. Survival curves for groups generated based on difference in first test



Considering that courses have different withdrawal ratios, we have analyzed also curves on a per-course basis. Results are shown in figure 2, where curves are shown including confidence intervals:

Fig. 2. Survival curves for individual courses



Statistical comparison for all groups results in relevant differences ($p<0.005$) in all cases, indicating both groups show different patterns regarding withdrawal. Table 3 reflects final withdrawal ratios for the generated groups and different courses, including the increase factor to facilitate comparison.

Table 3. Differences in withdrawal ratios at course end between students who pass and fail first assessment

	Pass 1st test	Fail 1st test	Factor
AAA	0.122	0.277	2.27
BBB	0.125	0.263	2.11
CCC	0.199	0.428	2.15
DDD	0.184	0.510	2.77
EEE	0.092	0.388	4.22
FFF	0.166	0.482	2.91
GGG	0.051	0.126	2.48

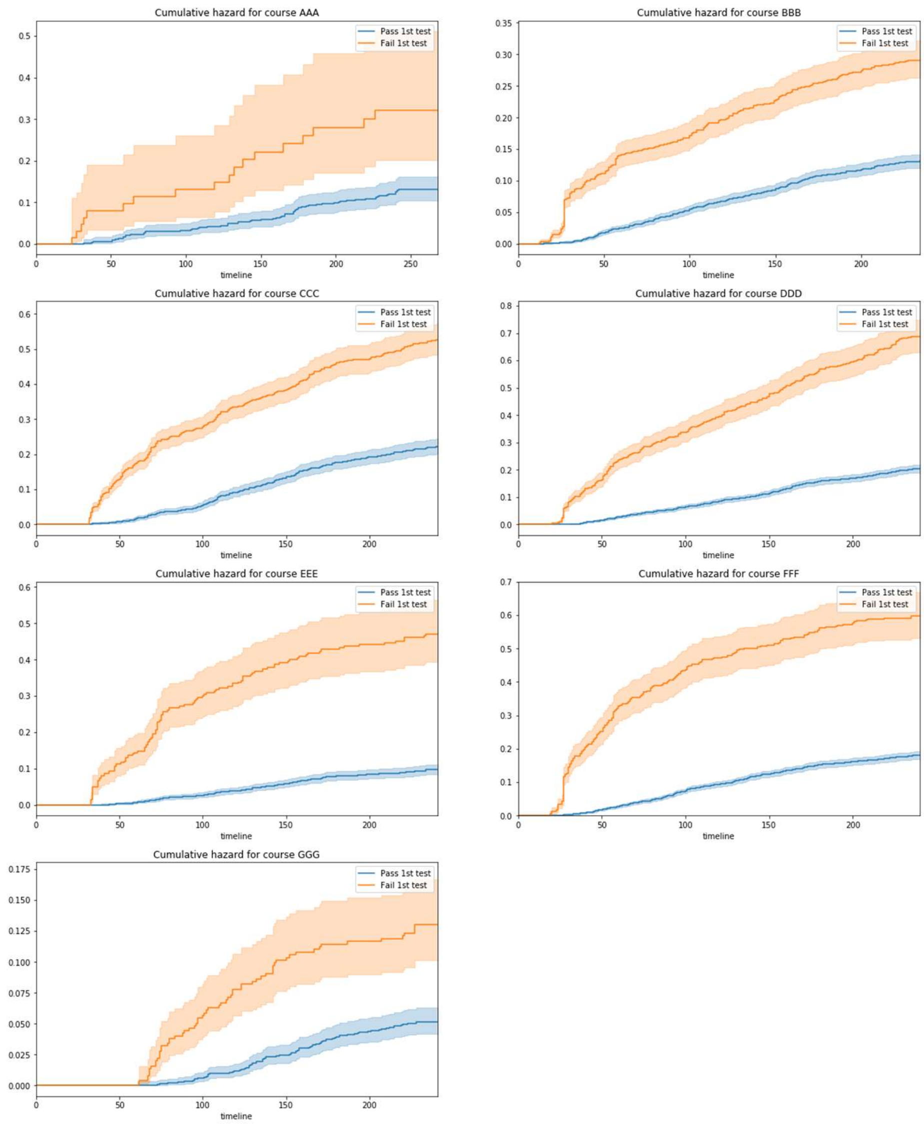
4.2 Higher and earlier withdrawal depending on the result of the first assessment

While survival curves provide group comparison in terms of survival probability, we can still get deeper into analyzing withdrawal hazard. Kaplan-Meier estimates do not provide this information, and we have to make use of specific methods to compute it. In particular, and considering that populations are segmented based on a single covariate, and we have no assumption about distribution, we can use a non-parametric estimator. In particular, we use Nelson-Aalen.

Nelson-Aalen is used to compute the cumulative hazard risk, understood as the probability of a student withdrawing from the course within a small interval of time, assuming she has survived up until the beginning of that interval. Cumulative hazard is preferred to point-wise estimations for being more stable.

In our case study, plots over time for individual courses have been plotted in Figure 3.

Fig. 3. Cumulative hazard curves for individual courses



As it can be seen, those students who fail the first test, show higher withdrawal rates in the long term, but also a relevant increase in early withdrawal.

5 Discussion

The research carried out contributes in two specific ways. First, it demonstrates the utility and potential application of survival analysis applications in e-learning. Second, it provides insight into the relevance of the first assessment in withdrawal.

Focusing on the RQ stated, results in section 4 clearly show that there are relevant differences in withdrawal among those students who fail the first test and those who pass it. This difference is reflected in Figures 2 and 3 for individual courses.

Regardless of the course, global withdrawal ratios are higher for those students failing the first test. As Figure 1 shows, the mean difference in survival at the end of the course is 2.57 higher for those students who pass the first test. When looking at individual courses, the increase in withdrawal at course end can be 4.22 times higher as reflected in Table 3. Besides this difference in final withdrawal, Figure 3 shows also an interesting insight. Much of the above difference is based on a much higher early withdrawal. Before going deeper into this fact, it must be pointed out that these tests are made in an early period when comparing to course duration. Data in Table 1 show that – except for course GGG - the test is performed around the first month of the course, in a course lasting for around 9 months. Also, some of the assessments do not even compute for global course grade and are under 20% in all cases.

With this data in mind, our results would indicate that the first assessment of a course has an interesting predictive power, regardless of its weight and even time. From a learning analytics perspective, the group of students who fail the first test are suitable for targeted interventions aimed to retain them in the course. This result is aligned with the literature reflecting the impact of early activity in a broader sense [22]–[24].

From a methodological perspective, the method exposed fits the suggestion to explore survival analysis in e-learning scenarios[20] with a specific application to withdrawal analysis. It is at least noticeable that a method that is common in medical research shows really few references in e-learning. The authors are open to collaborate in research lines following this idea, and in particular, linked to the analysis of the impact of different factors on withdrawal.

References

[1] A. Veiga, "Researching the bologna process through the lens of the policy cycle," in *European and Latin American Higher Education Between Mirrors*, Sense Publishers, 2014, pp. 91–108.

[2] E. S. Sanz-Pérez, "Students' performance and perceptions on continuous assessment. Redefining a chemical engineering subject in the European higher education area," *Educ. Chem. Eng.*, vol. 28, pp. 13–24, Jul. 2019, doi: 10.1016/j.ece.2019.01.004.

[3] V. Tinto, "Dropout from Higher Education: A Theoretical Synthesis of Recent Research," *Rev. Educ. Res.*, 1975, doi: 10.3102/00346543045001089.

[4] J. P. Bean, "Interaction Effects Based on Class Level in an Explanatory Model of College Student Dropout Syndrome," *Am. Educ. Res. J.*, 1985, doi: 10.3102/00028312022001035.

[5] A. P. Rovai, "In search of higher persistence rates in distance education online programs," *Internet High. Educ.*, vol. 6, no. 1, pp. 1–16, Jan. 2003, doi: 10.1016/S1096-7516(02)00158-6.
[6] M. Xavier and J. Meneses, "Dropout in Online Higher Education: A scoping review from 2014 to 2018," doi: 10.7238/uoc.dropout.factors.2020.
[7] D. Kember, "A Longitudinal-Process Model of Drop-Out from Distance Education," *J. Higher Educ.*, vol. 60, no. 3, pp. 278–301, May 1989, doi: 10.1080/00221546.1989.11775036.
[8] Y. Lee and J. Choi, "A review of online course dropout research: Implications for practice and future research," *Educational Technology Research and Development*, vol. 59, no. 5. pp. 593–618, Oct. 2011, doi: 10.1007/s11423-010-9177-y.
[9] M. S. Larsen, *Dropout Phenomena at Universities: What is Dropout? Why Does Dropout Occur? What Can be Done by the Universities to Prevent Or Reduce It? : a Systematic Review*. Danish Clearinghouse for Educational Research, 2013.
[10] J. M. Lim, "Predicting successful completion using student delay indicators in undergraduate self-paced online courses," *Distance Educ.*, vol. 37, no. 3, pp. 317–332, Sep. 2016, doi: 10.1080/01587919.2016.1233050.
[11] T. J. McClelland, "Why do they leave? an exploration of situational, dispositional, institutional, technological, and epistemological factors on undergraduate student withdrawal from online studies at an institute of technology in New Zealand," Nov. 2014.
[12] G. Packham, P. Jones, C. Miller, and B. Thomas, "E-learning and retention: Key factors influencing student withdrawal," *Educ. + Train.*, vol. 46, pp. 335–342, Aug. 2004, doi: 10.1108/00400910410555240.
[13] J. C. Juajibioy, "Study of University Dropout Reason Based on Survival Model," *Open J. Stat.*, vol. 06, no. 05, pp. 908–916, 2016, doi: 10.4236/ojs.2016.65075.
[14] Y. Min, G. Zhang, R. A. Long, T. J. Anderson, and M. W. Ohland, "Nonparametric survival analysis of the loss rate of undergraduate engineering students," *J. Eng. Educ.*, 2011, doi: 10.1002/j.2168-9830.2011.tb00017.x.
[15] D. Yang, T. Sinha, D. Adamson, and C. Rose, "'Turn on, Tune in, Drop out': Anticipating student dropouts in Massive Open Online Courses," in *Proceedings of the NIPS Workshop on Data Driven Education*, 2013.
[16] P. M. Moreno-Marcos, C. Alario-Hoyos, P. J. Munoz-Merino, and C. D. Kloos, "Prediction in MOOCs: A Review and Future Research Directions," *IEEE Trans. Learn. Technol.*, 2019, doi: 10.1109/TLT.2018.2856808.
[17] L. Aulck, N. Velagapudi, J. Blumenstock, and J. West, "Predicting Student Dropout in Higher Education," 2016.
[18] F. Emmert-Streib and M. Dehmer, "Introduction to Survival Analysis in Practice," *Mach. Learn. Knowl. Extr.*, vol. 1, no. 3, pp. 1013–1038, Sep. 2019, doi: 10.3390/make1030058.
[19] T. G. Clark, M. J. Bradburn, S. B. Love, and D. G. Altman, "Survival Analysis Part I: Basic concepts and first analyses," *British Journal of Cancer*. 2003, doi: 10.1038/sj.bjc.6601118.
[20] S. Ameri, M. J. Fard, R. B. Chinnam, and C. K. Reddy, "Survival analysis based framework for early prediction of student dropouts," in *International Conference on*

Information and Knowledge Management, Proceedings, 2016, doi: 10.1145/2983323.2983351.

[21] S. Laato, E. Lipponen, H. Salmento, H. Vilppu, and M. Murtonen, "Minimizing the Number of Dropouts in University Pedagogy Online Courses," doi: 10.5220/0007686005870596.

[22] J. A. Martínez-Carrascal, D. Márquez Cebrián, T. Sancho-Vinuesa, and E. Valderrama, "Impact of early activity on flipped classroom performance prediction: A case study for a first-year Engineering course," *Comput. Appl. Eng. Educ.*, vol. 28, no. 3, pp. 590–605, May 2020, doi: 10.1002/cae.22229.

[23] L. P. Macfadyen and S. Dawson, "Mining LMS data to develop an 'early warning system' for educators: A proof of concept," *Comput. Educ.*, 2010, doi: 10.1016/j.compedu.2009.09.008.

[24] V. A. Nguyen, "The Impact of Online Learning Activities on Student Learning Outcome in Blended Learning Course," *Journal of Information and Knowledge Management*. 2017, doi: 10.1142/S021964921750040X.

[25] S. Jiang, A. E. Williams, K. Schenke, M. Warschauer, and D. O. Dowd, "Predicting MOOC Performance with Week 1 Behavior," *Proc. 7th Int. Conf. Educ. Data Min.*, 2014.

[26] M. J. Bradburn, T. G. Clark, S. B. Love, and D. G. Altman, "Survival Analysis Part II: Multivariate data analysis- An introduction to concepts and methods," *British Journal of Cancer*. 2003, doi: 10.1038/sj.bjc.6601119.

[27] M. J. Bradburn, T. G. Clark, S. B. Love, and D. G. Altman, "Survival Analysis Part III: Multivariate data analysis - Choosing a model and assessing its adequacy and fit," *British Journal of Cancer*. 2003, doi: 10.1038/sj.bjc.6601120.

[28] T. G. Clark, M. J. Bradburn, S. B. Love, and D. G. Altman, "Survival analysis part IV: Further concepts and methods in survival analysis," *British Journal of Cancer*. 2003, doi: 10.1038/sj.bjc.6601117.

[29] S. Yang and R. Prentice, "Improved logrank-type tests for survival data using adaptive weights," *Biometrics*, 2010, doi: 10.1111/j.1541-0420.2009.01243.x.

[30] J. Kuzilek, M. Hlosta, and Z. Zdrahal, "Data Descriptor: Open University Learning Analytics dataset," *Sci. Data*, vol. 4, no. 1, pp. 1–8, Nov. 2017, doi: 10.1038/sdata.2017.171.