A Data Mining Approach to Understanding Financial Literacy Knowledge and Behavioral Patterns among Tertiary Students

Maricris M. Usita[1]* and Virgie Liza Luriban[2]
[1]College of Arts, Sciences, and Technology, Occidental Mindoro State College
San Jose, Philippines 5100
[2]School of Accountancy, Occidental Mindoro State College
San Jose, Philippines 5100
[1]maricrisusita954@gmail.com; [2]virgieliza@omsc.ph.education

Received: 22nd September 2024/ **Revised:** 24th December 2024/ **Accepted:** 2nd January 2025

How to Cite: Usita, M. M., & Luriban, V. L. (2025). A data mining approach to understanding financial literacy knowledge and behavioral patterns among tertiary students. *ComTech: Computer, Mathematics and Engineering Applications, 16*(1), 13−22. https://doi.org/10.21512/comtech.v16i1.12221

***Abstract* -** The research sought to data mine the financial literacy of tertiary students to evaluate and pinpoint deficiencies in their financial knowledge, measure the degree of financial goal-setting and budgeting practices, and determine their primary sources of financial advice. The data were gathered through validated questionnaires and disseminated through surveys. The research focused on tertiary students with 316 valid responses for analysis. The research used data mining techniques under the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology to identify students' financial literacy. Furthermore, a comprehensive analysis using Ms Excel, Statistical Package for Social Sciences (SPSS), and Waikato Environment for Knowledge Analysis (WEKA) reveals significant disparities in financial literacy among various fields of study and course levels. The investigation highlights essential financial behaviors, such as credit card utilization, saving patterns, and budgeting strategies, while revealing deficiencies in formal financial education. The analysis highlights the necessity for specialized financial literacy initiatives in educational programs to bridge knowledge deficiencies and encourage proficient budgeting and goal-setting techniques. The results offer practical guidance for educators, policymakers, and higher education institutions to improve students' financial well-being, in line with Sustainable Development Goals (SDGs) focused on poverty alleviation and economic development. The research advocates for financial literacy programs in the school curriculum and emphasizes enhancing student participation in workshops. Higher education institutions must provide well-structured financial advice and support services. Lastly, Future studies should delve deeper into socioeconomic factors to improve predictive models and intervention strategies.

Keywords: financial literacy knowledge, behavioral patterns, tertiary students

I. INTRODUCTION

Finances are a natural part of society, as money makes the world turn. In today's rapidly changing financial landscape, the demand that young people, especially university students, become financially literate is relevant. Financial literacy involves a basic knowledge of money that helps people to make well-informed choices in life (Lusardi & Mitchell, 2017). However, many surveys have shown disturbing levels of financial illiteracy among young students, which often leads to poor monetary decisions and increased debt burdens, causing long-run economic instability (Feng et al., 2019). Financial literacy is more than knowing how to invest. It focuses on individuals' skills and ability to effectively utilize financial resources to enhance long-term financial well-being (Kumar et al., 2022). Furthermore, assessing these skill sets is a must as it helps people in the present and future.

This significance has been revealed through a need for more financial awareness in developing nations (Grohmann et al., 2021). Financial literacy is crucial for economic growth and personal financial stability where individuals are less likely to make informed decisions about saving, investing, borrowing, and managing debt. Current tools and methodologies for evaluating financial knowledge are often insufficient or not finely tuned to capture individual understanding across diverse populations. In addition, measuring and predicting financial literacy is limited, highlighting the need for accurate estimation to allocate intervention programs to less financially literate groups (Rudd et al., 2022). This limitation poses a challenge in designing and implementing targeted educational interventions. Without precise evaluation mechanisms, it becomes difficult to identify those most in need of support, leading to misallocation of intervention programs.

Sustainable Development Goal 8 also involves people who have financial knowledge, management, and economic development (Johri et al., 2023) to make intelligent decisions about what jobs they want and investment opportunities that ensure economic stability and growth. The research underlying objective is to support these global goals while at the same time promoting a financially insular society through enhancing financial literacy among students in higher learning institutions. As a result, the research findings help bridge the gap towards achieving financial literacy by aligning with key SDGs, leading to a well-informed generation of empowered youth. Including interdisciplinary perspectives, including data-driven approaches, can provide an in-depth knowledge-based understanding of complex issues such as financial literacy, thereby reinforcing the relevance and significance of this research (Koohang et al., 2023).

It requires some educational strategies because financial education is crucial for individual and societal well-being. Incorrect decisions can lead to economic and social issues, emphasizing the need for an adequate financial foundation from an early age (Ruiz-Palomo et al., 2023). The research impact should prioritize improving financial literacy to achieve the United Nations' Sustainable Development Goals (SDGs) (Zaimovic et al., 2023). The research examines SDG 1: No Poverty, focusing on teaching students strategies to avoid poverty traps and build financial stability for the future. This initiative fosters financial inclusion, narrows the gender gap, reduces poverty, enhances well-being, promotes equitable access to education, advances gender equality, drives economic growth, and encourages responsible consumption and production (Zaimovic et al., 2023). Furthermore, it eradicates obstacles to quality education in line with SDG 4, which requires financial literacy as part of a holistic curriculum that prepares students for adult life's survival challenges. It is emphasized that entrepreneurial finance should be included in educational programs to promote financial literacy (Pham et al., 2020).

Then, data and patterns can be seen in any aspect of human life and can harnessed in many ways. As the dawn of technology comes about, data mining soon follows. In the last three decades, data mining has been a method for detecting trends within extensive datasets and information (Da Costa & Cabral, 2022). Data mining has many branches: classification, clustering, regression, and prediction. A tried-and-true type of data mining is CRISPS-DM, better known as Cross-Industry Standard Process for Data Mining. CRISP-DM serves as a recognized and universally applicable standard process model for data mining initiatives, offering a systematic framework for the development and execution of projects across diverse industries (Schröer et al., 2021). Furthermore, the steps of CRISP-DM include understanding the research, data collection, data preparation, modeling, evaluation, and deployment (Martínez-Plumed et al., 2021).

The research aims to understand the problem of the low financial literacy rates. Data are collected using surveys to assess the financial literacy levels across different demographics, including age groups, gender, academic years, and courses of tertiary students. The research analyzes the data and creates an appropriate model for it. Drawing valuable conclusions can be used to model and create effective programs in the curriculum by considering students' knowledge, behavior, and attitude towards money.

II. METHODS

A data mining life cycle is the core methodology used in the research to analyze and extract meaningful insights from the data. CRISP-DM, a well-established data mining method, guides the process. The CRISP adapts the CRISP-DM process for idea mining, enabling idea generation from unstructured textual data using Dynamic Topic modeling, unsupervised machine learning, and statistical analysis (Ayele, 2020). Integrating other methodologies helps to understand financial literacy information better.

As shown in Figure 1, the research utilizes the CRISP-DM framework to analyze financial literacy among undergraduate students. The methodology comprises six iterative stages: business understanding, data understanding, data preparation, modeling, evaluation, and deployment. The CRISP-DM framework ensures a structured, repeatable, and reliable process for extracting meaningful insights from data.

The business understanding phase involves defining research objectives, identifying aspects of financial literacy, and determining desired outcomes. The investigation aims to explore the aspects of financial literacy when looking at how students budget their allowances, what information they have about financial products, and which of them prefer saving or investing. As asserted, accurate and dependable insights arise from proper data collection and handling, especially in challenging fields like finance learning (Naghib et al., 2023). Consequently, it facilitates the creation of preliminary stages that lay a foundation for the data mining process and enables the research to focus on relevant queries and cover significant gaps in financial literacy (Kamalov et al., 2023). Aligning artificial intelligence and data mining efforts with sustainable education goals requires a clear business

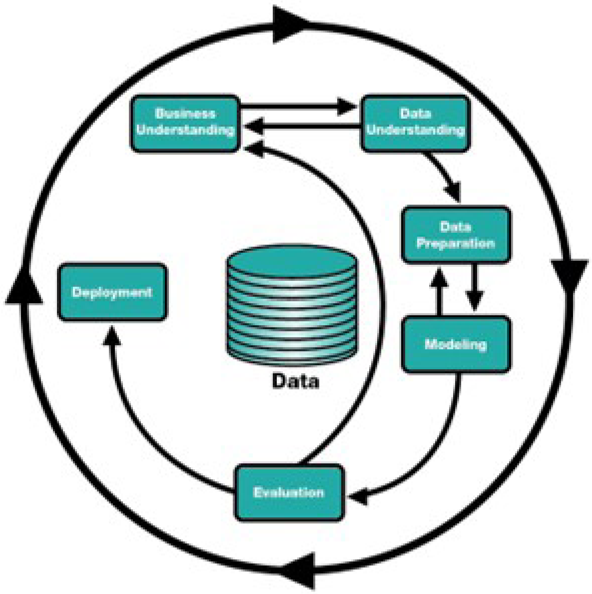

Figure 1 Data Mining Process in Cross-Industry Standard Process for Data Mining (CRISP-DM)

understanding (Kamalov et al., 2023).

The previous researchers have also shown that identifying important variables is necessary for developing reliable models and enhancing feature selection to improve machine learning performance across various fields (Schmidt et al., 2019). The data understanding phase involves collecting relevant data from undergraduate students, assessing its quality, and exploring potential variables influencing financial literacy. It also involves cleaning, transforming, and formatting the dataset to ensure compatibility with analytical tools and algorithms.

The data are collected using validated questionnaires shared through online survey platforms. The survey concentrates on essential aspects of financial literacy, including budgeting, savings, and making financial decisions. The process relies heavily on translating data-driven insights into actionable solutions for real-world issues (Naghib et al., 2023). The collected data undergo pre-processing to deal with missing values and normalize the data. The research stresses how to pre-process data for an enduring dataset, mainly when working through heterogeneous and intricate designs (Herzalla et al., 2023). The researchers follow ethical guidelines, ensuring informed consent and maintaining participant anonymity throughout the data collection. The researchers also conduct pilot testing on the questionnaire to confirm its reliability and alignment with the study's goals.

Additionally, the data collected and gathered at this stage follows the finance-related behavior of students, such as their knowledge and attitudes, through a survey. Data pre-processing should undergo a diligent examination to eliminate errors and biases that may jeopardize the outputs of subsequent stages in analysis, such as data mining, classification, and clustering, among others (Page et al., 2021). Identifying potential issues, such as missing values or biases in the data, which can affect the analysis, is definite properly, including its quality, structure, and patterns. It is an important step in forming initial hypotheses and guiding subsequent data mining. Finally, analyzing machine learning classifiers points out that understanding one's data impacts how predictive models work (Gupta et al., 2022). The research focuses on tertiary students from different programs. Researchers distribute 500 questionnaires online and retrieved 450 responses. After cleaning the data, the researchers retain 316 valid responses for analysis.

The modeling phase uses the Apriori Algorithm to uncover significant association rules, analyzing relationships among financial literacy behaviors and demographic factors. These processes involve cleaning datasets, converting them, arranging them during data preparation to make them suitable for analysis, normalizing the scores for financial literacy, categorizing their spending behavior, and encoding qualitative responses in terms of this study. Models must be suited to particular data types and required learning outcomes (Andersen, 2023). This phase is crucial because it ensures that model usage data is correct, consistent, and pertinent to research goals. The researchers clean data using Microsoft Excel (Ms Excel) before importing the Converted Comma Separated Values (CSV) file into Statistical Package for the Social Sciences (SPSS) to compute frequency, percentage, mean, standard deviation, and Analysis of Variance (ANOVA). The researchers apply the Apriori algorithm and K-Means clustering using Waikato Environment for Knowledge Analysis

(WEKA) for further analysis. Accurate and consistent data are crucial in disseminating implementation efforts, ensuring reliability in subsequent analyses and applications and aligning with the desired outcomes (Khadjesari et al., 2020).

The data modeling involves employing several data mining methods on the prepared data to determine patterns, trends, and relationships. The research aims to identify factors influencing financial literacy levels, predict individuals' financial literacy scores, and segment groups based on financial literacy behaviors. Therefore, the focus on introducing an ongoing review and modification of management strategy based on changing societal and educational expectations is valid (Tõnurist & Hanson, 2020). It focuses on achieving measurable outcomes, such as high prediction accuracy and actionable insights to guide targeted educational interventions. The process refers to as modeling mainly involves transforming raw data into actionable insights through which the researchers can identify latent factors contributing to financial literacy among higher education students. It demonstrates how modeling and data mining techniques evaluate financial literacy knowledge and behavioral patterns (Chambers et al., 2017).

The evaluation phase assesses the model's performance and reliability, aligning findings with the research objectives. The results are meticulously reviewed following the modeling process to determine their alignment with the research objectives. Researchers evaluate the performance of the financial literacy modeling using various metrics. This inclusive assessment adheres to recommendations on cognitive transformation technologies that ensure trustworthiness and actionable insights from the models (Vermesan et al., 2017).

The deployment phase provides actionable insights, identifying key areas for financial literacy improvement and recommending targeted educational interventions for tertiary students. The simplicity of data dimensions, identified by identifying the primary variables, is said to enhance machine learning performance (Cai et al., 2018). The researchers implement the financial literacy model into a user-friendly system or interface, such as a financial literacy assessment tool, to enable practical applications. Previous research has tried to map single-cell data (Lotfollahi et al., 2022). The research analyzes data to identify and understand the most important variables leading to high financial literacy levels. In contrast, researchers examine the cluster centroids to determine the attributes of different student groups.

The researchers continuously monitor the model's performance to ensure its effectiveness and relevance. By tracking its accuracy and relevance, the researchers identify and address the need for retraining. The researchers share insights derived from the model with stakeholders, enabling informed decision-making. These insights guide the creation of targeted interventions, such as financial education programs designed for specific demographic groups, ensuring that the modeling process directly improves financial literacy outcomes. According to Yeboah et al. (2024), education policy reforms significantly improve financial literacy scores in schools, with variations in effectiveness based on socioeconomic status and geographic regions. There is a need for effective use of research findings, particularly useful in international agencies with practical application, and can lead to significant policy shifts and amendments in education (Badache et al., 2023).

III. RESULTS AND DISCUSSIONS

Table 1 contains results for the classification of variables that include gender, age, course, and year level. The average classifying score is approximately 2.08 years old with a standard deviation of 0.39, indicating that ages are not too far apart. In addition, the mean gender score is 1.78 with a standard deviation of 0.44, which implies the narrowest range of financial literacy scores across genders. Moreover, based on these findings, it can be concluded that there are widespread financial literacy levels among courses, as indicated by a high mean score of 4.77 with a significant standard deviation (2.60).

The research also finds that differences in financial literacy levels can be observed between students at different stages within the year level. The analysis indicates that there may be differences in financial literacy across demographic categories, such as course, while demonstrating more variety than others, such as age and gender. The role of data mining for financial literacy involves helping people to make the right financial decisions and developing focused financial education programs customized to different sections of the population. Data mining greatly assists today's complex financial markets by simplifying and

Table 1 Financial Literacy Level among Different Age Groups, Gender, Courses, and Year Level

Group	N	Mean	Std. Deviation
Age	316	2.08	0.39
Gender	316	1.78	0.44
Course	316	4.77	2.60
Year Level	316	2.78	1.45

clarifying ideas about money. Their exhaustive study on generative AI for healthcare demonstrates how sophisticated data-driven approaches, such as data mining, among others, lead to better decision-making processes with directed solutions to diverse population needs (Sai et al., 2024). Like medicine, allowing the generation of specific educational interventions that best target unique poverty-related issues affects different demographics.

Table 2 presents an ANOVA test conducted to determine if there is difference in financial literacy levels across age groups. There are significant differences between groups ($F(14, 301) = 3.73$, $p < 0.001$). The research shows that age influences financial literacy considerably since Between Groups Sum of Squares (7.08) exceeds Within Groups Sum of Squares (40.79). Researchers find that gender significantly affects financial literacy scores ($F(14, 301) = 3.19$, $p < 0.001$). The large F-value shows that gender differences explain much variance in financial literacy concerning Between Groups Sum of Squares (7.92), which outweighs Within Groups Sum of Squares (53.45). The ANOVA test for the variable 'course' demonstrates significant differences in financial literacy levels across different courses ($F(14, 301) = 2.61$, $p = 0.001$). The Between Groups Sum of Squares (231.56) shows that differences among various courses are significant compared to the Within Groups Sum of Squares (1905.11), proving that course selection significantly impacts financial literacy levels. Year-level ANOVA results also show significant differences in financial literacy levels among years ($F(14, 301) = 2.46$, $p = 0.003$). A significant F-value implies that year level is important regarding financial literacy because the Between Groups Sum of Squares (67.93) is much larger than the Within Groups Sum of Squares (594.57).

Table 3 provides descriptive statistics about two main variables: students' understanding of basic finance concepts and involvement in financial education programs. The average score for knowledge about financial literacy stands at 3.92, with a standard deviation equal to 0.48. It indicates relatively high mean values, showing that most students have a good foundation on matters concerning basic finance concepts on an average basis. The variability of the mean is moderate, as indicated by its standard deviation, and implies that although many students have a good idea of financial literacy, a few do not know much about it. The lowest score is 1.80, while the highest is 5.00. The result represents the different financial literacy levels among these students since scores almost stretch from nearly one to five on this scale. Learners utilize the financial education program meagerly, as reflected by an average result of 1.30 and a standard deviation of 0.46. The low Standard Deviation (SD) value shows that responses are tightly clustered around the central tendency. Considering the minimum and maximum values (1 and 2), researchers conclude that participation in finance teaching is either low or medium.

Table 2 ANOVA Results for Demographic Profile of the Respondents

Group	**Source**	**Sum of Squares**	**Degrees of Freedom (df)**	**Mean Square**	**F-Value**	**Sig.**
Age	Between Groups	7.08	14	0.51	3.73	0.000
	Within Groups	40.79	301	0.14		
	Total	47.86	315			
Gender	Between Groups	7.92	14	0.57	3.19	0.000
	Within Groups	53.45	301	0.18		
	Total	61.37	315			
Course	Between Groups	231.56	14	16.54	2.61	0.001
	Within Groups	1905.11	301	6.33		
	Total	2136.67	315			
Year Level	Between Groups	67.93	14	4.85	2.46	0.003
	Within Groups	594.57	301	1.98		
	Total	662.49	315			

Table 3 Students' Understanding of Basic Financial Concepts and Their Participation in Financial Education

Variable	**Mean**	**Std. Deviation**	**Minimum**	**Maximum**
Financial Literacy	3.92	0.48	1.80	5.00
Financial Education	1.30	0.46	1.00	2.00

Table 4 reveals student respondents' distributions/variability in financial behaviors related to financial literacy and education issues. It shows an excellent level of understanding concerning finance, as depicted by the high mean for financial literacy knowledge. In contrast, a low mean score indicates limited formal education in this area. Respondents have different financial behaviors, as seen from the standard deviations showing moderate variability in both measures.

In Table 5, the Apriori algorithm identifies significant item sets and association rules in a database at a minimum support threshold of 0.45 and a confidence level of 0.9. The algorithm performs 11 cycles to refine the itemset to identify frequent patterns. The analysis reveals several notable association rules, including those with both credit card usage and a savings account, which are likely to have outstanding debts. Similarly, attending a financial workshop is linked to respondents' high financial knowledge and outstanding debts (Carpena et al., 2019). Association rules use metrics like confidence, lift, leverage, support, and conviction to measure the associations' strength, relevance, and reliability. High confidence values reveal strong connections, while lift values above 1 confirm positive correlations. Leverage and support quantify the dataset's applicability, while conviction values confirm the reliability of these relationships. Individuals with credit card usage have a 96% likelihood of having outstanding debts, indicating the need for targeted financial counseling. According to Białowolski (2019), consumer confidence influences household financial decisions, with higher confidence leading to increased borrowing for durables and mortgages and lower confidence increasing consumption savings. Financial education initiatives can benefit from focusing on specific age demographics. Additional associations, such as gender to financial knowledge and savings accounts to age, underscore demographic trends that school administrators can use to shape interventions. These results highlight the potential of data-driven insights to enhance financial literacy initiatives, develop targeted programs, and improve financial literacy and informed decision-making.

Additionally, financial workshops often trump comprehensive profitability analysis during decision-making. Hence, participants make financially unsafe choices without considering the costs involved (Kotsila et al., 2022). The same pattern points out that financial literacy programs can cause borrowing among people, just like anti-corruption campaigns that can bring about

Table 4 Financial Behaviors of the Respondents Using Association Rule Mining

Variable	Mean	Std. Deviation	Minimum	Maximum
Financial Literacy Knowledge	3.92	0.48	1.80	5.00
Financial Education	1.30	0.46	1.00	2.00

Table 5 Significant Association rules Identified Using the Apriori Algorithm.

Rule	Confidence	Lift	Leverage	Support	Conviction
Credit Card Usage=1 150 ==> Outstanding Debts=2 144	0.96	1.12	0.05	[15]	3.11
Important Financial Knowledge=5 Financial Engagement Workshop=2 160 ==> Age=2 148	0.93	1.04	0.02	[5]	1.32
Important Financial Knowledge=5 Financial Engagement Workshop=2 160 ==> Outstanding Debts=2 148	0.93	1.08	0.04	[11]	1.79
Age=2 Financial Decision Confidence=4 Financial Engagement Workshop=2 158 ==> Outstanding Debts=2 146	0.92	1.08	0.03	[10]	1.76
Financial Decision Confidence=4 Financial Engagement Workshop=2 178 ==> Outstanding Debts=2 164	0.92	1.08	0.04	[11]	1.72
Gender=2 Impt. Financial Knowledge=5 162 ==> Age=2 148	0.91	1.02	0.01	[3]	1.16
Financial Decision Confidence=4 Important Financial Knowledge=5 184 ==> Age=2 168	0.91	1.02	0.01	[3]	1.16
Savings Account=1 179 ==> Age=2 163	0.91	1.02	0.01	[3]	1.13
Important Financial Knowledge=5 214 ==> Age=2 194	0.91	1.02	0.01	[2]	1.09
Saving Knowledge=4 168 ==> Age=2 152	0.9	1.01	0.01	[2]	1.06
Credit Card Usage =1 150 ==> Outstanding Debts=2 144	0.96	1.12	0.05	[15]	3.11

unexpected results in organizational behavior (Cao et al., 2018). Furthermore, risk assessment during the evaluation of financial interventions should consider broader socioeconomic contexts because they have profound implications for the outcomes (Dwivedi et al., 2021). Making highly confident decisions on finance-related matters and attending seminars correlate with unpaid debts. Alternatively, people aged between 18 and 25 years who are confident about their financial decisions and attend workshops will likely have outstanding debts. High financial decision confidence and workshop attendance strongly correlate with outstanding debts.

Researchers also find a strong correlation between high financial knowledge and a specific age group, with 91% of cases involving this age group demonstrating high financial knowledge. The lift value of 1.02 suggests that individuals in this age group are likelier to exhibit high financial knowledge than random chance. This finding supports the need to explore age as a factor in financial literacy and design educational interventions tailored to specific age demographics.

Table 6 displays the findings of the K-Means clustering analysis of financial decision-making attributes with a focus on six key variables: Financial Decision Confidence (FDC), Importance of Financial Knowledge (IFK), Financial Anxiety (FA), Financial Management Ease (FME), Retirement Planning Importance (RPI) and Mandatory Financial Literacy (MFL). The dataset consists of 316 instances for analysis. This process is carried out through three iterations, resulting in two clear clusters. The first cluster, which consists of 181 instances or 57% of the whole data set, has relatively higher mean values for the IFK and MFL when compared to those in the entire dataset. The second cluster contains about 135 examples, which is approximately about 43%. For example, the average numbers for IFK and MFL values in cluster 0 are equal to 5, while those in cluster 1 are equivalent to 4. The clustering algorithm only takes 0 seconds to build a model from all these data sets. After analyzing different financial aspects, most people under cluster 0 have high scores on financial knowledge and literacy. In contrast, others under cluster 1 have a representable balanced distribution between them.

Table 7 provides information on model performance. It correctly classifies 77.53% of instances and achieves a Kappa statistic of 0.63, indicating substantial agreement between predicted and actual classifications. The Mean Absolute Error and Root Mean Squared Error values are relatively small, which implies that the predictions are accurate. By extension, the Relative Absolute Error percentage shows how much is wrong compared to what happens,

Table 6 K-Means Clustering Results for Financial Decision-Making

Attribute	**Complete Data (Mean)**	**Cluster 0 (Mean)**	**Cluster 1 (Mean)**
Financial Decision Confidence (FDC)	4	4	4
Importance of Financial Knowledge (IFK)	5	5	4
Financial Anxiety (FA)	4	4	4
Financial Management Ease (FME)	3	3	3
Retirement Planning Importance (RPI)	4	5	4
Mandatory Financial Literacy (MFL)	4	5	4

Table 7 Classification Performance Metrics for Financial-Goal Setting and Budgeting Practices

Group	**Std. Deviation**
Correctly Classified Instances	245 (77.53%)
Incorrectly Classified Instances	71 (22.47%)
Kappa Statistic	0.63
Mean Absolute Error	0.08
Root Mean Squared Error	0.25
Relative Absolute Error	33.61%
Root Relative Squared Error	70.78%
Total Number of Instances	316
Ignored Class Unknown Instances	1
Correctly Classified Instances	245 (77.53%)
Correctly Classified Instances	245 (77.53%)
Incorrectly Classified Instances	71 (22.47%)

while the same attribute for Root Relative Squared Error illustrates the proportion of the errors made by estimators.

Table 8 presents the distribution of responses to preferred sources of financial advice for a sample of 316 respondents. Most respondents prefer family and friends as the most frequently used sources (28.48% and 25.63%, respectively) as their primary source of financial advice. According to Afsar et al. (2018), financial literacy and parental socialization positively influence the saving behavior of university students, leading to increased willingness to save. Moreover, based on Prempeh et al. (2024), parental guidance and peer behaviors significantly affect students' saving habits. Students with strong parental financial socialization or positive peer influence tend to save more. Also, online resources such as websites, blogs, and videos are popular, with 25% citing digital content. Traditional print media, such as books, is still relevant with 12.97%, but only 7.91% cite professional financial advisors. The data suggest a strong dependence on informal and self-guided methods for acquiring financial knowledge. Targeted initiatives should be implemented, leveraging trusted relationships and promoting access to professional financial advisory services, to improve financial literacy. This approach can foster a more informed and financially resilient society.

IV. CONCLUSIONS

The research extensively investigates financial literacy and behavior among students using various statistical methods and data mining techniques. According to research findings, financial literacy varies significantly based on demographic factors such as course of study, gender, and year of study. Financial literacy levels vary between courses, thus pointing out the role of academic majors in acquiring financial knowledge. However, age and gender show insignificantly different results in that there are no substantial variations between these groups. ANOVA tests disclose that the level of financial literacy is influenced by age, gender, course, and year. According to the descriptive statistics, students have basic finance knowledge but must register for finance classes. Conversely, association rule mining and K-Means clustering analysis reveal more intricate relationships between financial activities and personal finances exhibiting significant patterns or clusters.

Last but not least, evaluation measures for classification performance indicate a high accuracy rate of 77.53%, which implies accurate predictions regarding setting goals for financial or budgeting activities. The research limitations include its small sample size, focus on demographic variables, lack of consideration of other factors, and reliance on self-reported data. The research needs to evaluate financial literacy comprehensively, and its classification accuracy of 77.53% may not accurately reflect the full predictive performance.

Additionally, the research's self-reported data may be subject to biases, and the findings may not apply to students in different locations or cultures. Future research should consider additional variables like family income, parental education, and access to financial resources to assess their impact on financial literacy and behavior. Intervention-based research can assess the effectiveness of financial education programs by developing mobile applications for students to learn more about financial literacy. Advanced data mining techniques can improve predictive models. Expanding research into decision-making models for students' financial planning can offer insights into factors influencing practical financial decisions beyond literacy.

The research on financial literacy has limitations, including a reliance on specific demographics or regions, a static data analysis, the Apriori Algorithm's inability to capture temporal changes, and the omission of external contextual variables like cultural, economic, or educational disparities. Future research should expand the dataset to include diverse demographics, regions, and temporal data, use advanced data mining and machine learning techniques, conduct longitudinal studies to track behavioral changes over time, use causal analysis methods to clarify relationships between variables, and design and empirically test financial

Table 8 Advice Source Distribution

Financial Advice Source	Frequency	Percent (%)	Valid Percent (%)	Cumulative Percent (%)
Family	90	28.48	28.48	28.48
Friends	81	25.63	25.63	54.11
Financial Advisors	25	7.91	7.91	62.02
Online Resources	79	25.00	25.00	87.02
Books	41	12.97	12.97	100.00
Total Valid Responses	316	100.00	100.00	

literacy interventions to ensure practical application and validate the study's findings. The research helps to understand better the long-term impact of financial education programs and the relationship between variables.

AUTHOR CONTRIBUTIONS

Conceived and designed the analysis, M. M. U. and V. L. L. Collected the data, M. M. U. and V. L. L.; Contributed data or analysis tools, M. M. U. and V. L. L.; Performed the analysis, M. M. U. and V. L. L.; Wrote the paper, M. M. U.; and Sought permission from different colleges and departments, V. L. L.

DATA AVAILABILITY

The data that support the findings of the research are openly available in Figshare at https://figshare.com/articles/dataset/dataset_Financial_Literacy_xlsx/29046950?file=54492719, reference number [10.6084/m9.figshare.29046950.v1].

REFERENCES

Afsar, J., Chaudhary, G. M., Iqbal, Z., & Aamir, M. (2018). Impact of financial literacy and parental socialization on the saving behavior of university level students. *Journal of Accounting and Finance in Emergency Economies, 4*(2), 133–140. https://doi.org/10.26710/jafee.v4i2.526

Andersen, J. P. (2023). Field-level differences in paper and author characteristics across all fields of science in Web of Science, 2000–2020. *Quantitative Science Studies*, *4*(2), 394–422. https://doi.org/10.1162/qss_a_00246

Ayele, W. Y. (2020). Adapting CRISP-DM for idea mining: A data mining process for generating ideas using a textual dataset. *International Journal of Advanced Computer Sciences and Applications*, *11*(6), 20–32. https://doi.org/10.14569/IJACSA.2020.0110603

Badache, F., Kimber, L. R., & Maertens, L. (2023). Rethinking methods of investigation and international organizations. In *International organizations and research methods: An introduction* (pp. 1–15). University of Michigan Press.

Białowolski, P. (2019). Economic sentiment as a driver for household financial behavior. *Journal of Behavioral and Experimental Economics*, *80*, 59–66. https://doi.org/10.1016/j.socec.2019.03.006

Cai, J., Luo, J., Wang, S., & Yang, S. (2018). Feature selection in machine learning: A new perspective. *Neurocomputing*, *300*, 70–79. https://doi.org/10.1016/j.neucom.2017.11.077

Cao, X., Wang, Y., & Zhou, S. (2018). Anti-corruption campaigns and corporate information release in China. *Journal of Corporate Finance*, *49*, 186–203. https://doi.org/10.1016/j.jcorpfin.2018.01.008

Carpena, F., Cole, S., Shapiro, J., & Zia, B. (2019). The ABCs of financial education: Experimental evidence on attitudes, behavior, and cognitive biases. *Management Science*, *65*(1), 346–369. https://doi.org/10.1287/mnsc.2017.2819

Chambers, D., Simpson, L., Neta, G., Schwarz, U. V. T., Percy-Laurry, A., Aarons, G. A., ... & Jones, E. (2017). Proceedings from the 9th annual conference on the science of dissemination and implementation. In *Implementation Science*, *12*, 1–55. https://doi.org/10.1186/s13012-017-0575-y

Da Costa, J. F. P., & Cabral, M. (2022). Statistical methods with applications in data mining: A review of the most recent works. *Mathematics*, *10*(6), 1–22. https://doi.org/10.3390/math10060993

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., ... & Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *57*. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Feng, X., Lu, B., Song, X., & Ma, S. (2019). Financial literacy and household finances: A Bayesian two-part latent variable modeling approach. *Journal of Empirical Finance*, *51*, 119–137. https://doi.org/10.1016/j.jempfin.2019.02.002

Grohmann, A., Hübler, O., Kouwenberg, R., & Menkhoff, L. (2021). Financial literacy: Thai middle-class women do not lag behind. *Journal of Behavioral and Experimental Finance*, *31*. https://doi.org/10.1016/j.jbef.2021.100537

Gupta, K., Kumar, P., & Upadhyaya, S. (2022). Comparative analysis of different machine learning classifiers for the prediction of chronic diseases. In *Intelligent Systems-Proceedings of ICIS-2022* (pp. 1–7). https://doi.org/10.13052/rp-9788770227667

Herzalla, D., Lunardi, W. T., & Andreoni, M. (2023). TII-SSRC-23 dataset: Typological exploration of diverse traffic patterns for intrusion detection. *IEEE Access*, *11*, 118577–118594. https://doi.org/10.1109/ACCESS.2023.3319213

Johri, A., Islam, M. U., & Kamal, M. (2023). Assessment of financial literacy and its impact on investor's decisions in Saudi Arabia: A study in the context of enabling financial planning to strengthen economic development. *Discrete Dynamics in Nature and Society*, *2023*(1). https://doi.org/10.1155/2023/9932444

Kamalov, F., Santandreu Calonge, D., & Gurrib, I. (2023). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. *Sustainability*, *15*(16), 1–27. https://doi.org/10.3390/su151612451

Khadjesari, Z., Boufkhed, S., Vitoratou, S., Schatte, L., Ziemann, A., Daskalopoulou, C., ... & Hull, L. (2020). Implementation outcome instruments for use in physical healthcare settings: A systematic review. *Implementation Science*, *15*, 1–16. https://doi.org/10.1186/s13012-020-01027-6

Koohang, A., Nord, J. H., Ooi, K. B., Tan, G. W. H., Al-Emran, M., Aw, E. C. X., ... & Wong, L. W. (2023). Shaping the metaverse into reality: A holistic multidisciplinary understanding of opportunities, challenges, and avenues for future investigation. *Journal of Computer Information Systems*, *63*(3), 735–765. https://doi.org/10.1080/08874417.2023.2165197

Kotsila, P., Anguelovski, I., García-Lamarca, M., & Sekulova, F. (2022). *Injustice in urban sustainability*. Routledge. https://doi.org/10.4324/9781003221425

Kumar, S., Rao, S., Goyal, K., & Goyal, N. (2022). Journal of Behavioral and Experimental Finance: A bibliometric overview. *Journal of Behavioral and Experimental Finance*, *34*, 1–19. https://doi.org/10.1016/j.jbef.2022.100652

Lotfollahi, M., Naghipourfar, M., Luecken, M. D., Khajavi, M., Büttner, M., Wagenstetter, M., ... & Theis, F. J. (2022). Mapping single-cell data to reference atlases by transfer learning. *Nature Biotechnology*, *40*, 121–130. https://doi.org/10.1038/s41587-021-01001-7

Lusardi, A., & Mitchell, O. S. (2017). How ordinary consumers make complex economic decisions: Financial literacy and retirement readiness. *The Quarterly Journal of Finance*, *7*(3), 1–31. https://doi.org/10.1142/S2010139217500082

Martínez-Plumed, F., Contreras-Ochando, L., Ferri, C., Hernández-Orallo, J., Kull, M., Lachiche, N., ... & Flach, P. (2021). CRISP-DM twenty years later: From data mining processes to data science trajectories. *IEEE Transactions on Knowledge and Data Engineering*, *33*(8), 3048–3061. https://doi.org/10.1109/TKDE.2019.2962680

Naghib, A., Jafari Navimipour, N., Hosseinzadeh, M., & Sharifi, A. (2023). A comprehensive and systematic literature review on the big data management techniques in the Internet of Things. *Wireless Networks, 29*(3), 1085–1144. Springer US. https://doi.org/10.1007/s11276-022-03177-5

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, *2021,* 1–36. https://doi.org/10.1136/bmj.n160

Pham, T. H., Ho, M. T., Vuong, T. T., Nguyen, M. C., & Vuong, Q. H. (2020). Entrepreneurial finance: Insights from english language training market in Vietnam. *Journal of Risk and Financial Management*, *13*(5), 1–23. https://doi.org/10.3390/jrfm13050096

Prempeh, A., Osei, B., & Osei, F. (2024). (2024). Financial literacy and saving behaviour among tertiary students. *Journal of Business Management and Accounting*, *14*(1), 115–142. https://doi.org/10.32890/jbma2024.14.1.5

Rudd, D. H., Huo, H., & Xu, G. (2022). Predicting financial literacy via semi-supervised learning. In *Australasian Joint Conference on Artificial Intelligence* (pp. 304–319). Springer International Publishing. https://doi.org/10.1007/978-3-030-97546-3_25

Ruiz-Palomo, D., Galache-Laza, M. T., Cisneros-Ruiz, A. J., & García-Lopera, F. (2023). Financial knowledge of pre-university students: Effects of age and gender. *Heliyon*, *9*(4), 1–10. https://doi.org/10.1016/j.heliyon.2023.e15440

Sai, S., Gaur, A., Sai, R., Chamola, V., Guizani, M., & Rodrigues, J. J. P. C. (2024). Generative AI for transformative healthcare: A comprehensive study of emerging models, applications, case studies, and limitations. *IEEE Access*, *12*, 31078–31106. https://doi.org/10.1109/ACCESS.2024.3367715

Schmidt, J., Marques, M. R. G., Botti, S., & Marques, M. A. L. (2019). Recent advances and applications of machine learning in solid-state materials science. *Npj Computational Materials*, *5*, 1–36. https://doi.org/10.1038/s41524-019-0221-0

Schröer, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. *Procedia Computer Science*, *181*, 526–534. https://doi.org/10.1016/j.procs.2021.01.199

Tõnurist, P., & Hanson, A. (2020). Anticipatory innovation governance: Shaping the future through proactive policy making. In *OECD Working Papers on Public Governance*. OECD Publishing. https://doi.org/10.1787/cce14d80-en

Vermesan, O., Eisenhauer, M., Sundmaeker, H., Guillemin, P., Serrano, M., Tragos, E. Z., ... & Bahr, R. (2017). Internet of things cognitive transformation technology research trends and applications. In O. Vermesan & J. Bacquet (Eds.), *Cognitive hyperconnected digital transformation* (pp. 17–95). River Publishers. https://doi.org/10.1201/9781003337584-3

Yeboah, C., Acquah, A., & Shelley, K. (2024). Evaluating the impact of education policy reforms on financial literacy: A meta-analysis of financial education initiatives in U.S. schools. *International Journal of Social Policy and Education, 6*(3), 60–72. https://doi.org/10.61494/ijspe.v6n3a7

Zaimovic, A., Torlakovic, A., Arnaut-Berilo, A., Zaimovic, T., Dedovic, L., & Nuhic Meskovic, M. (2023). Mapping financial literacy: A systematic literature review of determinants and recent trends. *Sustainability*, *15*(12), 1–30. https://doi.org/10.3390/su15129358