African Journal of Education Science and Technology (AJEST) Vol. 8 No.1 (October, 2024)
University of Eldoret, Kenya, Mount Kenya University, Kenya, Chukwuemeka Odemegwu Ojukwu University, Nigeria, Kyambogo University, Uganda and University of Makeni, Sierra Leone

Evaluating the Performance of Tree-Based Predictive Models as Programme Recommenders for University Entrants in Kenya.

**Kibuthi J. Kabiru[1], Ratemo C. Makiya[2], Anduvare, E.M[1].*
[1]Karatina University, P.O. Box 1957-10101, Karatina.
[2] Mama Ngina University College, P.O. Box 444-01030, Gatundu.

****Corresponding email: j.kabiru@outlook.com***

Abstract

Enrolling for the wrong programme by university students has, to an extent, contributed to the high rates of discontinuation on academic grounds, repeat year cases, change of programme after registration, interuniversity transfers, deferments to change programme, drop out cases, suspension over exam irregularities as well as to strikes. This study focused on finding a technological solution for reducing these cases by evaluating three tree-based predictive models and recommending the most predictive model to implement as a programme recommender. Data was collected in five selected public universities in Kenya using Google Forms. The respondents were 308 translating to 308 rows of data with 36 columns. Numpy, Pandas, Matplotlib, Sklearn, Seaborn, Scipy, Plotly python analytics libraries were deployed using Jupyter Notebook for Anaconda. The cleaned and processed dataset features had categorical variables thus one-hot-encoding technique was employed. Data was split for training and testing with the random_state set to 42. Gini index criteria was implemented. The three models were evaluated on their performance from the optimally split data for training and test with a 80:20 ratio. Random Forest (RF) came out the most predictive at 99.3% followed by Gradient Boosting (XG Boost) at 90% then Decision Tree (DT) at 80.93%. The testing accuracy score for RF was 81.72%, XGBoost was at 75.72% and DT was at 76.34%. Confusion matrix criterion was implemented to evaluate the performance of the three models. The results of this study have demonstrated the high accuracy level of RF as the most predictive tree-based model for this real-world University crisis. The model is recommended for development as a system to be integrated into the KUCCPS portal. The integrated system is dubbed Programme Recommender which if launched would highly predict the best programme of study for application by university entrants.

Keywords: *Random Forest, Gradient Boosting, Python, KUCCPS, Programme Recommender.*

INTRODUCTION

Predictive modeling is a mathematical technique that predicts future occurrences by examining pertinent past data. Machine learning is a predictive tool that utilizes algorithms that learn from primary/secondary data and recognize patterns to provide predictions. Education and research are among the main areas focused by Kenya’s Vision 2030 blueprint. The 2030 Agenda for Sustainable Development has the fourth goal focusing on quality education whose main aim is to ensure inclusive and unbiased quality education and endorse lifelong learning opportunities for all, UN DESA (2023). University entrants require relevant information as well as effective decision-making skills for them to achieve the fourth goal of the Sustainable Development Goals (SDGs).

Problem Statement

The objective of the study was to evaluate the performance of tree-based predictive models as programme recommenders for university entrants in Kenya as a solution to the progression crisis caused by studying wrong programmes in public universities. Maina (2020), Russell (2019), Thumiki (2019), Gathoni (2019), Sithole, et al (2017), and CSIR (2012), have shown that poor programme choice when joining a university, by university entrants, has to an extent contributed to the high rates of discontinuation on academic grounds, repeat year cases, change of programme after registration, interuniversity transfers, deferments to change programme, dropped out cases and suspension over exam irregularities. Students studying undesired programmes has led to strikes, Business Today Kenya (2016).

The factors that affect low rates of same cohort graduation in a university are lack of interest in the programme, financial challenges, family problems, health matters, maternity leave, job limitations, poor performance and death of a student (Nurmalitasari, et al, 2023).

The Kenya Universities and Colleges Central Placement Service (KUCCPS) portal which is the sole application for programmes in both public and private learning institutions does not have a guidance system yet. The available career guidance systems in Kenya are Intelligent Quotient (IQ) based *(Aptitude Tests)* which are implemented on web platform accessed by via internet browser or mobile application (AssessmentDay Ltd, 2023 & Humanmetrics Inc,1998-2023). These do not factor in the Kenya Certificate of Secondary Education (KCSE) performance and students interests thus cannot offer suitable advisory on programme of study to a university entrant in Kenya.

LITERATURE REVIEW

Gaps in the Existing University Programme Predictive Models

Review of literature suggests that this study is a classification problem. Findings suggests that there is a need to ensure students in transit from high-school education to university education apply for an appropriate programme in order to reduce the rates of the progression crisis, Giuseppe B. (2021) & Shikokoti et al (2023).

Ndung’u et al (2020) found out that 67.7% of the students concurred that they were not able to get information on career guidance from the KUCCPS website. KUCCPS portal does not feature a Programme Recommender page (tab) which could help students get to know their career inclinations. This page would prominently act as a guide for students to strategize their career aspirations in line with Ikigai concepts (Kumano, et al 2018) so as to narrow down on the specific programme that suits them.

The other machine learning models developed in line with career predictions like Neural Network (NN & ANN), Support Vector Machine (SVM), Adaboost, Expert Systems, Case Base Reasoning (CBR), Decision Support Systems (DSS) and Naïve Bayes have proved to work with each model giving a unique accuracy level.

University Programme Choice vis-à-vis Personality Traits, Interest and Qualifications

Nurmalitasari (2023) found out that the rates of incompletion by undergraduate students was due to dropouts by students studying programmes that they did not find interesting. According to Odour (2019), students were pursuing programmes that were not in line with their dream jobs. As a result, graduates ended up working in fields that don't align with their professional background or their preferred career path. Job dissatisfaction, a lack of motivation, subpar performance, a negative attitude toward work, high absenteeism, and high job turnover are the results of wrong programme choice. These have a detrimental impact on productivity within a company and could possibly lead to conflict between a staff and their employer. (Bonenberger et al., 2014).

Maina, (2020) recommended that in order to gain a better understanding of their interests, skills, and preferences, first-year undergraduate university students should be encouraged to do self-assessments. Kemboi et al (2016) study indicated that 26.7% were dissatisfied with their programme of study at Moi University. Atela (2020) reported that there was a significant relationship between personality types and career choice.

In 2016 Rebecca et al did a study on 399 students in Kenya and found a significant correlation between undergraduate students' personality types and profession choice. In this study the Ikigai calculator (Prajapati 2023 and Schippers 2019) has been incorporated for personality type data collection.

Rates of Inter-University Transfers, Change of Programme, Repeat Year, Drop Out, Discontinuation on Academic Grounds and Exams Irregularities

The Higher Education Student Statistics (HESA) in 2021 found out that during the 2018/2019 Academic Year, 2.39% of all first-time undergraduates did inter-university transfers University and 8.3% of the students dropped from the university.

In 2020, Maina found out that 23% changed their programmes of study and are likely to change once they join the University. According to Ayiro (2016), 65% of students from 22 East African universities expressed dissatisfaction with their academic programs. The study showed that about 20% to 50% of students placed by the Kenya Universities and Colleges Central Placement Service (KUCCPS) to join public Universities in Kenya requests for a change of programme.

Without proper vocational guidance and career counseling, many young people choose careers that are inappropriate for them because of ignorance, peer pressure, advice from friends and parents, or the prestige associated with particular jobs (Ryan et al 2019). For instance, according to a survey by Koech et al. (2016), 77.9% of undergraduate students if given an opportunity would wish to switch to another programme. Njoroge (2016) found out that 6% of female and 9% of male students postponed a semester at least once, with 1% of female students and 2% of male students having postponed twice due to uninterest in the programme they were pursuing at the university.

Lugulu and Katwa (2019) reported that examination results were used to determine students who progress to the next semester, sit special/supplementary examinations, repeat class or are discontinued from the program. In 2016, Njoroge did her research in Kenya on exam retakes,

ISSN: 2309:9240, All Rights Reserved for all authors in this Journal

semester deferments, and dropping out as factors influencing attrition rates. She found out that 12% of the male and 9% of the female students sampled had repeated an year of study.

According to Hanson's (2022) research, 32.9% of undergraduate students drop out of school. In a study on retention rates of undergraduate students in selected Universities in Kenya by Nyutu (2019), 12.1% of the student population, dropped out of their studies due to lack of enthusiasm to continue with the selected programme.

In 2022, due to academic dishonesty, twenty-seven students were forced to leave Harvard College during the 2020-2021 school year. Vivi and Leah (2022). In 2022, Wachira reported that students who were discontinued for failing examinations were among the 37.5% of the 30,000 deregistered students at the University of Nairobi.

METHODOLOGY

Data was collected via Google Forms circulated via class WhatsApp groups by various Head of Departments in the five selected universities. The targeted respondents were 200 for each university making the total population size to be 1,000. The sample size was 278 with a zscore of 1.96, standard deviation of 0.5 and a margin error of 0.05. The Google sheet was downloaded via Google sheets on a csv format ready for uploading after clean-up. The number of respondents were 308 thus a higher accuracy advantage.

Python analytics libraries (Numpy, Pandas, Matplotlib, Sklearn, Seaborn, Scipy, Plotly) were deployed using Python's Jupyter Notebook for Anaconda.

Data Preprocessing

Data preprocessing was done to ensure the data was properly formatted, categorized and augmented for training. Most of the variables in the dataset were categorical in nature. One-hot-encoding was therefore required. This resulted in having the new table to have 150 rows each assigned a binary indicator for nuanced predictions.

Table 1: Sample of the Categorical Dataset Description

Description	HardestSubject	TeacherSupport	GuardiansSupport	WrongProgramme
Count	308.00000	308.0000	308.0000	308.0000
Mean	4.0648356	4.503247	2.133117	2.009740
Std.Dv	5.0284186	3.585577	1.300892	1.190312
Min	1.000000	1.000000	1.000000	1.000000
25%	2.000000	2.000000	1.000000	1.000000
50%	3.000000	4.000000	2.000000	2.000000
75%	4.000000	5.000000	3.000000	3.000000
Max	31.00000	31.00000	5.000000	5.000000

Feature Selection for the Tree-Based Predictive Models

This study chose CART which utilizes Gini Index and Twoing Criteria (Breiman et al 1984) and handles both categorical and numeral variables. The splits in this research study used gini index and the obtained tree was pruned by Cost–Complexity Pruning through a probability distribution. It considered the intricacy of the tree as well as the quantity of mistakes. Every leaf's prediction was predicated on the node's weighted mean. Miheso (2020) and Ryan et al (2019) research studies support the features such as subjects done at KCSE, Ikigai concepts, teachers and peer influence, exposure to career guidance, parents support, KCSE performance, extra curricula activities selected for data input for these three algorithms.

Decision Tree Model Training and Testing

The researcher implemented the DT algorithm on the preprocessed data as well as with the manually prepared data with equal number of variables on a ratio of 80:20 for training and testing. In both cases the data was split for training and testing with the random state set to 42. Confusion matrix criterion was applied to assess the effectiveness of the DT model.

The researcher deduced that for the training score had an accuracy score of 80.93% for both manual cleaned and the one-hot-encoding data sets. However, with the preprocessed data the test results had a higher accuracy score of 76.34% whilst the manually labelled data derived an accuracy score of 72.04%.

Table 2: Classification Report for Decision Tree

Classification	One-hot-encoding Data Set		Manual Cleanup Dataset	
Classification	Train	Test	Train	Test
f1Score	80.93%	76.34%	80.93%	72.04%
Recall weighted avg	0.809300	0.7634006	0.809300	0.7204004

Gini index

The training phase incorporated gini index which is an impurity-based criterion that was used for splitting in order to develop the binary tree classifier. Measurement of the divergences between the probability distributions of the suitable programme values was done. The Gini index is defined with the formulae below.

$$Gini\ Index\ = 1 - \sum \left(P(x = k)\right)^{2}$$

where K= the proportion of the sample in class C

Equation 1: Gini Index function

Coding of the gini index was implemented using DecisionTreeClassifier function of the sklearn.tree module as well as the GridSearchCV of the sklearn.model_selection module in python. Training accuracy score was at 80.93% while the Testing accuracy score was at 76.34% with max_depth aset to be 4 and min_samples_leaf at 11.

Random Forest Algorithm for Classification

Random Forest is an ensemble learning algorithm implemented using sklearn.ensemble module by importing the RandomForestClassifier function. Ensemble learning does not rely on one decision tree but rather takes the prediction from each tree and based on the majority votes of predictions, it predicts the final output. The study implemented the RF algorithm yielding a training score of 99.3% and the test results having an accuracy score of 81.72%.
The bootstrapping value was true for the 1000 fits done on the 5 folds with min_samples_split of 5 and n-estimators of 400. Confusion matrix criterion was implemented for evaluating the performance of the RF model.

Table 3: Classification Report for Random Forest

Classification	Train	Test
f1Score	99.53%	81.72%
Recall weighted avg	0.995379	0.817204

Gradient Boosting Algorithm for Classification

According to Simplilearn (2024) and Tomonori M. (2022), this classifier operates by combining several weak learning models to output an effective predictive model. Gradient Boosting loss function is calculated with the formula below.

$$L = \frac{1}{n}\sum_{i=0}^{n}(y_i - \gamma_i)^2$$

Equation 2: Gradient Boosting loss function

Where L is the loss function, Gamma is the predicted value, yi is the observed value and gamma is the predicted value. $\frac{dL}{d\gamma} = -(\, y_i - \gamma_i) = -(Observed - Predicted)$
XGBClassifier of the xgboost module was utilized. The learning_rate was set to be 0.05 and the n_estimators to be 500 while the eval_metric was set to be "logloss" with early_stopping_rounds of 5 and n_jobs of -1.
The Gradient Boosting algorithm yielded a training accuracy score of 90% and the test results having an accuracy score of 75.72%.

Table 4: Classification Report for Gradient Boosting

Classification	Train	Test
f1Score	90%	75.72%
Recall weighted avg	0.90000	0.757210

Confusion Matrix Evaluation

The target variable set for this study was labelled SuitableProgramme. This was set to have two values depicted by P for Positive and N for Negative (Nisha 2023). From the

ISSN: 2309:9240, All Rights Reserved for all authors in this Journal

preprocessed data set, the columns represent the actual values of the target variable and the rows represent the predicted values of the target variable. The formula for evaluating the accuracy of the tree-based algorithms in this study was:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

Equation 3: Confusion matrix formula

TP depicts True Positive, True Negative (TN), False Positive (FP) *Type 1 error* and False Negative (FN), *Type 2 error*

Visualization of the Decision Tree

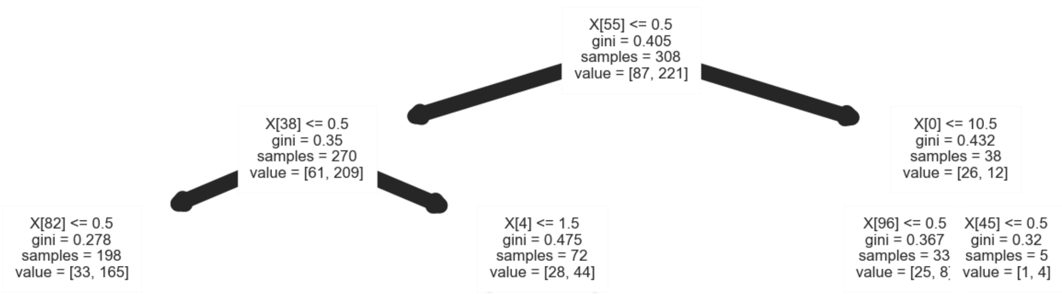

Figure 1: A section of the DT visualization showing gini index and set of values per sampled class

RESULTS AND DISCUSSION/FINDINGS

From the analysis of the three models the evaluation of their performance is shown in the **table** above. RF was found to be the most predictive for this case study.

Table 5: Training and Testing Accuracy of the 3 algorithms

S/No	Algorithm	Prediction Accuracy (Training)	Accuracy Score (Testing)
	Decision Tree	80.93%	76.34%.
	Random Forest	99.30%	81.72%

	Gradient Boosting	90.00%	75.72%

The confusion matrix results are indicated in the table below;

Table 6: Confusion matrix analysis for the 3 algorithms

S/No	Algorithm	Precision	Recall	F1 Scores
	Decision Tree	80.93%	80.93%	80.93%
	Random Forest	99.30%	99.30%	99.30%
	Gradient Boosting	90.00%	90.00%	90.00%

Jun, Mj. (2021) and Cha GW, et al (2021) discovered that RF had higher predictive power than XG Boost and Artificial Neural Networks (ANN), indicating that tree-based ensemble methods were more efficient method for forecasting.

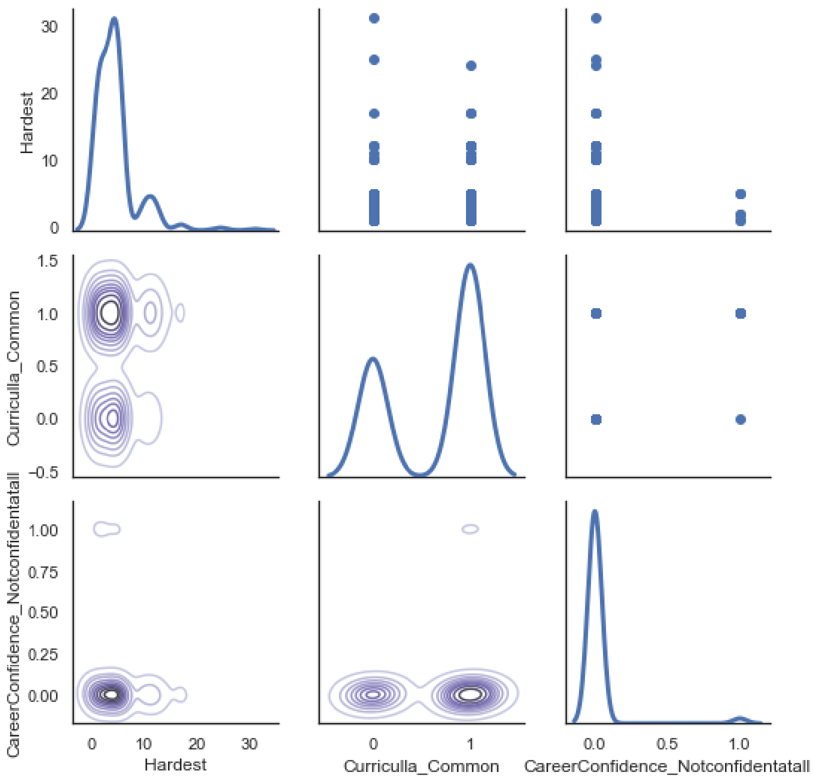

Figure 2: Scatter diagrams indicating relation between hardest subjects at high school and university's common courses.

The relationship between the hardest subjects in high school directly affects the performance of the common university courses leading to the desire of 20.5% of students lacking motivation to continue with their studies irrespective of the programme of study.

CONCLUSION AND RECOMMENDATIONS

Conclusion

The results of this performance evaluation study have demonstrated the high accuracy level of RF as a predictive model being a tree-based classification algorithm. Table 5 shows that RF accuracy was at 99.30%, XG Boost was at 90% and DT was at 80.93%.

This research deduces that, to a high extent, the rates of repeat year cases, discontinuation on academic grounds, deferment to change programme, exam irregularities, change of programme after year 1 of studies and dropped out cases, are due to selecting the wrong programme during application. The rates of these challenges can be reduced by introducing a Random Forest Programme Recommender system at the root level (KUCCPS Students Portal) which among other measures, would help cure this university progression crisis.

Recommendations

The Random Forest model is recommended for implementation by KUCCPS via development of a System to be integrated on the KUCCPS portal. Currently the portal is accessed via its webpage or via mobile App. KUCCPS Students Portal is a one stop shop to utilize this crucial component in its portal. KUCCPS should add a tab in the student portal titled Programme Recommender for engaging students to answer a specific question set based on their personality types and performance at KCSE and predict the best suitable programme for them to select during the application process.

The study also recommends that career guidance be done by Career Officers working in universities to Form 2 students before they select subjects. The purpose would be to enlighten Form 2 students where each and every subject is integrated in the curricula for each programme in the University.

Limitations

The results of this study were based on testing 3 tree-based algorithms. However, there are many other predictive algorithms like Neural Networks, Deep Learning, K-means Clustering, Prophet, K-Nearest Neighbor, Support Vector Machine among others that could derive higher predictability.

ISSN: 2309:9240, All Rights Reserved for all authors in this Journal

The work has focused on one of the causes that leads to repeat year cases, discontinuation on academic grounds, exam irregularities, dropping out cases, change of programme cases and deferment cases. It is important to understand that while this study offers a fundamental foundation in one of the technological solutions to the above phenomena *(programme recommender for university entrants)*, the other array of factors that causes progression crisis in Kenyan universities calls for independent scientific research work.

REFERENCES

AssessmentDay Ltd. Mercury House, Southwark, London, UK (2023) https://www.assessmentday.co.uk/.

Atela (2020) Relationship Between Personality Types and Career Choice Among Undergraduate Students of Maseno University, Kenya.
10.7176/JEP/11-14-14 *Journal of Education and Practice*. Vol 11, No 14 (2020) pp. 127-144

Ayiro, L. P. (2016*) "Career Choices: Dilemmas Facing East African*
Varsity Students." The East African. https://www.theeastafrican.co.ke/tea/ea-universities-guide/career-choices-dilemmas-facing-east-african-varsity-students-1349048.

Bonenberger, M., Aikins, M., Akweongo, P. et al. (2020). The effects of health worker motivation and job satisfaction on turnover intention in Ghana: a cross-sectional study. *Hum Resour Health 12, 43 (2014). https://doi.org/10.1186/1478-4491-12-43*

Briemann, J. Friedman, R. Olshen & C. Stone. (1984). CART: Classification and Regression Trees, EBelmont, s CA:Wadsworth Statistical Press,

Business Today Kenya (2016,) Laikipia University reopens amid confusion over legality of courses. https://businesstoday.co.ke/laikipia-university-reopens-amid-confusion-over-legality-of-courses/

Cha G.W, Moon HJ, Kim YC. Comparison of Random Forest and Gradient Boosting Machine Models for Predicting Demolition Waste Based on Small Datasets and Categorical Variables. Int J Environ Res Public Health. 2021 Aug 12;18(16):8530. doi: 10.3390/ijerph18168530. PMID: 34444277; PMCID: PMC8392226.

Gathoni, N. J., Sirera, M. A., & Olaly, W. (2019). Effectiveness of counselling services on retention rate of undergraduate students in selected universities in Kenya. *International Journal of Psychology and Counselling*, 11(4), 30-38. https://academicjournals.org/journal/IJPC/article-abstract/2E29C4560734 https://doi.org/10.5897/IJPC2019.0558

Giuseppe B. (2021), University Dropout Problems and Solutions, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3788701 Hanson, Melanie. "College Dropout Rates" *EducationData.org*, June 17, 2022, https://educationdata.org/college-dropout-ratespursuing.

Humanmetrics Inc (2023). This free personality test is based on Carl Jung's and Isabel Briggs Myers' personality type theory. https://www.humanmetrics.com/personality/test

Jun, Mj. (2021). A comparison of a gradient boosting decision tree, random forests, and artificial neural networks to model urban land use changes: the case of the Seoul metropolitan area. International Journal of Geographical Information Science. 35. 1-19. 10.1080/13658816.2021.1887490.

Kemboi, R.J., Kindiki, N.J., & Misigo, B. (2016). Relationship between Personality Types and Career Choices of Undergraduate Students: A Case of Moi University, Kenya. *Journal of Education and Practice*, 7, 102-112.

Kenya Vision 2030. (2022). Social Pillar. http://www.vision2030.go.ke/

Kumano, Michiko (2018). "On the Concept of Well-Being in Japan: Feeling
Shiawase as Hedonic Well-Being and Feeling Ikigai as Eudaimonic Well-Being". Applied Research in Quality of Life. 13 (2): 419–433. doi:10.1007/s11482-017-9532-9. ISSN 1871-2576. S2CID 149162906.

Lugulu J. and Katwa J. (2019). Impact of Deferments of Study by Undergraduate
Students in Public Universities: A Survey of Uasin Gishu County, Kenya. *Journal of African Studies in Educational Management and Leadership*. ISSN 2078-7650 Online# 2019, http://www.kaeam.or.ke

Maina, M.,W., K.(2020). Influence of Career Guidance Programmes on Change of Programme of Study among First-Year Undergraduate Students in Kenyan Universities. *Journal of Education*, 3(7), 28-41

Miheso, E. (2020). Determinants of students career choice in Tertiary Institutions in
Kakamega County, Kenya: a case of the Sigalagala National Polytechnic (Doctoral dissertation, University of Nairobi). Retrieved on March 10, 2020, from https:// bit.ly/3ifChvr

Nisha Arya Ahmed (2023) The Model Evaluation Tool Explained https://www.datacamp.com/tutorial/what-is-a-confusion-matrix-in-machine-learning

Ndung'u, Karanja & Obae, Rose. (2020). Effects of Career Placement by KUCCPS among
the Undergraduate Students in Public Universities in Kiambu County, Kenya. International Journal of Research and Innovation in Social Science. 04. 323-329. 10.47772/IJRISS.2020.41216.

Njoroge (2016). Examination Repeats, Semester Deferments and Dropping Out as

Contributors of Attrition Rates in Private Universities in Nairobi County Kenya. *International Journal of Education and Research* Vol. 4 No. 3 March 2016. pp 225-240. https://www.ijern.com/journal/2016/March-2016/17.pdf

Nurmalitasari, Zalizah Awang Long, Mohammad Faizuddin Mohd Noor, "Factors Influencing Dropout Students in Higher Education", *Education Research International*, vol. 2023, Article ID 7704142, 13 pages, 2023. https://doi.org/10.1155/2023/7704142

Prajapati, Aanchal. (2023). The Japanese Secret to A Long and Happy Life. 4. 38-40. 10.53361/dmejc.v4i01.06.Psychometric Success (2023). The Myers-Briggs Assessment Test (2023 Guide) https://psychometric-success.com/aptitude-tests/test-types/myers-briggsRebecca et al (2016), Relationship between Personality Types and Career Choices

Russell, Lisa & Jarvis, Christine. (2019). Student Withdrawal, Retention and Their Sense of Belonging; Their Experience in Their Words.

Research in Educational Administration & Leadership. 4. 494-525. 10.30828/real/2019.3.3.

Ryan, Yano & Rotich, Titus & Korir, Betty & Mutai, Kipkoech & Kosgei, Mathew &

Koech, Julius & Msc, & Koech, Mr. (2019). Factors Influencing the Choice of College Among Undergraduate Students in Public Universities in Kenya A Case Study of the University of Eldoret.

Schippers, Michaéla C.; Ziegler, Niklas (2019). "Life Crafting as a Way to Find Purpose and Meaning in Life". Frontiers in Psychology. 10: 2778. doi:10.3389/fpsyg.2019.02778. ISSN 1664-1078. PMC 6923189. PMID 31920827.

Shikokoti et al (2023). Factors Influencing Rate of Completion of Undergraduate Students

in Public Universities in Kenya. A Case of University of Nairobi, Faculty of Education. Journal of Education and Practice. 14. 2023. 10.7176/JEP/14-31-08.

Simplilearn (2024). Gradient Boosting Algorithm in Python with Scikit-Learn. https://www.simplilearn.com/gradient-boosting-algorithm-in-python-article

Sithole, Alec & Chiyaka, Edward & McCarthy, Peter & Mupinga, Davison & Bucklein,

Brian & Kibirige, Joachim. (2017). Student Attraction, Persistence and Retention in STEM Programs: Successes and Continuing Challenges. Higher Education Studies. 7. 46. 10.5539/hes.v7n1p46.

Thumiki, V. R. R. (2019). Student Dropout from Foundation Program at Modern College of Business & Science, Sultanate of Oman. *International Journal of Higher Education*, 8(5), 118-133, DOI: https://doi.org/10.5430/ijhe.v8n5p118

Tomonori M. (2022). All You Need to Know about Gradient Boosting Algorithm − Part 2. Classification. https://towardsdatascience.com/all-you-need-to-know-about-gradient-boosting-algorithm-part-2-classification-d3ed8f56541e

UN DESA. 2023. The Sustainable Development Goals Report 2023: Special Edition - July 2023. New York, USA: UN DESA. © UN DESA. https://unstats.un.org/sdgs/report/2023/

Vivi E. Lu & Leah J. Teichholtz, Crimson Staff Writers (2022) (November 8, 2022) In Six-Year High, 27 Undergraduates Forced to Withdraw from Harvard in 2020-2021 Due to Honor Code Violations. https://www.thecrimson.com/article/2022/11/8/honor-council-withdrawals/

Wachira Kigotho.(2022) University deregisters thousands of inactive students.https://www.universityworldnews.com/post.php?story=20220307125656824