



International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: <https://www.svedbergopen.com/>



Research Paper

Open Access

Educational Machine Learning For College Recommendation and Career Alignment Among Incoming University Students

Eric Santos Parilla

Northwestern University Philippines. Email: eric.parilla@nwu.edu.ph

Abstract

This work created and tested a machine learning approach to help incoming university students find colleges that fit both education plans and career goals. It used school records for 7,000 students at Northwestern University, covering the years 2020 through 2024. The analysis looked at demographic data, school results, admission details, and career-related items. These factors were treated as possible signals for predicting college placement. Six supervised learning models were trained and compared: Multinomial Logistic Regression, Decision Tree, Random Forest, Gradient Boosting, XGBoost, and LightGBM. To judge each model, the study used accuracy, precision, recall, and F1-score. It also used five-fold cross-validation to reduce the risk of one split giving a misleading result. Among the models, Random Forest reached the best holdout accuracy at 80.79%. Its F1-score was 80.77%, which was also the top value for that metric. Still, Multinomial Logistic Regression looked more stable across folds. It posted the top cross-validation F1-score at 79.13%. It also did not show clear signs of overfitting. For that reason, it was described as the most sound choice for use in the institution. The study also ran SHapley Additive exPlanations (SHAP) to interpret the drivers of the predictions. SHAP suggested that Senior High School strand carried the most useful signal. Next came admission scores for English, then Science, then Mathematics. Overall, the results claim that educational machine learning can turn existing student data into clearer and evidence-based suggestions. The framework is meant to act as decision support. It should work alongside professional academic advising, respect student choices, and encourage careful use of artificial intelligence in higher education. It is presented as responsible, understandable, and centered on students.

Keywords: educational machine learning; college recommendation; career alignment; explainable artificial intelligence; SHAP.

Introduction

Universities have in the recent past been using more data-driven methods that are related to machine learning, artificial intelligence (AI) and educational analytics. In most cases, institutions are at the moment better placed for management of better admissions, academic advisement, curriculum and student success processes. As always, colleges are advising and encouraging students based upon processed aptitude test results, in academic counseling, and from the student's point of view. These processes are significant and necessary, but they also fail to assess many variables in a single shot. Only academic readiness, career aspirations, socioeconomic concerns, plus the track records of the institution and other data, can get left out, leading to wrong placement of students in college, academic mismatch, and eventually student attrition (Delen et al., 2023; Sekeroglu et al., 2021).

Data Mining in Education (EDM) and Learning Analytics, along with their tremendous capacity to learn, create substantial means of retrieving such knowledge. Output of the data combined with it can be very useful in enlightening the institution in terms of what kinds of pattern to look for, seeing future student performances, and supporting fine-grained recommendations to improve evidence-based decision-making within the institution. In one recent and comprehensive systematic review, such findings were brought out when application was made in areas such as prediction of student performance, forecasting of enrollments, and predicting dropouts. Commonly used plus consistently improving performance across educational applications is the imputation through different machine-learning models such as logistic regression, decision trees, random forests, gradient boosting, XGBoost, and LightGBM (Sekeroglu et al., 2021; Tjandra et al., 2022). Expansion of

educational datasets also offers such possibilities for universities to fully engage data-driven all-round institution mining of student support in prospects of enhancing academic success and institutional sustainability.

The career counseling systems have undergone a massive transformation due to their junction with artificial intelligence. The AI-based career guidance systems of earlier days primarily worked as expert systems, which matched students and careers as per the responses to questionnaires and rule-based logic systems (Crowley 1992). The integration of educational attainment, personal interests, competencies, and labor market information has significantly improved these systems by predictive recommendation engines using contemporary machine learning techniques. For Guleria and Sood (2023), it was demonstrated that transparent accounts in AI improve the educational data mining inspired career counseling work by combining high predictive performance with transparent explanations. Similarly, a career guidance model based on machine learning was developed by José-García et al. (2022) to allow skills assessment and suggest career pathways in line with industry demand. Such progress effectively depicts the evolution of AI from simple recommendation systems towards highly intelligent decision support capabilities that can now personalize educational and career planning.

However, in the realm of educational machine learning, many models retain their "black-box" structures—they conduct decision-making processes the user cannot understand, and the dialogues of such a model itself are rather light and elegant. While a highly accurate predictive model is believed to fulfill all educational designs, it seems to be insufficient and ineffective as educational decisions directly influence the futures of the respective students since they require transparency, responsibility, and trust. Hence, for this reason, Explainable Artificial Intelligence (XAI) is really needed in this direction, which will come up with understandable explanations in order to illustrate the predictions produced by a machine learning model. Such explanation techniques as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) offer school leaders, counselors, administrators and students the opportunity to realize which variables account the most in any one of the predictions to be made direct learning (Delen et al., 2023). In sum, therefore, Explainable AI will gradually lead up toward a more ethically transparent and responsible use of machine learning in education while adding value to automatic recommendations in the long run by increasing stakeholder confidence.

Recent literature stresses that predictive analytics should not stop at identifying at-risk students instead providing proactive assistance to guide them before their academic journeys begin. College recommendation systems are trained using Machine based learning, which aids in the integration of other sources of deriving information such as academic performance, entrance examination score, psychometric assessments, career interests, and demographic characteristics. The recommendation systems suggest the academic program to which the student fits in best. A recent machine learning framework envisaged student career prediction and college recommendation synthesized that predictive analytics combined with personalized career counseling would enable students to make informed decisions about their education with the least subjective bias, while preventing career mismatches (SSRN, 2025). These systems also aid universities in improving enrollment planning, student retention, and curriculum alignment to labor market demands.

In light of these developments, the current study posits a College Recommendation Model for Career Alignment and Explanation through ML among Incoming University Students. The research prospect offers an appropriate model for machine learning under the supervision to give recommendations concerning a forthcoming university on the basis of students' entrance examination results, state and high school results, and career aspirations, interests, and socioeconomic variables. This research is aimed at those who are looking for suitable colleges. Unlike the contemporary pair of recommendation systems where prediction alone could be powerful, the new model champions transparency as it explains why and how students' career decisions are molded. Since the new model is capable of identifying the most stretchable of the stretch to spot the top five people best matched with the information under study, the application of this framework would link college and university students. Moreover, it cannot be denied that the proposed scheme of everyday grade management and advising could program realignment, as it makes new decisions binding into a system that would trigger an actual FYMCF!-Algo! upgrade. By doing so, the emerging recordable term from the field of explainable educational analytics would gain solid support in designing academic tracing in positive student-college matches, in penning proper advising, curbing program mismatch, and improving career readiness.

Research Framework

Theoretical Underpinning

This study is mainly based on the Person-Environment Fit Theory, which was formulated by Kurt Lewin and John L. Holland. P-E Fit Theory is one of those theories which has been developed fundamentally through the

inspiring thoughts of John L. Holland's (1959, 1997) Theory of Vocational Personalities and Work Environments. Truly, differences revealed in this regard could very well result from increased employability of workers in job accomplishment therefore increased satisfaction and motivation in the job assigned. This is further explained by the visualization that human behavior is much influenced by the environment in which one is placed economically or generally; for instance, it is generally believed that poor environmental waste management policies bring negative effects and some poor behavior in members of the society since they are witnesses of their contemporaries, perhaps the parents, an industrial sector, or any specific individual demonstrating such conduct. Holland defined it, and in his words, seeking to explain how individuals gravitate toward environmental endogamists-meaning those individuals tend to choose occupational and nonoccupational environments that are stable, predictable, and free from distraction. So, different types of P-E fit, according to Holland, can be seen in various environments and so can the impact of such fits on a host of other variables ranging from academic and vocational success to career satisfaction in the long run. From the views of previous researchers outlined above, the theory is seemingly not yet without boundaries. However, in the context of career decision making, it seems to assume scope that is broad and meaningful for students in terms of the choice of pathway across a variety of competencies or personal interests rather than just academic performance.

This very concept underlies the Person-Environment Fit Theory and it explains why precise college recommendation should be made, especially for incoming university students. The proposed clear machine learning model predicts the appropriate most college for a particular student with multiple dimensions of student information input such as test scores on admission examination, performance in senior high school and academic goals, career aspirations, interests, preferred courses, and demographic characteristics. Instead of advising in a conventional way or solely admission testing results, these machine learning models work by examining the complex relationships among these variables into the course recommendations that most suit the profile of the student in question. It is a form of prediction that comes in line with the spine of P-E Fit Theory and asserts that optimum educational outcomes require an agreement in student characteristics and the academic environment. The addition of a very much intricate analysis of student attributes heightens the chances of matching students to programs aligned well with their capacity and goals, thus reducing mismatching within the academe and increasing long-time educational success.

This research aims to augment practical applications of theory by integrating principles of Explainable Artificial Intelligence (XAI) into predictive modeling. The educational decision-making context demands significant transparency and accountability in recommending a policy that would significantly affect the future of students' academic careers, since both the decision and the students are at stake. Explainable AI, in the form of SHapley Additive exPlanations (SHAP), particularly assists in highly interpretable ways a predictor's contribution is made in relation to each recommendation, thereby enabling students, parents, academic advisers, and university administrators to understand what underpins the model's decisions. A demonstration of what recommendation is made as well as why a particular college represents the best fit for that person are the features that clarify the theoretical assumptions in the theory of fitting a person to an environmental context. The assimilation of 'explainable machine learning in Person-Environment Fit Theory results in an evidence-based and transparent framework that is student centered, which helps academic advisement, aligning courses at the college level, and facilitates educational planning informed by the option of higher education.

Related Studies

Artificial Intelligence and Machine Learning for College and Degree Recommendation. Artificial intelligence and machine learning increasingly guide students when selecting the right program according to academic possibilities, interests, and career interests. Recommender systems for education aim at analyzing the data provided for students to generate suggestions to help them overcome some of the uncertainty and over-information during decision-making'. These often use metrics such as GPAs, standardized tests, or admission patterns in most educational institutions. But there are growing ideas about considering multiple aspects in choosing academic programs, hence it is revealed that the selection of academic programs is composed of good academic performance, personal preferences, socio-economic background, career perception, etc., thus necessitating multiple dimensions in analysis. Urdaneta-Ponte et al, doing systematic reviews of educational recommender systems, declared that the application areas of such systems range from advisories on coursework selection, learning-resource recommendations, academic planning, and personalized learning. Systems that adapted commerce-oriented algorithms had to address education settings and the learners about whom they cater. Scholarship empirical investigational results indicate the vast potential for supervised learning to recommend undergraduate majors. Alghamdi et al. (2021), in a line of similar research, utilized

academic marks, marks on the entrance test, work experience, and data about placement to find out and predict suitable fields of specialization that could be made interesting for helping students choose their fields of specialization. Their results showed that the Random Forest and Gradient Boosting models would be integrated into intelligent systems for recommending the undergraduate major in an appropriate way. The Higher Secondary marks, marks in university, and entry-test marks were also found as very influential predictors for the same. The study added that pre-university academic performance can provide signals for students' readiness in different fields. But, grades don't always explained so much about how each student chooses their college. This system was developed by Delahoz-Domínguez et al. (2024), incorporating standardized test scores with socioeconomic data to provide a recommendation for studying at the university. The authors also emphasized that many of the existing systems focused narrowly on examinations, plus grades while ignoring a wider framework within which some students make educational or career decisions. The recommendation of the work has shown that the importance of assessing academic readiness by combining it with contextual factors before academic programs can be suggested. Similarly, The authors of another research, Cha et al. (2024), found that students will allow a machine to suggest zero related to an artificial intelligence course inventory and are looking for an evaluating support requisite on-course suitability as well. This finding showed that suggestion systems should improve when it comes to ranking more than one program and making it understandable for learners how closely the indicated academic pathways suit their preparation, interests, and long-term goals.

AI-Enhanced Career Guidance and Student-Career Alignment. Career development is a hard nut to crack as it takes understanding by students and interest, ability, value, educational options, and employment outcomes. Traditional career services typically require involving counselors, interviews, aptitude testing, and self-report instruments. All these are great forms of human support, but are limited as compared to others because they can be influenced by the availability and amount of counseling, the time needed, and the level of difficulty experienced in assisting students who each have manifold characteristics. AI-enhanced learning analytics steps into place as a decision-support technology, one equipped with more research- and goal-scoring ability. Gedrimiene et al. (2024) conducted a study on a new tool, which is an AI-based learning analytics tool that directs and stimulates career decision-making. It involved 106 students interacting with the tool and then evaluating its usefulness and limitations. The investigation clearly identified that using the AI-enhanced learning analytics tool does bring certain interests to users; among these include that users found it faster to display information regarding careers; needed assistance with analyzing and researching alternatives; the news being very diverse in areas relating to possible career paths; adhered towards decision placements; and gave complimentary support to self-reflection. This system was considered useful during the first changeover, especially where uncertainty was experienced due to novel options being explored within sought-after spheres of study or considered for advanced education or reconsidered about where they are heading in life. The research-backed the view that the AI offers a possibility as an assisting tool that can widen for students even the noblest options to consider and in designing ambitious careers. Similar research, however, evoked a requirement for higher personalization, a more didactic interface, and a better support of self-knowledge, for example. The tool was used to support users in trading education options; however, the process lacked serious explanation to go along with the different decision-making stages of the vocational track. The elements-of-self such as interests, professionally relevant skills, and future goals were presented in less integrative form than information on available programs and occupations. Gedrimiene et al. (2024) added that people were presented with just the final recommendation list, if at all, lacking the possibility to inspect why one would be suggested as opposed to the others, something that now becomes hidden from public understanding. This hiddenness can block individuals' capacity to critically assess suggestions and their independency in developing career decision-making skills. The issues are primordial on account of the postulation or prediction of the upshot of the major a student is able to finish academically. The proposal must take into account the interest, aspiration, capability, and orientation toward a certain field of work of the student. Hence, AI-based career guidance should operate as a complement to professional counseling rather than any automatic substitute for human judgment. Extended recommendations may be given based on several factors in clear-cut, evidence-based choices, while still allowing counselors and students to construe such suggestions in the context of the individual and societal roles of a student. A more holistic recommendation could be given to incoming university students through the integration of results among the different career-interest scorings, career aspiration, senior high school organization test preparations, course placements, and admission results than those acting as yardsticks by virtue of basic grades or examination scores.

Explainable Artificial Intelligence in Educational Recommendation Systems. Concerns about raising the increase of sophisticated machine learning algorithms in education towards slight transparency, fairness, accountability, and trust necessitate studying more about potential legal and regulatory impacts within organizations and learning. Random Forest can crunch complex nonlinear relationships and interactions with educational data, while Gradient Boosting, XGBoost, and LightGBM are just some of the ensemble methods available. But have difficulty with the interpretation of how they collectively settled on a classification. When machine-generated feedback guides high-stakes selection decisions such as curriculum placements, educational treatments, or vocational guidance, particularly the problems of the issue bound up with these recommendations, it leads to the questioning of the intent and purpose of the decision process. It makes sense for Khosravi et al. (2022) to note that Explainable Artificial Intelligence in Education aims to make AI-directed educational decisions comprehensible to participants. The need for the kind of educational explanation that would suit different users arises out of the idea that different constituents, say, students, educators, administrators, and system developers, would necessarily require different levels and types of explanations. Explanation methods like the SHapley Additive exPlanations go some way in helping researchers understand what proportion of the behavior of the whole model, or individual recommendations, can be attributed to a particular predictor. For example, in a college-recommendation context, SHAP might suggest that the main things impacting a recommendation are the high school strand of the senior high school, admission scores, career aspiration, interest area, or course-preference rank. This information enables teachers or academic advisers to analyze the educational justification of the recommendation before agreeing with what the model suggests. Explainable recommendations also facilitate the understanding of the academic part of students paralleled with the requirements needed by different academic programs of the university. However, simple and straightforward explanation will not always lead to ethical or trustworthy decision making. According to Farrow (2023), the feature of explainable AI in education should be considered in the context of a broader sociotechnical understanding. So, 'an explanation can be technically correct but it cannot necessarily investigate whether the right combination of variables is used by the model, whether historical data contain procedural biases, or whether users can vis-a-vis the recommendation, say something about it. That is why, in addition to explainability, there should be sufficient human oversight, fairness audit, control of data, and, possibly, intent scrutiny related to the purposes for which predictions are employed. Baker and Hawn (2022), among other experts, made this point referring to prejudice in the algorithm used in educational systems and discussed the potential for a data-driven model to reproduce inequities from a historical era or institutional mode of practice. Therefore it is both technical and pedagogically necessary where recommending college careers. The model must produce the correct predictions while also explaining why an individual program is recommended. The recommendation must be advisory without an automatic blockage of students from their dream degree. The transparent system instead offers an informed basis for conversation among the students, their parents, counselors, admissions officers and academic advisors. This maintains student agency without taking away the benefits of institutions in predictive analytics.

Synthesis. The studies reviewed have proved that AI and machine learning are aiding the betterment of academic recommendation and career guidance by making a combination with the data of elective information with the historical patterns learned through previous studies. Studies on degree and course recommending that for well display academic achievement, the admission tests, previous educational preparation and the enrollment detail are some of the indicators that can always help predict well-suited academic pathways. Further, there is proof indicating that most students would value the AI systems in helping not only in generating recommendations, but also with respect to simplification in seeking and appraising academic choices. However most helpful use will likely be like in the first case: as decision support rather than directly replacing professional academic and career counseling. It still faces a literature that shows that most current systems only tackle a single part of the entire decision-making problem. Some studies advise on courses; others propose largely academic-based admitted majors, and career guidance apps are inclined towards information services and skills exploration. Fewer studies have an integrated and explanatory methodology created to come up with a suggested classification of a right college according to the academic readiness, admission results, previous junior high, course preferences, career interests, aspirations, and the socioeconomic status of students. Of the several recommendation tools, some extrapolate final suggestions with no explanation of their model logics. This gap in their research world is concerned with the generation of transparent systems for college recommendations that are both predictive in terms of end-to-end performance and interpretatively meaningful to student-centered understanding. The purpose of this particular research, then, is to model and test explanations through explaining supervised machine learning algorithms and explainable AI concerning

factors that play into college recommendation among upcoming enrolled students. In this way, an evidence-based framework could accompany academic guidance leading to career alignment and an informed student choice, all within the logic of enhancing but without effacing human judgment as well as student agency.

Statement of the Problem

The primary purpose of this study is to develop an Explainable Machine Learning Framework for College Recommendation and Career Alignment among Incoming University Students. Specifically, it seeks to answer the following questions:

General Problem

How can educational machine learning be assessed to accurately recommend the most appropriate college for incoming university students based on their academic performance, admission assessment results, career aspirations, interests, and demographic characteristics?

Specific Problems

1. What are the demographic and educational profiles of the incoming university students in terms of:
 - 1.1 Senior High School strand;
 - 1.2 monthly family income bracket;
 - 1.3 career aspiration;
 - 1.4 interest area;
2. What are the academic and admission profiles of the incoming university students in terms of:
 - 2.1 Senior High School Mathematics grade;
 - 2.2 Senior High School English grade;
 - 2.3 Senior High School Science grade;
 - 2.4 Admission Test Mathematics score;
 - 2.5 Admission Test English score;
 - 2.6 Admission Test Science score;
 - 2.7 Overall Admission Test score;
 - 2.8 Admission Interview score; and
 - 2.9 Career Interest score?
3. Based from analysis which supervised machine learning algorithm demonstrates the best predictive performance for college recommendation among the following models?
 - 3.1 Multinomial Logistic Regression;
 - 3.2 Decision Tree;
 - 3.3 Random Forest;
 - 3.4 Gradient Boosting;
 - 3.5 XGBoost; and
 - 3.6 LightGBM.
4. How do the developed machine learning models compare in terms of predictive performance using the following evaluation metrics?
 - 4.1 Accuracy;
 - 4.2 Precision;
 - 4.3 Recall;
 - 4.4 F1-score; and
 - 4.5 Cross-validation performance.
5. Which predictor variables contribute most significantly to the college recommendation generated by the best-performing machine learning model based on Explainable Artificial Intelligence (SHAP) analysis?

Research Paradigm

Independent Variable

Dependent Variable

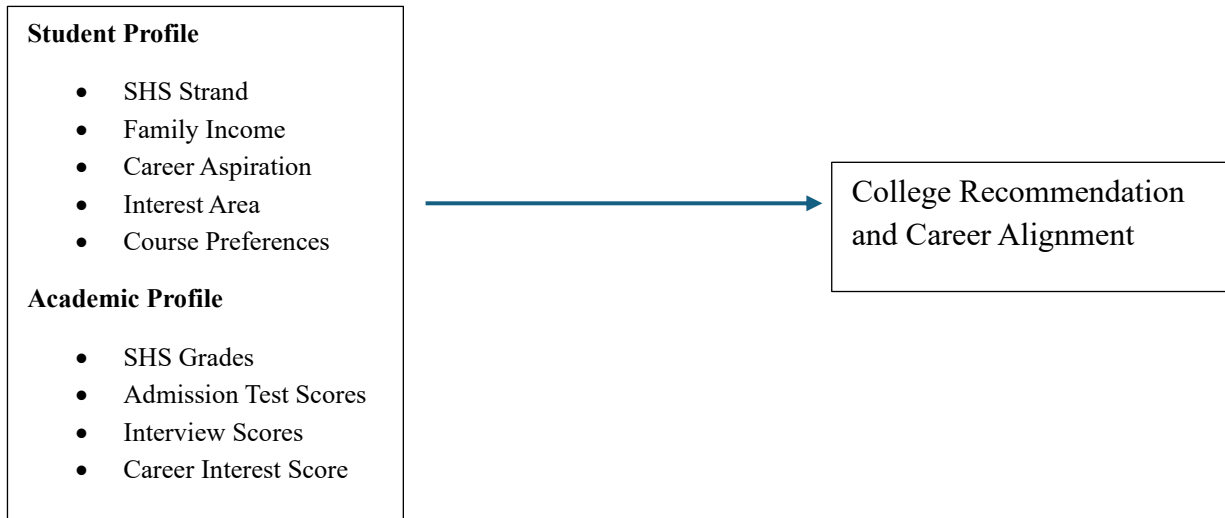


Figure 1. Research Paradigm

Research paradigm that aims at developing an explainable ML framework for college guidance and career alignment recommended to incoming university students. It makes use of students from all demographic, academic, and career patches of information as input; while on the process-it undergoes data preprocessing, application of a supervised machine learning algorithm for performance evaluation along a SHAP guided explanation that is more interpretable, arriving at the final production of the transparent decision-support framework that could recommend the most suitable college and give explanations for each prediction.

Methods

Research Design

This study used a numbers-based predictive analytics setup. It relied on document review. For the process plan, it followed the CRISP-DM method. The project built an educational machine learning system. It then tested the system to suggest a suitable college for new university students. It used inputs tied to grades, admission details, demographic factors, and career goals. To build the models, the researcher tried six supervised methods. These were Multinomial Logistic Regression, Decision Tree, Random Forest, Gradient Boosting, XGBoost, and LightGBM. After training, the researcher tuned settings to improve results. To understand what drove each recommendation, the study used SHapley Additive exPlanations, or SHAP.

The researcher moved through the main CRISP-DM steps. These steps were business understanding, data understanding, data preparation, modeling, evaluation, and deployment planning. Using this path, the study compared the models. It looked at accuracy, precision, recall, F1-score, and cross validation results. Then it chose the top model based on the scores. The researcher explained the final predictions with SHAP.

In the end, the study produced an explainable educational machine learning framework. It offered clear, evidence based support for college matching. It also supported career fit checks, academic advising, and institutional choices that rely on data.

Participants

The research used 7,000 first year applicants to Northwestern University. These students applied between 2020 and 2024. The author pulled records from the school's official databases. Only files with full and usable details were kept. Incomplete entries were left out. The dataset included several parts of each student profile. It covered background demographics. It also included Senior High School results. Admission test scores were included too. Interview ratings were part of the data. Career interest scores were included as well. Course choices and enrollment details were recorded. Academic performance records were also included. No new data was collected from students. The researcher relied on existing institutional records. There were no student surveys. No student interviews were run. No questionnaires were given. The Registrar's Office shared the academic, admission, and enrollment files. The Guidance Office shared the career interest results and the admission interview ratings. It also shared guidance notes tied to career plans and similar items. Before any

review, register removed direct personal identifiers. This was done to protect student privacy. After that, the researcher checked every record. The checks looked at completeness, internal match, and whether each item fit the machine learning plan.

Instrumentation

The study relied on a structured document review plan as its main tool. Data came from records kept at Northwestern University by the Registrar Office and the Guidance Office. The researcher chose document review because it offered a clear way to pull needed facts from existing files. This supported later prediction work, using what was already on hand. The Registrar Office shared several types of records. These included enrollment details, Senior High School grades, results from admission tests, and the student's enrollment history. It also included records related to changes over time, retention status, and other academic data. The Guidance Office shared its own set of materials. These were career aspiration profiles, results from career interest checks, scores from admission interviews, and the preferred degree tracks. It also included counseling notes and other guidance related items.

From those files, the study pulled these variables:

- Senior High School Mathematics, English, and Science grades;
- Admission test results in Mathematics, English, and Science;
- Total admission test score;
- Admission interview score;
- Career interest score;
- Senior High School strand;
- Monthly family income bracket;
- Career aspiration; and
- Interest area.

After pulling the data, the study cleaned and prepared it for machine learning work. It also checked the data for accuracy, then made sure the set was ready to use for model building and testing.

Data Gathering Procedure

The researcher got ethical approval from the Northwestern University Ethics Review Committee (NU ERC) before touching the institutional records. Once approval was granted, the researcher asked the Registrar's Office and the Guidance Office for access to the study datasets. Staff from those offices pulled the needed student records from the correct institutional systems.

Before the researcher received the data for analysis, the offices stripped out direct identifiers. This included student names, ID numbers, student numbers, home addresses, phone numbers, email addresses, and other personal details. After that, the academic and guidance files were joined using approved institutional keys. Throughout the study, the researcher kept student records anonymous and confidential. Next, the researcher worked on the merged dataset. The first steps were to check if the data was complete and consistent. Then duplicate rows were removed. Missing values were handled, categorical fields were encoded, and simple feature work was done as needed. After that, data quality checks were carried out to confirm the variables could support predictive modeling. After preprocessing, the dataset was split into training and testing parts. Stratified sampling was used so the target variable stayed balanced across both sets. The researcher then built and tested supervised machine learning models with the prepared data. Hyperparameter tuning and cross-validation were done, and SHapley Additive exPlanations (SHAP) were run to show how predictor variables affected the recommendations for college.

Data Analysis

The researcher worked with the gathered data in Python. They used supervised machine learning for the modeling step. Before training anything, the researcher checked the data with basic summary numbers. The researcher also did exploratory data analysis to see how the variables were shaped. This included looking for missing entries, possible outliers, and any odd values. They further scanned for trends and links across the student profile, academic details, admission factors, and career related fields. When building the college recommendation model, the researcher tried six supervised methods. The list was:

- Multinomial Logistic Regression;
- Decision Tree;
- Random Forest;
- Gradient Boosting;

- XGBoost; and
- LightGBM.

Next, the researcher tested how well each model predicted. They looked at accuracy, precision, recall, and F1-score. To reduce the chance of a lucky or unlucky split, they also ran cross-validation. That helped them see whether performance held up on other parts of the dataset. After reviewing the results, they picked the top model for the recommendation task. In the final step, they used SHAP to study the best model. The goal was to see how each predictor variable pushed the output. The researcher used this to judge the relative weight of the demographic, academic, admission, and career related characteristics. With this extra look, the model became easier to understand for advisers, guidance counselors, and school administrators who needed to know what drove the recommendations.

Ethical Considerations

Before starting the study, the researcher got approval from the Northwestern University Ethics Review Committee (NU ERC). During the work, the researcher followed the Philippine Data Privacy Act of 2012, Republic Act No. 10173. The researcher also followed the university rules on data handling. For analysis, the researcher relied only on anonymized institutional records. Before the dataset was accessed, the researcher removed any details that could identify a person. The Registrar's Office handled the request for student records. It approved what could be extracted. The office also worked with the University's Data Protection Officer (DPO). This coordination helped ensure the dataset included only items needed for the study. The researcher then limited who could view the dataset. Access was given only to approved staff. Security steps were put in place to keep student data private and protected. The machine learning model was used as a support tool, not as an automatic decision maker. It was used for college recommendation, career matching, academic advising, and planning within the institution. It did not take the place of advisers, guidance counselors, admissions staff, or university leaders. The researcher also did not rely on model outputs as the only basis for admission choices. The same limit applied to disciplinary matters, academic penalties, and other actions that could harm students. The researcher also respected student rights under the usual procedures. Students could request access to their data, ask for corrections, and request removal of personal data. These requests were handled through the Registrar's Office and the Data Protection Officer. Across the entire study, the researcher followed confidentiality rules, privacy practices, transparency steps, and accountability duties. The work also kept human review in place and used artificial intelligence in a responsible way.

Results

Table 1. Demographic and Educational Profile of Incoming University Students

Profile Variable / Category	Frequency (f)	Percentage (%)
Senior High School Strand		
GAS	2,047	29.24
STEM	1,851	26.44
HUMSS	1,205	17.21
ABM	1,040	14.86
TVL	857	12.24
Monthly Family Income Bracket		
Below ₱10,000	1,191	17.01
₱10,000–₱19,999	1,922	27.46
₱20,000–₱39,999	2,150	30.71
₱40,000–₱79,999	1,138	16.26
₱80,000 and above	599	8.56
Career Aspiration		

Profile Variable / Category	Frequency (f)	Percentage (%)
Tourism Officer / Travel Consultant / Events Coordinator	422	6.03
Hotel Manager / Food and Beverage or Guest Service Professional	420	6.00
Marine Deck Officer / Ship Navigation Officer	273	3.90
IT Specialist / Network Administrator / Systems Professional	273	3.90
Marine Engineer / Ship Engine Officer	272	3.89
Architect / Architectural Designer / Urban Designer	264	3.77
Medical Technologist / Laboratory Scientist	264	3.77
Human Resource Officer / Talent Development Specialist	263	3.76
Operations Supervisor / Supply Chain or Process Professional	263	3.76
Management Accountant / Cost Analyst / Budget Professional	260	3.71
Marketing Officer / Sales Manager / Brand Specialist	258	3.69
Certified Public Accountant / Auditor / Tax Professional	253	3.61
Radiologic Technologist / Medical Imaging Specialist	252	3.60
Computer Engineer / Embedded Systems or Hardware Professional	249	3.56
Elementary Teacher / Basic Education Learning Professional	249	3.56
Financial Analyst / Banking or Investment Associate	247	3.53
Software Developer / Data Analyst / Systems Developer	239	3.41
English Teacher / Language Instructor / Communication Professional	239	3.41
Mathematics Teacher / Academic Tutor / Quantitative Professional	237	3.39
Registered Nurse / Clinical Care Practitioner	235	3.36
Pharmacist / Clinical or Community Pharmacy Professional	234	3.34
Civil Engineer / Construction Project Engineer	234	3.34
Biologist / Research Assistant / Medical or Allied Science Professional	233	3.33
Media Practitioner / Communication Officer / Content Professional	227	3.24
Psychologist / Guidance Counselor / Human Resource Professional	217	3.10
Law Enforcement Officer / Criminologist / Public Safety Professional	215	3.07
Physical Education Teacher / Sports Coach / Fitness Professional	208	2.97
Interest Area		
Tourism, Travel, and Events Management	422	6.03
Hospitality, Culinary, and Service Management	420	6.00
Maritime Operations and Seafaring	273	3.90
Information Technology and Digital Systems	273	3.90
Maritime Engineering and Technical Operations	272	3.89
Architecture, Design, and Built Environment	264	3.77
Laboratory Science and Diagnostics	264	3.77
Business, Management, and People Development	263	3.76
Business Operations and Supply Chain	263	3.76
Accounting, Costing, and Management Control	260	3.71
Marketing, Sales, and Entrepreneurship	258	3.69
Accounting, Finance, and Business Analytics	253	3.61
Diagnostic Imaging and Allied Health	252	3.60
Computing, Engineering, and Technology	249	3.56

Profile Variable / Category	Frequency (f)	Percentage (%)
Teaching, Child Development, and Basic Education	249	3.56
Finance, Banking, and Investment	247	3.53
Computing, Programming, and Data	239	3.41
Teaching, Language, and Communication	239	3.41
Teaching, Mathematics, and Analytics	237	3.39
Nursing, Patient Care, and Health Services	235	3.36
Pharmaceutical and Health Sciences	234	3.34
Engineering, Infrastructure, and Design	234	3.34
Life Sciences and Health Preparation	233	3.33
Communication, Media, and Public Relations	227	3.24
Behavioral and Social Sciences	217	3.10
Criminal Justice and Public Safety	215	3.07
Physical Education, Sports, and Wellness	208	2.97
Total	7,000	100.00

Table 1 presents data on 7,000 incoming students. The group included students with varied school tracks, family situations, and planned careers. GAS made up the biggest portion at 29.24%. STEM came next with 26.44%. Together, these two SHS strands accounted for more than half of the entries. Family income was also uneven across students. Most fell under ₱20,000–₱39,999 with 30.71%. A large share also came from ₱10,000–₱19,999 at 27.46%. This pattern points to many applicants coming from lower to middle income households. When career goals were considered, the answers were spread across many areas. Still, tourism and hospitality stood out. Tourism reached 6.03% and hospitality was 6.00%. The wide distribution suggests that students arrived with different learning backgrounds and different reasons for enrolling. These figures can support machine learning work because they reflect several parts of each student profile. Even so, the summaries by themselves cannot prove which factors matter most. The later model testing and the SHAP results were used to check predictive importance.

Table 2 Academic and Admission Profile of Incoming University Students

Academic and Admission Variable	Mean	SD	Minimum	Maximum
SHS Grade in Mathematics	85.27	7.25	75.00	99.00
SHS Grade in English	86.11	7.32	75.00	99.00
SHS Grade in Science	85.29	7.27	75.00	99.00
Admission Score in Mathematics	82.28	10.22	55.00	100.00
Admission Score in English	82.98	10.01	55.00	100.00
Admission Score in Science	82.27	10.22	55.00	100.00
Overall Admission Score	82.51	7.93	55.00	100.00
Admission Interview Score	86.34	10.45	60.00	100.00
Career Interest Score	79.87	10.08	40.00	100.00

Table 2 indicates that the 7,000 students who came in tended to do well in school work and in admissions. Looking at SHS subject grades, English led the group with a mean of 86.11 and an SD of 7.32. Science came next, with $M = 85.29$ and $SD = 7.27$. Mathematics followed closely at $M = 85.27$, $SD = 7.25$. For the admission tests, the scores were a bit lower overall. English still had the top mean at 82.98, with $SD = 10.01$. The overall admission score was 82.51, with $SD = 7.93$. This suggests applicants showed fairly steady results rather than large swings. The interview results stood out with the highest overall mean, 86.34. But its SD was 10.45, so students varied more here than in the test and grade measures. Career interest, on the other hand, had the

lowest mean at 79.87. It also showed the widest spread, ranging from 40 to 100. In other words, career interest scores differed more from person to person than the academic and admissions scores did. Taken as a whole, the data show that many students joined the university with solid academic and admissions outcomes. Career interest appears more mixed across individuals.

Table 3 Predictive Performance of Machine Learning Algorithms for College Recommendation

Algorithm	Accuracy	Precision	Recall	F1-Score
Random Forest	0.8079	0.8108	0.8079	0.8077
Multinomial Logistic Regression	0.8071	0.8102	0.8071	0.8070
Gradient Boosting	0.8064	0.8094	0.8064	0.8063
XGBoost	0.8050	0.8071	0.8050	0.8048
LightGBM	0.7971	0.7991	0.7971	0.7969
Decision Tree	0.6636	0.6639	0.6636	0.6634

Table 3 indicates that Random Forest performed best. Its accuracy reached 0.8079, precision was 0.8108, recall was 0.8079, and the F1-score came to 0.8077. In practical terms, this model labeled about 80.79% of students to the correct colleges. Multinomial Logistic Regression was next. It had an F1-score of 0.8070. Gradient Boosting scored 0.8063, and XGBoost finished at 0.8048. LightGBM did a bit worse, with an F1-score of 0.7969. Decision Tree was the lowest performer. It showed accuracy of 0.6636 and an F1-score of 0.6634. These results suggest that a single-tree setup learned the multiclass structure less well. The top models also stayed close across accuracy, precision, recall, and F1-score. This points to steady predictive behavior. From the test set, Random Forest still looked like the top choice. However, its edge over Logistic Regression was small, so cross-validation and an overfitting check should be done before making a final decision.

Table 4 Per-College Predictive Performance of the Random Forest Model

College	Precision	Recall	F1-Score	Support
College of Allied Health Sciences	0.8696	0.7216	0.7887	194
College of Arts and Sciences	0.7874	0.7697	0.7784	178
College of Business Education	0.7686	0.8650	0.8139	311
College of Engineering, Architecture and Technology	0.8292	0.8257	0.8274	241
College of International Hospitality and Tourism Management	0.8351	0.8177	0.8263	192
College of Maritime Education	0.8558	0.8241	0.8396	108
College of Teacher Education	0.7650	0.7955	0.7799	176

Table 4 indicates that the Random Forest model stayed effective for all seven colleges. Still, the results were not the same across categories. The College of Maritime Education had the top F1-score at 0.8396, which reflects a good tradeoff between precision and recall. The College of Allied Health Sciences showed the highest precision at 0.8696, so its predictions for this college were often correct. Even so, its recall was 0.7216, meaning the model failed to catch a larger share of true Allied Health students. On the other hand, the College of Business Education reached the best recall at 0.8650. Its precision was weaker at 0.7686. This points to many correctly found COBE students, but also more cases where the model labeled students as COBE when they were not. Engineering, Architecture and Technology and International Hospitality and Tourism Management both had steady F1-scores above 0.82. In general, the findings suggest the Random Forest model worked in a consistent way across colleges. It did not depend on one standout category. This supports using it for multiclass college recommendation.

Table 5 Confusion Matrix of the Random Forest Model for College Recommendation

Actual College	CAHS	CAS	COBE	CEAT	CIHTM	CME	CTE	Total
College of Allied Health Sciences (CAHS)	140	5	21	14	3	4	7	194
College of Arts and Sciences (CAS)	3	137	21	4	1	2	10	178
College of Business Education (COBE)	4	6	269	5	13	2	12	311
College of Engineering, Architecture and Technology (CEAT)	4	7	8	199	9	4	10	241
College of International Hospitality and Tourism Management (CIHTM)	1	5	13	11	157	3	2	192
College of Maritime Education (COME)	1	6	4	5	1	89	2	108
College of Teacher Education (CTE)	8	8	14	2	4	0	140	176
Total Predicted	161	174	350	240	188	104	183	1,400

Table 5 indicates that the Random Forest system labeled 1,131 out of 1,400 students correctly, which is 80.79%. The remaining 269 students were labeled incorrectly. For the COBE group, 269 of 311 students were classified correctly. CEAT had 199 correct out of 241. CIHTM showed 157 correct out of 192. COME resulted in 89 correct predictions out of 108. CAHS had 140 correct out of 194. These counts suggest lower sensitivity than the other groups. A clear pattern of mistakes involved COBE. The model placed 350 students into COBE, even though 311 students were actually COBE. Within those errors, 21 students from CAHS and 21 students from CAS were put into COBE. Even with these issues, most cases formed a strong diagonal in the matrix. This indicates the model separated most students across the seven colleges. Overall, the confusion matrix backs the 80.79% accuracy for Random Forest and also points to specific college pairs where errors happened.

Table 6 Comparison of Weighted F1-Scores of Machine Learning Algorithms for College Recommendation

Rank	Machine Learning Algorithm	Weighted F1-Score
1	Random Forest	0.8077
2	Multinomial Logistic Regression	0.8070
3	Gradient Boosting	0.8063
4	XGBoost	0.8048
5	LightGBM	0.7969
6	Decision Tree	0.6634

Note. The weighted F1-score represents the harmonic mean of precision and recall while accounting for the number of observations in each college category. Higher values indicate better overall multiclass classification performance. Bold values indicate the highest-performing algorithm.

Table 6 lists the weighted F1-scores. Random Forest led the set with 0.8077. Right behind it was Multinomial Logistic Regression at 0.8070. Gradient Boosting reached 0.8063, and XGBoost scored 0.8048. The gaps among these top four models were small. They seemed to trade off precision and recall in a similar way across the seven college categories. LightGBM was a bit lower at 0.7969. Even so, it still showed reasonable predictive performance. Decision Tree finished last with 0.6634. That result points to weaker generalization. It also suggests it struggled with the more complex patterns in the predictors. Taken together, the outcomes support ensemble methods, especially Random Forest, for college recommendations. Still, since Random Forest and Logistic Regression were close, extra checks like cross-validation and overfitting review are still needed before choosing the final model.

Table 7 Five-Fold Cross-Validation Performance of Machine Learning Algorithms for College Recommendation

Algorithm	CV Accuracy	CV Precision	CV Recall	CV F1-Score	CV F1 SD
Multinomial Logistic Regression	0.7914	0.7934	0.7914	0.7913	0.0105
XGBoost	0.7889	0.7903	0.7889	0.7887	0.0122
Random Forest	0.7880	0.7902	0.7880	0.7879	0.0101
Gradient Boosting	0.7877	0.7898	0.7877	0.7876	0.0110
LightGBM	0.7813	0.7829	0.7813	0.7811	0.0104
Decision Tree	0.6469	0.6474	0.6469	0.6468	0.0092

Note. CV = cross-validation; SD = standard deviation. Higher accuracy, precision, recall, and F1-score indicate better predictive performance. A lower CV F1 SD indicates greater stability across folds.

Table 7 indicates that Multinomial Logistic Regression led the five-fold cross-validation results. It reached the top CV accuracy of 0.7914, with CV precision of 0.7934, CV recall of 0.7914, and an F1-score of 0.7913. XGBoost came in next. Its CV F1-score was 0.7887. Random Forest followed with 0.7879, and Gradient Boosting was close behind at 0.7876. LightGBM was lower at 0.7811. Decision Tree was much weaker, scoring 0.6468 on CV. The standard deviations were fairly small, from 0.0092 to 0.0122. That range points to steady results across the five folds. Also, Table 3 shows a different outcome on the test set. Random Forest had the best test-set F1-score at 0.8077. Yet Logistic Regression still had the best cross-validation F1-score at 0.7913. So the model choice should use both test results and cross-validation results. Relying only on Random Forest's test-set gain would not be the best approach.

Table 8 Overall Test-Set and Cross-Validation Performance of Machine Learning Algorithms for College Recommendation

Algorithm	Accuracy	Precision	Recall	F1-Score	CV Accuracy	CV Precision	CV Recall	CV F1-Score	CV F1 SD
Multinomial Logistic Regression	0.8071	0.8102	0.8071	0.8070	0.7914	0.7934	0.7914	0.7913	0.0105
XGBoost	0.8050	0.8071	0.8050	0.8048	0.7889	0.7903	0.7889	0.7887	0.0122
Random Forest	0.8079	0.8108	0.8079	0.8077	0.7880	0.7902	0.7880	0.7879	0.0101
Gradient Boosting	0.8064	0.8094	0.8064	0.8063	0.7877	0.7898	0.7877	0.7876	0.0110
LightGBM	0.7971	0.7991	0.7971	0.7969	0.7813	0.7829	0.7813	0.7811	0.0104
Decision Tree	0.6636	0.6639	0.6636	0.6634	0.6469	0.6474	0.6469	0.6468	0.0092

Table 8 indicates Random Forest led on the test set. Its accuracy was 0.8079 and its F1-score was 0.8077. Multinomial Logistic Regression did better under cross-validation. The CV accuracy was 0.7914. Precision was 0.7934. Recall was 0.7914. The CV F1-score was 0.7913. The spread for CV F1 was 0.0105. That suggests the scores stayed fairly steady across the five folds. After those two models, XGBoost, Random Forest, and Gradient Boosting were close behind. Their CV F1-scores were 0.7887, 0.7879, and 0.7876. The Decision Tree model was the lowest performer overall. In a direct comparison, Random Forest beat Logistic Regression by 0.0007 on the test F1-score. On the other hand, Logistic Regression beat Random Forest by 0.0034 on the CV F1-score. Taking both the holdout results and the repeated validation into account, Multinomial Logistic Regression offers the better case for generalization. It also fits as a more defensible final choice, even though Random Forest has a slightly higher single test-set score.

Table 9 Training and Testing Performance of Machine Learning Algorithms

Algorithm	Train Accuracy	Test Accuracy	Accuracy Gap	Train F1	Test F1	F1 Gap	Interpretation
Random Forest	1.0000	0.8079	0.1921	1.0000	0.8077	0.1923	Substantial overfitting

Algorithm	Train Accuracy	Test Accuracy	Accuracy Gap	Train F1	Test F1	F1 Gap	Interpretation
Multinomial Logistic Regression	0.7914	0.8071	-0.0157	0.7915	0.8070	-0.0155	Good generalization / no evident overfitting
Gradient Boosting	0.8198	0.8064	0.0134	0.8199	0.8063	0.0136	Good generalization / minimal overfitting
XGBoost	0.9221	0.8050	0.1171	0.9222	0.8048	0.1174	Moderate to substantial overfitting
LightGBM	1.0000	0.7971	0.2029	1.0000	0.7969	0.2031	Substantial overfitting
Decision Tree	1.0000	0.6636	0.3364	1.0000	0.6634	0.3366	Severe overfitting

Note. Positive gaps indicate that training performance exceeded testing performance and may signal overfitting. Larger gaps indicate stronger evidence of overfitting. A negative gap means the test set slightly outperformed the training set and does not, by itself, indicate a problem.

Table 9 made the generalization differences clear. Multinomial Logistic Regression performed best on new data. Its test accuracy and test F1 were a bit above its training values. The gaps were -0.0157 for accuracy and -0.0155 for F1. Gradient Boosting also held up well. It showed small positive gaps, 0.0134 for accuracy and 0.0136 for F1. Random Forest was different. It scored perfectly on training, yet it fell to 0.8079 on test accuracy. Test F1 was 0.8077. This points to strong overfitting. XGBoost showed a clear drop from training to testing. LightGBM and Decision Tree struggled more. Decision Tree had the largest issue, with an F1 gap of 0.3366. Overall, the results support Multinomial Logistic Regression as the more generalizable model. This also matches its strong cross validation results.

Table 10 Top 10 Predictors of College Recommendation Based on SHAP Analysis

Rank	Predictor	Mean Absolute SHAP Value
1	SHS Strand – ABM	0.5523
2	SHS Strand – STEM	0.4969
3	SHS Strand – TVL	0.4804
4	SHS Strand – HUMSS	0.4668
5	SHS Strand – GAS	0.3182
6	Admission Score in English	0.1994
7	Admission Score in Science	0.1497
8	Admission Score in Mathematics	0.1051
9	Admission Interview Score	0.0545
10	Monthly Family Income Bracket – ₱40,000–₱79,999	0.0534

Note. SHAP = SHapley Additive exPlanations; SHS = Senior High School; ABM = Accountancy, Business and Management; STEM = Science, Technology, Engineering and Mathematics; TVL = Technical-Vocational-Livelihood; HUMSS = Humanities and Social Sciences; GAS = General Academic Strand. Mean absolute SHAP values measure the average magnitude of each feature's contribution to model predictions. Higher values indicate greater predictive influence; they do not indicate whether a feature increases or decreases the probability of a particular college recommendation.

Table 10 indicates that the SHS strand provided the strongest predictive signal for college recommendation. ABM had the top mean absolute SHAP value at 0.5523. Next came STEM with 0.4969. After that were TVL at 0.4804 and HUMSS at 0.4668. GAS was much lower at 0.3182. Since all five strands appeared prominent, students' past academic track likely shaped how the model separated the colleges. For the continuous academic inputs, the Admission English Score was first at 0.1994. Science ranked next with 0.1497. Mathematics followed at 0.1051. This pattern suggests that admission test results also mattered for the recommendations. Admission

Interview Score added less at 0.0545. The ₱40,000–₱79,999 family income category was lower still, at 0.0534. The results in Table 10 also place that income category near the tenth position. In general, the SHAP output points to a model that depends most on SHS background and admission performance for its recommendations. At the same time, these SHAP numbers reflect predictive importance, not causality, and they do not show the direction of influence.

Discussion

The results indicate that educational machine learning can act as a decision aid for suggesting colleges to students entering university. The models did not all perform the same. Still, the best methods tended to label about 80% of students correctly. Random Forest led on the holdout set, with accuracy of 0.8079 and an F1-score of 0.8077. Multinomial Logistic Regression was close. Its test results were nearly the same, and it also had the top value in the five-fold cross-validation F1-score, reported as 0.7913. Logistic Regression did not show clear signs of overfitting. Random Forest reached a perfect score on the training data, but on the test set its F1-score dropped to 0.8077. For that reason, the overall pattern points to Multinomial Logistic Regression as the better choice when look at generalization, steadiness, and predictive strength at the same time.

These findings also match earlier work on educational recommendations using machine learning. Alghamdi et al. (2021) reported that supervised learning can help with undergraduate major selection. They found that prior study background and entry test results were strong predictors. Delahoz-Domínguez et al. (2024) similarly argued for using both academic indicators and contextual variables when suggesting university degrees.

In this study, the contribution is that the model predicted enrollment among seven colleges. It used a mix of demographic details, academic records, admission data, and career-related information. The results were fairly even across the colleges. That supports the idea that performance was not tied only to one college group. Overall, the outcomes suggest that recommender systems can support academic advising. They can help complement staff decisions, not take their place.

The SHAP results shed light on why the model produced certain recommendations. Among the factors, SHS strand carried the most weight. The top values were ABM (0.5523), STEM (0.4969), TVL (0.4804), HUMSS (0.4668), and GAS (0.3182). After that, the admission scores for English, Science, and Mathematics came in next. This setup fits Person–Environment Fit Theory. The idea is that better outcomes are more likely when a person’s traits match what an environment needs. Here, SHS specialization seems to reflect what students had learned before college. At the same time, admission performance points to academic readiness. So, the model likely learned how student study backgrounds relate to where they end up in college.

Even so, the results should be used carefully by the institution. SHAP importance shows predictive contribution, not direct cause. Also, the high impact of the SHS strand should not be used as a reason to automatically limit students to colleges that are usually linked to specific strands. Khosravi et al. (2022) stressed that AI-supported decisions should be easy to understand. Baker and Hawn (2022) noted that education algorithms may repeat the trends and unfairness already present in past data. Because of that, the system should work as an advisory tool that can be explained. It should help counselors, students, and administrators make better informed choices. Oversight by humans still matters. Student agency also matters, especially since these suggestions can shape educational access and later career options.

Implications and Future Research Directions

The findings indicate that educational machine learning can strengthen data-driven college recommendation and academic advising. The framework can help Northwestern University identify suitable college pathways while providing explainable recommendations. However, administrators and advisers should use model outputs as decision-support information alongside professional judgment, student preferences, and individual circumstances.

Next steps should test the framework with data from more than one university. It should also include different student cohorts. Long term work is also needed. Researchers should check whether the suggested placements connect to better outcomes. Outcomes to check include grades, retention, satisfaction, graduation, and career alignment. Future work should also look at fairness in how the algorithm performs across socioeconomic and demographic groups. It should compare other explainable AI approaches with SHAP.

Limitations

The study has limits, and these points should shape how to read the results and how to plan next work on the educational machine learning framework. First, the data came from historical institutional records. It covered 7,000 new students at Northwestern University from 2020 to 2024. Because of that, the findings might not hold up at other schools. Admission rules, student groups, and academic layouts can differ a lot. Second, the models used demographic data, academic data, admission data, and career related variables. Even so, the study did not add other factors that could matter. Examples include personality traits, learning preferences, parent expectations, money issues, and labor market conditions. Those omissions may affect what the models learn. Third, the study judged predictive ability by using historical college placement as the target. That choice can capture earlier institutional habits and student selection patterns. It might not show what an ideal college fit would look like on its own.

Conclusions

This work showed how educational machine learning could help Northwestern University match incoming students to colleges and career paths. The dataset used institutional records from about 7,000 students. The results suggested that demographic details, grades, admission factors, and career-related variables can help forecast suitable college placement. Six models were tested. Random Forest performed best on the holdout set. It reached 80.79% accuracy and an F1-score of 80.77%. Still, Multinomial Logistic Regression looked better on cross-validation. It scored 79.13% on F1 and did not show a clear overfitting pattern. Random Forest, LightGBM, and Decision Tree also showed big gaps between training and test results. Because of that, the study favored Multinomial Logistic Regression as the model that is more stable for use at the institution.

A SHAP review was added to make the model easier to interpret. The SHS strand was the top driver of predictions. Admission exam scores in English, Science, and Mathematics ranked next. The study read this as evidence that earlier educational track choices and academic readiness help separate students into better-aligned college routes. At the same time, the paper warned that SHAP scores reflect predictive value, not cause. They should not be treated like fixed rules for placement. In the end, the study argued that machine learning can support evidence-based advising. It can convert existing institutional data into recommendations that can be explained. Northwestern can use the proposed setup as decision support, not as an automatic decision tool. If predictive analytics are paired with advising staff input, student choices, and human judgment, the process can stay clearer and more centered on students.

Recommendations

Northwestern University should look at ways to embed educational machine learning into advising. The goal is to support decisions about college fit and career steps. This should happen without taking away professional judgment or student choice.

- 1) Create an AI-assisted college recommendation policy. Northwestern should write a clear policy for how predictive analytics may be used in admissions and advising. The tool may suggest college pathways. It must not make final calls on admission, placement, or enrollment.
- 2) Use a human-in-the-loop decision method. Advisers and guidance counselors should review model results with each student. The last decision should reflect model output plus the student's goals, interests, academic background, and stated preferences.
- 3) Start with the most supported model. The University should use Multinomial Logistic Regression at first. It showed the best cross-validation results and no clear overfitting in the study. The University should retrain and recheck the model when new cohorts are ready.
- 4) Build explainability into the process. The University should add SHAP or a similar method so advisers can see what drove each recommendation. Since the SHS strand was very influential, the University should bar using strand as an automatic limit on a student's college options.
- 5) Improve institutional data governance. The Registrar's Office, Guidance Office, Data Protection Officer, and the right academic units should set rules for data access and handling. This should cover anonymization, how long data stays, how to correct or delete data, how models are built, and who can use what. These steps should align with institutional policies and the Data Privacy Act of 2012.
- 6) Run annual model audits. Each year the University should check predictive accuracy and related metrics such as precision, recall, and F1-score. It should also review cross-validation results and results across colleges before the tool continues. It should test for gaps across key student groups to lower the chance of algorithmic bias.

- 7) Apply the framework in advising, not as a label. The University should use recommendations to prompt structured talks between students and advisers. The point is to discuss options, not just assign a predicted college. When student interests do not match the model, advisers should outline other pathways.
- 8) Set up oversight and accountability. Northwestern should form an interdisciplinary committee. It should include academic affairs, admissions, guidance, the Registrar, IT, research, and data-protection staff. This group should manage model validation, ethical checks, system updates, complaints, and rollout of the policy.

References

1. Albreiki, B., Rustamov, Z., Rustamov, J., & Zaki, N. (2024). Enhancing predictive performance in identifying at-risk students: Integration of topological features, node embeddings in machine learning models. In M. S. Khine (Ed.), *Machine learning in educational sciences: Approaches, applications and advances* (pp. 183–205). Springer. https://doi.org/10.1007/978-981-99-9379-6_10
2. Alghamdi, A. A., Alsubait, T. M., Alhakami, H., & Baz, A. (2021). Selection of the right undergraduate major by students using supervised learning techniques. *Applied Sciences*, 11(22), 10639. <https://doi.org/10.3390/app112210639>
3. Atalla, S., Al-Mousa, A., Al-Safadi, L., & Al-Khalifa, H. (2023). An intelligent recommendation system for automating academic advising based on curriculum analytics and performance modeling. *Mathematics*, 11(5), 1098. <https://doi.org/10.3390/math11051098>
4. Baker, R. S., & Hawn, A. (2022). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education*, 32, 1052–1092. <https://doi.org/10.1007/s40593-021-00285-9>
5. Bird, K. (2023). Predictive analytics in higher education: The promises and challenges of using machine learning to improve student success. *The AIR Professional File*, 161. <https://doi.org/10.34315/apf1612023>
6. Bognár, L. (2024). Predicting student attrition in university courses. In M. S. Khine (Ed.), *Machine learning in educational sciences: Approaches, applications and advances* (pp. 129–157). Springer. https://doi.org/10.1007/978-981-99-9379-6_8
7. Cha, S., Loeser, M., & Seo, K. (2024). The impact of AI-based course-recommender system on students' course-selection decision-making process. *Applied Sciences*, 14(9), 3672. <https://doi.org/10.3390/app14093672>
8. Chen, C.-H., Wang, J.-H., Chen, B., & Liao, C.-W. (2021). Predicting at-risk university students based on their e-book reading behaviours by using machine learning classifiers. *Australasian Journal of Educational Technology*, 37(4), 130–144. <https://doi.org/10.14742/ajet.6116>
9. Crowley, T. (1992). Computer-Aided Careers Guidance: An Investigation Involving an Artificial System. *British Journal of Guidance & Counselling*, 20(3), 344–351. <https://doi.org/10.1080/03069889208253631>
10. Delahoz-Domínguez, E. J., Zarthia-Sossa, J. W., & Montes-Hincapié, J. M. (2024). Recommender system for university degree selection. *Applied Sciences*, 14(18), 8311. <https://doi.org/10.3390/app14188311>
11. Delen, D., Davazdahemami, B., & Dezfouli, E. R. (2023). Predicting and mitigating freshmen student attrition: A local-explainable machine learning framework. *Information Systems Frontiers*. <https://doi.org/10.1007/s10796-023-10397-3>
12. DiStefano, C., Odero, G., & Starrett, A. (2024). An introduction to machine learning for educational researchers. In M. S. Khine (Ed.), *Machine learning in educational sciences: Approaches, applications and advances* (pp. 11–30). Springer. https://doi.org/10.1007/978-981-99-9379-6_2
13. Farrow, R. (2023). The possibilities and limits of explicable artificial intelligence in education. *Learning, Media and Technology*, 48(2), 266–279. <https://doi.org/10.1080/17439884.2023.2185630>
14. Gedrimiene, E., Silvola, A., Pursiainen, J., Rusanen, J., & Muukkonen, H. (2024). Artificial intelligence-enhanced learning analytics for supporting career decisions: Advantages and challenges from the user perspective. *Education and Information Technologies*, 29, 12725–12751. <https://doi.org/10.1007/s10639-023-12277-4>
15. Goren, O., Cohen, L., & Rubinstein, A. (2024). Early prediction of student dropout in higher education using machine learning models. In B. Paaßen & C. D. Epp (Eds.), *Proceedings of the 17th International Conference on Educational Data Mining* (pp. 349–359). International Educational Data Mining Society. <https://doi.org/10.5281/zenodo.12729834>

16. Guruge, D. B., Kadel, R., & Halder, S. J. (2021). The state of the art in methodologies of course recommender systems: A review of recent research. *Data*, 6(2), 18. <https://doi.org/10.3390/data6020018>
17. Kim, S., Choi, J., & Lee, Y. (2023). Student dropout prediction for university with high precision and recall. *Applied Sciences*, 13(10), 6275. <https://doi.org/10.3390/app13106275>
18. Kim, T.-Y., & Lim, J. (2023). Developing an intelligent recommendation system for non-information and communications technology major university students. *Applied Sciences*, 13(23), 12774. <https://doi.org/10.3390/app132312774>
19. Klačnja-Milićević, A., Ivanović, M., & Nanopoulos, A. (2015). Recommender systems in e-learning environments: A survey of the state-of-the-art and possible extensions. *Artificial Intelligence Review*, 44, 571–604. <https://doi.org/10.1007/s10462-015-9440-z>
20. Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 3, 100074. <https://doi.org/10.1016/j.caeai.2022.100074>
21. Kok, C. L., Ho, C. K., Chen, L., Koh, Y. Y., & Tian, B. (2024). A novel predictive modeling for student attrition utilizing machine learning and sustainable big data analytics. *Applied Sciences*, 14(21), 9633. <https://doi.org/10.3390/app14219633>
22. Li, Q., & Kim, J. (2021). A deep learning-based course recommender system for sustainable development in education. *Applied Sciences*, 11(19), 8993. <https://doi.org/10.3390/app11198993>
23. Ma, B., Lu, M., Taniguchi, Y., & Konomi, S. (2021). CourseQ: The impact of visual and interactive course recommendation in university environments. *Research and Practice in Technology Enhanced Learning*, 16, Article 18. <https://doi.org/10.1186/s41039-021-00167-7>
24. Pallipuram, V., Mammadov, G., Ho, B., Dhulipala, M., & Dziallas, S. (2024). Camelot: A council of machine learning strategies to enhance teaching. In M. S. Khine (Ed.), *Machine learning in educational sciences: Approaches, applications and advances* (pp. 47–70). Springer. https://doi.org/10.1007/978-981-99-9379-6_4
25. Seidel, E., & Kutieleh, S. (2017). Using predictive analytics to target and improve first year student attrition. *Australian Journal of Education*, 61(2), 200–218. <https://doi.org/10.1177/0004944117712310>
26. Sekeroglu, B., Abiyev, R., Ilhan, A., Arslan, M., & Idoko, J. B. (2021). Systematic Literature Review on Machine Learning and Student Performance Prediction: Critical Gaps and Possible Remedies. *Applied Sciences*, 11(22), 10907. <https://doi.org/10.3390/app112210907>
27. Urdaneta-Ponte, M. C., Méndez-Zorrilla, A., & Oleagordia-Ruiz, I. (2021). Recommendation systems for education: Systematic review. *Electronics*, 10(14), 1611. <https://doi.org/10.3390/electronics10141611>
28. Wu, B., & Liu, L. (2023). Personalized hybrid recommendation algorithm for MOOCs based on learners' dynamic preferences and multidimensional capabilities. *Applied Sciences*, 13(9), 5548. <https://doi.org/10.3390/app13095548>
29. Yağcı, M. (2022). Educational data mining: Prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments*, 9, Article 11. <https://doi.org/10.1186/s40561-022-00192-z>