A Comparative Analysis of Decision Tree, Random Forest, and Logistic Regression Models in Predicting Business Readiness for Digital Technology Integration

Gloria M. Ducut
College of Technology and Business, Pangasinan State University, Asingan Campus, Pangasinan, Philippines

***Abstract*—This study compared the performance of Decision Tree, Random Forest, and Logistic Regression models in predicting business readiness for digital technology integration using survey data from 400 business respondents in Pangasinan, Philippines. The analysis utilized variables related to technology utilization, perceived helpfulness, willingness to integrate technology, and challenges encountered in adopting digital marketing, e-commerce, and digital payment technologies. Business readiness was operationalized from respondents' willingness scores and classified into Ready and Less Ready categories. Descriptive results revealed very low technology utilization despite high perceived helpfulness, indicating a gap between awareness of digital technologies and their actual adoption. Lack of awareness emerged as the most frequently reported barrier, followed by data security concerns. Using an 80:20 train test split, the machine learning models achieved moderate predictive performance, with Decision Tree and Random Forest attaining the highest accuracy of 63.75%. Random Forest produced the best overall performance, achieving the highest weighted F1 score and demonstrating a more balanced classification capability than the other models. The findings highlight the potential of machine learning as a decision support tool for assessing business readiness and generating evidence-based insights that can support digital transformation planning, technology adoption strategies, and capacity building initiatives among businesses.**

***Keywords*—Business readiness; digital technology integration; machine learning**

I. Introduction

The rapid advancement of digital technologies has significantly transformed the business landscape, compelling organizations to adopt innovative tools and strategies to remain competitive. Digital transformation has become a critical driver of organizational growth, enabling businesses to improve operational efficiency, strengthen customer engagement, expand market reach, and enhance overall performance [1], [2]. As organizations increasingly integrate digital technologies such as social media platforms, e-commerce systems, and digital payment solutions into their operations, the concept of business readiness has gained considerable attention. Business readiness reflects an organization's preparedness to adopt, utilize, and sustain digital technologies through the availability of technological resources, organizational capabilities, employee competencies, and leadership support [3], [4].

Despite the recognized benefits of digital transformation, many organizations continue to face challenges in technology adoption. Small and medium-sized enterprises (SMEs), in particular, often encounter barriers related to limited awareness, inadequate digital skills, financial constraints, resistance to change, and insufficient organizational support [3],[5]. Consequently, understanding the factors that influence business readiness has become increasingly important for supporting successful digital transformation initiatives. Previous studies have emphasized that organizations with higher levels of digital readiness are better positioned to exploit digital opportunities, improve business processes, and achieve sustainable growth in dynamic business environments [1], [2], [4].

Theoretical perspectives on technology adoption provide valuable insights into the determinants of business readiness. The Technology Acceptance Model (TAM) suggests that perceived usefulness and perceived ease of use influence an individual's intention to adopt technology [6]. Similarly, the Unified Theory of Acceptance and Use of Technology (UTAUT) highlights the importance of performance expectancy, facilitating conditions, and organizational support in shaping technology adoption behavior [7]. Guided by these theoretical perspectives, technology utilization, perceived helpfulness, willingness to integrate technology, and challenges encountered are considered key determinants of organizational readiness for digital technology integration.

Recent developments in machine learning have created new opportunities for analyzing complex datasets and generating predictive insights for decision-making. Machine learning classification techniques have been widely applied in text classification, pattern recognition, predictive analytics, and decision-support systems because of their ability to identify hidden relationships and classify observations based on multiple predictor variables [8]-[11]. Various machine learning approaches, including neural networks, support vector machines, deep learning models, and ensemble learning methods, have demonstrated strong performance in classification tasks across diverse application domains [12]-[15],[18]-[21]. Furthermore, machine learning has been successfully applied to organizational and operational problems such as human resource management, production optimization, and decision-support systems, highlighting its potential for business analytics and strategic planning [16], [17], [22].

Among the commonly used classification algorithms, Decision Tree, Random Forest, and Logistic Regression have attracted considerable attention because they provide a balance between predictive performance and interpretability [23], [24], [25]. Decision Tree models generate transparent decision rules that facilitate understanding of classification outcomes, while Random Forest improves predictive accuracy through ensemble learning and reduced overfitting [25]. Logistic Regression remains a widely used statistical classification method capable of estimating the probability of categorical outcomes and explaining relationships between predictor variables and target classes [23]. These characteristics make the three algorithms suitable for analyzing survey-based datasets involving organizational and behavioral variables.

Although digital readiness and digital transformation have been extensively investigated in information systems and organizational research, the application of machine learning classification techniques for predicting business readiness remains limited. Furthermore, comparative evaluations of Decision Tree, Random Forest, and Logistic Regression models using survey-based organizational data are scarce. This limitation restricts the availability of data-driven tools that can support digital transformation planning and readiness assessment.

Addressing this gap, the present study compares the predictive performance of Decision Tree, Random Forest, and Logistic Regression models in classifying business readiness for digital technology integration. Using indicators related to technology utilization, perceived helpfulness, willingness to integrate technology, and challenges encountered, the study aims to identify the most suitable classification model and generate evidence-based insights that may support digital transformation planning and technology adoption strategies among businesses.

II. METHODOLOGY

A. Research Design

This study employed a quantitative predictive research design to evaluate business readiness for digital technology integration using machine learning techniques. Specifically, the predictive performance of Decision Tree, Random Forest, and Logistic Regression models was compared using survey-based data to determine the most accurate and reliable classification algorithm. The study focused on technology adoption in three key areas: digital marketing and social media technologies, e-commerce and online sales platforms, and digital payment solutions. Data were collected through a structured questionnaire that measured respondents' technology utilization, perceived helpfulness of digital tools, willingness to integrate technology into business operations, and challenges encountered during adoption. These variables served as predictors of business readiness and formed the basis for developing and evaluating the machine learning models, thereby providing a data-driven approach to assessing organizational preparedness for digital transformation.

B. Data Source

The dataset used in this study was derived from survey responses collected from business owners and operators in Pangasinan, Philippines. The structured questionnaire gathered information on the adoption and utilization of digital technologies, challenges encountered during technology use, willingness to integrate digital tools into business operations, and perceptions regarding the usefulness of technology in enhancing business performance. Specifically, the survey covered digital marketing technologies, including social media platforms (Facebook, Instagram, TikTok, and YouTube), blogs, Mailchimp, Search Engine Optimization (SEO), Pay-Per-Click (PPC) advertising, Content Management Systems (CMS), marketing analytics tools, and influencer marketing platforms. The dataset also included the use of e-commerce platforms such as Lazada, Shopee, Zalora, and Shein, as well as digital payment solutions including GCash, PayMaya, DragonPay, online banking transfers, and Coins.ph. These responses served as the primary data source for developing and evaluating the machine learning models used to predict business readiness for digital technology integration.

C. Respondents of the Study

The respondents of the study consisted of business owners, managers, and authorized business representatives from various enterprises in Pangasinan, Philippines. These individuals were selected because they possessed sufficient knowledge of their organizations' digital technology practices, including the adoption and utilization of digital marketing tools, e-commerce platforms, and digital payment systems. As key decision-makers or individuals directly involved in business operations, they were considered reliable sources of information regarding technology usage, challenges encountered during digital adoption, and the willingness of their businesses to integrate additional digital technologies. Only complete and valid responses were included in the analysis. Prior to model development, the dataset underwent a data cleaning process to identify and address incomplete, inconsistent, or erroneous entries, ensuring the quality and reliability of the data used for machine learning classification.

D. Variables of the Study

The dependent variable in this study was business readiness for digital technology integration. Since the survey instrument did not directly measure readiness, a readiness index was operationalized using respondents' willingness to integrate technology. This approach is grounded in the Technology Acceptance Model (TAM), which identifies behavioral intention as a key determinant of technology adoption, and the Unified Theory of Acceptance and Use of Technology (UTAUT), which emphasizes the influence of performance expectancy and facilitating conditions on technology acceptance and use [6], [7]. Accordingly, willingness to integrate technology was used as a proxy indicator of readiness because it reflects an organization's intention and preparedness to adopt digital technologies when enabling conditions are present. Guided by these theoretical frameworks, technology utilization, perceived helpfulness, and challenges encountered were selected as predictor variables because they represent important dimensions of technology acceptance, perceived value, and adoption barriers that may influence readiness for digital technology integration [6], [7]. These variables served as the primary inputs for the machine learning models used to classify business readiness and examine the factors associated with digital technology adoption.

TABLE I. VARIABLES USED IN THE MACHINE LEARNING MODELS

Variable	Description	Type	Coding
Technology Utilization	Level of use of digital technologies in business operations	Ordinal	1 = Not Utilized, 2 = Low Utilization, 3=Moderate Utilization, 4 = High Utilization, 5 = Heavily Utilized
Perceived Helpfulness	Perceived usefulness of digital technologies for business operations	Ordinal	1 = Not Helpful, 2 = Slightly Helpful, 3=Moderately Helpful, 4 = Helpful, 5 = Extremely Helpful
Willingness to Integrate Technology	Intention to adopt and integrate digital technologies	Ordinal	1 = Not Willing, 2 = Slightly Willing, 3=Moderately Willing, 4 = Very Willing, 5 = Extremely Willing
Challenges Encountered	Presence of barriers to technology adoption	Binary	0 = No, 1 = Yes
Business Readiness	Target variable derived from willingness scores	Categorical	Less Ready (< 3.00), Ready (≥ 3.00)

Table I outlines the variables used in the machine learning models for predicting business readiness for digital technology integration. Technology utilization, perceived helpfulness, and challenges encountered served as the predictor variables, while business readiness was used as the target variable. Since the survey instrument did not contain a direct measure of readiness, business readiness was derived from respondents' willingness to integrate technology and classified into two categories: Less Ready and Ready. The selection of these variables was guided by the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT), which emphasize the roles of perceived usefulness, behavioral intention, and facilitating conditions in influencing technology adoption.

E. Data Preparation

Prior to model development, the survey dataset underwent data preparation and preprocessing to ensure its quality, consistency, and suitability for machine learning analysis. The dataset was examined for missing values, duplicate records, inconsistent responses, and invalid entries, with incomplete responses excluded and minor missing values addressed through appropriate imputation techniques. Because the survey data primarily consisted of categorical and ordinal responses, these variables were transformed into numerical values to facilitate model training. Ordinal responses related to technology utilization, willingness to integrate technology, and perceived helpfulness were encoded using a five-point scale, while challenge-related variables were converted into binary values, with 1 indicating the presence of a challenge and 0 indicating its absence. Following data transformation, the dataset was reviewed to remove variables with excessive missing values, limited variability, or minimal analytical relevance. Business readiness was subsequently derived from respondents' willingness to integrate technology and classified into Ready and Less Ready categories based on the established readiness threshold. To avoid target leakage, willingness variables were used solely for the construction of the target variable and were not included as predictors during model training. Consequently, technology utilization, perceived helpfulness, and challenges encountered were retained as the final predictor variables used in the machine learning models.

F. Data Splitting

Following data preprocessing, the dataset was partitioned into training and testing sets to support model development and evaluation. The training dataset was used to train the Decision Tree, Random Forest, and Logistic Regression models, while the testing dataset was reserved for assessing their predictive performance using previously unseen observations. An 80:20 train test split was employed, with 80% of the data allocated for model training and 20% reserved for testing. This proportion is widely used in supervised machine learning studies because it provides sufficient data for model learning while maintaining an independent dataset for performance evaluation. Prior to model training, the distribution of the target variable was examined to identify potential class imbalance issues. Where necessary, appropriate measures were considered to ensure that the classification models were not overly influenced by the majority class, thereby improving the reliability and robustness of the prediction results.

G. Machine Learning Models

This study employed three supervised machine learning classification algorithms, namely Decision Tree, Random Forest, and Logistic Regression, to predict business readiness for digital technology integration. The Decision Tree model was selected because of its ability to generate clear and interpretable decision rules, making it useful for understanding how different predictor variables contribute to readiness classification. Random Forest was included as an ensemble learning technique that combines multiple decision trees to improve predictive performance and reduce the risk of overfitting. Logistic Regression was utilized as a statistical classification approach that estimates the probability of membership in a particular readiness category and provides interpretable relationships between predictor variables and classification outcomes. The inclusion of these three models allowed for a comprehensive comparison of classification performance, enabling the identification of the most suitable algorithm for predicting business readiness using survey-based data.

H. Model Training and Testing

Each classification model was trained using the training dataset, with technology utilization, perceived helpfulness, and challenges encountered serving as the predictor variables. The target variable was the business readiness classification derived from respondents' willingness to integrate technology. Following model training, the Decision Tree, Random Forest, and Logistic Regression models were evaluated using the testing dataset, which contained observations not included during the training process. This approach enabled the assessment of each model's predictive performance on previously unseen data and provided an indication of its ability to generalize beyond the training dataset. By separating model training from model evaluation, the study minimized the risk of overestimating model performance and ensured a more objective comparison of the classification algorithms.

I. Evaluation Metrics

The predictive performance of the Decision Tree, Random Forest, and Logistic Regression models was evaluated using standard classification metrics, including accuracy, precision, recall, F1 score, and the confusion matrix. Accuracy measured the overall proportion of correctly classified observations, while precision assessed the proportion of correct predictions within a given readiness category. Recall measured the ability of the model to correctly identify actual readiness classifications, and the F1 score provided a balanced assessment of precision and recall. The confusion matrix offered a detailed summary of correct and incorrect classifications across the readiness categories, enabling a more comprehensive evaluation of model behavior. Collectively, these metrics were used to compare the performance of the three machine learning algorithms and identify the most effective model for predicting business readiness for digital technology integration.

J. Comparative Analysis

Following model evaluation, a comparative analysis was conducted to assess the performance of the Decision Tree, Random Forest, and Logistic Regression models using accuracy, precision, recall, F1 score, and confusion matrix results. The model demonstrating the strongest overall performance across these metrics was identified as the most suitable approach for predicting business readiness for digital technology integration. In addition to predictive performance, model interpretability was also considered to support practical decision making. Accordingly, the models were compared not only in terms of classification accuracy but also in their ability to provide insights into how technology utilization, perceived helpfulness, and challenges encountered influence business readiness. This approach enabled the evaluation of both the predictive effectiveness and practical applicability of the machine learning models.

K. Ethical Statement

Participation was voluntary, and informed consent was obtained from all respondents prior to data collection. Respondents were assured that their responses would remain confidential and would be used solely for academic and research purposes. No personally identifiable information was collected or reported. All data were analyzed and presented in aggregated form to protect respondent privacy and comply with applicable ethical and data protection guidelines.

III. Results and Discussion

A. Descriptive Analysis of Business Readiness and Technology Adoption

The analysis was based on the raw survey dataset consisting of 400 valid respondent records collected from businesses in Pangasinan, Philippines. The survey measured key dimensions of digital technology adoption, including technology utilization, challenges encountered, willingness to integrate technology, and perceived helpfulness of digital marketing, e-commerce, and digital payment tools. These variables were aligned with the study's conceptual framework and served as the primary inputs for the machine learning models. Since the dataset did not contain a direct measure of business readiness, a readiness index was operationalized using respondents' willingness to integrate digital technologies. Consistent with technology adoption theories that identify behavioral intention as a strong predictor of actual technology use [6], [7], readiness was derived from the average willingness scores across the technology categories included in the survey. Respondents with an average willingness score of 3.00 or higher were classified as Ready, while those with scores below 3.00 were classified as Less Ready. This classification approach provided a practical and theoretically grounded method for generating the target variable used in the machine learning analysis.

TABLE II. Technology Utilization, Willingness to Integrate, and Perceived Helpfulness by Technology Category.

Technology Category	Mean Utilization	Mean Willingness	Mean Helpfulness
Digital Marketing and Social Media	1.12	2.79	4.38
E-commerce and Online Sales	1	3.03	4.56
Digital Payment Solutions	1.56	3.25	4.31

Table II presents the technology utilization, willingness to integrate, and perceived helpfulness of digital technologies across three major categories: digital marketing and social media, e-commerce and online sales, and digital payment solutions. The results revealed that digital payment solutions recorded the highest mean utilization score (1.56), indicating that these technologies were more widely adopted than the other categories. This finding suggests that businesses tend to prioritize digital technologies that provide immediate operational benefits, such as faster transactions, improved payment convenience, and enhanced customer service. The relatively high utilization of GCash and other digital payment platforms reflects the growing acceptance of cashless transactions and the practical value of digital payment systems in supporting day-to-day business activities.

In contrast, e-commerce and online sales technologies exhibited the lowest mean utilization score (1.00), indicating limited adoption among the surveyed businesses. Despite this low level of utilization, e-commerce received the highest perceived helpfulness score (4.56), demonstrating that respondents recognized its potential to expand market reach, improve customer access, and increase business opportunities. This disparity between perceived helpfulness and actual utilization suggests that businesses may be aware of the benefits of e-commerce but face barriers that prevent its full adoption. Similar observations have been reported in digital transformation studies, which emphasize that awareness of technology benefits does not automatically translate into implementation due to constraints related to resources, digital competencies, organizational readiness, and facilitating conditions [1]-[5].

Digital payment solutions also achieved the highest willingness to integrate score (3.25), followed by e-commerce technologies (3.03). These findings indicate that businesses are more prepared to adopt technologies that offer visible and immediate advantages. According to the Technology Acceptance Model, perceived usefulness significantly influences behavioral intention toward technology adoption [6]. Likewise, the Unified Theory of Acceptance and Use of Technology highlights the role of performance expectancy and facilitating conditions in shaping adoption behavior [7]. The relatively higher willingness scores observed for digital payment and e-commerce technologies suggest that businesses perceive these tools as beneficial and potentially valuable for improving operational efficiency and competitiveness.

The findings further reveal a notable gap between technology utilization and perceived helpfulness across all categories. Although respondents generally viewed digital technologies as useful, actual adoption remained limited. This pattern suggests that readiness for digital transformation is influenced by factors beyond perceived usefulness alone. Previous studies have identified organizational capabilities, digital literacy, leadership support, technological infrastructure, and strategic readiness as important determinants of successful digital transformation [1]–[5]. Consequently, initiatives aimed at improving awareness, digital skills, and access to technology resources may help bridge the gap between positive perceptions and actual technology adoption. Overall, the results indicate that while businesses recognize the value of digital technologies, further support and capacity-building efforts are needed to enhance readiness and promote broader digital technology integration.

TABLE III. MOST UTILIZED TECHNOLOGIES

Rank	Technology	Mean Utilization
1	GCash	3.71
2	Facebook	2.44
3	Reels	1.14
4	Online Banking Transfer	1.07
5	TikTok	1.04

Table III shows the most utilized technologies. The most utilized tool was GCash, with a mean utilization score of 3.71, followed by Facebook with 2.44. This suggests that respondents were most familiar with technologies that directly support payment transactions and customer communication. Most other tools had very low utilization, with several platforms recording a mean score of 1.00, including Lazada, Shopee, Zalora, Shein, PayMaya, DragonPay, and Coins.ph.

TABLE IV. HIGHEST WILLINGNESS TO INTEGRATE TECHNOLOGIES

Rank	Technology	Mean Willingness
1	GCash	4.64
2	Facebook	4.15
3	Online Banking Transfer	3.8
4	YouTube	3.73
5	Shopee / Lazada	3.5

Table IV presents the technologies with the highest willingness to integrate among the respondents. The respondents showed the highest willingness to integrate GCash, with a mean willingness score of 4.64, followed by Facebook with 4.15. This means that digital payment and social media tools were the most acceptable technologies among the respondents. On the other hand, lower willingness scores were observed for Coins.ph, Content Management Systems, Marketing Analytics, and Pay-Per-Click Advertising, suggesting that these tools may require more awareness, training, or support before businesses become ready to adopt them.

TABLE V. CHALLENGES ENCOUNTERED

Challenge	Frequency	Percentage
Lack of awareness	6,040	65.65%
Data security concerns	1,514	16.46%
Inadequate skills, training, and support	446	4.85%
Resistance to change among employees	443	4.82%
Lack of financial resources	439	4.77%
Connectivity or infrastructure issues	293	3.18%
Complexity of the technology/system	25	0.27%

[a] **Note:** Frequencies and percentages were calculated from the total number of coded challenge responses because respondents were allowed to select multiple challenges.

Table V illustrates the challenges encountered by businesses in adopting digital technologies. Lack of awareness emerged as the most frequently reported barrier, accounting for 65.65% of all coded responses, indicating that many businesses may have limited knowledge of available digital tools and their potential benefits for business operations. Data security concerns ranked second at 16.46%, suggesting that issues related to privacy, trust, and the protection of digital transactions also influence technology adoption decisions. In contrast, the low frequency of responses related to technology complexity indicates that businesses do not necessarily perceive digital tools as difficult to use. Rather, the findings suggest that limited awareness and concerns regarding security represent the primary obstacles to digital technology integration. These results highlight the importance of digital literacy initiatives, awareness campaigns, and cybersecurity education programs to improve business readiness and encourage wider adoption of digital technologies.

TABLE VI. BUSINESS READINESS CLASSIFICATION

Readiness Classification	Frequency	Percentage
Less Ready	240	60.00%
Ready	160	40.00%
Total	400	100.00%

Table VI presents the business readiness classification of the respondents. Based on the established readiness threshold, 240 respondents (60.00%) were classified as Less Ready, while 160 respondents (40.00%) were classified as Ready for digital technology integration. These findings indicate that although respondents generally recognized the benefits and usefulness of digital technologies, the majority had readiness scores below the threshold required for classification as Ready. The results suggest that positive perceptions of digital technologies do not necessarily translate into readiness for adoption, highlighting the need for greater awareness, capacity building, and support mechanisms to facilitate digital technology integration among businesses.

TABLE VII. MACHINE LEARNING MODEL RESULTS

Model	Accuracy	Weighted Precision	Weighted Recall	Weighted F1-Score
Decision Tree	0.6375	0.625	0.6375	0.6129
Random Forest	0.6375	0.6325	0.6375	0.6341
Logistic Regression	0.6125	0.6356	0.6125	0.6163

Table VII provides the performance of the machine learning models using an 80:20 train-test split. Decision Tree and Random Forest achieved the highest accuracy of 63.75%, while Random Forest obtained the highest weighted F1 score (0.6341), indicating the best overall classification performance. Logistic Regression achieved a slightly lower accuracy of 61.25% but recorded the highest weighted precision (0.6356). The moderate predictive performance of the models reflects the complexity of business readiness, which is influenced by multiple organizational and contextual factors beyond those captured in the survey. Despite relying on perception-based variables, the results demonstrate that machine learning, particularly Random Forest, can effectively identify readiness patterns and support data-driven decision-making for digital technology integration. Future studies may improve predictive performance by incorporating additional variables related to organizational capabilities, digital literacy, internet accessibility, and prior technology experience [1]-[5], [25].

TABLE VIII. CONFUSION MATRIX RESULTS

Model	Actual Less Ready Correctly Predicted	Actual Less Ready Predicted Ready	Actual Ready Predicted Less Ready	Actual Ready Correctly Predicted
Decision Tree	40	8	21	11
Random Forest	35	13	16	16
Logistic Regression	28	20	11	21

Table VIII summarizes the confusion matrix results of the three machine learning models. The Decision Tree model correctly classified most Less Ready businesses but showed limited effectiveness in identifying Ready businesses, indicating a tendency toward the majority class. Logistic Regression demonstrated better performance in identifying Ready businesses; however, it produced a higher number of false positive classifications, which reduced its overall accuracy. In contrast, the Random Forest model achieved a more balanced classification performance by accurately identifying both Less Ready and Ready businesses. Although its accuracy was comparable to that of the Decision Tree model, its higher weighted F1 score reflected a better balance between precision and recall across the two readiness categories. These findings suggest that Random Forest provided the most reliable and practical predictive performance for assessing business readiness for digital technology integration, making it the most suitable model for supporting data driven decision making and digital transformation planning.

B. Discussion

The findings revealed a notable gap between the perceived usefulness of digital technologies and their actual utilization among businesses in Pangasinan. Although respondents generally recognized the value of digital technologies in improving business operations, customer engagement, and market reach, actual technology utilization remained relatively low. This disparity suggests that positive perceptions alone do not guarantee technology adoption. According to the Technology Acceptance Model (TAM), perceived usefulness influences behavioral intention, which subsequently affects technology use [6]. Similarly, the Unified Theory of Acceptance and Use of Technology (UTAUT) highlights the importance of facilitating conditions and organizational support in translating intention into actual adoption behavior [7]. Thus, while businesses may acknowledge the benefits of digital technologies, existing barriers may prevent them from fully integrating these tools into their operations.

Among the technology categories examined, digital payment solutions demonstrated the highest level of adoption and readiness, largely driven by the widespread use of GCash and other mobile payment platforms. This finding suggests that businesses are more inclined to adopt technologies that offer immediate operational benefits, including transaction convenience, improved efficiency, and reduced cash handling risks. Previous studies have shown that organizations prioritize digital technologies that provide clear business value and contribute directly to operational performance and competitiveness [1],[2]. In contrast, the relatively low utilization of e-commerce platforms indicates that many businesses have not fully embraced online selling despite recognizing its usefulness. This finding may reflect gaps in digital competencies, awareness, organizational capabilities, and readiness, which have been identified as important determinants of successful digital transformation [3]-[5].

The challenge analysis further supports this interpretation. The predominance of lack of awareness as a reported barrier suggests that informational and knowledge-related constraints may be more significant than purely financial or technical limitations. This finding underscores the importance of digital literacy programs, technology demonstrations, business mentoring, and capacity building initiatives to enhance readiness for digital technology integration. Previous research has emphasized that digital readiness extends beyond technological infrastructure and includes leadership support, organizational capabilities, employee competencies, and the ability to adapt to changing business environments [3]-[5]. Consequently, efforts aimed at improving awareness and digital skills may substantially increase the likelihood of successful technology adoption among businesses.

With respect to predictive modeling, Random Forest emerged as the best-performing classification algorithm, achieving the highest overall accuracy and weighted F1 score among the evaluated models. The superior performance of Random Forest may be attributed to its ensemble learning mechanism, which combines multiple decision trees to improve predictive accuracy and reduce overfitting [25]. Although Decision Tree and Logistic Regression provided greater interpretability and clearer explanations of classification outcomes, Random Forest demonstrated a stronger ability to accurately classify both Ready and Less Ready businesses. These findings suggest that machine learning techniques can effectively predict business readiness using variables related to technology utilization, perceived helpfulness, and challenges encountered. The ability of the models to identify readiness patterns demonstrates the usefulness of machine learning as a decision support tool for assessing digital transformation readiness and technology adoption behavior. Consistent with previous studies, machine learning algorithms have proven effective in classification, prediction, and decision support applications across various organizational and operational contexts [16], [17], [22]. Although the predictive performance observed in this study was moderate, the results indicate that technology adoption characteristics can provide meaningful insights into business readiness and may support evidence-based planning, capacity building, and digital transformation initiatives among businesses.

The results further indicate that business readiness is influenced by factors beyond those captured in the present dataset. Although willingness to integrate technology served as a practical proxy for readiness, digital transformation literature recognizes readiness as a multidimensional construct that includes organizational resources, technological infrastructure, leadership commitment, digital competencies, and strategic orientation [1]-[5]. While TAM and UTAUT support the use of behavioral intention as a predictor of technology adoption [6], [7], future research should develop and validate more comprehensive readiness frameworks that incorporate technological, organizational, financial, and human resource dimensions. Such an approach would strengthen the conceptual measurement of readiness while improving the predictive capability of machine learning models.

Several limitations should be considered when interpreting the findings. First, the study employed a single 80:20 train-test split for model evaluation, which may influence performance estimates because classification results can vary depending on how the data are partitioned. Future studies should employ more robust validation approaches, such as k-fold cross-validation, stratified k-fold validation, or repeated random subsampling, to generate more stable and reliable performance estimates. Second, although the dataset consisted of 400 valid responses, only a subset was used for testing, which may affect the generalizability of the reported performance metrics. Third, the study focused on businesses within a specific geographical context, potentially limiting the applicability of the findings to other regions or business sectors. Future studies should therefore utilize larger and more diverse datasets to improve model robustness and external validity.

Another limitation is that feature importance analysis was not performed for the Random Forest model. Although Random Forest achieved the highest predictive performance, the relative contribution of individual predictor variables was not quantified. Consequently, the study focused primarily on model comparison and predictive accuracy rather than model explainability. Future research should incorporate feature importance analysis and model interpretation techniques such as permutation importance and SHAP (Shapley Additive Explanations) to identify the variables that most strongly influence business readiness. Such analyses would provide deeper insights into technology adoption behavior and support the development of more targeted and evidence based digital transformation strategies.

Guo (2025) [26] compared Decision Tree, Logistic Regression, AdaBoost, and Random Forest models in predicting corporate credit default and reported that the improved Random Forest model achieved superior predictive performance, with an accuracy of 94.2%, recall of 93.5%, and F1-score of 92.8%. The study attributed the strong performance of Random Forest to its ensemble structure, which enhances classification accuracy while reducing the risk of overfitting commonly associated with single-tree models. Although the predictive performance observed in the present study was more moderate due to the complexity and perception-based nature of business readiness data, the superior performance of Random Forest similarly demonstrates its capability to capture complex relationships among predictor variables and generate more balanced classifications than Decision Tree and Logistic Regression models. These findings further support the suitability of Random Forest as a practical decision-support tool for organizational assessment, technology adoption studies, and business analytics applications.

The findings of the present study are consistent with recent machine learning research that highlights the effectiveness of ensemble learning methods for classification tasks. Guo (2025) compared Decision Tree, Logistic Regression, AdaBoost, and Random Forest models in predicting corporate credit default and reported that the improved Random Forest model achieved superior predictive performance, with an accuracy of 94.2%, recall of 93.5%, and F1-score of 92.8%. The study attributed the strong performance of Random Forest to its ensemble structure, which enhances classification accuracy while reducing the risk of overfitting commonly associated with single-tree models. Although the predictive performance observed in the present study was more moderate due to the complexity and perception-based nature of business readiness data, the superior performance of Random Forest similarly demonstrates its capability to capture complex relationships among predictor variables and generate more balanced classifications than Decision Tree and Logistic Regression models. These findings further support the suitability of Random Forest as a practical decision-support tool for organizational assessment, technology adoption studies, and business analytics applications.

IV. CONCLUSION AND RECOMMENDATIONS

A. Conclusion

The study concluded that business readiness for digital technology integration among businesses in Pangasinan remains in the developmental stage. Although respondents generally perceived digital technologies as highly beneficial for improving business operations, customer engagement, and market expansion, actual utilization levels remained relatively low, indicating a gap between awareness of benefits and practical adoption. Digital payment solutions, particularly GCash, exhibited the highest levels of utilization and willingness to integrate, suggesting that businesses are more likely to adopt technologies that provide immediate and tangible operational advantages. In contrast, e-commerce platforms and several digital marketing technologies remained underutilized despite their perceived usefulness. Lack of awareness emerged as the primary barrier to adoption, followed by concerns regarding data security, highlighting the need for digital literacy initiatives, technology training, and business support programs. Among the machine learning models evaluated, Random Forest demonstrated the strongest predictive performance and was identified as the most suitable model for classifying business readiness, while Decision Tree and Logistic Regression provided valuable interpretability and baseline comparisons. Further, the findings demonstrate that machine learning can serve as an effective tool for predicting business readiness and generating evidence-based insights that support digital transformation planning and technology adoption strategies among businesses.

B. Recommendations

Businesses should strengthen digital awareness and skills through training and digital literacy programs, particularly in digital marketing, e-commerce, and digital payment systems. Given that lack of awareness was the primary barrier to adoption, local government units, business support organizations, and technology providers should collaborate in providing technical assistance and capacity-building initiatives. Future studies should utilize larger datasets, incorporate additional organizational and contextual variables, and explore advanced machine learning algorithms to improve the prediction of business readiness for digital technology integration. Future studies should also employ k-fold cross-validation and feature importance analysis to improve model robustness and enhance the interpretability of machine learning predictions.

ACKNOWLEDGMENT

The author gratefully acknowledges the Pangasinan State University for its support in the conduct and completion of this study. The author also utilized artificial intelligence-assisted tools to improve language clarity, readability, and editorial quality during manuscript preparation. These tools were not used to generate research data, perform data analysis, interpret results, or draw conclusions. All analysis, interpretations, and final manuscript content were reviewed, verified, and approved by the author, who assumes full responsibility for the accuracy and integrity of the work.

REFERENCES

[1] F. Li, "The digital transformation of business models in the creative industries: A holistic framework and emerging trends," Technovation, vol. 92–93, pp. 102012, 2020, doi: 10.1016/j.technovation.2017.12.004.

[2] M. Verhoef, T. Broekhuizen, Y. Bart, A. Bhattacharya, J. Dong, N. Fabian, and M. Haenlein, "Digital transformation: A multidisciplinary reflection and research agenda," Journal of Business Research, vol. 122, pp. 889–901, 2021, doi: 10.1016/j.jbusres.2019.09.022.

[3] A. A. AlNuaimi, K. Singh, S. Ren, V. Budhwar, and D. Vorobyev, "Mastering digital transformation: The nexus between leadership, agility, and digital readiness," Journal of Business Research, vol. 145, pp. 636–648, 2022, doi: 10.1016/j.jbusres.2022.03.038.

[4] M. North, T. Aramburu, and C. Lorenzo, "Promoting digitally enabled growth in SMEs: The role of digital readiness and organizational capabilities," Technology in Society, vol. 72, pp. 102178, 2023, doi: 10.1016/j.techsoc.2022.102178.

[5] A. Troise, E. Corvello, A. Ghobadian, and O. O'Regan, "How can SMEs successfully navigate VUCA environment: The role of agility in the digital transformation era," Technological Forecasting and Social Change, vol. 174, pp. 121227, 2022, doi: 10.1016/j.techfore.2021.121227.

[6] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," MIS Quarterly, vol. 13, no. 3, pp. 319–340, 1989, doi: 10.2307/249008.

[7] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view," MIS Quarterly, vol. 27, no. 3, pp. 425–478, 2003, doi: 10.2307/30036540.

[8] C. C. Aggarwal and C. Zhai, "A survey of text classification algorithms," in Mining Text Data, C. C. Aggarwal and C. Zhai, Eds. Berlin/Heidelberg, Germany: Springer, 2012, pp. 163–222.

[9] C. C. Aggarwal and C. X. Zhai, Mining Text Data. Berlin/Heidelberg, Germany: Springer, 2012.

[10] F. Miao, P. Zhang, L. Jin, and H. Wu, "Chinese news text classification based on machine learning algorithm," in Proc. Int. Conf. Intelligent Human-Machine Systems and Cybernetics (IHMSC), vol. 2, 2018, pp. 48–51.

[11] M. Ikonomakis, S. Kotsiantis, and V. Tampakas, "Text classification using machine learning techniques," WSEAS Transactions on Computers, vol. 4, no. 8, pp. 966–974, 2005.

[12] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[13] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," Neural Processing Letters, vol. 9, pp. 293–300, 1999.

[14] M. Heidarysafa, K. Kowsari, D. E. Brown, K. Jafari Meimandi, and L. E. Barnes, "An improvement of data classification using random multimodel deep learning (RMDL)," International Journal of Machine Learning and Computing, vol. 8, pp. 298–310, 2018.

[15] K. Kowsari, M. Heidarysafa, D. E. Brown, K. Jafari Meimandi, and L. E. Barnes, "RMDL: Random multimodel deep learning for classification," in Proc. Int. Conf. Information System and Data Mining, Lakeland, FL, USA, Apr. 9–11, 2018, doi: 10.1145/3206098.3206111.

[16] N. Pourkhodabakhsh, M. M. Mamoudan, and A. Bozorgi-Amiri, "Effective machine learning, meta-heuristic algorithms and multi-criteria decision making to minimizing human resource turnover," Applied Intelligence, 2022, doi: 10.1007/s10489-022-04294-6.

[17] Y. Li, S. Carabelli, E. Fadda, D. Manerba, R. Tadei, and O. Terzo, "Machine learning and optimization for production rescheduling in Industry 4.0," International Journal of Advanced Manufacturing Technology, vol. 110, no. 9–10, pp. 2445–2463, 2020, doi: 10.1007/s00170-020-05850-5.

[18] O. A. Montesinos-López, A. Montesinos-López, and J. Crossa, "Fundamentals of artificial neural networks and deep learning," in Multivariate Statistical Machine Learning Methods for Genomic Prediction, Cham, Switzerland: Springer International Publishing, 2022, pp. 379–425, doi: 10.1007/978-3-030-89010-0_10.

[19] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," IEEE Signal Processing Magazine, vol. 34, no. 6, pp. 26–38, 2017, doi: 10.1109/MSP.2017.2743240.

[20] S. Lai, L. Xu, K. Liu, and J. Zhao, "Recurrent convolutional neural networks for text classification," in Proc. Twenty-Ninth AAAI Conf. Artificial Intelligence, Austin, TX, USA, Jan. 25–30, 2015, vol. 333, pp. 2267–2273.

[21] S. Pal and S. Mitra, "Multilayer perceptron, fuzzy sets, and classification," IEEE Transactions on Neural Networks, vol. 3, no. 5, pp. 683–697, 1992, doi: 10.1109/72.159058.

[22] H. Lee, N. Aydin, Y. Choi, S. Lekhavat, and Z. Irani, "A decision support system for vessel speed decision in maritime logistics using weather archive big data," Computers & Operations Research, vol. 98, pp. 330–342, 2018, doi: 10.1016/j.cor.2017.06.005.

[23] S. Domínguez-Almendros, N. Benítez-Parejo, and A. Gonzalez-Ramirez, "Logistic regression models," Allergologia et Immunopathologia, vol. 39, no. 5, pp. 295–305, 2011, doi: 10.1016/j.aller.2011.05.002.

[24] J.-S. Chou and J. Thedja, "Metaheuristic optimization within machine learning-based classification system for early warnings related to geotechnical problems," Automation in Construction, vol. 68, pp. 65–80, 2016, doi: 10.1016/j.autcon.2016.03.015.

[25] L. Breiman, "Random forests," Machine Learning, vol. 45, pp. 5–32, 2001.

[26] K. Guo, "Corporate credit default prediction based on machine learning," in Proceedings of the 2025 International Conference on Economic Management and Big Data Application (ICEMBDA 2025), 2025.