



Developing a Framework for the Deployment of Predictive Analytics to Improve Postgraduate Student Throughputs at One Comprehensive South African University

Israel Kariyana¹ , Winter Sinkala²  & Neliswa Gqoli³ 

¹ Department of Continuing Professional Teacher Development, Faculty of Education, Walter Sisulu University, Mthatha Campus, South Africa.

² Department of Mathematics, Science and Computing, Walter Sisulu University, Mthatha Campus, South Africa.

³ Department of Adult, Foundation Phase & Education Foundations Education, Walter Sisulu University, Mthatha Campus, South Africa.

ABSTRACT

There is limited understanding of the opportunities available to universities through efficient deployment of predictive analytics. This study sought to develop a framework for the successful deployment of predictive analytics at one university to ensure high-quality postgraduate throughput rates. The study adopted a systematic literature review to elicit the opportunities presented by utilising predictive analytics in decision-making to promote postgraduate student throughput rates. It emerged that literature abounds on the manner big data analytics can be used to benefit universities and students. The study argued that the traditional, non-statistical approach which has long been used to address the unsatisfactory postgraduate throughput rates has failed to yield the required outcomes. It also noted the existing effort and support mechanisms to address postgraduate student retention and throughput rates which are necessary but not sufficient. A critical recommendation is that the proffered model should not be construed as a ‘perfect and single solution’ to capsize the poor postgraduate throughput rates at the university as different limitations exist. The study concluded that there is a clear call for the need to turn the current approach to the management and promotion of postgraduate student success. As such, the opportunities available are for those institutions that are committed to improving and magnifying their future practice by making meaning of the existing large data resources at their disposal.

Correspondence

Israel Kariyana

Email: ikariyana@wsu.ac.za

Publication History

Received 26th August, 2023

Accepted 13th October, 2023

Published online:

10th January, 2024

Keywords: Framework, Higher Education Institutions, Predictive Analytics, Throughput Rates

INTRODUCTION

Delivering the right curriculum to students can help any nation to achieve its goal of overall development in the area of education. For many nations, a stable and prosperous future still depends on high-quality education. A great education is influenced by many elements, such as actuality and information with specific goals.¹ According to Eduventures, "Predictive analytics can help schools accurately forecast student behaviour,

¹ Neha Kawchale and Rachana Satao, "Predictive Analytics in Education Context," *Int J Mod Trends Sci Technol* 2, no.11(2016):77–82.

especially when it comes to learning outcomes, recruiting, and retention”.² For instance, predictive analytics can use historical data to inform a school which candidates are most likely to enrol and, later in the student life cycle, which is most likely to continue with it and graduate. With the help of these data, institutions may help those who display indicators of distress before it is too late to take action. According to Kawchale and Satao, predictive analytics are frequently used in education to develop early warning systems that rely on data on student behaviour to spot students who are on the verge of dropping out and to direct interventions to assist them in finishing their course and or program of study.³ Additionally, predictive analytics can assist organisations in identifying and directing marketing efforts toward particular high schools that produce sizable percentages of the enrollees (or that the college or university would prefer to target for enrolment).

Institutions of higher learning are increasing their use of and capacity for learning analytics technologies. Institutions continue to advance even if more study is required to see whether this sociotechnical practice will result in appreciable benefits.⁴ Predictive analytics, as it is generally understood, is "an area of statistical analysis that deals with extracting information using various technologies that reveal relationships and patterns within large volumes of data that can be used to predict behaviour and events."⁵ In the field of education, predictive analytics are now more prevalent than ever before.⁶ Predictive Learning Analytics (PLAs) can detect students who are at risk of dropping out of school, but there has not been much research on how well they work in higher education.⁷

According to Kalechofsky “The natural survival instinct of humans is extended by predictive computing models. In today’s data intensive environment, predictive models are more crucial than ever to make sense of the world around us and to forecast, evaluate, or plan for potential future events.”⁸ Predictive analytics, according to Shankar in Kawchale and Satao, is a useful technique for bringing about good change across the student life cycle.⁹ It is crucial to keep students in school until they graduate because doing so will, among other things, increase student learning results, engagement and graduation rates, institutional return on investment (ROI) on recruitment expenditures, operational effectiveness, and the institution's capacity to meet the criteria of accrediting bodies and the federal government. It will also show that the institution is making much effort to improve student experiences. In this study, the phrases predictive analytics and predictive learning analytics are interchangeably used.

Inevitably by their nature, universities are custodians of large quantities of student data emanating from their many departments and relevant units. However, there are few or no known studies that have established the maximum utilisation of such data to maximise postgraduate throughputs. Wingfield argues that “universities are taking completion times very seriously, as they should, and faculties are being urged to improve their average throughput rates based on statistics that are generated annually.”¹⁰ However, this study argues that universities, especially those in Africa, have been generally slow to utilise data at their disposal to predict future trends and make critical decisions regarding postgraduate student throughputs. This study argues that the benefits of the full deployment of predictive analytics far outweigh the costs if the initiative is used to the optimum. However, optimum utilisation depends on whether there has been full commitment, acceptance and successful deployment within an all-systems-go operating environment. There is a need for no half-hearted approaches when real success is anticipated. In other words, all relevant stakeholders may need to buy this idea, fully cooperate from the onset and run with the idea. That way, everyone becomes part of the change within a continuous learning institution. Eduventures argues that “although gathering, analyzing, and interpreting data can be challenging,

² Eduventures, “Predictive Analytics in Higher Education: Data-Driven Decision-Making for the Student Life Cycle,” Eduventures, 2013, <https://www.sv-europe.com/wp-content/uploads/predictive-analytics-in-higher-education-data-driven-decision-making-for-the-student-life-cycle.pdf>.

³ Kawchale and Satao, “Predictive Analytics in Education Context.”

⁴ Kyle Jones et al., “Do They Even Care? Measuring Instructor Value of Student Privacy in the Context of Learning Analytics,” 2021.

⁵ Eduventures, “Predictive Analytics in Higher Education: Data-Driven Decision-Making for the Student Life Cycle,” 3.

⁶ Kelli A. Bird et al., “Bringing Transparency to Predictive Analytics: A Systematic Comparison of Predictive Modeling Methods in Higher Education,” *AERA Open* 7 (January 10, 2021): 233285842110376, <https://doi.org/10.1177/23328584211037630>.

⁷ Christothea Herodotou et al., “How Can Predictive Learning Analytics and Motivational Interventions Increase Student Retention and Enhance Administrative Support in Distance Education?,” *Journal of Learning Analytics* 7, no. 2 (September 19, 2020), <https://doi.org/10.18608/jla.2020.72.4>.

⁸ Hal Kalechofsky, *A Simple Framework for Building Predictive Models* (MSquared Consulting, 2016), 3.

<https://www.msquared.com/wp-content/uploads/2017/01/A-Simple-Framework-for-Building-Predictive-Models.pdf>.

⁹ Shankar M Patil, “Predictive Analytics in Higher Education,” *Life* 4, no. 12 (2015); Kawchale and Satao, “Predictive Analytics in Education Context.”

¹⁰ Brenda Wingfield, “Can We Improve Postgraduate Degree Throughput Rates?,” *South African Journal of Science* 107, no. 11/12 (November 1, 2011), 1. <https://doi.org/10.4102/sajs.v107i11/12.967>.

the undeniable payoff is the ability to make informed decisions that drive success, rather than leaving success to chance.”¹¹

Most governments globally had strained budgets to fund all their sectors even before the impact of the COVID-19 pandemic. Faced with this reality of persistently declining funding from governments while threatened by students at risk of failing courses or dropping out, universities worldwide are required to rethink their sustainability models. Consequently, there are growing calls for the need of accountability as well as soliciting third stream income generating projects for Higher Education Institutions (HEIs). Intrinsically, in many countries, models of funding HEIs place postgraduate students (Honours to Doctoral/PhD degrees), respectively, at the higher end of the income-generating continuum, and this is most profitable when the students graduate on time. It is, therefore, critical that universities deploy strategies that enhance postgraduate student throughputs. Ironically, the focal university is characterised by low throughput rates and publication records compared to other types of universities in South Africa. Consequentially and broadly speaking, the limited understanding of the opportunities available to universities through the efficient deployment of predictive analytics has prompted this study.

LITERATURE REVIEW

Selected Studies on Postgraduate throughput Rates in South Africa

According to Bird et al., state and federal policymakers are putting more and more pressure on institutions to raise the completion rates of both undergraduate and postgraduate students.¹² However, Botha contends that to improve institutional efficacy and realise national imperatives and goals, much more has to be understood about this issue.¹³ In recent years, it has become clear that university policies based on the "ivory tower" paradigm promote social exclusion and undercut the mission of public institutions to serve the public good. Universities are, therefore, more eager than ever to respond to student perspectives and success necessities by giving solutions and resources that are attentive to student requirements.¹⁴

According to the South African environment, funding is correlated with both the objectives of the national policy and the effectiveness of the institutions.¹⁵ Publicly, this asks these colleges to intentionally use means and purposes to institutionalize student-centred programs that improve students' success and quality of life.¹⁶ According to the Council on Higher Education (CHE), the University of Pretoria records dropout rates by cohort and makes distinctions between courses with varying lengths to better understand the process.

In one study based on postgraduate perspectives, Khauoe and Fore concluded that factors affecting the throughput rates of postgraduate students at a University of Technology included employment responsibilities, supervisor relationships, poor time management and the ambiguity of research.¹⁷ Considering these trends in Africa and other countries around the world, studies on the length of postgraduate studies and concerns about cutting the time it takes students to complete their postgraduate studies have turned into issues of extreme significance. These concerns are not only for students and higher education managers but also for governments, sponsors of postgraduate studies, and other stakeholders in higher education.¹⁸ There will eventually be specific circumstances where customised interventions are created for specific students.¹⁹ Cele cites Picciano and Avella et al. in support of their claim that big data analytics may help students and universities by:

- encouraging the use of resources and evidence-based choices,
- offering knowledgeable viewpoints on teaching strategies, student preparation, and their efficacy; increasing efficiency and organisational productivity,

¹¹ Eduventures, "Predictive Analytics in Higher Education: Data-Driven Decision-Making for the Student Life Cycle," 6.

¹² Bird et al., "Bringing Transparency to Predictive Analytics: A Systematic Comparison of Predictive Modeling Methods in Higher Education."

¹³ R J Botha, "Student Throughput Trends on Postgraduate Level: An African Case Study," *The Independent Journal of Teaching and Learning* 13, no. 2 (2018): 53–66.

¹⁴ N. Cele, "Big Data-Driven Early Alert Systems as Means of Enhancing University Student Retention and Success," *South African Journal of Higher Education* 35, no. 2 (May 2021), <https://doi.org/10.20853/35-2-3899>.

¹⁵ Mmatshuene Anna Segooa, Billy Mathias Kalema, and Ali Hussein Zolait, "Predictive Analytics to Improve Outcome-Based Funding for the Public Universities in South Africa through Big Data," in *2019 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)* (IEEE, 2019), 1–6.

¹⁶ Cele, "Big Data-Driven Early Alert Systems as Means of Enhancing University Student Retention and Success," 58-59.

¹⁷ Thakane Khauoe, "Factors Impacting Postgraduate Throughput Rates at a South African University of Technology" (Cape Peninsula University of Technology, 2020).

¹⁸ Botha, "Student Throughput Trends on Postgraduate Level: An African Case Study."

¹⁹ Cele, "Big Data-Driven Early Alert Systems as Means of Enhancing University Student Retention and Success."

- increasing the clarity with which student needs are recognized,
- enhancing high-level comparison and networking,
- providing predictive models for behaviour and performance, and
- enhancing holistic responsiveness to teaching and learning challenges.²⁰

Predictive Analytics in Higher Education

The use of predictive analytics has grown significantly in the field of education.²¹ Colleges and universities can use predictive analytics to determine which students are more at risk for decline, and armed with rich, historical data, craft segment-specific retention campaigns aimed at convincing them to keep working toward degree completion.²² Although higher education is a comparatively late adopter of predictive analytics as a management tool, predictive analytics are used by colleges and universities for a variety of purposes such as detecting students who may fail on their loans and focusing on alumni who are likely to make large donations to the institution.²³ Institutions are turning to predictive analytics due to a variety of contextual reasons.²⁴

The most typical application of predictive analytics is to pinpoint students who are at risk of failing classes or leaving school, and then to target these students with several student success strategies (such as intrusive advising, additional financial aid, etc.).²⁵ Predictive analytics allows for the timely making of informed decisions. In the real world, making informed decisions on time inevitably poses a competitive advantage for any institution or organisation. For instance, meeting or missing the recruitment or throughput targets by a number which has a significant impact can make all the difference. By examining the volume, veracity, velocity, diversity, and usefulness of massive amounts of data as well as interactive exploration, predictive analytics aims to provide pertinent information, actionable insight, better results, and smarter judgments as well as to forecast future events.²⁶

The use of predictive analytics in higher education has the potential to boost efficiency in how limited resources are allocated by concentrating on students who would benefit the most from additional intervention. Predictive analytics tactics have been widely and quickly adopted; a third of all institutions have made investments in the technology and spent hundreds of millions of dollars on it overall.²⁷ Predictive analytics enable a continual learning loop in which analysis guides choice by expanding the institutional management toolkit. These choices produce results which are then evaluated and coupled with new data to produce more educated choices. It is doubtful that one will perfectly get it the first time but making consistent progress through data-driven decisions ultimately pays off.²⁸ Reliable, consistent, and fair predictions from underlying models are necessary for efficiency improvements to be realised using predictive analytics. However, because most predictive analytics solutions utilised in higher education are proprietary and run by private entities, academics and college administrators have little to no capacity to evaluate predictive analytics software on these aspects. Institutions and students are exposed to several hazards because of this lack of transparency. The accuracy with which models identify at-risk students might differ significantly and this may result in an inefficient and wasteful use of institutional resources. Additionally, biased models may encourage institutions to intervene disproportionately with students from underrepresented backgrounds and may exacerbate psychological hurdles already experienced by students, such as feelings of fear and social isolation.²⁹

²⁰ Cele, "Big Data-Driven Early Alert Systems as Means of Enhancing University Student Retention and Success"; Anthony G Picciano, "The Evolution of Big Data and Learning Analytics in American Higher Education.," *Journal of Asynchronous Learning Networks* 16, no. 3 (2012): 9–20; John T Avella et al., "Learning Analytics Methods, Benefits, and Challenges in Higher Education: A Systematic Literature Review.," *Online Learning* 20, no. 2 (2016): 13–29.

²¹ Bird et al., "Bringing Transparency to Predictive Analytics: A Systematic Comparison of Predictive Modeling Methods in Higher Education."

²² Eduventures, "Predictive Analytics in Higher Education: Data-Driven Decision-Making for the Student Life Cycle."

²³ Manuela Ekowo and Iris Palmer, "Predictive Analytics in Higher Education," *Five Guiding Practices for Ethical Use*, 2017.

²⁴ Bird et al., "Bringing Transparency to Predictive Analytics: A Systematic Comparison of Predictive Modeling Methods in Higher Education."

²⁵ Rayane Alamuddin, Daniel Rossman, and Martin Kurzweil, "Interim Findings Report from the MAAPS Advising Experiment," June 27, 2019, <https://doi.org/10.18665/sr.311567>.

²⁶ Kawchale and Satao, "Predictive Analytics in Education Context."

²⁷ Jill Barshay and Sasha Aslanian, "Colleges Are Using Big Data to Track Students in an Effort to Boost Graduation Rates, but It Comes at a Cost," *The Hechinger Report*, 2019.

²⁸ Eduventures, "Predictive Analytics in Higher Education: Data-Driven Decision-Making for the Student Life Cycle," 6.

²⁹ Gregory M. Walton and Geoffrey L. Cohen, "A Brief Social-Belonging Intervention Improves Academic and Health Outcomes of Minority Students," *Science* 331, no. 6023 (March 18, 2011): 1447–51, <https://doi.org/10.1126/science.1198364>.

Improving Postgraduate Student throughputs through Predictive Analytics

For many years, various sectors have utilised predictive analytics, particularly when it comes to analysing consumer behaviour.³⁰ Literature posits that entities that fully utilise predictive analytics have a competitive advantage over others. Although there is preliminary evidence that PLAs can support learning, Herodotou et al. assert that little has been done to apply and evaluate them in the creation of motivational interventions or to take into account what this means in terms of institutional strategies to promote retention rates.³¹ Based on a previous study, Eduventures issues a call to action to all schools and institutions to think about incorporating predictive analytics into their arsenal of tools that support and facilitate evidence-based decision-making.³²

The experiences of postgraduate students, who frequently work either part-time or full-time while continuing their studies, are not well documented in the literature. The higher education industry is witnessing a rise in the number of postgraduate student workers. However, little study has been done to assess the expectations of postgraduate students who must simultaneously manage three overlapping role domains: work, personal life, and studies.³³ Institutions frequently hold extensive but unused databases. Institutions should first create a data strategy. This should include a list of the questions they have about the learning route and suggestions on how to use the data that is currently available. It might also serve as a guide for future data collection.³⁴

Universities can create general, shared, responsive solutions with the aid of data analytics, which also creates models that may predict future patterns and trends while allowing for the development of specific types of tailored interventions.³⁵ The university must identify possible at-risk students early in the first year of study to implement institutional support and intervention initiatives and increase retention rates in the students' second year of study. The practical use of real-time scoring and the provision of a list of students in danger of leaving their studies by the second year serve to highlight the necessity of this investigation. A list of students that need extra academic support is produced by using a predictive analytics tool like KNIME and importing/exporting the data and resulting in a management information system like HEDA.³⁶

The current study was triggered by the overwhelming backlog of postgraduate students who have overstayed in the system in the studied university's biggest faculty, the Faculty of Education. Students who fail to graduate on time face exclusion as the University considers them as liabilities. However, there is no known model that has purely focused on postgraduate students. At the university under study, the data provided by the Student Tracking Unit (STU) only relate to the examination-based modules and nothing about postgraduate research-based projects. The evidence of this is that there is no immediate data available about the attrition rates or retention strategies for research masters and doctoral students. Thus, they are left at the peril of manual interventions without any computerised interventions. What makes this more daunting is the fact that the students have diverse challenges which require a significant amount of time to compile, process and utilise. More to that, the challenges of postgraduates are not only limited to the students but also span from supervisors to institutional culture and support mechanisms. For instance, Kariyana and Marongwe conclude that participants recognized the current bad institutional research culture and strongly recommended, among other things, that the university become a learning institution to encourage the growth of strong campus research will.³⁷

METHODOLOGY

This study adopted a systematic review research design. Tawfik et al. define a systematic review as a review using a systematic method to summarize evidence on questions with a detailed and comprehensive plan of

³⁰ Eduventures, "Predictive Analytics in Higher Education: Data-Driven Decision-Making for the Student Life Cycle."

³¹ Herodotou et al., "How Can Predictive Learning Analytics and Motivational Interventions Increase Student Retention and Enhance Administrative Support in Distance Education?"

³² Eduventures, "Predictive Analytics in Higher Education: Data-Driven Decision-Making for the Student Life Cycle."

³³ Ingrid le Le Roux, "Coaching as Support for Postgraduate Students: A Case Study," *SA Journal of Human Resource Management* 16, no. 1 (2018): 1–7.

³⁴ Universities UK, "Analytics in Higher Education," 2016, www.universitiesuk.ac.uk.

³⁵ Cele, "Big Data-Driven Early Alert Systems as Means of Enhancing University Student Retention and Success."

³⁶ A Lourens and D Bleazard, "Applying Predictive Analytics in Identifying Students at Risk: A Case Study," *South African Journal of Higher Education* 30, no. 2 (June 2016), <https://doi.org/10.20853/30-2-583>.

³⁷ Kariyana Israel and Marongwe Newlin, "Developing Research Willpower At One South African University As A Strategy To Transform Higher Education Institutions: Experiences Of The Mentees.," *E-BANGI Journal* 18, no. 10 (2021).

study.³⁸ According to Littell et al., “A systematic review aims to comprehensively locate and synthesize research that bears on a particular question, using organized, transparent, and replicable procedures at each step in the process.”³⁹ Khan et al., also posit that a review earns the adjective systematic if it is based on a clearly formulated question, identifies relevant studies, appraises their quality and summarizes the evidence by use of explicit methodology.⁴⁰ Uman argues that systematic reviews typically involve a detailed and comprehensive plan and search strategy derived a priori, to reduce bias by identifying, appraising, and synthesizing all relevant studies on a particular topic.⁴¹ Pigott and Polanin, assert that “Systematic reviews analyze and synthesize a body of literature in a logical, transparent, and analytical manner.”⁴² Impellizzeri and Bizzini also add that “By conducting a properly performed systematic review, the potential bias in identifying the studies is reduced, thus limiting the possibility of the authors to select the studies arbitrarily considered the most ‘relevant’ for supporting their own opinion or research hypotheses. Systematic reviews are considered to provide the highest level of evidence.”⁴³ Considering that this was a feasibility study aimed at proposing the deployment of a predictive analytics model, the view was that a systematic literature review was the most suitable design.

FINDINGS AND DISCUSSION

Developing a Framework for the Deployment of Predictive Analytics in HE

According to Kalechofsky, “A model is when data is used to make decisions. In other words, it entails employing models that are statistically and empirically valid to make judgments and execute actions using data.”⁴⁴ Most essentially, organisations should use question- or curiosity-driven techniques when implementing and developing analytics. A data-driven strategy is not likely to be automatically successful. It frequently fails to generate a robust analytical framework and does not promote institutional buy-in. Data must be manipulated so that novice users may use it. It needs to result in practical insights.⁴⁵ Figure 1 below is the framework being proposed.

The test for collinearity and data partitioning is the most crucial procedure before modelling. Building a model often involves building it on training data and evaluating it on testing data. In an iterative process known as predictive analytics, the model's settings, parameters, and/or inputs are modified; the model is rebuilt using training data, and the new model is then evaluated using testing data. Until the ideal “fit” is produced, this process is repeated. A third data set (validation data) is used to provide a final estimate of the prediction model's performance or accuracy after it is implemented, preventing over-fitting in the process.⁴⁶ Every time data is utilised to train a predictive modelling technique, predictive models are produced. To put it another way, data with a strategy for predictive modelling equals a predictive model. A predictive model is developed using data and mathematics, and it incorporates learning by developing a mapping function between a collection of input data fields and a response or goal variable.⁴⁷

Kariyana and Sonn warn that “An infinite number of models can be created and analysed in logically and mathematically correct ways, but most are built on unrealistic assumptions and are therefore useless.”⁴⁸ Learning analytics “...requires bringing people with high levels of technical expertise together with others who understand pedagogy and educational processes.”⁴⁹ This, in the researchers’ view, is the first relevant step

³⁸ Gehad Mohamed Tawfik et al., “A Step by Step Guide for Conducting a Systematic Review and Meta-Analysis with Simulation Data,” *Tropical Medicine and Health* 47, no. 1 (December 1, 2019): 46, <https://doi.org/10.1186/s41182-019-0165-6>.

³⁹ Julia H Littell, Jacqueline Corcoran, and Vijayan Pillai, *Systematic Reviews and Meta-Analysis* (Pocket Guide to Social Work Re, 2008), 1.

⁴⁰ Khalid Khan et al., “Five Steps to Conducting a Systematic Review,” *Journal of the Royal Society of Medicine* 96, no. 3 (2003): 11821.

⁴¹ Lindsay S Uman, “Systematic Reviews and Meta-Analyses,” *Journal of the Canadian Academy of Child and Adolescent Psychiatry* 20, no. 1 (2011): 57.

⁴² Terri D. Pigott and Joshua R. Polanin, “Methodological Guidance Paper: High-Quality Meta-Analysis in a Systematic Review,” *Review of Educational Research* 90, no. 1 (February 25, 2020): 24, <https://doi.org/10.3102/0034654319877153>.

⁴³ Franco M Impellizzeri and Mario Bizzini, “Systematic Review and Meta-analysis: A Primer,” *International Journal of Sports Physical Therapy* 7, no. 5 (2012): 495.

⁴⁴ Kalechofsky, *A Simple Framework for Building Predictive Models*, 3.

⁴⁵ Universities UK, “Analytics in Higher Education.”

⁴⁶ Lourens and Bleazard, “Applying Predictive Analytics in Identifying Students at Risk: A Case Study.”

⁴⁷ Kalechofsky, *A Simple Framework for Building Predictive Models*, 3.

⁴⁸ I. Kariyana and R.A. Sonn, “A Case Scenario of the Dynamism and Fragility of the Feasibility of Deconceptualising Theoretical Models,” in *The Dynamics of Changing Higher Education in the Global South*, ed. B. Mpofo and S. Ndlovu-Gatsheni, 2019, 254.

⁴⁹ Niall Sclater, *Learning Analytics Explained* (Taylor & Francis, 2017), 16.

towards a successful postgraduate student predictive analytics system development and deployment at the institution.

Assumptions

The study argues that the traditional, non-statistical approach which has long been used to address the unsatisfactory postgraduate throughput rates has failed to yield the required outcomes. It also notes the existing effort and support mechanisms to address postgraduate student retention and throughput rates which are necessary but not sufficient. The main reason for the lack of such sufficiency is the fact that they seem to be selective and have been designed specifically for predicting quantitative-based, non-research modules. Thus, a postgraduate predictive analytics model becomes paramount due to its specificity. It is built upon the following assumptions.

- Postgraduate students remain a critical element of the national development imperative.
- Postgraduate throughput rates remain a determinant of university competence.
- Postgraduate students remain at the higher income end of the DHET funding model.
- There is a commitment by the universities to improve postgraduate throughput rates.
- Significant information critical for postgraduate students’ progress is missing in university records.

Predictive analysis and modelling can be broadly classified into the following three categories: plan, develop, and implement.

A. Model ideation and planning

This step includes team selection, scoping and dataset preparation. Approximately 40% of the total time may be spent on this portion of the exercise. The datasets that will be used must first be put together in order to develop a predictive model. You must establish specific goals, clean up and arrange the data, process the data by fixing missing values and outliers, do a descriptive analysis of the data using statistical distributions, and produce data sets that will be utilised in model construction.

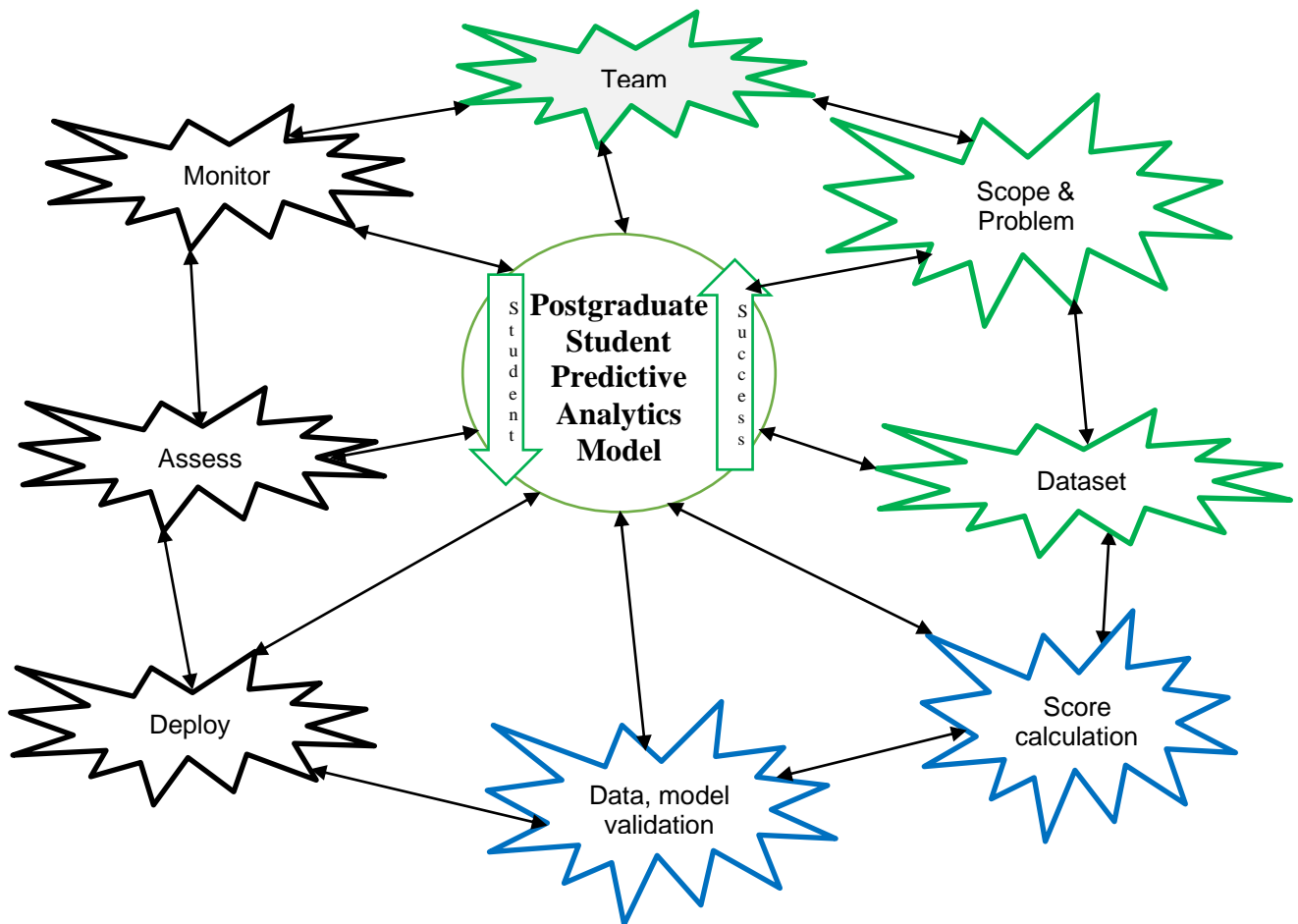


Figure 1: A proposed framework for the deployment of postgraduate student predictive analytics by Kariyana et. al.

Team selection

Team selection is the most critical step. The selection must be informed by such factors including commitment and willingness to professionally serve, relevant expertise, etc.

Scope and problem identification

The problem and the scope are identified. A clear specific objective for the model is formulated.

Dataset preparation

- data identification,
- data collection: document analysis (departmental records, progress reports, Higher Education Management Information System (HEMIS) office statistics etc, interviews, questionnaires (for students and supervisors)
- data analysis
- dataset creation

Regarding data collection, these interventions should be thorough and incorporate time for faculty consultation, online learning assistance, peer-led activities, tutorials, one-on-one advising, counselling, and mentoring, as well as student participation in the intervention design process.⁵⁰ Eduventures claims that even if done in departmental silos and even if an institution's efforts to converge data are in their infancy, there is no reason why solid, reliable data cannot be obtained right away.⁵¹ The researchers contend that ensuring that data are captured right away is the first step toward the numerous advantages of predictive analytics. This is crucial since the foundation of predictive analytics is the analysis of past trends. The path to future operational insights is thus paved by data collection at any time.

B. Model Development

- Model: includes writing the model code, building the model, calculation of scores,
- Validate: involves validation of the data.

This could consume 20% or so of the whole time. The technical components of developing models can be left to the data scientists or technical analysts for this portion.

An example from the generated data set will be used to develop the model. A portion of the original list of variables examined for the model will be included in the final model. Since some of the variables taken into account for the model will be connected, this should be okay. Other factors will have been eliminated since they have little to no predictive value.

Calculate a Score

When used on the postgraduate student population, the created model will be an equation that will assign a score to each student. Depending on what the model is set up to predict in the given time frame, the score may represent a student's likelihood to succeed or fail.

Validate the Model

To validate the model, a holdout group will be used. Students who were not involved in the model's development make up this group. As a result, they stand for a group of never-before-seen students who represent all postgraduate students.

C. Implement the Model

The model will be used to rank postgraduate students, forecast model performance, assess and track the model, and direct activities based on the model. Deploy, Assess, and Monitor are the three sub-phases in this process. There will be a need to think about how data will be accessed, used, and stored. This might take up to 40% of the overall time when the model is regularly utilised by the institution. This needs to be a continuous operational activity backed by the IT department.

⁵⁰ Cele, "Big Data-Driven Early Alert Systems as Means of Enhancing University Student Retention and Success."

⁵¹ Eduventures, "Predictive Analytics in Higher Education: Data-Driven Decision-Making for the Student Life Cycle."

Deploy the Model

It would be time to put the model into practice. A portion of the data will be used to build the model. The model will be applied to the focus case basis once it has been finished and confirmed.

Assess

Rankings will be generated from the model and interpreted according to the original determination of whether the prediction is to ascertain the likelihood of either persisting or dropping out. Even if these predictions might not be entirely correct, the ranking list offers a great starting point for tailoring the care, and consequently, the degree of service provided to different postgraduate student groups.

Monitor

It will be taken into account how frequently to run the model against the base and provide scores, as well as how to continuously assess the model's performance and maybe make use of new information as it becomes available.

D. Modelling Longevity and Considerations

It is helpful to consider the application of predictive models in the context of ongoing improvement. The construction of models cannot be abandoned. The model might need feeding, watering, watching, and caring for, but with time and as conditions change and more data are collected, it might get better. The researchers contend that accurate predictions can be made by models of moderate complexity.

Lourens and Bleazard have argued that:

Predictive modeling has the benefit of being quickly deployed over live data to score the data in real time after a model has been finalized. Numerous papers on student retention focus exclusively on the variables that affect retention rather than the most important stage, live scoring. Accessing or extracting data from an institutional database is simple, but writing the results back to the database so that decision-makers may see them immediately is not always feasible. In order for the prediction models to stay flexible and helpful, it is also crucial to continuously enhance them by periodically changing them by adding more factors.⁵²

RECOMMENDATIONS

A Cautionary Position

The proffered model should not be construed as a 'perfect and single solution' to capsize the poor postgraduate throughput rates at the university as different limitations exist. Rather, as the researchers picture this initiative, they share a cautionary position as posited by Wingfield in which she raised the question of whether a 'one size fits all' model is really appropriate to apply to postgraduate student throughput rates.⁵³ She said, "Instead of defining a hard figure that really means very little, we should more carefully investigate what exactly it is we seek to achieve and how best to reach this goal. To generate the greatest number of graduates who are educated, well-rounded, and experienced, we must take into account the diversity of our potential students, their interests, and personal circumstances."⁵⁴ As such, care must be taken to address the possible lack of transparency in predictive analytics in higher education to promote accurate data as the sole basis for making informed and wholesome decisions about improving postgraduate throughput rates.

CONCLUSION

The study concludes that there is a clear call for the need to turn the existing approach to the management and promotion of postgraduate student success. It is evident that at the national level, the prerequisite is that HEIs ought to turn around themselves and devise means to enhance their sustainability through uncompromised and flawless third stream income generating projects. Literature is clear on what is possible and has been tried and tested with regard to enhancing postgraduate throughput rates across various institutions. As such, the opportunities available are for those institutions that are committed to improving and magnifying their future practice by making meaning of the existing large data resources at their disposal.

⁵² Lourens and Bleazard, "Applying Predictive Analytics in Identifying Students at Risk: A Case Study."

⁵³ Wingfield, "Can We Improve Postgraduate Degree Throughput Rates?"

⁵⁴ Wingfield, "Can We Improve Postgraduate Degree Throughput Rates?" 2.

BIBLIOGRAPHY

- Alamuddin, Rayane, Daniel Rossman, and Martin Kurzweil. "Interim Findings Report from the MAAPS Advising Experiment," June 27, 2019. <https://doi.org/10.18665/sr.311567>.
- Avella, John T, Mansureh Kebritchi, Sandra G Nunn, and Therese Kanai. "Learning Analytics Methods, Benefits, and Challenges in Higher Education: A Systematic Literature Review." *Online Learning* 20, no. 2 (2016): 13–29.
- Barshay, Jill, and Sasha Aslanian. "Colleges Are Using Big Data to Track Students in an Effort to Boost Graduation Rates, but It Comes at a Cost." *The Hechinger Report*, 2019.
- Bird, Kelli A., Benjamin L. Castleman, Zachary Mabel, and Yifeng Song. "Bringing Transparency to Predictive Analytics: A Systematic Comparison of Predictive Modeling Methods in Higher Education." *AERA Open* 7 (January 10, 2021): 233285842110376. <https://doi.org/10.1177/23328584211037630>.
- Botha, R J. "Student Throughput Trends on Postgraduate Level: An African Case Study." *The Independent Journal of Teaching and Learning* 13, no. 2 (2018): 53–66.
- Cele, N. "Big Data-Driven Early Alert Systems as Means of Enhancing University Student Retention and Success." *South African Journal of Higher Education* 35, no. 2 (May 2021). <https://doi.org/10.20853/35-2-3899>.
- Eduventures. "Predictive Analytics in Higher Education: Data-Driven Decision-Making for the Student Life Cycle." Eduventures, 2013. <https://www.sv-europe.com/wp-content/uploads/predictive-analytics-in-higher-education-data-driven-decision-making-for-the-student-life-cycle.pdf>.
- Ekowo, Manuela, and Iris Palmer. "Predictive Analytics in Higher Education." *Five Guiding Practices for Ethical Use*, 2017.
- . "The Promise and Peril of Predictive Analytics in Higher Education: A Landscape Analysis." *New America*, 2016.
- Herodotou, Christothea, Galina Naydenova, Avi Boroowa, Alison Gilmour, and Bart Rienties. "How Can Predictive Learning Analytics and Motivational Interventions Increase Student Retention and Enhance Administrative Support in Distance Education?" *Journal of Learning Analytics* 7, no. 2 (September 19, 2020). <https://doi.org/10.18608/jla.2020.72.4>.
- Impellizzeri, Franco M, and Mario Bizzini. "Systematic Review and Meta-analysis: A Primer." *International Journal of Sports Physical Therapy* 7, no. 5 (2012): 493.
- Israel, Kariyana, and Marongwe Newlin. "Developing Research Willpower at One South African University as a Strategy to Transform Higher Education Institutions: Experiences of the Mentees." *E-BANGI Journal* 18, no. 10 (2021).
- Jones, Kyle, Amy VanScoy, Kawanna Bright, and Alison Harding. "Do They Even Care? Measuring Instructor Value of Student Privacy in the Context of Learning Analytics," 2021.
- Kalechofsky, Hal. *A Simple Framework for Building Predictive Models*. MSquared Consulting, 2016. <https://www.msquared.com/wp-content/uploads/2017/01/A-Simple-Framework-for-Building-Predictive-Models.pdf>.
- Kariyana, I., and R.A. Sonn. "A Case Scenario of the Dynamism and Fragility of the Feasibility of Deconceptualising Theoretical Models." In *The Dynamics of Changing Higher Education in the Global South*, edited by B. Mpofo and S. Ndlovu-Gatsheni, 237–52, 2019.
- Kawchale, Neha, and Rachana Satao. "Predictive Analytics in Education Context." *Int J Mod Trends Sci Technol* 2, no. 11 (2016): 77–82.
- Khan, Khalid S, Regina Kunz, Jos Kleijnen, and Gerd Antes. "Five Steps to Conducting a Systematic Review." *Journal of the Royal Society of Medicine* 96, no. 3 (2003): 118–21.
- Khauoe, Thakane. "Factors Impacting Postgraduate Throughput Rates at a South African University of Technology." Cape Peninsula University of Technology, 2020.
- Littell, Julia H, Jacqueline Corcoran, and Vijayan Pillai. *Systematic Reviews and Meta-Analysis*. Pocket Guide to Social Work Re, 2008.
- Lourens, A, and D Bleazard. "Applying Predictive Analytics in Identifying Students at Risk: A Case Study." *South African Journal of Higher Education* 30, no. 2 (June 2016). <https://doi.org/10.20853/30-2-583>.
- Patil, Shankar M. "Predictive Analytics in Higher Education." *Life* 4, no. 12 (2015).
- Picciano, Anthony G. "The Evolution of Big Data and Learning Analytics in American Higher Education." *Journal of Asynchronous Learning Networks* 16, no. 3 (2012): 9–20.
- Pigott, Terri D., and Joshua R. Polanin. "Methodological Guidance Paper: High-Quality Meta-Analysis in a Systematic Review." *Review of Educational Research* 90, no. 1 (February 25, 2020): 24–46. <https://doi.org/10.3102/0034654319877153>.
- Roux, Ingrid le Le. "Coaching as Support for Postgraduate Students: A Case Study." *SA Journal of Human*

Resource Management 16, no. 1 (2018): 1–7.

Sclater, Niall. *Learning Analytics Explained*. Taylor & Francis, 2017.

Segooa, Mmatshuene Anna, Billy Mathias Kalema, and Ali Hussein Zolait. “Predictive Analytics to Improve Outcome-Based Funding for the Public Universities in South Africa through Big Data.” In *2019 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)*, 1–6. IEEE, 2019.

Tawfik, Gehad Mohamed, Kadek Agus Surya Dila, Muawia Yousif Fadlelmola Mohamed, Dao Ngoc Hien Tam, Nguyen Dang Kien, Ali Mahmoud Ahmed, and Nguyen Tien Huy. “A Step by Step Guide for Conducting a Systematic Review and Meta-Analysis with Simulation Data.” *Tropical Medicine and Health* 47, no. 1 (December 1, 2019): 46. <https://doi.org/10.1186/s41182-019-0165-6>.

Uman, Lindsay S. “Systematic Reviews and Meta-Analyses.” *Journal of the Canadian Academy of Child and Adolescent Psychiatry* 20, no. 1 (2011): 57.

Universities UK. “Analytics in Higher Education,” 2016. www.universitiesuk.ac.uk.

Walton, Gregory M., and Geoffrey L. Cohen. “A Brief Social-Belonging Intervention Improves Academic and Health Outcomes of Minority Students.” *Science* 331, no. 6023 (March 18, 2011): 1447–51. <https://doi.org/10.1126/science.1198364>.

Wingfield, Brenda. “Can We Improve Postgraduate Degree Throughput Rates?” *South African Journal of Science* 107, no. 11/12 (November 1, 2011). <https://doi.org/10.4102/sajs.v107i11/12.967>.

ABOUT AUTHORS

Prof. Israel Kariyana is an Associate Professor in the Department of Continuing Professional Teacher Development, Faculty of Education, Walter Sisulu University, Mthatha Campus, South Africa. His research focuses on Educational Management, Mathematics Education and Sustainable Education.

Prof. Winter Sinkala is an Associate Professor of Mathematics at Walter Sisulu University, Mthatha Campus, South Africa. His research interests include Lie symmetry analysis, fixed point theory, and mathematics education.

Dr. Neliswa Gqoli is a Senior Lecturer at Walter Sisulu University (WSU) in the Department of Adult and Educational Foundations where she teaches Psychology of Learning and Mathematics Teaching. She obtained her Master’s Degree at Walter Sisulu University and PhD (Early Childhood Development) at the University of Free State. Her research focus is on Mathematics Education. As the Senior Researcher in the Faculty, she has published articles and book chapters nationally and internationally. She taught mathematics for 25 years in Eastern Cape Primary Schools. She is a Faculty Board Member and a coordinator for the Post Graduate Certificate in Education at WSU. She is also the chairperson of the Faculty Social Committee. She is a member of the South African Research Association for Early Childhood Education (SARAECE).