A Causal Inference Study on the Effects of First Year Workload on the Dropout Rate of Undergraduates

Marzieh Karimi-Haghighi[1]([✉]) [ID], Carlos Castillo[1,2] [ID],
and Davinia Hernández-Leo[1] [ID]

[1] Universitat Pompeu Fabra, Roc Boronat 138, Barcelona, Spain
{marzieh.karimihaghighi,carlos.castillo,davinia.hernandez-leo}@upf.edu
[2] ICREA, Pg. Lluís Companys 23, Barcelona, Spain

Abstract. In this work, we evaluate the risk of early dropout in undergraduate studies using causal inference methods, and focusing on groups of students who have a relatively higher dropout risk. We use a large dataset consisting of undergraduates admitted to multiple study programs at eight faculties/schools of our university. Using data available at enrollment time, we develop Machine Learning (ML) methods to predict university dropout and underperformance, which show an AUC of 0.70 and 0.74 for each risk respectively. Among important drivers of dropout over which the first-year students have some control, we find that first year workload (i.e., the number of credits taken) is a key one, and we mainly focus on it. We determine the effect of taking a relatively lighter workload in the first year on dropout risk using causal inference methods: Propensity Score Matching (PSM), Inverse Propensity score Weighting (IPW), Augmented Inverse Propensity Weighted (AIPW), and Doubly Robust Orthogonal Random Forest (DROrthoForest). Our results show that a reduction in workload reduces dropout risk.

Keywords: University dropout · Machine learning · Causal inference · Average treatment effect

1 Introduction

Research on actionable indicators that can lead to interventions to reduce dropout has received increased attention in the last decade, especially in the Learning Analytics (LA) field [18,29,31,32,34]. These indicators can help provide effective prevention strategies and personalized intervention actions [17,27]. Machine Learning (ML) methods, which identify patterns and associations between input variables and the predicted target [25], have been shown to be effective at this predictive task in many LA studies [1,4,10,15,23,26].

Dropout is a serious problem especially in higher education, leading to social and financial losses impacting students, institutions, and society [7]. In particular, the early identification of vulnerable students who are prone to fail or drop

their courses is necessary to improve learning and prevent them from quitting and failing their studies [20].

We remark that among students who discontinue their studies, some subgroups are over-represented, something that needs to be considered when designing dropout-reduction interventions. For example, in the UK, older students at point of entry (over 21 years) are more likely to drop out after the first year compared to younger students who enter university directly from high school [17], something that we also observe in our data. In the US, graduation rates among ethnic minority university students are lower than among White students [30]. Disparities in dropout risk have been studied in previous work [11,13,14,16]. Recent studies [8,21,22,24] look at the influence on student's performance and dropout of factors such as having a scholarship or being employed. In our work, we consider the increased dropout risk of older students and of students who do not enter university immediately after high school, and we study the effects of some features such as age and workload (i.e., number of credits taken on the first year).

Research Contribution. In this work, we use causal inference methods to study the effects of several features on the risk of early dropout in undergraduates students. We consider students enrolled between 2009 and 2018 in eight centers at our university. The average dropout rate we observe among these students is 15.3%, which is lower than the European average (36%) [35]. The originality of our contribution relies on its focus on students who have higher risk, the combination of features, the use of causal inference methods, and the size and scope of our dataset.

Specifically, we predict the risks of early dropout (i.e., not enrolling on the second year) and underperformance (failing to pass two or more subjects in the first year in the regular exams[1]) using Machine Learning (ML) methods. ML models are created using features available at the time of enrolment and the predictive performance of the models is evaluated in terms of AUC-ROC (Area Under ROC Curve). For the sake of space, we focus our exposition on dropout.

Among features available at the time of enrolment, we obtain the most important features for predicting dropout in our setting, which are the workload (number of credits taken) in the first year, admission grade, age, and study access type. Focusing on the workload, which is the most important feature and one over which first-year students have some level of control (only a minimum number of credits is established), we compute its effect on dropout risk in different age and study access type groups. We use causal inference methods to test the effects of combinations of theses features, and calculate the average treatment effect on dropout; the methods we use are the most used in the literature [2,3] including Propensity Score Matching (PSM) [28], Inverse-Propensity score Weighting (IPW) [6], Augmented Inverse-Propensity Weighted (AIPW) [12], and Doubly Robust Orthogonal Forest Estimation (DROrthoForest) [5] methods.

---

[1] These students have an opportunity of taking a resit exam which may finally result in passing or failing the subject, but given that passing the regular exam at the end of the course is expected, we consider failing the regular exam as underperforming.

The rest of this paper is organized as follows. After outlining related work on Sect. 2, the dataset used in this study is described and analysed in Sect. 3. The methodology is presented in Sect. 4. Results are given in Sect. 5, and finally, the results are discussed and the paper is concluded in Sect. 6.

2   Related Work

Machine Learning (ML) methods have been used to predict dropout and detect students at risk in higher education and play essential roles in improving the students' performance [1]. In a reference [4], the impact of ML on undergraduate student retention is investigated by predicting students dropout. Using students' demographics and academic transcripts, different ML models result in AUCs between 0.66 and 0.73. Another reference [7] develops a model to predict real-time dropout risk for each student during an online course using a combination of variables from the Student Information Systems and Course Management System. Evaluating the predictive accuracy and performance of various data mining techniques, the study results show that the boosted C5.0 decision tree model achieves 90.97% overall predictive accuracy in predicting student dropout in online courses. In a study [23], early university dropout is predicted based on available data at the time of enrollment using several ML models with AUCs from 0.62 to 0.81. Similarly, in a recent study [10], several ML methods are used to predict the dropout of first-year undergraduate students before the student starts the course or during the first year.

Some studies look at the features driving dropout. A reference [9] identifies factors contributing to dropout and estimates the risk of dropout for a group of students. By presenting the computed risk and explaining the reasons behind it to academic stakeholders, they help identify more accurately students that may need further support. In a research [33], the potential relationship between some features (academic background, students' performance and students' effort dimensions) and dropout is investigated over time by performing a correlation analysis on a longitudinal data collected spanning over 11 years. The results show that the importance of features related to the academic background of students and the effort students make may change over time. On the contrary, performance measures are stable predictors of dropout over time. Influential factors on student success are identified in a reference [19] using subgroup discovery; this uncovers important combinations of features known before students start their degree program, such as age, sex, regional origin or previous activities.

Recent work uses sophisticated statistical methods including causal inference. In a very recent paper [21], using propensity score matching (PSM) it is investigated whether university dropout in the first year is affected by participation in Facebook groups created by students. The estimated effect indicates that participation in social media groups reduces dropout rate. Another recent paper [24], implements an uplift modeling framework to maximize the effectiveness of retention efforts in higher education institutions, i.e., improvement of academic performance by offering tutorials. Uplift modeling is an approach for estimating

the incremental effect of an action or treatment on an outcome of interest at the individual level (individual treatment effect). They show promising results in tailoring retention efforts in higher education over conventional predictive modeling approaches. In a study, the effect of grants on university dropout rates is studied [22]. The average treatment effect is estimated using blocking on the propensity score with regression adjustment. According to their results, grants have a relevant impact on the probability of completing college education.

In our paper, we carefully measure the effect of the most important features (the number of credits in the first year, age, and study access type) on the early risk of dropout in undergraduate studies. This effect is obtained for combinations of these features. The Average Treatment Effect (ATE) is measured using multiple causal inference methods [2,3] as discussed in the introduction. It is noteworthy that according to a recent survey, the methods we use in this paper have not been applied in related studies so far [1].

3   Dataset

The anonymized dataset used in this study has been provided by Universitat Pompeu Fabra and consists of 24,253 undergraduate students who enrolled between 2009 to 2018 to 21 different study programs offered by eight academic centers. From this population, about 5% of cases were discarded for various reasons: 54 had an external interruption in their education between the first and second study year, 469 students did not have grade records (dropped out before starting), 560 students were admitted but did not enroll for the first trimester, and 74 cases did not have a study access type. Finally, 23,096 cases remained.

Students were admitted to university through four access types: type I students took a standard admission test (81%), type II students moved from incomplete studies in another university or were older than 25 (10%), type III students completed vocational training before (7%), and type IV students completed a different university degree before (2%). First year courses add up to a total of 60 credits across all study programs, this is also the median number of credits taken by first year students. However, students are also free to take additional credits out of different educational offers at the university such as languages, sports, and solidarity action.

The main studied outcome is dropout and consists of students who enroll in the first year but not in the second year. We also studied underperformance, which we defined as failing two or more subjects of the first year in the regular exams. Out of 23,096 cases, 3,531 students drop out (15.3%) and 6,652 students underperform (28.8%). Per-center dropout, underperformance, and other features are shown in Table 1. There are various differences among centers.[2] The students in the School of Engineering and Faculty of Humanities have the highest dropout and underperformance rates and the Faculty of Communication has

---

[2] ENG: Engineering, HUM: Humanities, TRA: Translation and Language Sciences, POL: Political and Social Sciences, HEA: Health and Life Sciences, ECO: Economics and Business, COM: Communication.

Table 1. Per-center statistics: number of students, drop-out rate, underperformance rate, percentage of national students, percentage of men, average age, average first year credits, average grade on the first year, and percentage of students in access type I.

Center	N	Dropout rate	Underperf. rate	National %	Male %	Avg. age	Avg. credits	Avg. grade	Access type I
ENG	2,444	41%	56%	89%	79%	19.4	63.4	4.6	65%
HUM	1,749	22%	33%	90%	32%	20.3	63.1	5.9	76%
TRA	2,292	16%	28%	88%	18%	19.3	62.9	6.3	83%
POL	1,683	14%	27%	94%	55%	18.8	63.1	6.2	87%
HEA	1,206	14%	16%	93%	25%	19.0	60.2	7.2	82%
LAW	5,479	12%	32%	92%	33%	19.3	62.5	6.0	79%
ECO	5,707	9%	26%	93%	47%	18.5	62.9	6.3	88%
COM	2,536	7%	7%	96%	27%	18.8	61.7	7.5	84%
All	**23,096**	**15%**	**29%**	**92%**	**40%**	**19.1**	**62.6**	**6.2**	**81%**

the lowest dropout rate and the best performance. In the Faculty of Communication, which has the lowest dropout and underperformance rates, there are more national students compared to other schools. In the School of Engineering, with the highest dropout and underperformance rates, males are in the majority. The average age in the two centers with the highest dropout and underperformance rates (School of Engineering and Faculty of Humanities) is higher compared to other faculties. In these two centers, the percentage of students admitted through a standard test (study access type I) is lower than other centers, and we can observe higher average number of credits and lower average grades in their first year compared to others. In the Faculty of Humanities, 22% of the students drop out (that includes 38% of those who underperform), while in the Faculty of Law, with almost the same underperformance rate, only 12% of the students drop out (including 18% of those who underperform). This might be partially explained because in the Faculty of Law, students are one year younger (19.3 vs 20.3 years old on average) and are also slightly more likely to come directly from high school (study access type I: 79% vs 76%).

4   Methodology

Our study focuses on modeling dropout and underperformance risks using data available at the time students enrol. The feature set for our two models consists of demographics (gender, age, and nationality), study access type, study program, number of first year credits, and average admission grade. Different ML algorithms: logistic regression (LR), multi-layer perceptron (MLP), and decision trees are used to predict the risks. Both ML models are trained using students enrolled between 2009 to 2015 (16,273 cases), and tested on students enrolled in 2016, 2017, and 2018 (6,823 cases). Due to space consideration and because of the severity of dropout, we mainly focus on this risk. Using a feature selection

Table 2. Dropout rate (%) across groups defined by age, workload (number of credits), and access type. Differences of ten percentage points or more appear in **boldface**.

Center	ENG	HUM	TRA	POL	HEA	LAW	ECO	COM	**All**
Age > Avg. age	45	28	24	**21**	13	**21**	16	12	**26**
Age ≤ Avg. age	39	21	15	**12**	15	**10**	8	6	**13**
Access types III/IV	44	28	**27**	23	16	18	**21**	11	**24**
Access types I/II	40	22	**16**	14	14	11	**9**	6	**14**
Credits > 60	47	29	22	22	21	19	11	7	18
Credits ≤ 60	39	20	15	13	13	10	9	6	14
Age > Avg. age & credits > 60	**53**	29	**33**	**29**	13	**33**	18	13	**32**
Others	**39**	22	**16**	**13**	14	**11**	9	6	**14**
Acc. types III/IV & credits > 60	**51**	**36**	**61**	**27**	15	**33**	**23**	10	**30**
Others	**40**	**22**	**16**	**14**	14	**11**	**9**	7	**15**

method based on decision trees (CART), we find that among the features available at the time of enrolment, the most important features in predicting dropout risk are the number of credits in the first year (workload), admission grade, age, and study access type.

In Table 2, we compare the dropout rate of different student groups in terms of these features and some of their combinations (due to space constraints, we omit some combinations). This comparison shows the following results. Students older than the average age have higher rate of dropout than younger students, across all centers except the Faculty of Health and Life Sciences (HEA). Students admitted through study access types III and IV have a higher dropout rate compared to the cases admitted through access types I and II; and students taking more credits than the median also have higher dropout rate. Considering combinations of these features, we can see that mostly older students with a number of credits larger than the average, as well as students admitted through access types III and IV who take a larger number of credits than the average have higher dropout rates. Results for underperformance (omitted for brevity) are similar, except in two senses: they do not hold for Engineering (ENG) and Humanities (HUM), possibly in part due to the overall lower grades in these centers compared to all others (Table 1), and they do not hold for credits alone, but for credits in combination with other features.

We aim to determine the causal effects on dropout of the features we studied by the following intervention: taking a workload in the first year of less credits than the median. The number of credits taken is a feature over which students have some degree of control at the enrolment time. Since higher dropout rates are observed among older students and students with access types III and IV, we are interested in the following scenarios:

– Scenario 1: in this scenario, the study group is limited to the first-year students who are older than the mean. Among these, those with less workload (credits < median) are considered as treated and those with more workload (credits ≥ median) are regarded as a control group.

– Scenario 2: in this scenario, the study group are all students. Older students taking less workload (credits < median) plus all younger students are considered as treated, and older students with more workload (credits ≥ median) are regarded as a control group.
– Scenario 3: in this scenario, the study group is limited to students from access types III and IV. Among these, students with less workload (credits < median) are considered as treated and students with more workload (credits ≥ median) are regarded as a control group.

The propensity of treatment is estimated in each scenario using Machine Learning (ML) models and input features including demographics (gender and nationality), study programs, and average admission grade. In scenarios 1 and 2, study access type is also added as a feature, and in scenario 3, age is added as a feature. We compute the Average Treatment Effect (ATE) of each treatment on the dropout probability using various causal inference methods:

– The propensity score matching method [28], in which data is sorted by propensity score and then stratified into buckets (five in our case). In our work, we obtain ATE by subtracting the mean dropout of non-treated (control) cases from treated ones in each bucket.
– Inverse-Propensity score Weighting (IPW) [6]: The basic idea of this method is weighting the outcome measures by the inverse of the probability of the individual with a given set of features being assigned to the treatment so that similar baseline characteristics are obtained. In this method, the treatment effect for individual i is obtained using the following equation:

$$TE_i = \frac{W_i Y_i}{p_i} - \frac{(1 - W_i)Y_i}{1 - p_i} \tag{1}$$

W_i shows treatment (1 for treated and 0 for control cases), p_i represents probability of receiving treatment (propensity score of treatment), and Y_i shows dropout (1 if drop out and 0 if not drop out) for individual i.
– Augmented Inverse-Propensity Weighted (AIPW) [12]: This method combines both the properties of the regression-based estimator and the IPW estimator. It has an augmentation part $(W_i - p_i)\widehat{Y_i}$ to the IPW method, in which $\widehat{Y_i}$ is the estimated probability of dropout using all features applied to the propensity score model plus the treatment variable. So, this estimator can lead to doubly robust estimation which requires only either the propensity or outcome model to be correctly specified but not both. We can compute the treatment effect on individual i as:

$$TE_i = \frac{W_i Y_i - (W_i - p_i)\widehat{Y_i}}{p_i} - \frac{(1 - W_i)Y_i - (W_i - p_i)\widehat{Y_i}}{1 - p_i} \tag{2}$$

– Causal forests from EconML package [5]: This method uses Doubly Robust Orthogonal Forests (DROrthoForest) which are a combination of causal forests and double machine learning to non-parametrically estimate the treatment effect for each individual.

Table 3. AUC-ROC of the prediction of dropout and underperformance across centers. Centers are sorted left-to-right by decreasing dropout rate.

Center	**All**	ENG	HUM	TRA	POL	HEA	LAW	ECO	COM
Dropout	**0.70**	0.72	0.72	0.68	0.67	0.57	0.64	0.67	0.68
Underperformance	**0.74**	0.82	0.80	0.73	0.69	0.53	0.64	0.69	0.76

Table 4. AUC-ROC of propensity score prediction.

	Scenario 1	Scenario 2	Scenario 3
N	3,866	23,096	1,963
Model	MLP	MLP	LR
AUC	0.75	0.91	0.75

In IPW, AIPW, and DROrthoForest, we obtain the individual treatment effect TE_i, which is the difference between the outcomes if the person is treated (treatment) and not treated (control). In other words, this effect is the difference of dropout probability when the student is treated and not treated; a negative value shows a reduced dropout risk and a positive value indicates an increased dropout risk. The resulting ATE is the average over individual treatment effects.

5   Results

The ML-based models of dropout and underperformance obtained using an MLP (Multi-Layer Perceptron) with 100 hidden neurons show the best predictive performance, with AUC-ROC of 0.70 and 0.74 for each risk respectively. Table 3 shows the AUC-ROC per center, and we observe that the AUC-ROC is in general higher for centers with higher dropout and underperformance rates. We also observe that dropout and underperformance predictions are not reliable for some centers, particularly Health and Life Sciences (HEA), and Law, where the AUC is less than 0.65.

For the three scenarios introduced in Sect. 4, the best predictive performance results obtained for the propensity score of the related treatment are shown on Table 4 in terms of AUC-ROC. Propensity is better predicted for scenarios 1 and 2 with the Multi-Layer Perceptron (MLP) and for scenario 3 with the Logistic Regression (LR). In each scenario, we removed study programs with relatively low predictive performance. According to the AUC values, ML models show accurate results in all of the scenarios, especially in scenario 2. In all scenarios, there is an overlap in the distribution of the propensity scores of treatment and control groups to find adequate matches (figure omitted for brevity). This is a necessary condition to be able to apply some of our methods.

Our goal is to determine whether these "treatments," which have a common feature of involving less workload, reduce dropout rate. The Average Treatment

Table 5. ATE obtained using Propensity Score Matching with five buckets.

Propensity	1. Low	2. Med-low	3. Med	4. Med-high	5. High
Scenario 1	0.18	0.04	−0.05	0.02	−0.08
Scenario 2	−0.04	0.03	0.00	0.02	−0.42
Scenario 3	0.04	−0.08	−0.17	0.30	−0.22

Table 6. IPW, AIPW, and DROrthoForest results estimating the Average Treatment Effect (ATE) and its 95% confidence interval [lower-ci, upper-ci] in three scenarios.

Scenario	IPW			AIPW			DROrthoForest		
	lower-ci	ATE	upper-ci	lower-ci	ATE	upper-ci	lower-ci	ATE	upper-ci
Scenario 1	−0.06	0.02	0.11	−0.01	0.07	0.15	−0.07	−0.06	−0.05
Scenario 2	−0.03	0.03	0.09	−0.06	0.01	0.08	−0.04	−0.04	−0.03
Scenario 3	−0.12	−0.01	0.10	−0.10	0.01	0.12	−0.07	−0.05	−0.03

Effect (ATE) obtained using propensity score matching is shown on Table 5. Across all three scenarios we can see mixed results, as in some propensity buckets the treatment increases the risk of dropout (scenario 1, bucket "1. Low"; scenario 3, bucket "4. Med-high") while in other cases the results are neutral or large reduction. In general, the results suggest that in high propensity to treatment conditions (bucket "5. High" i.e., students who are already likely to take less workload) there is a substantial reduction of the probability of dropout, particularly in scenarios 2 and 3.

The ATE values obtained from IPW, AIPW, and DROrthoForest methods are shown in Table 6 for all scenarios. In the case of IPW and AIPW, we can see that the 95% confidence intervals (from "lower-ci" to "upper-ci" in the table) contain the value zero. This means that the uncertainty in these methods is large and we cannot establish with them whether there is a change in the dropout risk due to the treatment. However, the results with the DROrthoForest method, which is a combination method of causal forests and doubly robust learner, are all negative with confidence intervals that do not contain the zero; indeed, they show a reduction of the probability of dropout of about 5% points in all three scenarios because of the treatment.

6   Discussion, Conclusions, and Future Work

In this study, we first created ML models to predict dropout (students who enroll in the first year but do not show up in the second year) and underperformance (failing two or more subjects in the regular exams of the first year), using only information available at the time of enrollment. The obtained AUC-ROC of our models were 0.70 and 0.74 for dropout and underperformance risks respectively, which shows a relatively reliable prediction of students at risk. This is particularly true for centers having large risk of dropout or underperformance, while

the performance of the same models for centers having lower risk is lower. This is to some extent expected and in those cases we are modeling a phenomenon that is more rare.

Next, we focused in dropout risk prediction and found that workload (first year credits) was an important feature. We also compared dropout risk across various groups of students. The comparison showed that to a large extent there is higher probability of dropout in older students (age > average-age), in students taking a higher workload (more first year credits than the established minimum and the median), and in students admitted through access types III and IV.

We considered three scenarios using a combination of these features. In these scenarios, interventions were designed having the common characteristic of a reduced workload for students. In each scenario, the propensity score of the treatment was obtained with AUC-ROC of 0.75–0.91 using ML-based models. Then, for each scenario, the Average Treatment Effect (ATE) on dropout was computed using causal inference methods. The results suggest a negative effect, i.e., a reduction of risk of dropout, following a lower number of credits taken on the first year. An actionable recommendation that these results suggest is to ask students at risk (in this study, older students and students admitted through access types III and IV) to consider taking a reduced workload (e.g., the minimum established), or to ask educational policy makers to consider revising the regulations that establish the minimum number of credits (e.g., to reduce the current minimum).

In addition to creating ML models for early prediction of dropout and underperformance risks that exhibit high predictive performance, the originality of this contribution is focusing on the vulnerable groups of students prone to dropout, studying combinations of different features such as workload, age, and study access type, and using different causal inference models to calculate the effects of these features on dropout in terms of ATE. Causal inference methods such as the ones we used provide a path towards effectively supporting the students. They also allow to perform observational studies, as education is a domain in which some types of direct experimentation might be unethical or harmful. We also used a large dataset and our results hold across substantially diverse study programs. We stress that the methodology we described is broadly applicable. Our findings are likely to be specific to this particular dataset, but show the general effectiveness of the methodology in this setting.

More scenarios can be defined in terms of other combinations of the relevant features, to determine their effects on dropout or underperformance. Additionally, the causal inference methods used in this study can also be applied to other risks faced by higher education students.

References

1. Albreiki, B., Zaki, N., Alashwal, H.: A systematic literature review of student' performance prediction using machine learning techniques. Educ. Sci. **11**(9), 552 (2021)
2. Athey, S.: Machine learning and causal inference for policy evaluation. In: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 5–6 (2015)
3. Athey, S., Wager, S.: Estimating treatment effects with causal forests: an application. Observational Stud. **5**(2), 37–51 (2019)
4. Aulck, L., Velagapudi, N., Blumenstock, J., West, J.: Predicting student dropout in higher education. arXiv preprint arXiv:1606.06364 (2016)
5. Battocchi, K., et al.: EconML: A Python Package for ML-Based Heterogeneous Treatment Effects Estimation, version 0.x (2019). https://github.com/microsoft/EconML
6. Bray, B.C., Dziak, J.J., Patrick, M.E., Lanza, S.T.: Inverse propensity score weighting with a latent class exposure: estimating the causal effect of reported reasons for alcohol use on problem alcohol use 16 years later. Prev. Sci. **20**(3), 394–406 (2019). https://doi.org/10.1007/s11121-018-0883-8
7. Bukralia, R., Deokar, A.V., Sarnikar, S.: Using academic analytics to predict dropout risk in E-learning courses. In: Iyer, L.S., Power, D.J. (eds.) Reshaping Society through Analytics, Collaboration, and Decision Support. AIS, vol. 18, pp. 67–93. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-11575-7_6
8. Choi, Y.: Student employment and persistence: evidence of effect heterogeneity of student employment on college dropout. Res. High. Educ. **59**(1), 88–107 (2018). https://doi.org/10.1007/s11162-017-9458-y
9. Chounta, I.A., Uiboleht, K., Roosimäe, K., Pedaste, M., Valk, A.: From data to intervention: predicting students at-risk in a higher education institution. In: Companion Proceedings 10th International Conference on Learning Analytics & Knowledge (LAK20) (2020)
10. Del Bonifro, F., Gabbrielli, M., Lisanti, G., Zingaro, S.P.: Student dropout prediction. In: Bittencourt, I.I., Cukurova, M., Muldner, K., Luckin, R., Millán, E. (eds.) AIED 2020. LNCS (LNAI), vol. 12163, pp. 129–140. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-52237-7_11
11. Gardner, J., Brooks, C., Baker, R.: Evaluating the fairness of predictive student models through slicing analysis. In: Proceedings of the 9th International Conference on Learning Analytics & Knowledge, pp. 225–234 (2019)
12. Glynn, A.N., Quinn, K.M.: An introduction to the augmented inverse propensity weighted estimator. Polit. Anal. **18**(1), 36–56 (2010)
13. Hutt, S., Gardner, M., Duckworth, A.L., D'Mello, S.K.: Evaluating fairness and generalizability in models predicting on-time graduation from college applications. International Educational Data Mining Society (2019)

14. Karimi-Haghighi, M., Castillo, C., Hernandez-Leo, D., Oliver, V.M.: Predicting early dropout: calibration and algorithmic fairness considerations. In: ADORE Workshop at the International Conference on Learning Analytics & Knowledge (LAK) (2021)

15. Kemper, L., Vorhoff, G., Wigger, B.U.: Predicting student dropout: a machine learning approach. Eur. J. High. Educ. **10**(1), 28–47 (2020)

16. Kizilcec, R.F., Lee, H.: Algorithmic fairness in education. arXiv preprint arXiv:2007.05443 (2020)

17. Larrabee Sønderlund, A., Hughes, E., Smith, J.: The efficacy of learning analytics interventions in higher education: a systematic review. Br. J. Edu. Technol. **50**(5), 2594–2618 (2019)

18. Leitner, P., Khalil, M., Ebner, M.: Learning analytics in higher education—a literature review. In: Peña-Ayala, A. (ed.) Learning Analytics: Fundaments, Applications, and Trends. SSDC, vol. 94, pp. 1–23. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-52977-6_1

19. Lemmerich, F., Ifl, M., Puppe, F.: Identifying influence factors on students success by subgroup discovery. In: Educational Data Mining 2011 (2010)

20. Márquez-Vera, C., Cano, A., Romero, C., Noaman, A.Y.M., Mousa Fardoun, H., Ventura, S.: Early dropout prediction using data mining: a case study with high school students. Expert. Syst. **33**(1), 107–124 (2016)

21. Masserini, L., Bini, M.: Does joining social media groups help to reduce students' dropout within the first university year? Socioecon. Plann. Sci. **73**, 100865 (2021)

22. Modena, F., Rettore, E., Tanzi, G.M.: The effect of grants on university dropout rates: evidence from the Italian case. J. Hum. Cap. **14**(3), 343–370 (2020)

23. Nagy, M., Molontay, R.: Predicting dropout in higher education based on secondary school performance. In: 2018 IEEE 22nd International Conference on Intelligent Engineering Systems (INES), pp. 000389–000394. IEEE (2018)

24. Olaya, D., Vásquez, J., Maldonado, S., Miranda, J., Verbeke, W.: Uplift modeling for preventing student dropout in higher education. Decis. Support Syst. **134**, 113320 (2020)

25. Pal, S.: Mining educational data to reduce dropout rates of engineering students. Int. J. Inf. Eng. Electron. Bus. **4**(2), 1 (2012)

26. Plagge, M.: Using artificial neural networks to predict first-year traditional students second year retention rates. In: Proceedings of the 51st ACM Southeast Conference, pp. 1–5 (2013)

27. Romero, C., Ventura, S.: Guest editorial: special issue on early prediction and supporting of learning performance. IEEE Trans. Learn. Technol. **12**(2), 145–147 (2019)

28. Rosenbaum, P.R., Rubin, D.B.: The central role of the propensity score in observational studies for causal effects. Biometrika **70**(1), 41–55 (1983)

29. Sclater, N., Peasgood, A., Mullan, J.: Learning analytics in higher education. Jisc, London, p. 176 (2016). Accessed 8 Feb 2017

30. Shapiro, D., et al.: Completing college: a national view of student completion rates-fall 2011 cohort (2017)

31. Siemens, G.: Learning analytics: the emergence of a discipline. Am. Behav. Sci. **57**(10), 1380–1400 (2013)

32. Syed, M., Anggara, T., Lanski, A., Duan, X., Ambrose, G.A., Chawla, N.V.: Integrated closed-loop learning analytics scheme in a first year experience course. In: Proceedings of the 9th International Conference on Learning Analytics & Knowledge, pp. 521–530 (2019)

33. Tanvir, H., Chounta, I.-A.: Exploring the importance of factors contributing to dropouts in higher education over time. Int. Educ. Data Min. Soc. (2021). ERIC
34. Viberg, O., Hatakka, M., Bälter, O., Mavroudi, A.: The current landscape of learning analytics in higher education. Comput. Hum. Behav. **89**, 98–110 (2018)
35. Vossensteyn, J.J., et al.: Dropout and completion in higher education in Europe: main report (2015)