Predictive Effects of Core Subject Grades on Senior High School Strand Selection

Gerlie D. Bornea., Joeteddy B. Bugarin

College of Development and Management, University of Southeastern Philippines

DOI: https://dx.doi.org/10.47772/IJRISS.2026.10100222

Received: 24 December 2025; Accepted: 29 December 2025; Published: 31 January 2026

ABSTRACT

This study examined whether Grade 10 core-subject grades (English, Mathematics, and Science) can be used to provide a preliminary, grades-only screening of Senior High School (SHS) strand options (ABM, HUMSS, STEM, and TVL). Using multinomial logistic regression on 400 student records, English emerged as the strongest predictor, with Mathematics and Science contributing more modestly in selected comparisons. In classification, the grades-only model correctly classified 215 of 400 cases (overall accuracy = 53.75%). Preliminary grades-only screening achieved an overall accuracy of 53.75%, constrained by grade profile overlap across strands. Strand-level performance varied substantially (ABM = 17.4%, HUMSS = 49.2%, STEM = 75.0%, TVL = 74.6%), indicating limited separability for strands with similar grade patterns. These results support the use of grades as an initial filter to narrow counseling conversations, but not as a stand-alone placement or decision-making tool without complementary measures such as interest inventories, aptitude assessments, and structured guidance inputs.

Keywords: Multinomial logistic regression, Senior High School ABM (Accountancy, Business and Management Strand), HUMSS (Humanities and Social Sciences), STEM (Science, Technology, Engineering and Mathematics), TVL (Technical-Vocational Livelihood track), Predictive Model

INTRODUCTION

The United Nations Sustainable Development Goal 4 (SDG4) highlights the importance of equitable, quality, and inclusive education. The attainment of these goals is influenced by interrelated components, such that quality education is built on an input-process-output framework (Pärli, 2023). This study aims to advance this goal by providing a mechanism that assists students during the critical decision-making stage, which may significantly influence their future. These studies look at how students' grades in core subjects like English, Math, and Science affect their choice of Senior High School strand. The studies will not only give ideas about this relationship but will also help in making a predictive tool that could help the school with guidance, counseling, and student advising.

Deficiency of such schools' support system, especially in choosing an SHS track, exacerbates misalignment that could eventually lead to career disillusionment. Sarona-Pedro (2025) explained that this happens because students often rely on peer influence rather than structured aptitude assessments. Addressing this issue will reduce global concerns regarding the misalignment of learning outcomes, essential competencies, and student pathways (UNESCO, 2025).

In the Philippines, the revised curriculum is one of the major changes made by the Department of Education. It focuses on making content delivery more efficient, expanding elective course options, ensuring closer alignment with industry standards, and helping students better prepare for higher education, starting their own businesses, and getting a job. This revised curriculum puts forth a "doorway option" where students are given autonomy to take electives of their choice, rather than being trapped in a predetermined career path (Department of Education, 2025). However, this may pose additional challenges, especially in school logistics, particularly in distributing teacher workloads, a lack of specialization for new subjects, difficulty in organizing class schedules, and efficiently managing constrained resources (Ignacio & Bajet, 2025; Adarlo & Jackson, 2017; Alinea et al., 2024;

Acosta & S. Acosta, 2016; EDCOM 2 Communications, 2025; Zhang et al., 2025). Moreover, this "doorway option" could also lead to pressure-induced strand selection. Granting students more autonomy may lead to poor choices that are influenced by their friends rather than their own ability, which might result in failure, wasted time, and career dissatisfaction (Malipot, 2025; Dangoy & Madrigal, 2020). Several studies discussed the negative impact of being compelled to pressure-induced strand selection, like blockage of motivation, dissatisfaction, adaptation challenges, and potentially hindering a student's overall success in life and future career path (Sadjail et al., 2022; Shankhdhar et al., 2020; Sidek et al., 2023).

Proactive planning can mitigate the likelihood of setbacks. Therefore, forecasting the probability of students selecting a specific strand based on their academic achievements in Junior High School may help guide students and reduce the risk of shortages. Holland`s theory of career choice serves as the guiding principle of the study, which posits that students tend to choose their educational and career choices because they are aligned with their personal characteristics and capabilities (Adlya & Zola, 2022). It suggests that students naturally gravitate towards subjects and strands that align with their personality and academic strengths(Desvikayati et al., 2025).

Furthermore, this study seeks to address that gap by developing a predictive model that could assist in forecasting students' choices and guide them in making thoughtful decisions. According to Bejar (2024), predictive analytics tools are both an opportunity and a necessity in the educational realm. In the current context, it assists the schools in making plans, such as preparing and training teachers for high-demand subjects, allocating classroom and learning resources effectively, and directing students toward electives that align with their strengths. Previous studies revealed the functionality, dependability, and accuracy of predictive analytics in identifying predictors, understanding schooling outcomes, and making informed decisions (Vargas, 2024; Balontong, 2024; Yu & Upah, 2018; Borghosh, 2024). Currently, since the program has not been implemented, no predictive studies have been conducted to forecast students' elective choices in the new senior high school curriculum, especially in a local context.

Specifically, this study aims to answer the following research questions:

1. To what extent do English, Math, and Science grades in Grade 10 predict the likelihood of students choosing specific Senior High School strands?

2. Which subject (English, Math, or Science) serves as the strongest predictor of students' elective or strand choices?

3. What predictive model can be developed to determine the likelihood of students choosing a specific strand?

Null Hypothesis (Ho):

Ho_1: Grade 10 academic performance in Mathematics, Science, and English does not significantly predict the likelihood of students choosing a specific senior high school strand.

Ho_2: Neither Mathematics, Science, nor English grades predicts the likelihood of students choosing a specific senior high school strand.

Ho_3: There is no predictive model that could be developed to determine the likelihood of students choosing a specific strand

Conceptual Diagram

It is shown in the conceptual diagram that the independent variables, which are grades in English, Mathematics, and Science, are being used to investigate whether or not they have an impact on the probability of students selecting a Senior High School strands Multinomial Logistic Regression program. Following that, the findings will serve as the foundation for the creation of a prediction model, which will be used to estimate or anticipate the outcomes of possible strand selections.

INTERNATIONAL JOURNAL OF RESEARCH AND INNOVATION IN SOCIAL SCIENCE (IJRISS)
ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

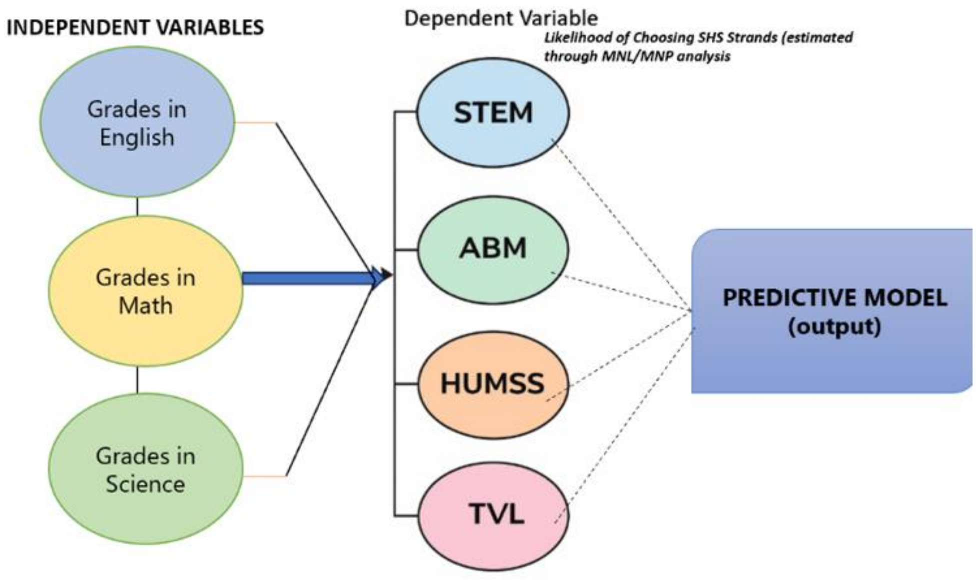

Figure 1. Conceptual diagram of the influence of core subject grades on strand selection.

METHODOLOGY

Research Design

This study employed a quantitative-predictive research design, specifically employing multinomial logistic regression. The use of multinomial logistic regression was justified since the dependent variable, strand choice, is categorical with four nominal outcomes (ABM, HUMSS, STEM, and TVL).

The design was suitable, as the main aim of the study was to assess the degree to which students' academic performance in English, Mathematics, and Science forecasts their strand selection in Senior High School (Lee, 2025). The predictive design enables the researcher to analyze the impact of independent variables (grades in English, Math, and Science) on the probability of a student choosing a particular Senior High School track—specifically, Academic (ABM, HUMSS, STEM) or Technical-Vocational-Livelihood (TVL). The model provides predictive and inferential insights, enabling the computation of odds ratios and anticipated probabilities for evaluating classification accuracy.

Research Instrument

The study used documentary data obtained from students' official school records, particularly their final grades in English, Mathematics, and Science. These grades were the independent variables for the regression analysis. The strand in which each student enrolled was the dependent variable. The use of the official school records ensured the objectivity, reliability, and validity of data, since they were checked by the school registrar and class advisers. There will be no survey questionnaires or interviews since the study requires only secondary data and is purely quantitative. Nonetheless, this technique ensured minimum measurement bias and maximum accuracy in capturing the academic performance.

Data Sources, Sample Size, and Sampling Technique

This study used purposive sampling to select the students' records for analysis. Purposive sampling was considered suitable because the study was limited to respondents who had full and valid academic records in English, Mathematics, and Science in Grade 10 with documented Senior High School strand choices. This process ensured that only relevant and analyzable data got included in the predictive modeling. A sample of 400 student records was taken from the population for data collection. The sample size was intentionally selected to be larger than the minimum required to increase the precision and reliability of the results. Selecting a larger sample helped to increase the statistical power of the model, improve the stability of parameter estimates, and decrease sampling error to increase the reliability of the results Clinical Research Centre, Sarawak General

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

Hospital, Ministry of Health, Kuching, Malaysia et al., 2018). The criteria in choosing the 400 respondents included: availability of complete final grades in English, Mathematics, and Science; having a clear record of the selected Senior High School strand (ABM, HUMSS, STEM, or TVL); and being enrolled in the same school year of the study. Records that were incomplete, inconsistent, or had no strand information were excluded from the sample. This procedure ensured that the final sample would be representative of the target population and that the sample was suitable for predictive modeling.

Data Gathering Procedure

1. Written permission was secured from the school administration to access student records
2. Academic grades and SHS strand choices were extracted from the school file and database, respectively.
3. Records were anonymized using unique respondent codes to ensure confidentiality.

Data Preparation

1. Perform descriptive analysis to have
2. Run Multinomial Logistic Regression and cross-validate it with Multinomial Probit Regression
3. Compare the Akaike Information criterion (AIC) and determine which has the lowest value. The result with the lowest value will be used in the analysis of the study
4. Test the Independence of Irrelevant Alternatives (IIA) assumption (Hausman–McFadden, Small–Hsiao) to know if MNL holds the assumption of IIA or has a p-value greater than p>.0
5. Interpretation of Coefficients and Marginal Effects.
6. Classification accuracy check
7. Develop the predictive model equation

Ethical Consideration

This study is guided by the Data Privacy Act 2012 and anonymity. Students' names are coded to ensure confidentiality. Before data collection, the researcher obtained approval from the school administration to access school files.

RESULTS

Descriptive Statistics of Study Variables

Table 1 presents the descriptive analysis of strand selection among Senior High School students. The Technical-Vocational-Livelihood (TVL) strand is the most frequently selected strand, with 126 students (31.5%), followed by Humanities and Social Sciences with 122 students (30.5%), Accountancy, Business and Management (ABM) with 92 students (23.0%), and Science, Technology, Engineering and Mathematics (STEM) with 60 students (15.0%).

Table 1. Summary of Strand Choice

Strand Choice	Frequency	Percentage (%)
ABM	92	23.0
HUMSS	122	30.5

STEM	60	15.0
TVL	126	31.5
Total	400	100%

Note. Percentages (%) are based on the total sample size

Similarly, Table 2 presents the descriptive results of grades in the three core subjects. It shows that English grades had a mean of 84.56 and a standard deviation of 5.137, while the Math grades had a mean of 86.37 and a standard deviation of 4.590. Science grades had a mean of 84.68 and a standard deviation of 5.553. This strong academic performance of students was fairly consistent in three core subjects.

Table 2. Descriptive Statistics for Predictor Variables: Grade 10 Core Subject Grades (N = 400)

Predictors	**N**	**Mean**	**Standard Deviation**
English	400	84.67	5.137
Math	400	86.37	4.590
Science	400	84.68	5.553

Note. **N**= total sample size. Values reflect Grade 10 grades in English, Math, and Science

Model Comparison and Selection

The Akaike information criterion was calculated and evaluated to identify the most suitable model for the data that would guarantee strength and reliability. The model with the lowest AIC value will be favored. Table 3 presents the results of the Akaike Information Criterion value. Multinomial Probit has an AIC value of 833.83, which is higher than the multinomial logit model, which obtained the lowest AIC value (797.254). This suggests that the multinomial logit model is a superior model fit to be used for the study.

Table 3.

Comparison of Model Fit Indices Using AIC

Model	BIC	AIC	Conclusion
Model 1: Multinomial Regression (Probit)	**881.7297**	**833.8321**	**Higher AIC – rejected**
Model 2: Multinomial Regression (Logit)	**845.152**	**797.254**	**Lower AIC – preferred**

Note: **BIC**=Bayesian Information Criterion**, AIC** =Akaike`s information criterion.
Lower AIC and BIC indicate a better-fitting and more parsimonious model.

Independence of Irrelevant Alternatives (IIA)

The Independence of Irrelevance Alternatives (IIA) using the Hausman-McFadden test was conducted to test whether the multinomial logit is appropriate for the structure of the variable. The result of the Hausman test, as shown in Table 4, the multinomial logit model holds the assumption of IIA assumption, with indicates that all categories have p-value > 0.05 (Outcome 1: $\chi^2(6) = -0.60$, p = 1.000; Outcome 2: $\chi^2(6) = 11.84$, p = .066; Outcome 3: $\chi^2(7) = -0.85$, p = 1.000; Outcome 4: $\chi^2(7) = -116.54$, p = 1.000). Given that all p-values exceeded

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

the .05 threshold, the null hypothesis of IIA holding was accepted. Thus, the multinomial logit model is appropriate for the data.

Table 4. Hausman-McFadden tests for Independence of irrelevant alternatives

Removed alternative	**Test type**	**χ^2**	**df**	**p>chi^2**	**Decision**
ABM	Hausman–McFadden	-0.600	6	1.000	Fail to reject IIA
HUMSS	Hausman–McFadden	11.841	6	0.066	Fail to reject IIA
STEM	Hausman–McFadden	-0.845	7	1.000	Fail to reject IIA
TVL	Hausman–McFadden	-116.542	7	1.000	Fail to reject IIA

Note: p-values greater than .05 indicate no violation of IIA. Negative χ^2 values reported by Stata default to p = 1.000.

Model Fit and Predictor of Significance (MNL)

Table 6 shows the Likelihood Ratio Test. It revealed that the predictors (English, Math, and Science grades) are significant predictors and contributed improvement of the model compared to a single intercept model ($\chi^2(9)$ = 271.110, p <.001). This means that at least one predictor contributes to the significance of the model fit in predicting strand choice based on the Grade 10 grades.

Table 6 Model fitting information for the multinomial logistic regression

Model	**AIC**	**BIC**	**−2 Log Likelihood**	**χ^2**	**df**	**P**
Intercept-only	1050.36	1062.34	1044.36	—	—	—
Final model	797.25	845.15	773.25	271.11	9	< .001

Note: The likelihood ratio chi-square compares the final model to the intercept-only model.

a significant p-value indicates that the predictors improve model fit

Pseudo R-Square Values

The pseudo-R-squared was examined to further evaluate the amount of the variance reported. Table 7 shows an approximation of how much variance in the strand choice is accounted for by the predictors- English, Math, and Science grades. The Cox and Snell R^2 result (0.492, upper limit < 1.0) suggests that 49.2% of the variance in the dependent variable is explained by the predictors. While the Nagelkerke R^2 result (0.528, > 0.3 to 0.5) revealed that 52.8% is a more interpretable estimate of the model`s explanatory power and implies moderate to strong predictors. In addition, the result of the most conservative McFadden R^2 (0.251) indicates good model fit, meaning the predictors substantially improve the prediction of the dependent variables compared to a null model. Overall, the values and models confirm a good fit to the data, as indicated by the pseudo-R^2 values: Cox and Snell R^2 = .492, Nagelkerke R^2 = .528, and McFadden R^2 = .251. Around 49%–53% of the variance in the dependent variable is explained by the predictors, reflecting a moderately strong model fit within social science standards.

Table 7 Pseudo R-Square Indices for the Multinomial Logistic Regression

Measure	**Value**
Cox and Snell	0.492

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

Nagelkerke	0.528
McFadden	0.251

Note. Values < .10 = weak, .10–.30 = moderate, and > .30 = strong (McFadden, 1974).

Likelihood Ratio Test for Individual Predictors of Strand Choice

To determine which predictors (English, Math, and Science grades) have the most impact on the students' selection of electives or strands, the Likelihood Ratio Tests for the individual predictors were performed. Table 8 presents the results of three predictors, viz., English, Math, and Science grades. The results showed that the English grades had the greatest impact, x^2 (3) = 41.807, p < .001, followed by Math, x^2 (3) = 16.960, p = .001, and Science, x^2 (3) = 11.848, p = .008. These results suggest that students' academic performance in these subjects plays a significant role in their choice of the strand in the Senior High School. Of the three predictors, English performance seems to be the most important factor in deciding which strand to prefer.

Table 8 Likelihood Ratio Test for Predictors

Effect	-2 Log Likelihood of Reduced Model	Chi-square (X^2)	Degree of freedom	Sig	Description
Intercept	1006.704	233.45	3	0	
English	815.061	41.807	3	<.001	Strong, significant predictor
Math	790.214	16.96	3	0.001	Significant predictor
Science	785.102	11.848	3	0.008	Significant predictor

Note. All predictors are statistically significant at p < .05.

Multinomial Logistic Regression Coefficients (Reference Category: TVL)

Table 9 displays the multinomial logistic regression results and provides a clear understanding of how the variables—grades in English, math, and science—influence the chance of students selecting ABM, HUMSS, or STEM in reference to the TVL category. The coefficient, exponentiated beta, and confidence intervals showed the differential power of three core subjects across all academic strands. English appeared as a consistent predictor of strand choice across strands. For ABM, for every one-point increase in English grades, there is a twenty-seven percent (27%) likelihood that students will choose ABM strand (Exp(B) = 1.27, 95% CI [1.13, 1.43], p < .001) relative to TVL strands; a thirty-eight percent (38%) likelihood to choose HUMSS over TVL (Exp(B) = 1.38, 95% CI [1.23, 1.55], p < .001); and a fifty-one percent (51%) likelihood to choose STEM compared to TVL (Exp(B) = 1.51, 95% CI [1.28, 1.78], p < .001).

In terms of Math grades, it only exhibits a significant predictive power capacity for STEM (Exp(B) = 1.35, 95% CI 1.12-1.64, p-value = .002) and ABM (Exp(B) = 1.12, 95% CI 1.01-1.23, p = .028) but not for HUMSS (Exp(B) = 0.99, p = .848). Furthermore, this implies that for every one-point increase in Math grade increases the odds of choosing the ABM strand increase by 12% compared with TVL. For STEM, a one-point increase in Math increases the odds of choosing it by **35%** over TVL. For Science grades, it only significantly predicted STEM selection (Exp(B) = 1.27, 95% CI 1.12–1.64, p = .003) but did not significantly predict ABM (Exp(B) = 1.02, p-value = .707) and HUMSS (Exp(B) = 1.01, p = .777). This means further that a one-point increase in science grades increases the odds of choosing STEM by **27%**.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

Table 9 Parameter estimates for the multinomial logistic regression predicting strand choice (reference category = TVL)

Outcome (vs. TVL)	Predictor	B	SE	Wald	p	Exp(B)	95% CI for Exp(B)
	Intercept	-30.931	3.998	59.862	.000		
ABM	English	0.237	0.061	15.28	< .001	1.270	1.13–1.43
	Math	0.11	0.05	4.8	0.028	1.120	1.01–1.23
	Science	0.019	0.051	0.14	0.707	1.019	0.92–1.13
HUMSS	English	0.322	0.059	30.16	< .001	1.380	1.23–1.55
	Math	−0.009	0.046	0.04	0.848	0.990	0.91–1.09
	Science	0.014	0.048	0.08	0.777	1.014	0.92–1.12
STEM	English	0.409	0.084	23.7	< .001	1.510	1.28–1.78
	Math	0.303	0.099	9.39	0.002	1.354	1.12–1.64
	Science	0.24	0.08	9.04	0.003	1.271	1.09–1.49
Note. TVL serves as the reference category. Significant predictors (p < .05) indicate increased odds of choosing the specified strand relative to TVL.							

Derived Predictive Model Equations

Based on the regression coefficients, the final predictive model equations were derived to compute the probability of each strand being selected. The exponential terms are defined as:

$$(Eq.1)\ \boldsymbol{e^{ABM} = e^{(-30.931+0.237English+0.110Math+0.019Science)}}$$

$$(\text{Eq. }2)\ \boldsymbol{e^{HUMSS} = e^{(-27.070+0.322English-0.009Math+0.014Science)}}$$

$$(Eq.3)\ \boldsymbol{e^{STEM} = e^{-83.425+0.409English+0.303Math+0.240Science}}$$

The corresponding probabilities for each strand are computed as:

FINAL PREDICTIVE MODEL

1. $P(ABM) = \frac{e^{ABM}}{1+e^{ABM}+e^{HUMSS+eSTEM}}$
2. $P(HUMSS) = \frac{e^{HUMSS}}{1+e^{ABM}+e^{HUMSS+eSTEM}}$
3. $P(STEM) = \frac{e^{STEM}}{1+e^{ABM}+e^{HUMSS+eSTEM}}$
4. $P(TVL) = \frac{1}{1+e^{ABM}+e^{HUMSS+eSTEM}}$

These equations represent the final predictive model, which can estimate the likelihood that a student with specific English, Math, and Science grades will choose each Senior High School strand.

In summary, the multinomial logistic regression model significantly predicted students' strand choices based on English, Math, and Science grades. English emerged as the strongest predictor. The model explained approximately 25–53% of the variance and achieved an overall accuracy of 53.8%. STEM and TVL strands were best predicted. The final predictive model provides equations that can estimate probabilities for each strand choice, potentially guiding academic advising and career planning for Grade 10 students.

Based on the result of multinomial logistic regression, the following log-odds equation was derived. These equations show how increases in subject grades affect the log-odds of choosing ABM, HUMSS, or STEM relative to TVL.

Classification Accuracy of the Model

Table 10 showed a cross-tabulation of the actual choices of strands by students against the modelled predicted choices of strands. This process aimed to determine the validity of the model of classification in assigning students into their respective Senior High School strands. The model obtained the correct classification in 215 out of the 400 cases, which means that the overall model's accuracy is 53.75 %. When analyzed by predictors, the analysis showed that the STEM and TVL strands had the highest correct prediction rate of 75.0 percent and 74.6 percent, respectively, indicating acceptable model performance for technical and vocational tracks.

Conversely, the accuracy of the model was quite low for the HUMSS strand (49.2 percent) and ABM in particular (17.4 percent), with a significant fraction of ABM students being misclassified (17.4 percent). These results imply that the model was better at distinguishing strands that have unique characteristics, like STEM and TVL, than strands that have overlapping interests and competencies, such as ABM and HUMSS. The moderate overall accuracy suggests that, although the model perhaps could be a useful tool in determining strand choice, there is still room for improvement in refining the predictors or the classification algorithm.

Table 10 Classification Accuracy of the Model Based on Strand Choice and Predicted Response Category

	Predicted Response Category					Total
STRAND CHOICE		ABM	HUMSS	STEM	TVL	
ABM	Count	16	44	13	19	92
	% within STRANDCHOICE	17.4%	47.8%	14.1%	20.7%	100.0%
HUMSS	Count	10	60	17	35	122
	% within STRANDCHOICE	8.2%	49.2%	13.9%	28.7%	100.0%
STEM	Count	2	12	45	1	60
	% within STRANDCHOICE	3.3%	20.0%	75.0%	1.7%	100.0%
TVL	Count	7	24	1	94	126
	% within STRANDCHOICE	5.6%	19.0%	0.8%	74.6%	100.0%
	Count	35	140	76	149	400
	% within STRANDCHOICE	8.8%	35.0%	19.0%	37.3%	100.0%

Effect Size Considerations (Cohen's d)

Cohen's d effect sizes are useful for expressing the practical magnitude of grade-profile differences. However, the current results tables report regression coefficients/odds ratios (Table 9) and a confusion matrix of actual versus predicted strands (Table 10), but do not report the distributions of grades within the predicted groups that are required to compute Cohen's d for actual-versus-predicted grade profiles. Accordingly, Cohen's d values are not presented here. Future replications should retain the same modelling framework while additionally reporting

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

mean grades and standard deviations for each predicted strand group, including misclassified cases, to enable transparent effect-size interpretation.

DISCUSSION

The results of this study showed that Grade 10 academic performance in English, Mathematics, and Science is found to be a statistically significant predictor of SHS strand choices of the students. The model provides an acceptable fit, indicating moderate-to-strong pseudo-R2 values, which provide support for the substantive predictive value of these academic domains, despite failing to account for some factors that could influence the student decision-making process.

The findings confirmed the theoretical paradigm of Holland's Theory of Vocational Choice (1997), which states that individuals tend to be attracted to professional settings that are compatible with their talents and interests. Improved academic grades indicate skills that are associated with certain strands, such as Mathematics and Science for STEM; English for HUMSS. Similarly, Super's (1990) Life-space Theory, and Life-Span Theory of Super, and the Social Cognitive Career theory (Lent, Brown, & Hackett, 1994) made a claim that self-concept, hence self-efficacy, constructed through academic experiences, mediates career and educational dispositions. Ajzen's Theory of Planned Behavior (1991) also postulates, and research supports, that students' perceptions of control over competency with regard to a given subject affect their intentions to pursue related courses. English was the most powerful and reliable predictor across all strands, highlighting its importance in the facilitation of academic achievement. This finding is consistent with Cummins' (1979) proposition of cognitive-academic language proficiency, in which linguistic proficiency will further facilitate knowledge acquisition between and across disciplinary boundaries. Mathematics and Science filled the role of strand-specific predictors, having a major impact on STEM choices and, to a smaller degree, on Accounting-Business-Management (ABM). Such a trend is not unexpected as the strands are entirely dependent on quantitative and analytical skills. The predictiveness of the model was higher for STEM and TVL strands, most likely because these have more clearly delineated academic profiles. On the other hand, the lower accuracy that the model achieved with ABM and HUMSS could be due to overlapping competences, specifically in the English subject, as well as to the inclusion of non-academic factors such as personal interests, the socio-economic context, or the school guidance strategies. Collectively, the results provide support for the fact that academic performance is a relevant, yet not exclusive, determinant of strand choice. Simultaneously, other significant influences, such as interests, aspirations, and contextual conditions, exert their considerable influence but should be methodically incorporated into guidance practices.

The findings confirmed Holland's Vocational Choice Theory (1997), which posits that students choose environments that match their abilities and interests (Hartmann et al., 2021). High grades signal competencies that align with specific strands, such as Math and Science for STEM, or English for HUMSS. Likewise, the result that English revealed to be the strongest predictor across strands aligns with Cummins' (1979) idea of cognitive-academic language proficiency, where language ability supports learning across disciplines (Nyoni, 2018). Specifically, Math and Science were strand-specific predictors significantly influencing STEM and ABM. This pattern is expected as these strands require logic and analytical skills, reflecting the critical role of these subjects in academic success.

Consequently, the model predicted STEM and TVL strands with higher accuracy, likely because these strands have more distinct academic profiles. In contrast, ABM and HUMSS showed lower accuracy, possibly due to overlapping competencies, particularly in English, and other non-academic influences like interests, socioeconomic factors, or school guidance practices. Overall, the results confirm that academic performance is a key but not exclusive factor in strand selection. Other variables—such as interests, aspirations, and context—also play important roles and should be integrated into guidance systems.

Limitations of a Grades-Only Predictive Model

A central constraint of the present model is its exclusive reliance on Grade 10 English, Mathematics, and Science marks. While these subjects are foundational across SHS curricula, they do not uniquely encode strand-specific competencies such as business numeracy and applied accounting for ABM; humanities and social science writing demands for HUMSS; laboratory-based scientific reasoning for STEM; or technical skill acquisition and performance tasks for TVL. As a result, strands with overlapping curricular demands can exhibit similar grade profiles, limiting separability in a grades-only classifier.

This limitation is evident in the strand-level classification results. The model's overall accuracy was 53.75% (215/400 correct classifications), but accuracy varied widely by strand: ABM = 17.4%, HUMSS = 49.2%, STEM = 75.0%, and TVL = 74.6%. The very low ABM accuracy (17.4%) reflects substantial misclassification into HUMSS (47.8%) and TVL (20.7%), consistent with English and Mathematics grade patterns that are insufficiently distinct for ABM relative to neighboring non-STEM strands. Similarly, HUMSS accuracy (49.2%) indicates persistent overlap with other strands (notably TVL at 28.7% and STEM at 13.9%), underscoring that grades alone cannot reliably capture the strand-specific competency boundary when subject requirements and assessment structures are shared across programs.

Grades-Only Prediction in Prior Orientation Research

Prior orientation and pathway-prediction studies indicate that models relying primarily on academic marks can produce uneven multi-class performance when options share similar achievement patterns. In an SHS-focused strand prediction study using student grades, Nazareno et al. (2019) reported that predictive performance can vary across strands, reflecting the degree of separability of grade signals across options. More broadly, study-path selection research emphasizes that prediction quality improves when academic history is complemented with additional learner attributes like motivational or contextual indicators, rather than treating grades as a complete proxy for readiness or fit (Dirin & Saballe, 2022). In the Philippine context, track/strand prediction using deep learning similarly demonstrates the value of incorporating richer input features beyond a minimal grade set, supporting the view that transcript-only approaches should be interpreted as screening-level aids rather than stand-alone placement mechanisms (Hernandez & Atienza, 2021).Consistent with the present study's strand overlap issue, marks-driven multi-class modelling in adjacent educational settings has also reported moderate accuracy bands in some model configurations when categories are not cleanly separated by academic signals (Alsayed et al., 2021).

Practitioner Disclaimer

For educators and policymakers, the present grades-only model should be used only as an initial screening or filtering aid to support counseling conversations. Because grade profiles can be similar across strands, prediction accuracy varies widely by strand (approximately 17.4%–75.0% in this sample). High-stakes placement decisions should triangulate grades with student interests, aptitude measures, and structured guidance processes.

CONCLUSION

This study demonstrates that Grade 10 English, Mathematics, and Science grades can support a preliminary, grades-only screening of SHS strand options, but with clear boundaries. The multinomial logistic regression model achieved an overall accuracy of 53.75%, reflecting substantial overlap in core-subject grade profiles across strands. Strand-level performance ranged from 17.4% (ABM) to 75.0% (STEM), indicating that grades alone are least informative for strands whose achievement patterns are highly similar. Accordingly, the model is not proposed as a definitive placement mechanism; instead, it can be used to narrow initial options and to prompt targeted guidance discussions. Future work should integrate complementary predictors (e.g., interest/aptitude measures, performance-task indicators, and contextual variables) and report additional diagnostics that quantify practical separation between strands, thereby improving the interpretability and responsible use of predictive screening in SHS advising.

RECOMMENDATIONS

In light of the conclusions, the following recommendations are offered:

1. Institutionalize a blended, ethical strand-guidance framework. Use the grades-only model as preliminary screening to support and enhance their career guidance counseling programs and activities, but not a

placement standard, because accuracy varies widely by strand (≈17.4%–75.0%) due to grade-profile overlap. Final strand decisions should be made through triangulation of: (a) student choice and interests, (b) counseling interviews, and (c) aptitude/readiness indicators, with special attention to ABM and HUMSS, where misclassification is highest.

2. Strengthen counseling protocols using model outputs as diagnostic signals. Guidance offices may use predicted probabilities (not only predicted classes) to identify students with "borderline" profiles—especially those predicted with low confidence or those in strands with low predictive separation (ABM, HUMSS). These cases should trigger deeper counseling, parent consultation, and exploration of learner goals and constraints.

3. Integrate complementary predictors to improve strand discrimination. Future implementations should extend beyond English, Mathematics, and Science grades by including structured interest inventories such as career/strand preference measures, aptitude indicators, and performance-task evidence aligned with strand-specific demand. For example, applied business tasks for ABM, writing/argumentation indicators for HUMSS, laboratory/problem-solving evidence for STEM, and competency-based assessments for TVL. This is essential for strands whose competencies are not adequately captured by core-subject marks alone.

4. Develop strand-specific competency measures aligned with curriculum standards**.** Policy and program developers should consider designing standardized, strand-aligned diagnostic tools (short performance assessments or readiness indicators) that complement grades. Such tools would help reduce grade-overlap ambiguity and better reflect real strand readiness

5. For Future Researchers to explore additional predictors beyond core grades that could enhance the prediction quality and interpretability, such as including strand-level diagnostics, socioeconomic status, parental influence, career interest inventories, and psychological factors, to enhance the predictive accuracy of models. In addition, to replicate the study with larger and more diverse samples across different schools and regions to validate and generalize the findings. Consider comparing traditional statistical models with machine learning approaches such as decision trees, random forests to improve classification accuracy.

REFERENCES

1. Adarlo, G., & Jackson, L. (2017). For Whom Is K-12 Education: A Critical Look into Twenty-First Century Educational Policy and Curriculum in the Philippines. In S. Choo, D. Sawch, A. Villanueva, & R. Vinz (Eds.), Educating for the 21st Century (pp. 207–223). Springer Singapore. https://doi.org/10.1007/978-981-10-1673-8_11
2. Adlya, S. I., & Zola, N. (2022). Holland's Theory of Guiding Individual Career Choices. Jurnal Neo Konseling, 4(4), 30. https://doi.org/10.24036/00698kons2022
3. Alinea, J. M., Garcia, C., Lagado, S., & Noblefranca, S. M. (2024). Employability Skills of Senior High School Tech-Voc Students: A Measure of Industry Preparedness. Interdisciplinary Research Review, 19(4), 63–79.
4. Alsayed, A. O., Rahim, M. S. M., AlBidewi, I., Hussain, M., Jabeen, S. H., Alromema, N., Hussain, S., & Jibril, M. L. (2021). Selection of the Right Undergraduate Major by Students Using Supervised Learning Techniques. Applied Sciences, 11(22), 10639. https://doi.org/10.3390/app112210639
5. Balontong, R. (2024). C4.5 Decision Tree Algorithm and Linear Regression in Guidance and Counseling Decision Support System. Journal of Innovative Technology Convergence, 6(1). https://doi.org/DOI:%2520https://doi.org/10.69478/JITC2024v6n2a12
6. Bejar, R. Jr. (2024). Precision in Progress: Leveraging Data Mining Technique to Empower Career Path Selection for Incoming Senior High School Students. RSIS International.Org, 178–191. https://doi.org/DOI:%2520https://dx.doi.org/10.47772/IJRISS.2024.801014
7. Borghosh, P. R. M. (2024). Predictive Analytics Techniques in Education by Artificial Intelligence Tools for Enhancing Academic Assessment: Systematic Review. https://doi.org/10.21608/ijel.2024.401345

8. C. Acosta, I., & S. Acosta, A. (2016). Teachers' Perceptions on Senior High School Readiness of Higher Education Institutions in the Philippines. Universal Journal of Educational Research, 4(10), 2435–2450. https://doi.org/10.13189/ujer.2016.041024
9. Clinical Research Centre, Sarawak General Hospital, Ministry of Health, Kuching, Malaysia, Bujang, M. A., Sa'at, N., Biostatistics Unit, National Clinical Research Centre, Ministry of Health, Kuala Lumpur, Malaysia, Tg Abu Bakar Sidik, T. M. I., Biostatistics Unit, National Clinical Research Centre, Ministry of Health, Kuala Lumpur, Malaysia, Chien Joo, L., & Clinical Research Centre, Sarawak General Hospital, Ministry of Health, Kuching, Malaysia. (2018). Sample Size Guidelines for Logistic Regression from Observational Studies with Large Population: Emphasis on the Accuracy Between Statistics and Parameters Based on Real Life Clinical Data. Malaysian Journal of Medical Sciences, 25(4), 122–130. https://doi.org/10.21315/mjms2018.25.4.12
10. Dangoy, J. E., & Madrigal, D. V. (2020). Career Preferences and Factors Influencing the Career Choice of Senior High Students of a Catholic School. Philippine Social Science Journal, 3(2), 95–96. https://doi.org/doi:10.52006/main.v3i2.235
11. Department of Education. (2025). The Strengthened Senior High School Program Shaping Paper. Department of Education. https://www.deped.gov.ph/wp-content/uploads/SHAPE-PAPER-BCD-a_o-May-22.pdf?utm
12. Desvikayati, I., Daharnis, & Ifdil. (2025). Analysis Of Holland's Theory In Career Decision Making Of Vocational High School Students. Mahir : Jurnal Ilmu Pendidikan Dan Pembelajaran, 4(1), 1–6. https://doi.org/10.58432/mahir.v4i1.1213
13. Dirin, A., & Saballe, C. A. (2022). Machine Learning Models to Predict Students' Study Path Selection. International Journal of Interactive Mobile Technologies (iJIM), 16(01), 158–183. https://doi.org/10.3991/ijim.v16i01.20121
14. EDCOM 2 Communications. (2025). DepEd, CHED should review proposed SHS electives, college subjects to ensure they are implementation-ready. EDCOM 2. The Second Congressional Commission on Education. https://edcom2.gov.ph/deped-ched-should-review-proposed-shs-electives-college-subjects-to-ensure-they-are-implementation-ready/
15. Hartmann, F., Heine, J.-H., & Ertl, B. (2021). Concepts and Coefficients Based on John L. Holland's Theory of Vocational Choice—Examining the R Package holland. Psych, 3(4), 728–750. https://doi.org/10.3390/psych3040047
16. Hernandez, R., & Atienza, R. (2021). CAREER TRACK PREDICTION USING DEEP LEARNING MODEL BASED ON DISCRETE SERIES OF QUANTITATIVE CLASSIFICATION. Applied Computer Science, 17(4), 55–74. https://doi.org/10.35784/acs-2021-29
17. Ignacio, A. A., & Bajet, A. P. (2025). Assessing the Implementation of Senior High School Mathematics Curriculum: Teachers' and Students' Perceptions in Calbayog City Division, Philippines. European Journal of Education and Pedagogy, 6(1), 61–65. https://doi.org/10.24018/ejedu.2025.6.1.918
18. Khan, W. U., Ghazi, S. R., Ali, A., & Khan, F. (2025). Quality of Primary Education in Bannu Division: An Analysis of Quality of Input, Process, and Output Indicators. ProScholar Insights, 4(3), 112–122. https://doi.org/10.55737/psi.2025c-43109
19. Lee, F. (2025). What is logistic regression? https://chatgpt.com/c/68bc99db-9dc0-8325-b896-e70e9315d0ad
20. Malipot, M. (2025). K to 12 Program: 11 years of transforming Philippine education. MB125. https://125.mb.com.ph/k-to-12-program-11-years-of-transforming-philippine-education/?utm
21. Nazareno, A. L., Lopez, M. J. F., Gestiada, G. A., Martinez, M. P., & Roxas-Villanueva, R. M. (2019). An artificial neural network approach in predicting career strand of incoming senior high school students. Journal of Physics: Conference Series, 1245(1), 012005. https://doi.org/10.1088/1742-6596/1245/1/012005
22. Nyoni, E. (2018). Instructional Interventions Used To Develop Cognitive Academic Language Proficiency in the Teaching of a Content Area Subject Language Learners. Rethinking Teaching and Learning in the 21st Century, 6.
23. Pärli, R. (2023). How input, process, and institutional factors influence the effects of transdisciplinary research projects. Environmental Science & Policy, 140, 80–92. https://doi.org/10.1016/j.envsci.2022.11.019

24. Sadjail, S.-I. A., Sansawi, D. J., & Matolo, M.-A. L. (2022). Factors Influencing Students in Choosing their College Course. Psychology and Education: A Multidisciplinary Journal, 3(8), 697–700.
25. Sarona-Pedro, D. S. (2025). Mixed Method Study on Course Shifting among College Students.
26. Shankhdhar, A., Agrawal, A., Sharma, D., Chaturvedi, S., & Pushkarna, M. (2020). Intelligent decision support system using decision tree method for student career. 2020 International Conference on Power Electronics & IoT Applications in Renewable Energy and Its Control (PARC), 140–142.
27. Sidek, M. R. E. B., Impil, J. A. D., Marcos, M. J. B., Punzalan, E. E. M., Ramos, J. T., & Ador, Z. I. (2023). Exploring Course Mismatch among Former STEM Students: A Narrative Research. Cognizance Journal of Multidisciplinary Studies, 3(7), 78–90. https://doi.org/10.47760/cognizance.2023.v03i07.008
28. Vargas, P. C. (2024). Predictive Analysis Application for Academic Student Advancement in Secondary Education Institutions. Journal of Innovative Technology Convergence, 6(1). https://doi.org/DOI:%2520https://doi.org/10.69478/JITC2024v6n2a04
29. Varoquaux, G. (2018). Cross-validation failure: Small sample sizes lead to large error bars. NeuroImage, 180, 68–77. https://doi.org/10.1016/j.neuroimage.2017.06.061
30. Wang (汪茂华), M., & Liu (刘华聪), H. (2025). School-to-School Collaboration, Student Learning Outcomes, and Mediating Factors: Evidence From Shanghai’s Strong School Project. ECNU Review of Education, 20965311251330932. https://doi.org/10.1177/20965311251330932
31. YU, C., & Upah, S. (2018). Data Analytics and STEM Student Success: The Impact of Predictive Analytics-Informed Academic Advising Among Undeclared First-Year Engineering Students. Sage Journals, 22(3). https://doi.org/10.1177/1521025118772307
32. Zhang, B., Li, S., Hounie, I., Bastani, O., Ding, D., & Ribeiro, A. (2025). Alignment of large language models with constrained learning (Version 2). arXiv. https://doi.org/10.48550/ARXIV.2505.19387
33. Гриценок, І. (2025). ЗМІСТ ПРОФЕСІЙНОЇ ОРІЄНТАЦІЇ ТА КОНСУЛЬТУВАННЯ З КАР’ЄРИ ЗДОБУВАЧІВ ФАХОВОЇ ПЕРЕДВИЩОЇ ОСВІТИ. B U L L E T I N OF OLEKSANDR DOVZHENKO HLUKHIV NATIONAL PEDAGOGICAL UNIVERSITY, 2(58), 114–123. https://doi.org/10.31376/2410-0897-2025-2-58-114-123