C4.5 Decision Tree Algorithm and Linear Regression in Guidance and Counseling Decision Support System

Rea P. Balontong [1], Jake R. Pomperada [2]

Abstract: This study employs a combination of quantitative and experimental approaches to explore the relationship between family dynamics, psychological test results, and academic performance among students in specific courses. Quantitative data collection was conducted using existing student profiles from the guidance office, while an experimental research method utilized the C4.5 algorithm to determine these relationships. The participants were students enrolled in various courses over a specific academic year. The decision tree generated by the C4.5 algorithm was used to predict academic performance based on family dynamics and psychological test results. The study also employed Simple Linear Regression to identify the predictors of academic performance. The accuracy of the predictive model was tested using a confusion matrix, which yielded an accuracy rate of 97%. Additionally, a correlation matrix and regression analysis were performed to identify significant correlates and predictors of academic performance. The findings of this study suggest the potential for positive changes in guidance services, particularly in identifying predictors of academic performance and informing decision-making processes. These findings contribute to our understanding of schooling outcomes in relation to family dynamics and parental migration status.

Keywords: C4.5, Linear Regression, Decision Support System, Guidance Counseling, Confusion Matrix

1. Introduction

Technological advancements, like multimedia, have heightened the effectiveness and importance of employing information and communications technology (ICT) in both educational and occupational guidance. ICT's role in guidance can be viewed as a tool, an alternative, or a catalyst for change, with the potential to greatly enhance access to guidance services by overcoming constraints of time and space. Systems integrating ICT can be designed for guidance counseling, akin to what could be termed a decision support system (DSS). A survey was conducted among a sample of school guidance counselors

[1] College of Computer Studies, Central Philippine University, Jaro, Iloilo City, Philippines
Email: rpbalontong@cpu.edu.ph
[2] Department of Information Technology, Technological University of the Philippines Visayas, Bacolod City, Philippines
Email: jakerpomperada@gmail.com

Received [February 13, 2024]; Revised [March 16, 2024]; Accepted [April 11, 2024]

Published by InnoCon Publishing
ISSN 2704-4440

in Metro Manila, the Philippines, to determine their level of awareness, attitudes, and extent of practice of ICT in the conduct of their work. The respondents showed a higher level of awareness of communication technologies than of information technologies. They spent an average of 1.2 hours per day using a computer in the workplace. Their primary sources of guidance-related information were print-based, but a few also cited the Internet. The respondents had a positive attitude toward the use of ICT in guidance [1]. Decision Support System are a class of computerized information systems that support decision-making activities and are interactive computer-based systems and subsystems intended to help decision-makers use communications technologies, data, documents, knowledge, and/or models to complete decision process tasks. Information may be presented graphically and may include an expert system or artificial intelligence. It may be aimed at business executives or some other group of knowledge workers [2]. DSS can be applied in different areas, some of which are clinical, business and management, agriculture production, and marketing, to name a few. There are also different types of DSS, namely, communication-driven DSS, data-driven DSS, document-driven DSS, knowledge-driven DSS, and model-driven DSS [3].

The Guidance office faces challenges in managing student records, especially in collecting and interpreting psychological test results, which is time-consuming. Given the importance of timely decisions in student guidance, there is a pressing need to computerize these processes to handle large volumes of data efficiently. By leveraging Information and Communication Technology (ICT), the office can not only streamline operations but also monitor student performance and offer guidance counseling effectively. This proactive approach involves early intervention for students facing academic difficulties, achieved through the design and development of a decision support system tailored for guidance and counseling purposes.

1.1 Statement of the Problem

Recognizing the importance of DSS in decision-making, it is therefore vital for counselors to have complete and accurate data when performing the process of assessment, consultation, and monitoring. For this reason, the researcher would like to:

Recognizing the importance of DSS in decision-making, it is therefore vital for counselors to have complete and accurate data when performing the process of assessment, consultation, and monitoring. For this reason, the researcher would like to:

1. Determine the effectiveness of the C4.5 algorithm in determining the relationships that exist between variables and their predictability;
2. Test whether C4.5 was able to determine the relationship that exists between students' family dynamics, psychological test results, and students' academic performance and to identify predictors of students' academic performance;
3. Determine the functionality, reliability, usability, efficiency, maintainability, and portability of the system features and functionalities.

1.2 Hypothesis

In this research, based on the problems stated, the researcher hypothesized that the C4.5 algorithm can be used to determine the relationships that exist between variables and their predictability.

1.3 Theoretical Background

In this research, it represents the concepts and theories that explain and support the study conducted. It also discusses how, in each study, these concepts and theories were applied.

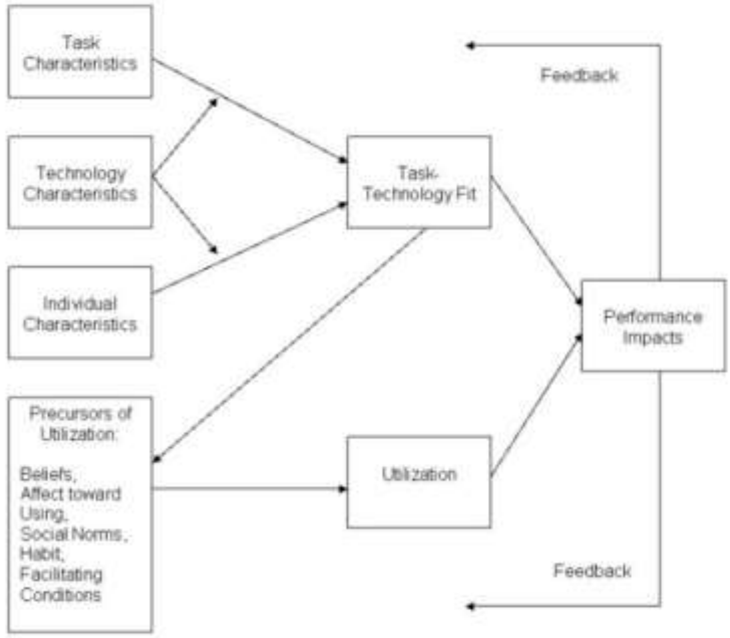

Figure 1. Task-Technology Fit

Based on Figure 1, the Task-Technology Fit (TTF) Theory, first introduced by Goodhue and Thompson in 1995 [4], asserts that the performance of individuals using technology is shaped by how well the features of the technology align with the demands of the tasks they engage in. This theory underscores the significance of assessing the compatibility and efficacy of both the task and the technology, highlighting their interplay in determining user performance and satisfaction. When there is a good fit between the tasks that individuals need to accomplish and the features of the technology available to them, they are more likely to experience higher levels of productivity and satisfaction.

Cognitive theory was utilized in this study. Cognitive learning theories propose that learning entails integrating events into an active storage system called schemata. This framework offers insight into the workings of the mind, with schemata serving as mental representations formed through experience. The mind uses these schemata to selectively organize and process incoming information. The processing of information for storage involves several key cognitive components [5].

The concept of cognitive learning theory was applied in the DSS feature of the study. Looking at how the whole theory was formulated is similar to how a computer works. Data are being entered into the system; the system processes this data and transforms it into useful information. In some cases, this data is being analyzed to give it more meaning and be used in decision making. As part of the analysis, it illustrates that the system is looking at patterns of similarity in the set of data it stores.

2. Methodology and Design

2.1 Quantitative Methodology

A combination of quantitative and experimental approaches was used in this study. The process of collecting and analyzing numerical data. It can be used to find patterns and averages, make predictions, test causal relationships, and generalize results to wider populations [6].

Quantitative data was gathered based on the existing student profile from the Guidance office, which includes family dynamics, psychological test results, and the academic performance of students. For the determination of the system features and functionalities, a system evaluation was conducted by the guidance director and guidance counselors. Experimental research was used to determine the relationship between family dynamics, psychological test results, and academic performance using the C4.5 algorithm.

2.1.1 Participants

The participants for this study are composed of students enrolled from 1st year for School Year (SY) 2015-2016, 1st year to 2nd year for SY 2016-2017, and 1st year to 3rd year for SY 2017-2018 in the following courses:

1. Bachelor of Science in Computer Science
2. Bachelor of Science in Information Systems
3. Bachelor of Science in Information Technology
4. Bachelor of Science in Digital Media and Interactive Arts

The students enrolled in the specified semesters should have completed the following requirements:

1. They have filled out the family dynamics survey form.
2. They have taken the psychological test.

2.1.2 Procedure

C4.5 was used as the algorithm that determines the relationship of family dynamics and psychological test results to academic performance. It is a decision tree-based algorithm popularly used to perform data mining tasks [7]. It is designed mainly to classify data sets based on the defined rules in decision trees, which could contain categorical and numerical attributes [8]. The technique in the decision tree family can produce both decision trees and rule sets. The model is easy to understand as rules derived from the technique have a straightforward interpretation [9].

2.1.3 Data Gathering

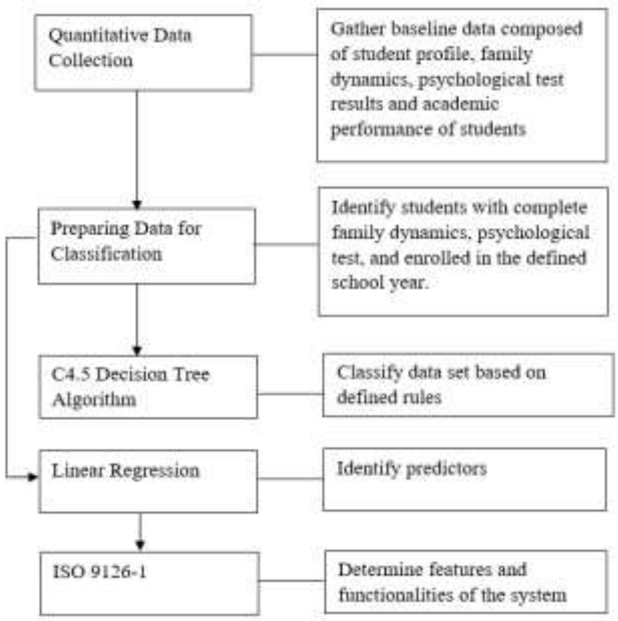

Figure 2. Process for Data Gathering

As shown in Figure 2, the first set of data gathered includes the student profile, family dynamics, psychological test results, and academic performance of students. The data is composed of students enrolled from 1st year for School Year (SY) 2015-2016, 1st year to 2nd year for SY 2016-2017, and 1st year to 3rd year for SY 2017-2018. The data were classified in which students who have a complete family dynamic, were able to take the psychological test, and are currently enrolled in the college were identified. Based on the number of students enrolled in the specified year, only one hundred and twenty-two (122) students were considered part of the data set.

2.1.4 Procedure

Building the Decision Tree. In reference to problem 1, a decision tree is a flowchart-like tree structure where each internal node (none leaf) denotes a test on the attribute, each branch represents an outcome of the test, each leaf node (or terminal node) holds a class label, and the topmost node is the root node [10].

The decision tree is a divide-and-conquer approach to splitting the data set based on a given question. In the study, the researchers wanted to predict students who would have good and poor academic performance based on variables. Such as the family dynamics, which describe the family structure of a student, the result of the psychological test, and the academic performance of a student, whether probationary or not.

To classify the group of students based on the given data set, first identify the splitting criterion, which uses the concepts of Entropy and Information Gain.

The following computations are being performed:

1. Equation 1 is used to compute the Entropy of the whole training data set.

$$E(D) = \sum_{i=1}^{m} -P_i \, log_2 \, (P_i) \tag{1}$$

where: E(D) = entropy of the whole training data set

m = is the number of outcomes

i = set of outcomes both positive and negative

P_i = probability of an outcome

2. Equation 2 and Equation 3 are used to compute the Entropy and Information Gain of each attribute, respectively.

$$E_A(D) = \sum_{j=1}^{y} (\frac{D_j}{D})(-P_i \, log_2 \, (P_i)) \tag{2}$$

where: E_A(D) = entropy of a single attribute

v = is the number of outcomes of an attribute

j = set of outcomes both positive and negative of an attribute

Pi = probability of an outcome

$$Gain\ (A) = E(D) - E_A(D) \tag{3}$$

where: Gain (A) = information gain of an attribute

After the information gain of each attribute is computed, the gain with the highest value is considered the splitting attribute and becomes the root node of the decision tree. The same concept was applied to each branch until a leaf was identified.

2.1.5 Accuracy Test

In order to test the accuracy of the rules generated using the decision tree, the confusion matrix was used. The confusion matrix is a useful tool for analyzing how well a classifier can recognize tuples of different classes [10].

Misclassification or Error Rate. This refers to how often it is wrong in an overall evaluation. In the study, it represents the percentage of the algorithm that classifies the data set incorrectly. Equation 4 is used to compute the Error Rate.

$$Error\ Rate = \frac{FP+FN}{Total\ Number\ of\ Dataset} \tag{4}$$

Misclassification or Error Rate = (1+2)/122

= 0.0246

= 2.46 %

True Positive or Sensitivity Rate. This refers to how often an algorithm predicts yes when the actual value is yes. The percentage that this rate represents in the study is the number of values found in the data set wherein the student was classified to have an academic problem, both in the actual and predicted result. The True Positive can be computed using Equation 5.

$$True\ Positive = \frac{TP}{Actual\ Value\ of\ Yes} \tag{5}$$

True Negative or Specificity Rate. This refers to how often it predicts no when the actual value is no. The percentage that this rate represents in the study is the number of values found in the data set wherein the students were classified as having no academic problems, both in the actual and predicted results. The True Negative can be computed using Equation 6.

$$True\ Negative = \frac{TN}{Actual\ Value\ of\ No} \tag{6}$$

True Negative or Specificity = 117/118

= 0.9915

= 99.15 %

The following rates are computed where,

TP = True Positive (2),

TN = True Negative (117),

FP = False Positive (1),

FN = False Negative (2),

Actual Yes = 4,

Actual No = 118,

Predicted Yes = 3,

Predicted No = 119

Total number of data set = 122.

Accuracy. This refers to how often the classifier is correct in an overall evaluation. The percentage that this rate represents is the number of correctly classified instances, which shows the effectiveness of the algorithm in terms of prediction. For the study, it can be defined that the algorithm used in predicting the academic status of a student is accurate.

Based on the given dataset, the computed accuracy percentage is high, which can be computed using Equation 7.

$$Accuracy = \frac{TP + TN}{Total\ Number\ of\ Dataset} \tag{7}$$

Accuracy = (2+117)/122

= 0.9754

= 97.54 %

The rates that were computed from the confusion matrix, such as the Accuracy, True Positive Rate, True Negative Rate, and Misclassification Rate based on the values stated, support the accuracy of the predictions generated by the algorithm used in the research.

To validate the result of the computed accuracy, the researcher used third-party software known as Waikato Environment for Knowledge Analysis (WEKA). The WEKA is a collection of different machine learning algorithms for data mining and tools that can be used for data pre-processing, classification, regression clustering, and visualization [11].

Identify predictors. In reference to problem 2, to test whether family dynamics and psychological tests are good predictors of academic performance, Linear Regression was used.

In Linear Regression, the scores were predicted on one variable from the scores on a second variable. The variable we are predicting is called the criterion variable and is referred to as Y. The variable we are basing our predictions on is called the predictor variable and is referred to as X. When there is only one predictor variable, the prediction method is called simple regression [12].

Identify features and functionalities for the system. In reference to problem 3, a survey was done with the guidance counselors of the university and they were asked to rate the different modules using a scale of 1 – 5 where 1= Very Poor, 2 = Poor, 3 = Good, 4 = Very Good and 5 = Excellent. The criteria were based on ISO 9126-1 [13], where the software quality model identifies six (6) main quality characteristics: functionality, reliability, usability, efficiency, maintainability, and portability.

2.2 Experimental Methodology

The experimental method was used in the study to test the C4.5 algorithm in determining the relationship that exists between students' family dynamics, psychological test results, and students' academic performance. The developed software was used to simulate the said algorithm. A series of tests were conducted on the data set using the software until the desired accuracy test result was achieved.

3. Results and Discussion

3.1 Accuracy Test of C4.5 Predictive Model

The accuracy of the rules generated by the model C4.5 was tested using the confusion matrix for a data set of 122 students, which was later used to test the accuracy of the model to predict the academic performance of the sample respondents. Table 1 shows the results of the test using a confusion matrix. Given the following values for the variables: Number of data sets (N), True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN), Actual Yes, and Actual No.

Table 1. Confusion Matrix

N = 122	Result: No	Result: Yes	
Actual: No	TN = 117	FP = 1	118
Actual: Yes	FN = 2	TP = 2	4
	119	3	

The computation of accuracy is as follows:

Accuracy = (TP+TN)/N

Where: TP = 2, TN = 117, N = 122

Accuracy rate = (2+117)/122 = 0.97 x 100 = 97%

The computed accuracy rate is equal to 97% in terms of the classification and the rules it generates.

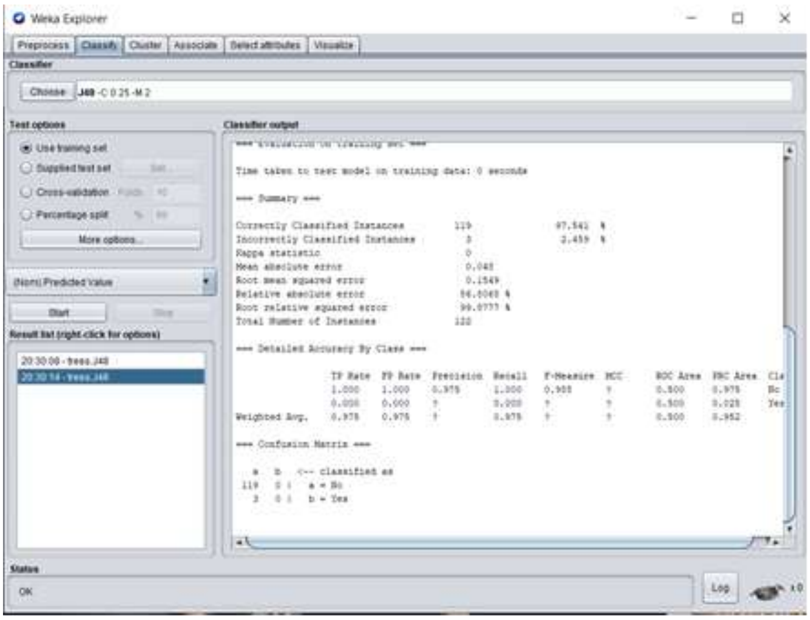

Figure 3. WEKA Computed Screenshot

The same set of data was run on the WEKA workbench to validate the result that was computed by the system, and the same result was derived as shown in Figure 3. Wherein it was able to correctly classify 119 number of instances equivalent to 97% and incorrectly classify 3 number of instances equivalent to 3% out of 122 data sets.

3.2 Correlation Matrix for the Study Variables

A Pearson's correlation coefficient matrix was generated to test whether significant correlations exist between study variables.

Table 2. Correlations Matrix among Study Variables (N=122)

	AcadPer.	Probil	Gender	NE Domain	Separated	ParBoth	ParSP	Deceased	BOFW	FFOFW	FMOFW
AcadPer.	1.000	.601	-.198	.025	.096	-.076	.032	.027	-.259	.050	.149
Probil	**.601**	1.000	-.161	-.085	.334	-.150	-.015	-.042	-.042	.046	.079
Gender	**-.198**	-.161	1.000	.217	-.153	-.005	-.011	.104	-.006	-.228	.187
NE Domain	.025	-.085	.217	1.000	-.064	-.136	.031	.123	.078	-.213	.064
Separated	.096	.334	-.153	-.064	1.000	-.379	-.058	-.033	-.033	.089	.079
ParBoth	-.076	-.150	-.005	-.136	-.379	1.000	-.517	-.291	-.291	.150	-.373
ParSP	.032	-.015	-.011	.031	-.058	-.517	1.000	-.045	-.045	-.150	-.089
Deceased	.027	-.042	.104	.123	-.033	-.291	-.045	1.000	-.025	-.085	-.050
BOFW	**-.259**	-.042	-.006	.078	-.033	-.291	-.045	-.025	1.000	-.085	-.050
FFOFW	.050	.046	-.228	-.213	.089	.150	-.150	-.085	-.085	1.000	-.168
FMOFW	.149	.079	.187	.064	.079	-.373	-.089	-.050	-.050	-.168	1.000
AcadPerf	.	.000	.014	.390	.146	.201	.364	.383	.002	.291	.051
Probil	**.000**	.	.038	.175	.000	.050	.435	.325	.325	.308	.194
Gender	**.014**	.038	.	.008	.047	.477	.451	.126	.472	.006	.020
NE Domain	.390	.175	.008	.	.241	.067	.369	.088	.197	.009	.241
Separated	.146	.000	.047	.241	.	.000	.262	.360	.360	.165	.193
ParBoth	.201	.050	.477	.067	.000	.	.000	.001	.001	.049	.000
ParSP	.364	.435	.451	.369	.262	.000	.	.312	.312	.049	.165
Deceased	.383	.325	.126	.088	.360	.001	.312	.	.391	.177	.292
BOFW	**.002**	.325	.472	.197	.360	.001	.312	.391	.	.177	.292
FFOFW	.291	.308	.006	.009	.165	.049	.049	.177	.177	.	.032
FMOFW	.051	.194	.020	.241	.193	.000	.165	.292	.292	.032	.

*NED = Number of elevated domains
Confidence level = 95%

Results were presented in Table 2 of the ten (10) variables considered in this study: the number of times under probationary, gender, and both parents working abroad were found to be significant correlates of average grade. Specifically, probationary status has a significant moderate positive relationship with academic performance (r=0.601, p-value = 0.000), while gender and both parents working abroad have significant negligible relationships with academic performance (r=-0.198, p-value = 0.014, r=0.259, p-velue = 0.002). This means that these three variables may be good predictors of academic performance.

A linear regression analysis was used to further test whether the above three significant variables could be good predictors of the academic performance of students. Results are shown in Table 3. The data shows that regression model 1 has the highest predictive ability (R2 = 0.427, F-change p-value = 0.000) at the 95% level of significance. This model includes the three variables (probationary status, gender, and whether both parents are working abroad), as compared to Model 2, which consists of the variables probationary status and whether both parents are working abroad. This means that a combination of variables – probationary status, gender, and both parents working abroad – can explain 42.7% of the variance in the academic performance of the students. The regression goodness of fit test result (F-p-value = 0.000) also shows that the model is a good predictor of students' academic performance.

Table 3. Regression Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics R Square Change	F Change	df1	df2	Sig. F Change
1	0.654[a]	.0.427	0.413	0.408	0.427	29.325	3	118	.000
2	0.645[b]	0.416	0.406	0.410	0.055	11.195	1	119	.001

Model 1	Sum of Squares	df	Mean Square	F	Sig
Regression	14.654	3	4.885	29.325	0.000
Residual	19.655	118	0.167		

a. Predictors: (Constant), BOFW. Gender, Probi1

b. Predictors: (Constant), Probe1, BOFW

To determine which of the three variables mentioned above is considered the strongest predictor of academic performance, the beta coefficient was generated. Table 4 shows the results where the variable of both parents working abroad is the strongest predictor of academic performance (beta coefficient = -0.808, p-value = 0.001), followed by probationary status (beta coefficient = 0.571, p-value = 0.000). Gender has a beta coefficient of -0.119 with a p-value of 0.131, which means that when taken alone, its effect is not statistically significant. This means that for a unit change in the variable that both parents are working abroad, there is a corresponding -0.808 change in the academic performance of the student, while 0.571 is probationary status.

Table 4. Beta Coefficients

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	2,821	0.114		24.656	0.000
	Probe1	0.571	0.070	0.574	8.119	0.000
	Gender	-0.119	0.078	-0.107	-1.521	0.131
	BOFW	-0.808	0.230	-0.236	-3.384	0.001

Dependent Variable: Academic performance

The features and functionalities of the decision support system were subjected to a user's satisfaction evaluation in terms of Functionality, Reliability, Usability, Efficiency, Maintainability and Portability. The following table shows the results. Table 5 shows the summary of the rating of each respondent for each module. There were six (6) respondents and eight modules.

Table 5. Sum of Rating for Each Module per Respondent

Modules	R1	R2	R3	R4	R5	R6
1	30	27	28	28	27	27
2	29	30	28	27	28	30
3	27	28	28	28	27	30
4	30	28	26	28	27	30
5	29	28	28	27	28	30
6	30	26	27	27	27	30
7	30	24	28	27	28	30
8	30	30	28	28	27	30

Table 6 shows the overall rating for each module based on its characteristics. The sum of the ratings in Table 5 was divided by six (6), which represents the characteristics. As a result of the survey, all modules were rated excellent based on a rating scale of 1= Very Poor, 2 = Poor, 3 = Good, 4 = Very Good and 5 = Excellent.

Table 6. Average Rating

Modules	R1	R2	R3	R4	R5	R6	Over-all
1	5	4.5	4.67	4.67	4.5	4.5	4.64
2	4.83	5	4.67	4.5	4.67	5	4.78
3	4.5	4.67	4.67	4.67	4.5	5	4.67
4	5	4.67	4.33	4.67	4.5	5	4.7
5	4.83	4.67	4.67	4.5	4.67	5	4.72
6	5	4.33	4.5	4.5	4.5	5	4.64
7	5	4	4.67	4.5	4.67	5	4.64
8	5	5	4.67	4.67	4.5	5	4.81

3.3 Implications

The study's findings suggest the potential for positive change in enhancing the services offered by the Guidance office to both students and their academic performance. The theories employed in the study served as the framework for designing the entire system, defining the role of technology in task execution and data transformation into useful information. By integrating these theories, the study identified opportunities for utilizing technology and developing applications to provide valuable information in the field of guidance and counseling.

The result reveals that, for the most part, there is no statistically significant difference in the schooling outcomes of children whose biological parents are together in the Philippines and those with parents working internationally. Any observed differences tend to be positive, particularly when only the mother is abroad. The negative effects of migration are evident only when both parents are overseas. This information is crucial, as it suggests that having a mother or father abroad implies that the dropout rate in school is not significantly higher compared to other factors, such as financial needs or domestic responsibilities. The study also indicates gender differences, where the presence of both parents abroad is associated with a lower level of educational attainment for sons compared to daughters [14]. The same result regarding both parents working abroad was also derived from the study of Albarico *et al.* [15].

It was found that school-related factors, home-related factors, and personal circumstances significantly influence students' academic performance. Specifically concerning home-related aspects, students show slight agreement across six different indicators. Notably, two of these indicators suggest that students with both parents working abroad and not residing with them may be affected [15].

Students who received social support from their families were more likely to be satisfied with their lives, have a positive mood, and have a negative mood than those who did not. As a result, the students should have full support from their families and those around them while they study, as it has a positive impact on the students' academic performance [16].

The same result was also derived from the study of Briones *et. al.* [17], wherein one of the factors identified was the parenting style, in which the respondents chose the supportive style of parenting, which got the highest number of responses. Which means that having a supportive parent who is described as having a warm and supportive-child relationship is a crucial determinant of positive outcomes in school and for socioeconomically diverse youth, as are parents who provide moral support and give motivation to the children [17].

The results of the study conducted by Husaini and Shakur [16] revealed a significant impact on students' performance. This is because male and female learners exhibit different learning rates and behaviors. It was observed that the majority of female students display more positive behavior than male students and are more efficient in completing tasks [16]. Similar findings were reported in the study of Al-Alawi *et al.* [18], which identified six (6) factors affecting student academic achievement, including gender. According to the experimental results, these factors negatively affect academic performance and significantly contribute to students' academic probation [18].

4. Conclusion and Recommendations

The study concludes that the model, particularly the C4.5 model, effectively determines the relationship between family dynamics, psychological test results, and academic performance, as demonstrated by the accuracy of the test results. Additionally, Linear Regression analysis confirms the variables' suitability as predictors of academic success, supporting the study's hypothesis. The survey assessing the features and functionalities of the decision support system against ISO 9126-1 standards yielded excellent ratings across all criteria, with respondents suggesting additional modules like statistical reporting. Consequently, the study effectively addresses the identified issues.

The researchers recommend implementing a decision support system in the Guidance office to enhance services and aid in student guidance and monitoring. This system would facilitate information generation for decision-making, enabling earlier interventions and proactive service provision. During the study, notable areas for improvement were identified. First, enhance the counseling process by incorporating common observations gathered after interpreting psychological test results into the system. This data could help identify common issues among students and inform program development. Second, create a checklist of symptoms related to the psychological test domains to enhance classification. Lastly, consider additional factors that may influence student academic performance as predictors.

Acknowledgement

We extend our heartfelt gratitude to the Director and Guidance Counselors of Central Philippine University for their unwavering support, invaluable assistance, and remarkable cooperation throughout the research process.

We extend our deepest gratitude to Prof. Anrem J. Balontong for his unwavering support, invaluable feedback, and tireless assistance in the technical aspects of the research. His continuous motivation and encouragement from the inception of the research conceptualization to its completion have been instrumental in its success.

Dr. Mary O' T. Penetrante for her invaluable guidance and assistance throughout the study, particularly in the quantitative aspect of the research. Her expertise and support have been invaluable in shaping the methodology and analysis of this study.

We express our sincere appreciation to Prof. Guy V. Arellano and Resty S. Galido for graciously sharing their expertise and insights in the technical aspects of the research, which greatly enriched its quality and depth.

We also extend our gratitude to the students of the College of Computer Studies for their active participation in the research activities, which significantly contributed to the success of this endeavor.

Their support, assistance, and contributions have been instrumental in the realization of this research endeavor.

References

[1] L. R. Vinluan, "*The Use of ICT in School Guidance: Attitudes and Practices of Guidance Counselors in Metro Manila, the Philippines*," International Journal for the Advancement of Counselling, vol. 33, no. 1, March 2011, pp. 22–36, doi: 10.1007/s10447-010-9110-4.

[2] D. J. Power, "*Decision Support Systems: Concepts and Resources for Decision Support Systems: Concepts and Resources for Managers Managers*," Quorum Books, Westport, Connecticut, United States, 2002, ISBN: 9781567204971, www.scholarworks.uni.edu/facbook/67.

[3] D. J. Power, F. Burstein, R. Sharda, "*Reflections on the Past and Future of Decision Support Systems: Perspective of Eleven Pioneers*," Decision Support - An Examination of the DSS Discipline, 2011, pp. 25–48. doi: 10.1007/978-1-4419-6181-5_2.

[4] D. Goodhue, "*The Model Underlying the Measurement of the Impacts of the IIC on the End-users*," Journal of the American Society for Information Science, vol. 48, no. 5, May 1997, pp. 449–453, doi: 10.1002/(SICI)1097-4571(199705)48:5<449::AID-ASI10>3.0.CO;2-U.

[5] C. Grider, "Foundations of Cognitive Theory: A Concise Review," www.eric.ed.gov/?id=ED372324 (Accessed March 10, 2024).

[6] P. Bhandari, "*What Is Quantitative Research? Definition, Uses & Methods*" www.scribbr.com/methodology/quantitative-research/ (Accessed March 10, 2024).

[7] M. M. Mazid, A. B. M. S. Ali, and K. S. Tickle, "*Improved C4.5 Algorithm for Rule Based Classification*," in Proc. 9th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases, University of Cambridge, Cambridge, UK, February 2010, pp. 296-301.

[8] V. S. Moertini, "*Towards The Use Of C4.5 Algorithm For Classifying Banking Dataset*," Integral, vol. 8, no. 2, October 2003, pp. 105-116, ISSN: 1410-1335.

[9] H. Jantan, A. R. Hamdan, Z. A. Othman, "*Human Talent Prediction in HRM using C4.5 Classification Algorithm*," International Journal on Computer Science and Engineering, vol. 2, no. 8, November 2010, pp. 2526-2534.

[10] J. Han, M. Kamber, J. Pei, "*Data Mining: Concepts and Techniques*," 3rd Edition, The Morgan Kaufmann Series in Data Management Systems," Elsevier, Waltham, MA, USA, 2011.

[11] R. Ratra, P. Gulia, and N. S. Gill, "*Performance Analysis of Classification Techniques in Data Mining using WEKA*," in Proc. Proceedings of the International Conference on Innovative Computing & Communication, SSRN, July 2021, doi: 10.2139/ssrn.3879610.

[12] D. M. Lane, D. Scott, M. Hebl, R. Guerra, D. Osherson, H. Zimmer, "*Introduction to Statistics*," www.onlinestatbook.com/Online_Statistics_Education.pdf (Accessed March 10, 2024).

[13] ESSI-SCOPE Quality Characteristics and their application, "*ISO 9126: The Standard of Reference*," www.win.tue.nl/~wstomv/edu/2ip30/references/9126ref.html (Accessed March 10, 2024)

[14] M. J. B. Arguillas, L. Williams, "*The Impact of Parents' Overseas Employment on Educational Outcomes of Filipino Children*," International Migration Review, vol. 44, no. 2, July 2018, pp. 300-319, 2010, doi: 10.1111/j.1747-7379.2010.00807.x.

[15] A. G. Albarico, R. P. Blas, A. P. Cruz, G. M. Enriquez, S. Ildefonso, "*Factors Affecting Senior High School Students Poor Academic Performance*," International Research Journal of Modernization in Engineering Technology and Science, vol. 5, no. 4, April 2023, pp. 5565–5573, doi: 10.56726/IRJMETS37089.

[16] Y. Al Husaini, N. Syufiza, A. Shukor, Y. N. Said, A. Husaini, "*Factors Affecting Students' Academic Performance: A Review*," Social Science Journal, vol. 12, no. 6, 2022.

[17] S. K. F. Briones, R. J. R. Dagamac, J. D. David, C. A. B. Landerio, "*Factors Affecting the Students' Scholastic Performance: A Survey Study*," Indonesian Journal of Educational Research and Technology, vol. 2, no. 2, January 2022, pp. 97-102, doi: 10.17509/ijert.v2i2.41394.

[18] L. Al-Alawi, J. Al Shaqsi, A. Tarhini, and A. S. Al-Busaidi, "*Using Machine Learning to Predict Factors Affecting Academic Performance: The Case of College Students on Academic Probation*," Education and Information Technologies, vol. 28, no. 10, March 2023, pp. 12407-12432, Oct. 2023, doi: 10.1007/s10639-023-11700-0.