Detecting At-Risk Students in an Animal Science Module with Performance Data (Test 1 and 2) from Two Different Cohorts in a Rural-Based University

F. Nde Fon(✉) and M. Sibanda

Department of Agriculture, Faculty of Science and Agriculture, University of Zululand, KwaDlangezwa, Empangeni 3886, South Africa
FonF@unizulu.ac.za

Abstract. Students' dropout which is often reflected as at-risk students remains a major problem in Higher Education worldwide not excluding South Africa. Numerous factors have been associated with student's dropout such as first-generation student pressure, choosing the wrong institution and course, social status at the institution, academic readiness, financial and family commitments and poor academic support. A couple of researchers have also established that early identification of at-risk students with mitigation strategies can decrease dropout. Test-1 (T1) has often been used for quick identification, but this study investigated the potential of adding Test 2 (T2) early to improve exam failure prediction. The study reflected on the role of T1 and T2 in predicting at-risk students (exam failure) in a small class (Applied Animal Nutrition module at a rural-based University). A total of 51 students belonging to cohort-1 (23) and cohort-2 (28) participated in the study for 14 weeks. A hybrid mode of lecture (half online and half face-to-face) was adopted after the relieved of COVID-19 pandemic. A double blinded study was used to avoid biasing the study. T1 and T2 predicted that 19.6% and 35.3% of students were at-risk of failing their exams, respectively. However, the exam failure rate showed that T1 and T2 were able to predict only 15.7% and 29.4%, respectively. T2 was 15.5% more accurate in predicting at-risk students than T1. It was also noticed that cohor-1 performed better than cohort-2 hence need for further investigation.

Keywords: At-risk · student performance · Test · predictor

1 Introduction

Higher education (University) is a whole new ball game for learners who comes from different High School settings (disadvantaged or advantaged) yet treated the same way especially those that met the minimum admission requirement and go straight into a degree programme [1, 2]. A lot of studies show that these students become overwhelmed with their new environment and struggle a lot to cope with their academic core functions hence leading to high dropouts or low throughput [1, 2]. Dropout has been reported to

M. Makua et al. (Eds.): TFC 2022, ASSEHR 732, pp. 18–26, 2023.
https://doi.org/10.2991/978-2-38476-006-0_3

be highest among first year students [3, 4] but it continuous to in-crease throughout the degree programme and even more serious at the postgraduate levels [5]. It was worsened by COVID-19 pandemic where all Universities moved from traditional face-to-face teaching and learning to modes such as emergency remote learning and online learning [6]. Prior COVID-19 pandemic, only few Universities were engaged with Blended teaching and learning and even fewer offering full online [7]. In Africa, rural-based Universities suffered the impact of COVID-19 pandemic most and South Africa was not exempted [8]. The rate of dropout and those that suspended university studies during the pandemic was even higher despite all the efforts from universities to promote student learning. Students that are more vulnerable to dropout at universities are often classified as at-risk students. Therefore, understanding these students and what makes them at-risk is critical in developing any prediction model or support programmes.

2 At-Risk Students

At-risk is generally used to describe students that do not perform well in a traditional educational environment and are vulnerable to failing or dropping a module or programme. Several characteristics associated with at-risk students have been discussed in different conceptual models of student's retention and dropout [27]. Tinto argue that each student is unique and their characteristics will depend on previous academic environment, family history, own abilities and values, academic intentions and commitment to the educational process [9]. Several other authors modified the retention and dropout models by adding new dimensions for better understanding such as external factors (friends) [10], human capital requirement influences [11] and amount of energy dedicated in studying [12].

3 At-Risk Students' Identification Models or Methods

Identification is key in reducing student's dropout which has been described by many world-wide [13, 27] as well as in South Africa as alarming, using the leaking pipeline where students failed to complete a programme within the minimum required time or dropout [9, 14–16]. Among the several methods used for identification of at-risk students are; 1. Predictive formula: an example of such formula was seen in a study where students at-risk to fail their National Medical Licensure Examination were identified using the predictive pass rate formula [17, 21]. The study used 7 variables before admissions and 10 variables after admissions which included both demographic data and performance data regressed. The method was very useful as the predictive formula was able to predict the five students that failed but there was a 58% false positive prediction, 2. Iterative Logistic Regression methods: This method used historical student record to try and predict if they are at-risk of failing or dropping out in future semesters [13]. Their results should that student identification was possible, and support was rendered, 3. Administrative student data and Machine learning methods: it uses regression analysis, neural networks, decision trees, and the AdaBoost algorithm to identify students at-risk. The results showed better predictive values for private universities than public [18]. It was interesting to find out that accuracy improved from 85 to 95% over four semesters in the private University while using the model, 4. Using formative assessment in higher education: Here, the students were given formative assessment task and during the process, tutors were assigned to the groups to assist struggling students [19]. The results demonstrated that at risk students' groups achieved a summative grade average of 8% higher than the class average grade without the implementation of the early warning system. These are very basic methods that are potentially cheaper and available for many if not all to implement [19]. In similar study, Veerasamy, et al. [20] classified the formative assessment method as simple, cheap and non-machine model available to all.

The study reported that personalized feedback automatically generated for individual students reduced the gap between at-risk students and those that were not [6]. A predictive model trained to use machine learning and deep learning algorithms: The authors depended on a machine to pick up student assessments scores and engagement intensity (learning behaviours) to produced results that were accurate precise, able to recall and support in the identification of at-risk students [22]. In a similar model, Karalar, et al. [23] were able to predict students at-risk of academic failure during the COVID-19 pandemic. Quiz score, degree, number of lecture notes down-load, number of other course materials downloaded, and total time spent watching recorded course videos were said to be the most important variables in the predictive model [23]. Early assessment: This method and the formative assessment method above were not much different but the timing (2 weeks) of the assessment makes it very unique in terms of early identification of at-risk students. The study showed a predictive accuracy of 60% in the identification of at-risk students as well as established that students with less than 25% marks were potential at-risk [24]. The non-machine models where key is stimulating the current study, as its look for alternatives testing that will continue to improve the predictive accuracy to higher percentages.

In most of the predictive models especially the machine learning models, the variable collected composed of both intrinsic and extrinsic factors that can play a key role in at-risk students learning. As good as all the models are, it is a very complex issue when it comes to individual students' intrinsic characteristics because development is natural although can be natured to an extent. In a study by Tokan and Imakulata [25] they were able to demonstrate that intrinsic motivation and extrinsic motivation does play a key role in students' learning. If this is true, it implies that every predictive model design or conclusion should consider these factors. Despite all these models, universities continue to see high numbers of at-risk students hence the need for continuous research.

4 Conceptual Framework for the Identification of At-Risk Students

Universities provided, tablets, computers, online learning management systems that were zero data rated, extra online classes, online tutorials, printed study guides and capacity building programmes for both staff and students, yet dropout is still high [1]. Many studies showed that the major problem of at-risk students emanated from missing assessments (test), high absenteeism [2], no participation in class activities, illness, drugs and alcohol abuse, peer influence and family related problems (Fig. 1). Therefore, there is a need to identify these students timeously and put programmes in place that will assist them. Figure 1 demonstrates a conceptual framework that is used to identify at-risk students. In most rural-based Universities, the first test is often used to identify at-risk students or

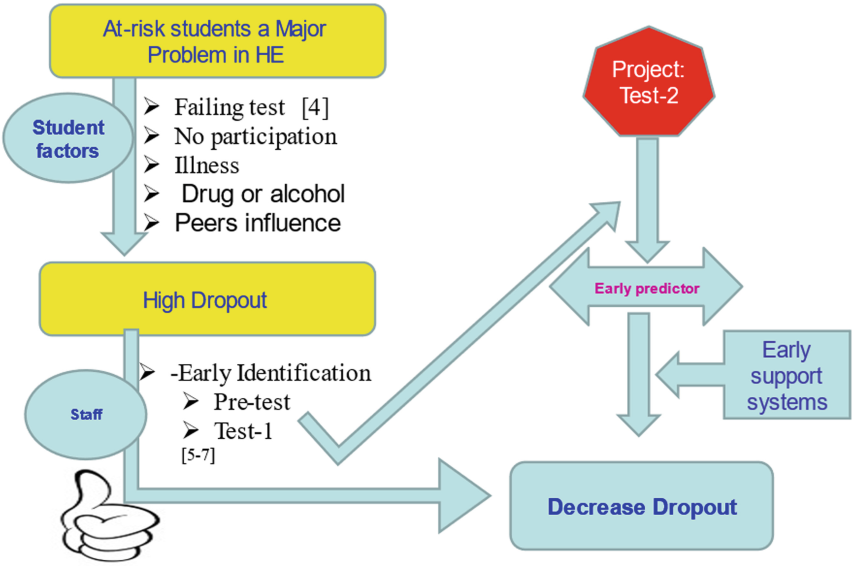

Fig. 1. Conceptual framework for the determination of At-risk students.

some diagnostic test [28]. This study is suggesting the concept of using test two instead of only test one to identify at-risk students. The major reason is that test one is often written so early in the semester which means that less course material to memorize and reproduce. Therefore, there might be chances of missing students that are at-risk hence the need to use early test two to try as an attempt to improve the identification of potential at-risk students. The rural-based University in question, uses test one as one of its main tool for the identification for at-risk students so as to provide quick help to avoid or reduce dropouts. Despite this early identification, the number of students that keeps on failing the exam or the module remains almost the same and sometimes even increased.

Therefore, the question asked was, is test one too early to predict students at-risk? Based on this, another question was asked, can test two be a better predictor of at-risk students than test 1? Hence, the main aim of this study was to compare the predictive potential of test 1 and test 2 in an Applied Animal Nutrition module to predict students at-risk (potential to fail their final year exams or module) between two different cohorts in the same class at a rural-based University. It was hypothesized that test one will be a better predictor of at-risk student than test 2.

5 Materials and Methods

The study was carried out at rural-based University using an Applied Animal Nutrition module class in the Province of KwaZulu-Natal South Africa. This study used test one (T1) and Test 2 (T2) as one of the main predictors of at-risk students (those that have the potential to fail the main exam in a small class (51 students). Within the class, it was also noticed that the student population were from a diverse background and communities around the University. Secondly, the class was coincidentally composed of students from two different cohorts namely, cohort-1 (23 students) and cohort-2 (28 students). Co-hort-1 students were admitted a year before Cohort-2 but could not take the module for one reason or another hence their presence in the current class. A hybrid mode of teaching (half online and the other face-to-face) was adopted after the relieved of COVID-19 pandemic. All tests and exams were summative in nature and T1, T2 and exams were written after 4, 6 and 13 weeks, respectively. All tests were written on Moodle Learning management system and the marks were extracted and downloaded in an excel sheet and prepared for analysis. A double blinded approach was used to avoid biasing the study.

5.1 Statistical Analysis

For descriptive statistics (frequency and averages of students that passed or failed T1, T2 or exams) SPSS version 22 software was used while comparison of the effect of those that failed T1 and T2 on exam failure rates were compared using chai-square. Differences between T1 and T2 on exam failure rate (at-risk students) predictions were confirmed when $P < 0.05$.

6 Results

The results showed that all three forms of summative assessment were successfully written as shown in Fig. 1a–d. In Fig. 1a the percentage of students that passed T1 was higher ($P < 0.05$) than those that past T2. It was also noted that cohort-2 students performed better ($P < 0.05$) than Cohort-1 students in both T1 and T2 (Fig. 1a). In the exams, less ($P < 0.05$) than 50% of the class did pass when the results were combined for both cohorts. However, cohort-2 score more than 50% exam passed prate while chort-1 did not. Figure 1d demonstrated the usefulness of T1 and T2 in predicting students that will fail the main exam. T2 predicted a higher ($P < 0.05$) number of students will fail the main exams than T1 and this was truly reflected in the exam results. However, false positive prediction was higher in T1 prediction than in T2. The overall predictive value showed that T2 was a better ($P < 0.05$) predictor than T1.

7 Discussion

Using assessments for the identification or tracking of at-risk students has always been used by many institutions as well as researchers [3] but the accuracy of predictions is still questionable because University dropouts are still high or throughput still low. However, most researchers through different concepts and theories have improved the efficiency of this method by using complex models to input beyond test marks for the prediction of at-risk students [4, 5]. Although using these models seems to be moving towards a positive direction, there are still setbacks because most of the models are still at the preliminary stages, expensive and complex. Some rural-based universities may not be well equipped and ready for such complex systems but are still wining by using methods from first principle to identify at-risk students. Diagnostic testing or test one is being

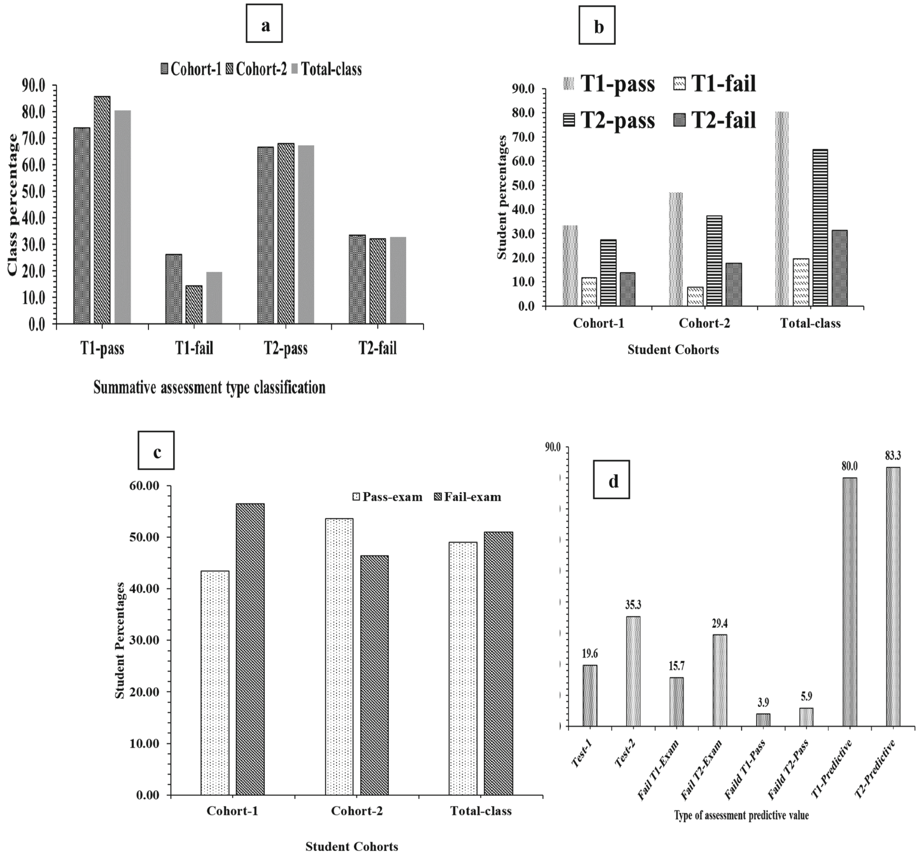

Fig. 2. 1a, b, c and d. Student percentage performances in Test 1, Test 2 and exams. 1a demonstrates the students' percentages pass and failure in all three assessments while 1b looks at cohort performance and 1c total passes and lastly 1d predicting exam failure potential using T1 and T2

used by many for identification of at-risk students, but the failure rate is still high hence the reason for T2 in this study as a potential improvement strategy (Fig. 2).

The results in this study showed that most students did very well in the first test. This was associated to the fact that T1 is always early in the semester and students turn to memorize the fewer pages of notes and regurgitate during assessment. However, it does not dispute the fact that there might still be students that will be struggling. It was also noticed that, the students in this class were from two different cohorts which added more value to the study. Cohort-1 was the batch prior COVID-19 pandemic while cohort-2 was a COVID-19 batch hence embraced online studies more that the resistant chort-1 that was used to face-to-face teaching and learning. Cohort-2 was consistently performing better than cohort-1 in both T1 and T2. This brought in the complexity of intrinsic factors that are very difficult to take care of [25]. However, this study resorted in separating cohorts into two and reported on their predictability of at-risk students separately. The results from the exams showed that less than 50% of the class passed the exam which was pushed down by cohort-1 where their overall pass percentage was less than 50%. In terms of at-risk students' predictability, T2 showed a higher predictive accuracy than T1. This was very clear when T2 showed a 15.5% more accuracy in identifying students that will fail the main exam than T1. It was also supported by T2 lower false positive predictor of students that will fail the exam than T1. The hypothesis was therefore, rejected that T1 was a better predictor of students that have the potential to fail the main exam than T2.

8 Conclusion

This study demonstrated that both test one and test two had a key role in predicting whether a student will fail an exam or not. Test two seems to be a better predictor with more accuracy (15.5%) than test one, hence can add more value among the different parameters being used for the identification of at-risk students. It was also noted that extrinsic and intrinsic value still have a key role to play as demonstrated by the differences in percentage pass rates observed between the cohorts.

Acknowledgements. The study would like to acknowledge the students and staff members who participated in the study and the Department of Agriculture and University of Zululand for providing the resources for carrying out this research. This study demonstrates that both test one and test two have a key role in predicting whether a student will fail an exam or not. Test two seems to be a better predictor (15.5%) than test one hence can add more value among the different parameters for the identification of at-risk students. It was also noted that extrinsic and intrinsic value still have a key role to play as demonstrated by the differences in percentage pass rate between the cohorts.

References

1. Williams, W. M. & Dahan, T. A. Exploring the Nursing Universal Retention and Success Model for Student Success in a School of Nursing. Nurse Educator, 10.1097 (2022).
2. Smith, B. I., Chimedza, C. & Bührmann, J. H. Individualized help for at-risk students using model-agnostic and counterfactual explanations. Education and Information Technologies 27, 1539-1558. https://doi.org/10.1007/s10639-021-10661-6 (2022).
3. Opazo, D., Moreno, S., Álvarez-Miranda, E. & Pereira, J. Analysis of First-Year University Student Dropout through Machine Learning Models: A Comparison between Universities. Mathematics 9, 2599 (2021).
4. Fourie, C. M. Risk factors associated with first-year students' intention to drop out from a university in South Africa. Journal of Further and Higher Education 44, 201-215. https://doi.org/10.1080/0309877X.2018.1527023 (2020).
5. Chinyakata, R., Raselekoane, N. R. & Mphahlele, R. R. C. Factors contributing to postgraduate students working and studying simultaneously and its effects on their academic performance. African Renaissance 16, 65-81. https://doi.org/10.31920/2516-5305/2019/v16n1a4 (2019).
6. Tsolou, O., Babalis, T. & Tsoli, K. The Impact of COVID-19 Pandemic on Education: Social Exclusion and Dropping out of School. Creative Education Vol. 12 No. 03, 16. https://doi.org/10.4236/ce.2021.123036 (2021).

7. Makhado, L. et al. An Exploratory-Descriptive Study on the Impact of COVID-19 on Teaching and Learning: The Experiences of Student Nurses in the Rural-Based Historically Disadvantaged University of South Africa. INQUIRY: The Journal of Health Care Organization, Provision, and Financing 59, 00469580221093191. https://doi.org/10.1177/00469580221093191 (2022).
8. Mthalane, P. P., Agbenyegah, A. T. & Dlamini, B. I. Reflection on student drop-out against the backdrop of COVID-19 in the South African Educational context amongst marginalised groups of students. African Sociological Review/Revue Africaine de Sociologie 25, 194-217 (2021).
9. Productivity Commission. The Demand Driven University System: A Mixed Report Card-Research Paper (2019).
10. Le, K. N. & Tam, V. W. How to retain postgraduate students in engineering programmes: a practical perspective. European Journal of Engineering Education 33, 511-524 (2008).
11. Ampaw, F. D. & Jaeger, A. J. Completing the three stages of doctoral education: An event history analysis. Research in Higher Education 53, 640-660 (2012).
12. McQueen, H. Integration and regulation matters in educational transition: A theoretical critique of retention and attrition models. British Journal of Educational Studies 57, 70-88 (2009).
13. Zhang, L. & Rangwala, H. in Artificial Intelligence in Education. (eds Carolyn Penstein Rosé et al.), 613-626 (Springer International Publishing).
14. Letseka, M. Student Retention and Graduate Destinations: Higher Education and Labour Market Access and Success (2010).
15. Mphekgwana, P. M., Mabila, T. E., Tirivangasi, H. M. & Makgopa, H. M. Analysis of survival rates among postgraduate students at a historically disadvantaged university in South Africa. Gender & Behaviour 18 (2020).
16. Johann, M. et al. (Department of Science and Technology, South Africa, 2015).
17. Tsunekawa, K., Suzuki, Y. & Shioiri, T. Identifying and supporting students at risk of failing the National Medical Licensure Examination in Japan using a predictive pass rate. BMC Medical Education 20, 419. https://doi.org/10.1186/s12909-020-02350-8 (2020).
18. Berens, J., Schneider, K., Görtz, S., Oster, S. & Burghoff, J. Early detection of students at risk–predicting student dropouts using administrative student data and machine learning methods. Available at SSRN 3275433 (2018).
19. Cassells, L. The effectiveness of early identification of ‘at risk’ students in higher education institutions. Assessment & Evaluation in Higher Education 43, 515-526, https://doi.org/10.1080/02602938.2017.1376033 (2018).
20. Veerasamy, A. K., Laakso, M.-J. & D’Souza, D. Formative assessment tasks as indicators of student engagement for predicting at-risk students in programming courses. Informatics in Education 21, 375-393. https://doi.org/10.15388/infedu.2022.15 (2022).
21. Azcona, D., Hsiao, I. H. & Smeaton, A. F. Detecting students-at-risk in computer programming classes with learning analytics from students’ digital footprints. User Modeling and User-Adapted Interaction 29, 759-788. https://doi.org/10.1007/s11257-019-09234-7 (2019).
22. Adnan, M. et al. Predicting at-Risk Students at Different Percentages of Course Length for Early Intervention Using Machine Learning Models. IEEE Access 9, 7519-7539. https://doi.org/10.1109/ACCESS.2021.3049446 (2021).
23. Karalar, H., Kapucu, C. & Gürüler, H. Predicting students at risk of academic failure using ensemble model during pandemic in a distance learning system. International Journal of Educational Technology in Higher Education 18, 63. https://doi.org/10.1186/s41239-021-00300-y (2021).
24. Veerasamy, A. K., D’Souza, D., Apiola, M.-V., Laakso, M.-J. & Salakoski, T. in 2020 IEEE Frontiers in Education Conference (FIE), 1-9 (IEEE).

25. Tokan, M. K. & Imakulata, M. M. The effect of motivation and learning behaviour on student achievement. South African Journal of Education 39 (2019).
26. Chung, E., Subramaniam, G. & Christ Dass, L. Online Learning Readiness Among University Students in Malaysia Amidst Covid-19. Asian Journal of University Education, 45-58%V 16. https://doi.org/10.24191/ajue.v16i2.10294 (2020).
27. Woldehanna, T., Endale, K., Hamory, J. & Baird, S. Absenteeism, dropout, and on-time school completion of vulnerable primary school students in Ethiopia: exploring the role of adolescent decision-making power in the household, exposure to violence, and paid and unpaid work. The European Journal of Development Research 33, 1349-1389, https://doi.org/10.1057/s41287-021-00454-5 (2021).
28. Kong, X., Liu, X., Gu, J., Qiao, Y. & Dong, C. in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 6002-6012.