Research

This open-access article is distributed under Creative Commons licence CC-BY-NC 4.0.

Decision tree classification of academic success in a first-year human anatomy course at a South African university

E Bruwer, PhD; **N Jooste,** PhD

Department of Human Anatomy and Physiology, Faculty of Health Sciences, University of Johannesburg, South Africa

Corresponding author: *N Jooste (njooste@uj.ac.za)*

Background. Human anatomy is a fundamental subject in health sciences education, yet academic success in this topic is often hindered by various challenges, including the extensive module content, complex concepts and reduced instructional time. Identifying students at risk of underperformance early in the module can facilitate timely interventions.

Objective. To develop a classification scheme to predict academic success in a first-year human anatomy module in South Africa using admission-stage data, including South Africa's National Senior Certificate (NSC) subjects and qualification enrolment.

Methods. Data from 196 students enrolled in Optometry, Nursing and BSc Life and Environmental Sciences programmes (2017 - 2018) were analysed. A decision tree classification model (J48 algorithm) was applied to identify key predictors of student success, categorised as failed, passed or distinction.

Results. Results indicated that qualification enrolment and the NSC subjects Home Language, Mathematics and History were the most informative predictors. The model demonstrated a 55.6% overall accuracy, with the highest classification success for the 'passed' category (93/165) but limited predictive power for 'failed' and 'distinction' classifications. The kappa value (0.1083) and mean absolute error (0.1606) indicated moderate agreement and acceptable accuracy. These findings highlight the complexity of predicting academic performance in the study of human anatomy and suggest that although admission data may provide some insights, additional factors (e.g. learning strategies and motivation) could play crucial roles.

Conclusion. Mathematics at secondary-school level played a key role in distinguishing students likely to pass human anatomy modules, across all degree programmes that were analysed. Furthermore, the study findings underscore the need for supplementary assessment tools and support mechanisms to enhance student success. Future research should explore the integration of cognitive and behavioural predictors alongside academic metrics to improve classification models.

Keywords: academic success, admissions criteria, human anatomy

Afr J Health Professions Educ 2026;18(3):3213. https://doi.org/10.7196/AJHPE.2026.v18i3.3213

The study of human anatomy provides essential foundational information for qualifications within the health sciences[1,2]; however, there are often challenges associated with academic success in this subject. These include large volumes of content[3] with up to 1.25 million new terms[4]; complex concepts[3] and the three-dimensional nature of the content[5]; stress associated with encountering the deceased during practical learning[3,6]; and increased pace associated with the trend of decreasing time for anatomy modules.[5,7,8] Depending on the duration and timing of the anatomy module (i.e. semester/term-based module), the identification of students who have difficulties grasping the anatomy content may occur too late to introduce meaningful interventions.[9] Lecturers can take up to 3 months to review enough in-module assessment data to identify underperforming students.[9]

Thus, anatomy educators may find it beneficial to consider criteria other than in-module assessment data that act as an early warning system. Some suggested indicators of success in anatomy education include students' secondary school science grades, their grade point average (GPA), the university's mathematics and English placement test scores, and medical college admission test results.[7,10] Indicators not based on previous performance can also be informative, such as the Learning and Study Strategies Inventory (LASSI) strategy, which provides information about skill, will and self-regulation.[7] Khalil *et al.*[7] found that performance in anatomical sciences was significantly correlated with eight of the ten LASSI subscales (anxiety, concentration, information processing, motivation, self-testing, selecting main ideas, time management and test strategies).

Even when various types of admissions data were used to predict student success (e.g. GPA, science GPA, science credits, Otis-Lennon Mental Abilities Test, Nelson-Denny Reading Test and type of school attended), only approximately half (46%) of the variance was accounted for by these data,[11] indicating the multifaceted and complex nature of predicting student success. This is supported by studies that indicate that the motivation and test strategies that students employ make a significant difference to their success in anatomy, regardless of previous performance or entrance criteria.[7] Thus, the admission-stage metrics could serve as a tool to identify students who may benefit from additional assistance and mentoring to ensure student success by leveraging other factors such as test strategies.

The aim of this study was to build a classification scheme that would classify academic success, defined as passed, failed or distinction, in a first-year human anatomy module, based on information captured during student admission (qualification enrolment and subjects according to the South African Department of Basic Education's National Senior Certificate (NSC)).

Methods

The focus of the study was the data of 196 students enrolled at a South African university in a first-year human anatomy module for the

qualification programmes Optometry (B. Optom, *n*=79), Nursing (B. Cur, *n*=82) and BSc Life and Environmental Sciences (*n*=35), for the student cohorts enrolled in 2017 and 2018. Four students were excluded from the Optometry cohort, as their final secondary school results were unavailable on the enrolment system.

Most students entered the course directly after finishing secondary school, and any students repeating the module were removed from the data set to prevent the duplication of student data. Although first-year human anatomy was offered as two separate semester modules, only the data from semester 1 were included.

Students were subject to differing entrance criteria depending on the degree programme. BSc students enrolled in the first-year human anatomy course as an elective without entrance criteria. Optometry and Nursing students were required to have matriculated with the following secondary school subjects: Life Sciences, Mathematics, Physical Sciences and English.

Data were requested through a Promotion of Access to Information Act (PAIA) application to the institution in question, after which access to de-identified student global records was permitted. In addition, ethics clearance for the study was obtained from the university's research ethics committee (ref. no. REC-241112-035). In line with ethics and PAIA permissions, no personal information was obtained for individual students, and no students could be identified using the research data.

Admissions data collected from the students' global records included the NSC secondary-school subjects studied and marks obtained, their admission point score (APS) and the qualification (i.e. degree programme) in which the student enrolled. In addition to these data, the first-year final marks for the module Human Anatomy, as the dependent variables, were obtained from the students' global records for classification purposes.

When finishing secondary school, all students matriculated with at least two language subjects – at least one designated Home Language and one designated First Additional Language. Language subjects were taken by the students in the frequencies reported in Table 1.

A minimum of five non-language subjects were additionally required at secondary school level. These subjects varied among students (Table 1). However, most students had Life Orientation, as well as the required subjects specific for entrance to the health sciences qualifications (Optometry and Nursing) as stipulated above.

Individual secondary school subject marks were not evaluated in this study's model, as most students who gain entrance to a degree programme will have similar admission metric requirements (i.e. grade symbols and APSs). Instead, the presence or absence of secondary school subjects was considered. If a subject was outside of the elected secondary school subjects for a student, the data cell was left blank (i.e. treated as missing data). The requirements for BSc Life and Environmental Sciences students to enter Human Anatomy were far fewer. Thus, the data of this student grouping could provide a contrasting background to the more uniform data of the students accepted into the health sciences programmes.

For the first-year human anatomy course, students from both qualifications and the BSc cohort were required to prove competency in the same content, taught by the same lecturer and assessed using the same assessments. The final semester module results for first-year human anatomy included both the final period mark (i.e. semester mark) and the exam mark. This mark was converted to three success categories: failed (final mark <50%), passed (final mark ≥50% and <75%) and distinction (final mark ≥75%). The discrete categories created for both secondary school subjects and the human anatomy module meant that error rates were reduced, and greater accuracy of classification was achieved during data analysis.[12]

The data described above were received on a Microsoft Excel spreadsheet and saved as a comma delineated file, after conversion to discrete categories was completed. The comma delineated file was further transformed into an Attribute – Relation File Format (.arff) to allow the import of data into the data mining tool – WEKA (Waikato Environment for Knowledge Analysis; available at https://waikato.github.io/weka-wiki/downloading_weka/). WEKA is open-source software for data mining and machine learning developed by the University of Waikato, New Zealand.

Table 1. Subjects on National Senior Certificates for student cohorts during the admission stage. All students completed a Home Language and a First Additional Language, and at least five non-language subjects

Subject name	Number of students (*N*=196)
Afrikaans First Additional	54
Afrikaans Home Language	11
English First Additional	139
English Home Language	57
sePedi Home Language	39
seSotho Home Language	8
seTswana Home Language	12
siSwati Home Language	12
Tshivenda Home Language	1
isiXhosa Home Language	8
xiTsonga Home Language	8
isiZulu First Additional	2
isiZulu Home Language	41
Life Orientation	195
Life Sciences	194
Mathematics	194
Physical Sciences	192
Geography	99
Agricultural Sciences	32
Accounting	27
Computer Applications Technology	13
Business Studies	10
History	5
Tourism	5
Applied Mathematics	4
Dramatic Arts	4
Visual Arts	4
Economics	3
Information Technology	3
Engineering Graphics and Design	2
Electrical Technology	2
Mathematical Literacy	2
Religion Studies	2
Agricultural Management Practices	1
Civil Technology	1
Consumer Studies	1
Hospitality Studies	1
Mechanical Technology	1
Music	1

The WEKA version 3.8.6 data mining tool analysed patterns using admission data to construct a classification scheme. WEKA classification schemes are generated decision trees with rules that can potentially be used as predictions without the need for algorithm coding.

Informative classification was conducted using the decision tree algorithm J48 (based on the C4.5 algorithm of Quinlan[13]) with 10-fold cross-validation to prevent overfitting of the data in the model. The decision tree method classifies data into distinct groups that are generated using the strongest separation in the values of the dependent variable. Each internal node indicates a test using an independent variable, each branch from this node represents an outcome from the test, and each leaf node is assigned a dependent variable class label. The J48 algorithm was used as it can handle missing data, as well as both continuous and discrete input data.

Results

The information that contributed to an informative classification model, as selected by the J48 decision tree algorithm, included the qualification that students were enrolled in, the NSC Home Language subject that students matriculated with, and whether they studied the NSC subjects of Mathematics and/or History (Fig. 1). As can be seen from the classification matrix (Table 2), the algorithm had the best results when classifying the 'passed' category (93/165) and the poorest when classifying 'failed' (0/1) or 'passed with supplementary' (1/2). No leaves on the decision tree were categorised as 'passed with supplementary'. Passed with supplementary refers to a situation where the student is granted an additional opportunity to demonstrate competence by completing a new assessment. If successful, the student is considered to have passed the module with a minimum required mark of 50%. The correct classification for this decision tree classification scheme was 55.6%. The correct classification percentage appears low; however, the decision tree model prevents data from being overfitted, which is often observed with traditional models (e.g. regression analysis). Thus, this model provided a minimum correct classification of 55.6%, with potential improvement dependent on future test samples. The kappa value indicated agreement of 0.1083 and a mean absolute error of 0.1606, while the weighted average F-measure indicated acceptable accuracy of the model dataset at 0.486.

From the resultant decision tree (Fig. 1), we observed that students who did not take Mathematics at secondary-school level were classified into the failed Human Anatomy category. Students who completed Mathematics were further classified into their qualification programmes (i.e. Nursing, Optometry and BSc Life and Environmental Sciences). The Life and Environmental Sciences students were not classified into further branches but achieved a passing grade. Nursing students were further divided into students who took History at secondary school level and students who did not take History, with students with History on their NSC more likely to procure a distinction in Human Anatomy, and students without the subject only passing the course. A similar pattern was observed with Optometry, considering the NSC Home Language subjects, with students who took English and Afrikaans more likely to procure a distinction in Human Anatomy, while students who took any other Home Language more likely only to pass.

Thus, only Mathematics distinguished a failed grade from a passed or distinction grade. All other subjects were qualification specific and provided classifications of a distinction v. a pass grade. Although some clear patterns and trends were observed from the decision tree classification scheme, it is important to note that some categories (e.g. History, Mathematical Literacy as opposed to pure Mathematics, and some of the other African Home Languages) had very small sample sizes representing them (Table 1). However, we were able to draw some tentative conclusions that may assist lecturers in human anatomy with the early identification of students who may require additional assistance. In addition, the results indicated that the type of interventions that may assist students in achieving a passed v. failed,

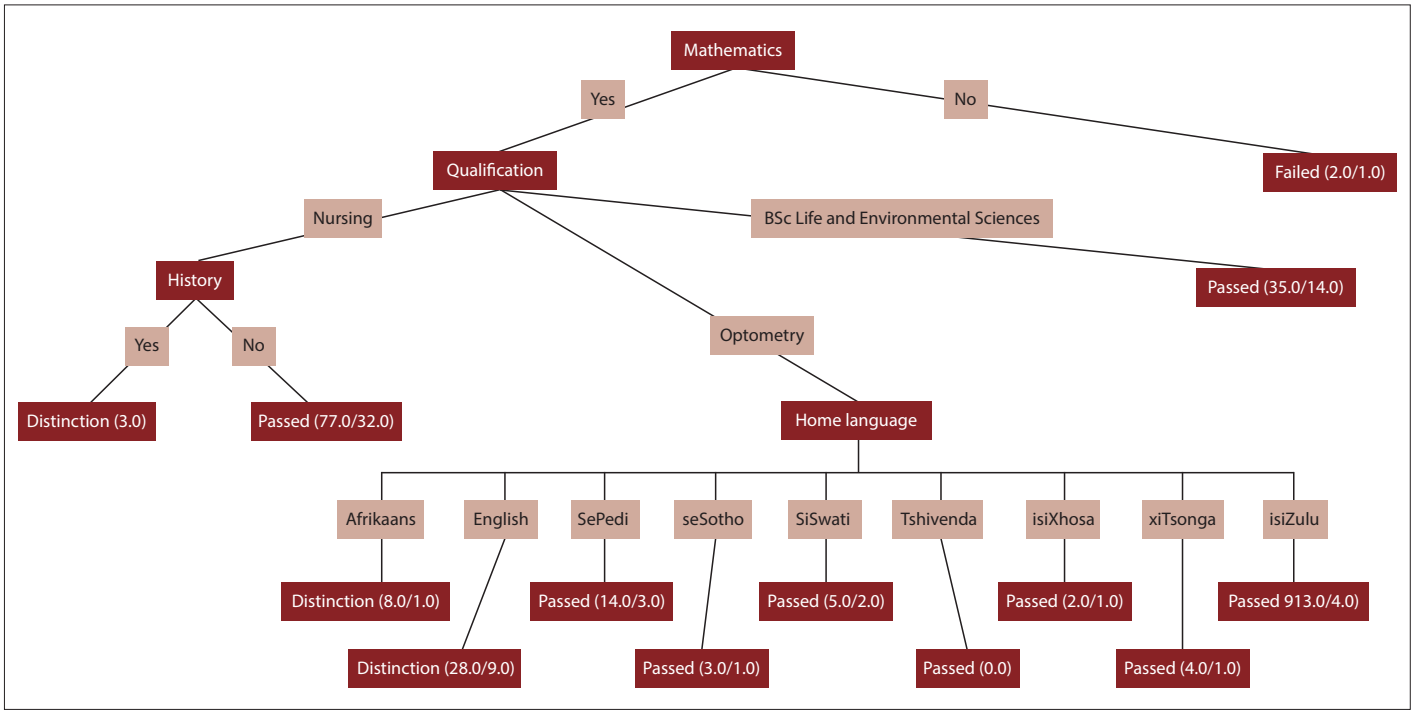


Fig. 1. Decision tree classifying students into success categories based on secondary school subjects completed and qualifications enrolled in.

Table 2. Classification matrix indicating the classification according to the decision tree algorithm (columns) v. the actual categories according to the data (rows)

	Failed	Passed with Supplementary	Passed	Distinction	Total
Failed	0	0	13	1	14
Passed with Supplementary	0	1	11	1	13
Passed	1	1	93	11	106
Distinction	0	0	48	15	63
Total	1	2	165	28	*N*=196

or distinction v. passed, classification may depend on both their qualifications and their final secondary-school subjects.

Discussion

The fast pace and short time often allocated to human anatomy modules at tertiary institutions can introduce difficulties when educators need to identify underperforming students through in-module assessments and introduce interventions.[9] This study's findings have shown that admission criteria identified with the decision tree classification scheme at the enrolment stage can act as an early identifier of underperforming students.

This study's results highlighted the core function of mathematics at secondary school when distinguishing students who are likely to succeed in human anatomy studies. Students who were enrolled in NSC Mathematical Literacy (as opposed to pure Mathematics) at secondary school were more likely to fail a human anatomy module. This aligns with literature that indicates that university students with higher scores associated with mathematical skills (at secondary school or as part of benchmarking admission tests) were more inclined to pass anatomy courses.[10] As calculations are relatively few in the subject of anatomy, transferable skills may include categorising information and ordering or sequencing structures or events in the body, as indicated by Gultice *et al.*[10] In addition, as the study of human anatomy requires two-dimensional and three-dimensional visuospatial orientation,[5] we contend that the geometry component of secondary school Mathematics may prepare students to better visualise the complex spatial images present in human anatomy studies.

All other indicators distinguished between students who attained a passed classification v. those who achieved a distinction in first-year Human Anatomy. These indicators were specific to the qualification in which the students were enrolled. Students enrolled in the BSc Life and Environmental Sciences degree programme were sorted into a passed category with no further indicators consulted.

Students enrolled in the Nursing degree programme were categorised according to the presence or absence of History on their NSC. Students who completed History were allocated to the distinction category, whereas those without simply passed Human Anatomy. History is a content-heavy subject requiring memorisation of many dates and events for success, with an often-used assessment style that requires students to rote-learn content.[14-16] Students who use this approach in secondary school may be better prepared for the rigours of learning complex concepts and the many new terms that need to be memorised in the study of human anatomy. This is corroborated by literature that suggests memorisation and rote learning are key components of mastering human anatomy, especially in the preliminary stages of understanding, which requires information management and accurate recall.[17] However, there is consensus in the literature that rote learning in isolation does not promote quality learning and that the ultimate goal is understanding through deep learning approaches.[17]

Students enrolled in the Optometry degree programme were categorised according to the Home Language subject they completed for their NSC. Students who completed Afrikaans or English as their Home Language for the NSC were allocated to the distinction category. By contrast, students who completed any other Home Language subject (various other African languages in this instance) were allocated to the passed category for first-year Human Anatomy. Lucas *et al.*,[4] similarly found that English was the greatest predictor of success (when compared with biology and chemistry). This suggests that the Germanic language base of Afrikaans and English contributes to student success. This could be beneficial in two ways. Firstly, as English is the language of tuition, students with a foundation in the Germanic languages may adapt more easily. An additional link between good English skills and success, as highlighted by Lukas *et al.*,[4] is that students proficient in English could write more within the time limit of an assessment. Early studies[18,19] on the topic have shown that where English is not the primary language of tuition in secondary school, English ability is significantly positively correlated with results in the health sciences when they are taught in English.

Secondly, we tentatively propose that the Germanic base of English and Afrikaans may be more readily transferable to the Latin and Greek roots that provide the base for the anatomical terminology that is introduced in the study of human anatomy. However, this requires further investigation.

It is important to note that the choice of Home Language is likely limited by the options available at students' respective secondary schools. This may also be associated with the quality of the secondary school and its education programme, which we could not assess in this study. Schools located in rural areas or that are disadvantaged owing to previous political influence within the country (i.e. classified in the Department of Basic Education's quintiles 1 to 3) may provide limited access to a greater variety of subjects, including Home Language subjects. In contrast, literature shows that students from affluent schools have the advantages of better access to study materials, equipment and highly qualified teachers.[20] Although an exploration of the influence of school affluence and location was outside the scope of this study, our findings based on Home Language subject choice may align with Mabizela and Chivandi,[21] who found that students admitted from rural areas had lower mean scores in medical physiology and biochemistry, foundational science courses in the health sciences that are similar to human anatomy.

Interestingly, this study's decision tree algorithm did not include the NSC subject of Life Science (i.e. Biology) at any level as an indicator of success. This corroborates previous findings by Lukas *et al.*,[4] who found that A-level Biology grade was not a predictor of final preclinical grades, which included anatomy. However, taking A-level Biology but not achieving higher than grade E was a significant indicator associated with a fail grade in preclinical grades.[4]

The APS score also provided no input to this decision tree algorithm. We consider that the narrow range of APS scores linked to students who obtain successful admission may have negated effects that might have been

seen if more variation was introduced for this criterion. A similar finding was reported when considering ENTER scores, an admission criterion that ranks overall student achievement in a similar way to APS. ENTER score could not predict success in regional anatomy studies; however, combined systemic anatomy and physiology outcomes had some prediction success using this metric.[22]

Although this method provided some good insights, correct classification was only possible 55.6% of the time. Specifically, classification of a failed category was difficult. This may have been owing to other factors not considered in the scope of this study, such as reading skills, study methods, stress levels and students' ability to adapt to a new environment. Gultice *et al.*[10] also found that the complex and interrelated nature of influences obscured outcomes. Thus, factors in the time between admission and final human anatomy results could still provide notable contributions toward student success, as suggested by Leonard and Niebuhr,[11] such as teaching methods and learning strategies, as emphasised by Harmon *et al.*[5] This early prediction classification scheme could thus be used to identify students who require these specific interventions.

Conclusions

This study explored the potential of using admissions data to classify academic success in a first-year human anatomy module. The J48 decision tree algorithm identified qualification enrolment, and NSC Home Language, Mathematics and History as key predictors of student outcomes. However, the overall classification accuracy was 55.6%, with the model performing best in categorising students who passed, while struggling to accurately predict failed categories. The results indicated that Mathematics was associated with higher success rates, likely reflecting the importance of logical reasoning and problem-solving in the study of anatomy. The subject of History as a predictor suggests that students with previous experience in memorisation and recall may be better prepared for the theoretical load of human anatomy. The role of Home Language highlights the potential effect of language proficiency on comprehension and engagement with complex anatomical terminology. However, further research with expanded data sets and additional variables, such as study habits and engagement levels, may improve predictive accuracy and inform targeted interventions to enhance student success in learning anatomy.

Data availability statement. Data are available from the corresponding author on reasonable request, subject to ethical and institutional requirements.

Declaration. None.

Acknowledgements. The authors would like to acknowledge the institution that granted us access to the admissions data and Human Anatomy grades. We would also like to acknowledge the students whose data contributed to the results presented here. Finally, we would like to acknowledge our colleagues who provided advice that contributed to the final manuscript.

Author contributions. EB: conceptualisation and study design, data collection, contribution to manuscript; NJ: study design, data analysis, drafting and editing of manuscript.

Funding. None.

Conflict of interest. None.

1. Turney BW. Anatomy in a modern medical curriculum. Ann Royal Coll Surg Engl 2007;89(2):104-107. https://doi.org/10.1308/003588407X168244
2. Losco CD, Grant WD, Armson A, Meyer AJ, Walker BF. Effective methods of teaching and learning in anatomy as a basic science: A BEME systematic review: BEME guide no. 44. Med Teach 2017;39(3):234-243. https://doi.org/10.1080/0142159x.2016.1271944
3. Lai HL, Lee YF, Lai KC. The influence of humanised anatomical pedagogy on psychophysiological responses and academic achievement in nursing students. J Prof Nurs 2020;36(4):245-250. https://doi.org/10.1016/j.profnurs.2019.11.005
4. Lucas P, Lenstrup M, Prinz J, Williamson D, Yip H, Tipoe G. Language as a barrier to the acquisition of anatomical knowledge. Med Educ 1997;31(2):81-86. https://doi.org/10.1111/j.1365-2923.1997.tb02463.x
5. Harmon DJ, Attardi SM, Waite JG, Topp KS, Smoot BJ, Farkas GJ. Predictive factors of academic success in neuromusculoskeletal anatomy among doctor of physical therapy students. Anat Sci Educ 2023;16(2):323-333. https://doi.org/10.1002/ase.2202
6. Bahşi İ, Topal Z, Çetkin M, Orhan M, et al. Evaluation of attitudes and opinions of medical faculty students against the use of cadaver in anatomy education and investigation of the factors affecting their emotional responses related thereto. Surg Radiol Anat 2021;43:481-487. https://doi.org/10.1007/s00276-020-02567-8
7. Khalil MK, Williams SE, Hawkins HG. Learning and study strategies correlate with medical students' performance in anatomical sciences. Anat Sci Educ 2018;11(3):236-242. https://doi.org/10.1002/ase.1742
8. Fitzgerald JE, White MJ, Tang SW, Maxwell-Armstrong CA, James DK. Are we teaching sufficient anatomy at medical school? The opinions of newly qualified doctors. Clin Anat 2008;21(7):718-724. https://doi.org/10.1002/ca.20662
9. Hortsch M, Mangrulkar RS. When students struggle with gross anatomy and histology: A strategy for monitoring, reviewing, and promoting student academic success in an integrated preclinical medical curriculum. Anat Sci Educ 2015;8(5):478-483. https://doi.org/10.1002/ase.1519
10. Gultice A, Witham A, Kallmeyer R. Are your students ready for anatomy and physiology? Developing tools to identify students at risk for failure. Adv Physiol Educ 2015;39(2):108-115. https://doi.org/10.1152/advan.00112.2014
11. Leonard DL, Niebuhr BR. Admissions variables as predictors of performance in basic science coursework. Occup Ther J Res 1986;6(2):105-113. https://doi.org/10.1177/153944928600600204
12. Rajalakshmi A, Vinodhini R, Bibi KF. Data discretisation technique using WEKA tool. Int J Sci Eng Comput Technol 2016;6(8):293-298.
13. Quinlan JR. Induction of decision trees. Mach Learn 1986;1(1):81-106. https://doi.org/10.1007/BF00116251
14. Kabombwe YM, Mulenga IM. Implementation of the competency-based curriculum by teachers of History in selected Secondary Schools in Lusaka district, Zambia. Yesterday Today 2019;(22):19-41. https://doi.org/10.17159/2223-0386/2019/n22a2
15. Ndomondo E, Mbise A, Katabaro J. History teachers' conceptualization of competency-based curriculum in transforming instructional practices in lower secondary schools in Tanzania. Soc Sci Humanit Open 2022;6(1):100331. https://doi.org/10.1016/j.ssaho.2022.100331
16. Ndomondo E. Promoting history subject skills through competency-based assessment in Tanzania rural secondary schools: Teachers' understanding and practices. Soc Sci Humanit Open 2024;9:100876. https://doi.org/10.1016/j.ssaho.2024.100876
17. Pandey P, Zimitat C. Medical students' learning of anatomy: Memorisation, understanding and visualisation. Med Educ 2007;41(1):7-14. https://doi.org/10.1111/j.1365-2929.2006.02643.x
18. Ahmed B, Ahmed LB, Al-Jouhari MM. Factors determining the performance of medical students of the Faculty of Medicine, University of Kuwait. Med Educ 1988;22(6):506-508. https://doi.org/10.1111/j.1365-2923.1988.tb00794.x
19. Hassan MO, Gumaa KA, Harper A, Heseltine GFD. Contribution of English language to the learning of basic medical sciences in Sultan Qaboos University. Med Teach 1995;17(3):277-282. https://doi.org/10.3109/01421599509008317
20. McManus IC, Dewberry C, Nicholson S, Dowell JS. The UKCAT-12 study: Educational attainment, aptitude test performance, demographic and socio-economic contextual factors as predictors of first year outcome in a cross-sectional collaborative study of 12 UK medical schools. BMC Med 2013;11:244. https://doi.org/10.1186/1741-7015-11-244
21. Mabizela SE, Chivandi E. Influence of admission categories, programme, race and residence status on MBBCh and BHSc student performance: Academic performance of 2021 students in Physiology module 2024. Health SA Gesondheid 2025;30:a2933. https://doi.org/10.4102/hsag.v30i0.2933
22. Green R, Brown E, Ward A. Secondary school science predictors of academic performance in university bioscience subjects. Anat Sci Educ 2009;2(3):113-118. https://doi.org/10.1002/ase.82

Received 3 March 2025; accepted 13 June 2026.