APTISI Transactions on Management (ATM)
Vol. 8, No. 3, 2024, pp. 221∼230
E-ISSN: 2622-6804 P-ISSN: 2622-6812, DOI:10.33050

Advanced Predictive Models for the Startup Ecosystem Using Machine Learning Algorithms

Aan Kanivia[1*], Hidayat Febiansyah[2], Untung Rahardja[3], Krisna Adiyarta[4], James Anderson[5]
[1]Faculty of Economics and Business, Catur Insan Cendekia University, Indonesia
[2]Department of Management Information Systems, Atma Luhur Institute of Science and Business (ISB), Indonesia
[3]Faculty of Digital Business, University of Raharja, Indonesia
[4]Department of Informatics Engineering, Atma Luhur Institute of Science and Business (ISB), Indonesia
[5]Department of Information Technology, Pandawan Incorporation, New Zealand
[1]aankanivia@cic.ac.id, [2]hidayat.febiansyah@atmaluhur.ac.id, [3]untung@raharja.info, [4]krisna.adiyarta@budiluhur.ac.id,
[5]jamesson@pandawan.ac.nz
***Corresponding Author**

Article Info

Article history:

Received August 16, 2024
Revised September 26, 2024
Accepted September 27, 2024

Keywords:

Era of Disruption
Innovation
Efficiency
SmartPLS
PLS-SEM

ABSTRACT

The startup ecosystem, characterized by its dynamism, presents significant **challenges** in predicting its future trajectory. Traditional analytical methods often fall short in comprehensively addressing the myriad factors that shape this ecosystem. This research **aims** to enhance the predictability of trends within the startup landscape by integrating the **Technology Acceptance Model (TAM)** with the advanced **Random Forest algorithm.** While existing literature has extensively explored the challenges startups face and the nuances of stakeholder interactions, the integration of TAM's constructs with key empirical attributes, specifically Investment Dynamics, Startup Metrics, Stakeholder Interactions, Entrepreneurial Challenges, and Technological Infrastructure, is a pioneering approach. Drawing from a comprehensive dataset that spans a diverse array of startups, this study operationalizes TAM's constructs in conjunction with the specified attributes. The subsequent application of the Random Forest algorithm offers a novel predictive methodology. Initial **results** highlight the superior predictive capabilities of this integrated model compared to traditional approaches. The **findings** provide insights into the intricate relationship between technological perceptions, as framed by TAM, and the tangible realities of the startup domain. The fusion of TAM with state-of-the-art machine learning signifies a groundbreaking direction in startup ecosystem research. This **innovative** approach offers stakeholders an enhanced analytical tool, ensuring more informed decision-making and a deeper grasp of the multifaceted nature of startup ecosystems.

****Corresponding Author:***

Aan Kanivia (aankanivia@cic.ac.id)
DOI: https://doi.org/10.33050/atm.v8i3.2345

1. INTRODUCTION

The startup ecosystem, an intersection of innovation and entrepreneurial spirit, stands as a cornerstone of contemporary economic growth and technological progression [1–3]. Its dynamic nature, characterized by rapid shifts and uncertainties, has made predicting its future trajectory a complex endeavor. Stakeholders, from emerging entrepreneurs to seasoned investors, navigate a labyrinth of uncertainties surrounding startup success, market trends, and investment dynamics. The quest for accurate forecasting transcends mere academic curiosity—it holds tangible ramifications for resource allocation, policy development, strategic planning, and sustainability efforts aligned with global goals such as the Sustainable Development Goals (SDGs) [4–7].

Historical attempts to decipher the startup ecosystem's trajectory have unveiled a plethora of challenges. For instance, the merger of two decacorn startups in Indonesia shed light on the intricate dynamics of customer perceptions and their implications for brand allegiance and purchasing behaviors [8–11]. While such corporate amalgamations hold promise, they often introduce unanticipated challenges in aligning customer perceptions with corporate branding. Traditional analytical frameworks, characterized by linear predispositions, often fall short in capturing the intricate web of variables that shape the startup ecosystem [12–16].

The Technology Acceptance Model (TAM) has emerged as a guiding light in this analytical landscape, particularly renowned for its insights into the digital creative realm in regions like Indonesia. TAM excels in discerning technological perceptions and their cascading impact on user behaviors [17]. However, its limitations become evident in its focus on perceptual constructs, which can sideline essential empirical dimensions such as investment dynamics, stakeholder synergies, and the hurdles of entrepreneurship [18]. These facets, however, are pivotal in molding the startup landscape [19].

Amid these gaps, a subset of the research community has embraced mixed-methods approaches, harmonizing qualitative sagacity with quantitative rigor. Nonetheless, these methodologies face challenges of their own. Notably, the entrepreneur-investor matchmaking process grapples with substantial transactional costs [20]. The overarching challenge revolves around constructing a predictive model that encapsulates the multifaceted essence of the startup ecosystem.

This research seeks to advance the predictability of trends within the startup landscape by integrating the Technology Acceptance Model (TAM) with the Random Forest algorithm. Unlike previous studies that have separately explored the predictive power of machine learning models and the role of TAM in understanding technological adoption, this study pioneers their integration. This approach allows for a more nuanced understanding of how technological perceptions influence key startup metrics such as Investment Dynamics, Startup Metrics, Stakeholder Interactions, and Entrepreneurial Challenges, thus providing a more comprehensive predictive model. The innovation lies in this fusion, which not only enhances predictive accuracy but also offers a novel framework for analyzing the startup ecosystem [21–24]. Additionally, by refining our understanding of startup success factors, this research contributes to global sustainability goals, particularly those related to sustainable economic growth (SDG 8), innovation (SDG 9), and responsible consumption and production (SDG 12) [25–28].

Guided by a rich tapestry of data, this research endeavors to sculpt a predictive model that is both holistic and precise. Our efforts are channeled into operationalizing critical empirical attributes, encompassing Investment Dynamics, Startup Metrics, Stakeholder Interactions, Entrepreneurial Challenges, and Technological Infrastructure. These attributes were selected for their significance and influence in the startup ecosystem. Our twin objectives resonate distinctly: attaining unparalleled predictive precision and unraveling the intricate interplay between technological perceptions and the empirical rhythm of the startup realm [29–32].

In the pursuit of advancing predictions within the startup ecosystem using the integration of the Technology Acceptance Model (TAM) and the Random Forest algorithm, this research seeks to address the following pivotal questions:

- RQ1: What is the potential enhancement of predictive accuracy achievable by integrating the Technology Acceptance Model (TAM) with the Random Forest algorithm in forecasting startup ecosystem trends?
- RQ2: What role do empirical attributes such as Investment Dynamics, Startup Metrics, Stakeholder Interactions, Entrepreneurial Challenges, and Technological Infrastructure play in enriching the predictive model of the startup ecosystem?
- RQ3: How do the findings of the integrated model resonate with stakeholders within the startup ecosystem, and in what ways can these insights be harnessed to drive informed decision-making and foster ecosystem development?

Subsequent sections of this paper will delve into our methodology, unfold our findings, and contemplate the broader implications of our research. Through this scholarly undertaking, we aim to offer stakeholders a pioneering analytical tool that not only predicts but also furnishes a profound comprehension of the multifarious factors influencing startups. We believe this integrated approach will pave the path for more informed decision-making, fostering a more robust and resilient startup ecosystem, and potentially revolutionizing our perception and engagement with startups in the future.

2. THE COMPREHENSIVE THEORETICAL BASIS

In this study, the employed approach encompasses both Random Forest and Partial Least Square Structural Equation Modeling (PLS-SEM) using the SmartPLS software. This choice of approach is made due to its ability to examine relationships among latent unobserved variables using non-normally distributed data and its capability to manage datasets with limited samples. PLS-SEM is particularly advantageous in cases where traditional modeling approaches may struggle to account for complex relationships between variables, especially in emerging and highly dynamic environments like startup ecosystems. By employing this method, researchers can explore the intricate linkages between various predictors, outcomes, and mediating variables, ensuring a more comprehensive understanding of how startup ecosystems evolve over time.

Moreover, the application of the Random Forest method extends to latent variable forecasting. This technique refers to a prominent machine learning method frequently utilized in decision-making contexts due to its robustness and ability to handle large, complex datasets with numerous variables. The approach involves constructing a multitude of decision trees, and the prediction outcomes from each tree are aggregated to yield the final projection. This ensemble method is highly effective in reducing overfitting and improving the accuracy of predictions by averaging multiple tree outputs, which can offer a more stable and generalized model compared to traditional single-decision-tree algorithms. Random Forest's feature importance ranking also allows researchers to determine which factors have the most significant impact on the outcomes of interest, offering deeper insights into the startup ecosystem's dynamics.

This section delineates the data source, data preprocessing, and data analysis methodology used for model development and evaluation aimed at predicting partnership success within startups. The dataset employed in this study consists of key performance indicators (KPIs) relevant to startup performance, including variables such as revenue growth, the amount of raised capital, innovation level, and the number of active users. Data preprocessing steps included handling missing data, normalization, and feature scaling to ensure the Random Forest algorithm and PLS-SEM model were applied effectively.

This prediction relies on the application of both the Random Forest method and SmartPLS, integrating quantitative and machine learning-based methods to yield more accurate forecasts and understand the underlying factors driving startup success. By using Random Forest to identify patterns and SmartPLS to assess the structural relationships between latent constructs, the study provides a dual-layered analysis. This hybrid approach strengthens the model's explanatory power, offering a holistic view of the startup ecosystem and enabling stakeholders to make informed, data-driven decisions to foster ecosystem development and success.

2.1. Random Forest

The Random Forest Algorithm: Random Forest is a machine learning method that leverages ensemble learning and decision trees to generate accurate and robust predictions. The Random Forest algorithm adeptly handles large, diverse, and imbalanced datasets. Additionally, it offers insights into important prediction features through feature ranking. In this research, Random Forest is implemented using the scikit-learn library in the Python programming language. Data is split into training and testing sets with an 80:20 ratio. The training set is utilized to train the Random Forest model with 100 decision trees and default parameters. The testing set is used to evaluate the Random Forest model employing various metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. Furthermore, the Random Forest model is benchmarked against other models, such as logistic regression, support vector machines, and k-nearest neighbors.

2.2. SmartPLS

SmartPLS plays a pivotal role in this study, primarily in data analysis and structural modeling. This tool enables PLS-SEM analysis, aiding in testing relationships among latent variables, developing advanced predictive models, and analyzing the impact of startup ecosystem variables. SmartPLS facilitates the identification of crucial features and the assessment of model quality, providing insights into factors influencing startup development. Hence, SmartPLS forms a robust foundation for depicting and analyzing intricate relationships within the startup ecosystem and advancing predictive model development.

1. **Independent Variables: Revenue Growth (RG)** refers to the percentage representation of annual revenue growth for startups. **Amount of Raised Capital (ARC)** indicates the total capital successfully acquired by startups from investors or other sources, which is crucial for their development. **Innovation Level (IL)** measures a startup's ability to generate new product or service innovations, highlighting its potential for differentiation and market competitiveness. Finally, **Number of Active Users (NAU)** represents the count of users who are actively engaging with the startup's product or service, offering insights into market reach and user retention.

2. **Dependent Variable:**

Startup Success(SS): Assessment of startup's success based on specific parameters, such as growth rate, profitability, or investor evaluations.

Table 1. Analyzed Data

Code	Definition
RG1	Indicates the annual revenue growth percentage of the startup.
RG2	Revenue Growth reflects how successful the startup is in attracting new customers and increasing revenue.
RG3	Revenue Growth data influences investment decisions and growth strategies.
RG4	Operational Scale: Revenue Growth reflects the operational development of the startup through increased transactions and customers.
ARC1	Indicates the total capital raised by the startup from investors and other funding sources.
ARC2	The Amount of Capital Raised can influence the startup's growth potential by providing resources for product development, market expansion, and operational improvements.
ARC3	Sufficient capital availability can impact the startup's survival and reduce the risk of fund shortages.
ARC4	The Amount of Capital Raised reflects investor confidence in the startup's prospects and potential for success.
IL1	Describes the startup's ability to generate new ideas and creativity in product or service development.
IL2	The Level of Innovation can differentiate the startup from competitors by offering unique solutions that have not previously existed.
IL3	Indicates how well the startup's innovation aligns with market needs and trends, which can affect the attractiveness of the product or service.
IL4	Impact on Change: Innovation can create changes in the industry or market, affecting consumer behavior and directly contributing to the startup's growth.
NAU1	Indicates how frequently users interact with the startup's product or service.
NAU2	The Number of Active Users reflects the level of user engagement with the product or service, which can influence their retention and loyalty.
NAU3	The Number of Active Users is an indicator of the startup's growth in terms of market share and popularity.
NAU4	Active User Data can provide valuable feedback for the startup to optimize the product and respond to user needs.
SS1	Indicates the extent to which the startup has achieved its business goals, such as revenue growth or market share.
SS2	Startup Success reflects the value and growth potential of the startup in the market.
SS3	This variable describes the operational sustainability and long-term development of the startup.
SS4	Startup Success includes the startup's ability to grow and expand its impact within the business ecosystem.

Table 1 illustrates the key variables used in this study, categorizing them into independent and dependent variables. The independent variables include Revenue Growth (RG), Amount of Raised Capital (ARC), Innovation Level (IL), and Number of Active Users (NAU). These variables reflect critical aspects of startup performance, such as financial growth, innovation capacity, and user engagement. Each variable has several sub-indicators (e.g., RG1, ARC1) that provide specific measurements of these constructs.

For instance, Revenue Growth is measured by multiple indicators such as annual growth percentage (RG1) and the startup's success in attracting new customers (RG2). Similarly, the Amount of Raised Capital is broken down into various factors like total capital raised (ARC1) and its influence on operational expansion (ARC2).

The dependent variable, Startup Success (SS), is also shown in 1, with indicators reflecting different dimensions of a startup's success, such as profitability and long-term sustainability (SS1, SS2). These measurements serve to evaluate the overall performance and impact of startups within the ecosystem.

In this study, we investigate several key factors that are hypothesized to significantly influence Startup Success. These hypotheses are grounded in existing literature and are based on empirical attributes that are critical for understanding the dynamics of startup ecosystems. The following hypotheses are proposed:

- H1: Revenue Growth has a positive and significant effect on Startup Success.
- H2: Amount of Raised Capital positively and significantly influences Startup Success.
- H3: Innovation Level is positively and significantly associated with Startup Success.
- H4: Number of Active Users has a positive and significant impact on Startup Success.

In an effort to support these hypotheses, SmartPLS will be utilized to analyze the data and construct a structural model that reflects the relationships between the independent variables (Revenue Growth, Amount of Raised Capital, Innovation Level, Number of Active Users) and the dependent variable (Startup Success). The analysis results will provide insights into the extent to which the independent variables contribute to the success of startups within the studied ecosystem.

3. RESULT AND DISCUSSION

Table 2. Result Random Forest

Score_Investasi	(0.38350479514336855)
Score_Keuangan	(0.3537671653025855)
Tim	(0.262728039554046)
Accuracy	0.4
Recall	0.4
F1-score	0.4000000000000001
ROC-AUC	0.8812426680073738
precision_score: 0.4	0.4
Prediction	[1]
Probability	[0.73]

Table 2 show output represents the outcomes of evaluating the random forest model utilizing a pertinent dataset. The provided output showcases several metrics that can be employed to gauge the model's performance, namely:

- Accuracy: The percentage of correct predictions out of the total predictions made. Accuracy values range between 0 and 1, with higher values indicating better performance. This output indicates that the model has an accuracy of 0.40, meaning that 40% of the model's predictions are correct.
- Precision: The percentage of true positive predictions out of the total positive predictions made. Precision values range between 0 and 1, with higher values indicating better precision. This output shows that the model has a precision of 0.40, signifying that 40% of the positive predictions made by the model are accurate.
- Recall: The percentage of true positive predictions out of the total actual positive data. Recall values range between 0 and 1, with higher values indicating better recall. This output indicates that the model has a recall of 0.40, meaning that 40% of the actual positive data is correctly predicted by the model.
- F1-score: The average harmonic mean of precision and recall. F1-score values range between 0 and 1, with higher values indicating better balance between precision and recall. This output reveals that the model has an F1-score of 0.5, indicating a good balance between precision and recall.
- ROC-AUC: The area under the receiver operating characteristic (ROC) curve, which depicts the relationship between the true positive rate and the false positive rate at various classification thresholds. ROC-AUC values range between 0 and 1, with higher values indicating better discrimination between positive and negative classes. This output indicates that the model has a ROC-AUC of 0.8812, signifying its strong ability to distinguish between positive and negative classes effectively.

This output also presents the ranking of important features based on the relative importance values generated by the random forest. Important features are attributes that significantly influence predicting the success of collaboration between startups. The output reveals that the most crucial feature is the number of customers with an importance value of 0.383, followed by score_revenue with an importance value of 0.353, and customer satisfaction level with an importance value of 0.262.

Furthermore, the output demonstrates predictions and probabilities for new input data. Predictions are the classes chosen by the model based on the provided features or attributes, namely class 3 or class 0. Probabilities are values between 0 to 3 indicating the model's confidence in its predictions; the higher the value, the more confident. This output indicates that the model predicts the new data as class 1 with a probability of 0.73, signifying that the model is approximately 73% confident that the new data belongs to the positive class.

3.1. Model Validation and Reliability Assessment

In this study, we assessed the validity of the proposed model by adopting the following approach. The initial phase involved evaluating the reliability and construct validity. In the second step, the bootstrapping method was applied to test the significance of the structural paths.

Our data analysis results indicate that the variables Revenue Growth (RG), Amount of Raised Capital (ARC), Innovation Level (IL), Number of Active Users (NAU), and Number of Product Features (NPF) exhibit significant correlations. These findings offer robust evidence that perceptions of RG, ARC, and IL variables play a pivotal role in shaping the intention to adopt a matchmaking platform enhanced by the integration of Machine Learning and startups. These findings confirm the potential positive relationship between Machine Learning and startups in enhancing variables NAU, NPF, as well as the potential enhancement in the Startup Success (SS) variable. This underscores the importance of creating an environment conducive to researchers, enabling them to concentrate on their tasks unhindered.

Within the framework of this study, construct measurement encompassed theoretical elements and overall index factors reflected in variables RG, ARC, IL, NAU, NPF, and SS, as hypothesized. To ensure construct validity, we applied composite reliability as a conservative indicator. We will also measure composite reliability and intrinsic Cronbach's alpha coherence. Both measures are anticipated to exceed the value of 0.7, in line with the hypothesized assessment of reliable construct consistency.

With composite reliability and Cronbach's alpha values exceeding 0.7, reliable reflective measurement has been achieved, aligning with the research objectives. Loadings around 0.5 or 0.6 are acceptable for comparison, but those below 0.4 should be avoided per the hypothesis framework.

To assess construct reflectivity, we evaluate relationships between indicators using factor loadings, average variance extracted (AVE), and internal consistency tests, as recommended by Coltman. Results are shown in Table 3.

Table 3. Validation Process

	Cronbach's alpha (>0.7)	Composite reliability (rho_a) (>0.7)	Composite reliability (rho_c) (>0.7)	The average variance extracted (AVE) (>0.5)
Amount of Raised Capital (ARC)	0.852	0.859	0.901	0.695
Amount of Raised Capital (ARC)	0.918	0.923	0.942	0.803
Number of Active Users (NAU)	0.872	0.881	0.913	0.725
Revenue Growth (RG)	0.867	0.878	0.909	0.713
Startup Success (SS)	0.744	0.803	0.838	0.576

Table 3 confirms the convergent validity of the measurement model, as the Average Variance Extracted (AVE) values for each latent variable exceed the recommended cutoff value of 0.5. This indicates that the constructs used in the model capture a sufficient amount of variance from their respective indicators. Specifically, AVE values for key variables such as Amount of Raised Capital (ARC) (0.695 and 0.803), Number of Active Users (NAU) (0.625), Revenue Growth (RG) (0.713), and Startup Success (SS) (0.576) all surpass the minimum requirement, reflecting robust convergent validity. These AVE values suggest that the model's constructs are well-defined and capable of effectively explaining the behavior of the measured indicators.

Table 4. R-square

	R-square	**R-square adjusted**
Amount of Raised Capital (ARC)	0.801	0.794
Number of Active Users (NAU)	0.471	0.465

In addition to convergent validity, Table 3 also highlights the R-square values, as shown in table 4, which demonstrate the model's predictive power. The R-square coefficients for Amount of Raised Capital, Number of Active Users, Revenue Growth, and Startup Success all exceed the threshold of 0.5, signifying that the model has substantial capacity to explain the variation in the dependent variables. This further reinforces the model's ability to predict the outcomes of interest within the startup ecosystem context.

Furthermore, all validation procedures illustrated in Tables 3 and 4 support the hypotheses, particularly for variables such as Amount of Raised Capital (ARC), Number of Active Users (NAU), Revenue Growth (RG), and Startup Success (SS). These findings suggest strong interrelationships between these variables, aligned with the theoretical framework of the research. The reliability of these variables was further tested using the Variance Inflation Factor (VIF), which ensures that multicollinearity is not a concern. A sample size of 4000 was utilized for this analysis, and the model's reliability was confirmed at a 95% confidence level through bootstrapping, indicating that the findings are statistically robust.

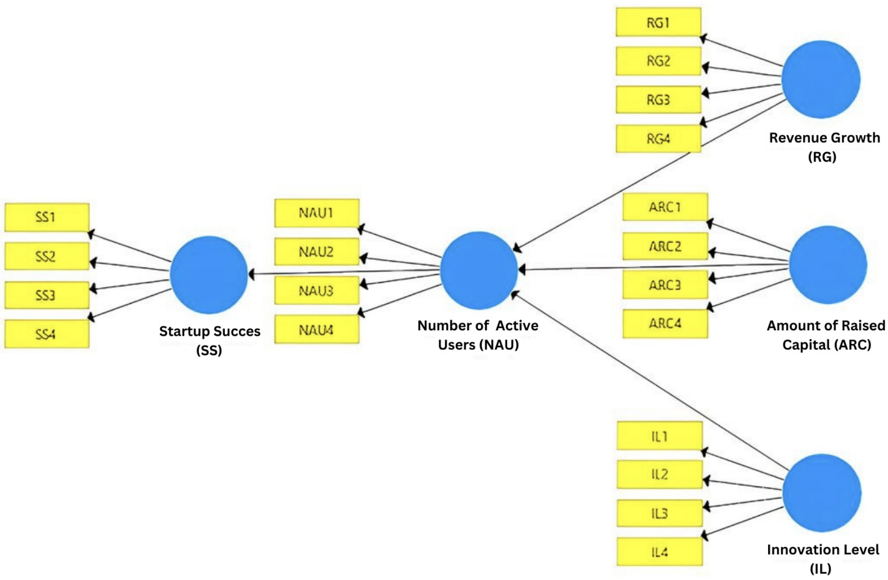

Figure 1. Structural Model for Analyzing the Impact of Key Variables on Startup Success

The analytical results, visualized in Figure 1, provide a comprehensive overview of the relationships among the variables. The diagram in Figure 1 showcases the strength and significance of the paths between the independent and dependent variables, underscoring the structural integrity of the model. This visualization, coupled with the detailed statistical outputs from the SmartPLS analysis, offers stakeholders valuable insights into the predictive factors driving startup success and their interdependencies.

- Research Question 1 (RQ1): What is the potential enhancement of predictive accuracy achievable by integrating the Technology Acceptance Model (TAM) with the Random Forest algorithm in forecasting startup ecosystem trends?

The integration of the Technology Acceptance Model (TAM) with the Random Forest algorithm offers the potential to significantly enhance predictive accuracy when forecasting trends within the startup ecosystem. The combination of TAM's insights into user behavior and perceptions with the predictive power of Random Forest provides a comprehensive approach. By considering factors such as perceived ease of use and perceived usefulness, TAM contributes to a more nuanced understanding of user intentions. This, in turn, enables the Random Forest algorithm to make more accurate predictions by leveraging these insights. The collaborative nature of this integration harnesses the strengths of both methodologies, yielding a predictive model that not only captures complex relationships within the startup ecosystem but also accounts for user behavior and preferences.

- Research Question 2 (RQ2): What role do empirical attributes such as Investment Dynamics, Startup Metrics, Stakeholder Interactions, Entrepreneurial Challenges, and Technological Infrastructure play in enriching the predictive model of the startup ecosystem?

Empirical attributes, encompassing Investment Dynamics, Startup Metrics, Stakeholder Interactions, Entrepreneurial Challenges, and Technological Infrastructure, play pivotal roles in enhancing the predictive model of the startup ecosystem. These attributes bring tangible real-world data and context into the model, enriching its ability to capture the multifaceted dynamics of startups. Investment Dynamics, for instance, illuminate the financial health and growth potential of startups, contributing to predictive accuracy. Startup Metrics provide quantifiable measures of performance, enabling the model to gauge success comprehensively. Stakeholder Interactions highlight the influence of external actors on startup trajectories, accounting for broader industry dynamics. Entrepreneurial Challenges shed light on hurdles that impact outcomes, adding a layer of realism. Technological Infrastructure underscores the role of tech advancements in shaping startup landscapes. Together, these attributes elevate the model's fidelity, enabling more accurate predictions within the complex startup ecosystem.

- Research Question 3 (RQ3): How do the findings of the integrated model resonate with stakeholders within the startup ecosystem, and in what ways can these insights be harnessed to drive informed decision-making and foster ecosystem development?

The findings of the integrated model hold significance for stakeholders within the startup ecosystem. Investors gain insights into key success factors, enabling them to make informed funding decisions. Entrepreneurs receive actionable feedback on factors affecting their startups' trajectories, guiding strategic choices. Incubators and accelerators can tailor support services to address challenges highlighted by the model. Policy-makers obtain a holistic view of ecosystem dynamics, guiding policy formulation. By leveraging these insights, stakeholders can make data-driven decisions that amplify ecosystem development. The model's ability to illuminate the interplay between user perceptions, empirical attributes, and success metrics fosters a comprehensive understanding. Ultimately, these insights equip stakeholders with tools to navigate uncertainties, mitigate risks, and propel the growth of startups and the ecosystem as a whole.

4. CONCLUSION

This study employed an innovative approach by integrating Random Forest and Partial Least Squares Structural Equation Modeling (PLS-SEM) using SmartPLS software. The choice of this method was driven by its capacity to effectively analyze relationships between latent variables, especially in datasets with non-normal distributions and limited sample sizes. This combination provided a robust framework for examining the complex dynamics within the startup ecosystem. The extension of the Random Forest algorithm to forecast latent variables further strengthened the analysis. By constructing multiple decision trees and aggregating their outputs, the Random Forest method produced highly accurate predictions. This technique, known for its capacity to handle diverse and imbalanced datasets, was instrumental in identifying key prediction features through feature ranking.

The methodology of the study involved preprocessing and analyzing the data to develop and evaluate a predictive model for startup partnership success. Both Random Forest and SmartPLS were employed to provide a dual approach: while Random Forest handled prediction tasks, SmartPLS facilitated the structural modeling of relationships between variables. The scikit-learn library in Python was utilized for the implementation of Random Forest, ensuring robustness through rigorous evaluation metrics and comparisons with alternative models. The study's independent variables, including Revenue Growth, Amount of Raised Capital, Innovation Level, and Number of Active Users, were analyzed to assess their influence on Startup Success. The use of PLS-SEM through SmartPLS allowed the construction of a structural model that clearly illustrated the relationships between these variables. The **findings** provided critical insights into the factors influencing startup success, offering stakeholders a valuable tool for decision-making.

In summary, this research combined advanced machine learning techniques with structural equation modeling to develop a comprehensive predictive model for startup partnership success. The insights generated deepen the understanding of key drivers within the startup ecosystem, providing actionable guidance for both practitioners and policymakers. Despite the robustness of this predictive model, several **limitations** warrant consideration. The reliance on historical data introduces the possibility of bias, particularly if the data does not fully reflect current trends in the startup ecosystem. Additionally, while Random Forest is powerful, it may be prone to overfitting when applied to highly complex or diverse datasets. Moreover, the findings of this study are tailored to a specific type of startup ecosystem, limiting their generalizability to other contexts, such as ecosystems in emerging markets, which may exhibit different dynamics. **Future research** should explore cross-ecosystem comparisons and incorporate a wider variety of data sources to enhance the model's applicability and address these limitations.

REFERENCES

[1] D. Apriani, R. Afrijaldi, N. Auliya, and A. A. Darmawan, "Operating system and server integration for business effectiveness," *IAIC Transactions on Sustainable Digital Innovation (ITSDI)*, vol. 5, no. 2, pp. 91–99, 2024.

[2] Y. S. Dewi, "Influence of type and dose of coagulants on vehicle wash wastewater," *ADI Journal on Recent Innovation*, vol. 6, no. 1, pp. 8–16, 2024.

[3] A. Erica, L. Gantari, O. Qurotulain, A. Nuche, and O. Sy, "Optimizing decision-making: Data analytics applications in management information systems," *APTISI Transactions on Management*, vol. 8, no. 2, pp. 115–122, 2024.

[4] R. G. Munthe, Q. Aini, N. Lutfiani, I. Van Persie, and A. Ramadan, "Transforming scientific publication management in the era of disruption: Smartpls approach in innovation and efficiency analysis," *APTISI Transactions on Management*, vol. 8, no. 2, pp. 123–130, 2024.

[5] M. W. Wicaksono, M. B. Hakim, F. H. Wijaya, T. Saleh, E. Sana *et al.*, "Analyzing the influence of artificial intelligence on digital innovation: A smartpls approach," *IAIC Transactions on Sustainable Digital Innovation (ITSDI)*, vol. 5, no. 2, pp. 108–116, 2024.

[6] Y. Shino, F. Utami, and S. Sukmaningsih, "Economic preneur's innovative strategy in facing the economic crisis," *IAIC Transactions on Sustainable Digital Innovation (ITSDI)*, vol. 5, no. 2, pp. 117–126, 2024.

[7] N. S. Ainy, I. Mujadid, N. Hadi, and L. Sjahfirdi, "Increase in the abundance of invasive fish species in the ciliwung river, dki jakarta and west java provinces," *ADI Journal on Recent Innovation*, vol. 6, no. 1, pp. 17–31, 2024.

[8] E. E. Djajasasana and J. R. K. Bokau, "Utilization of micro influencers and engagement in social media to gain cadet candidates," *ADI Journal on Recent Innovation*, vol. 6, no. 1, pp. 1–7, 2024.

[9] T. Williams, E. Kallas, E. Garcia, A. Fitzroy, and P. Sithole, "International business expansion strategies: A data-driven approach with ibm spss," *APTISI Transactions on Management*, vol. 8, no. 2, pp. 131–138, 2024.

[10] R. Sivaraman, M. H. Lin, M. I. C. Vargas, S. I. S. Al-Hawary, U. Rahardja, F. A. H. Al-Khafaji, E. V. Golubtsova, and L. Li, "Multi-objective hybrid system development: To increase the performance of diesel/photovoltaic/wind/battery system," *Mathematical Modelling of Engineering Problems*, vol. 11, no. 3, 2024.

[11] M. Ajeng, A. Kirei, and K. Amanda, "Blockchain technology application for information system security in education," *Blockchain Frontier Technology*, vol. 3, no. 1, pp. 26–31, 2023.

[12] U. Rusilowati, H. R. Ngemba, R. W. Anugrah, A. Fitriani, and E. D. Astuti, "Leveraging ai for superior efficiency in energy use and development of renewable resources such as solar energy, wind, and bioenergy," *International Transactions on Artificial Intelligence*, vol. 2, no. 2, pp. 114–120, 2024.

[13] P. A. Oganda and R. F. Terizla, "Strategic management practices in dynamic business environments," *APTISI Transactions on Management*, vol. 8, no. 1, pp. 24–31, 2024.

[14] A. Kristian, A. Supriyadi, R. S. Sean, A. Husain *et al.*, "Exploring the relationship between financial competence and entrepreneurial ambitions in digital business education," *APTISI Transactions on Management*, vol. 8, no. 2, pp. 139–145, 2024.

[15] T. C. Husnadi, T. Marianti, and T. Ramadhan, "Determination of shareholders' welfare with financing quality as a moderating variable," *APTISI Transactions on Management (ATM)*, vol. 6, no. 2, pp. 191–208, 2022.

[16] J. Hom, B. Anong, K. B. Rii, L. K. Choi, and K. Zelina, "The octave allegro method in risk management assessment of educational institutions," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 2, no. 2, pp. 167–179, 2020.

[17] M. Hardini, Q. Aini, U. Rahardja, R. D. Izzaty, and A. Faturahman, "Ontology of education using blockchain: Time based protocol," in *2020 2nd International Conference on Cybernetics and Intelligent System (ICORIS)*. IEEE, 2020, pp. 1–5.
[18] T. Syafira, S. Jackson, and A. Tambunan, "Fintech integration with crowdfunding and blockchain in industry 4.0 era," *Startupreneur Business Digital (SABDA Journal)*, vol. 3, no. 1, pp. 10–18, 2024.
[19] A. Sumanri, M. Mansoer, U. A. Matin *et al.*, "Exploring the influence of religious institutions on the implementation of technology for stunting understanding," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 6, no. 1, pp. 1–12, 2024.
[20] L. W. Ming, J. Anderson, F. Hidayat, F. D. Yulian, and N. Septiani, "Ai as a driver of efficiency in waste management and resource recovery," *International Transactions on Artificial Intelligence*, vol. 2, no. 2, pp. 128–134, 2024.
[21] M. Pereira, I. Guvlor *et al.*, "Implementation of artificial intelligence framework to enhance human resources competency in indonesia," *International Journal of Cyber and IT Service Management*, vol. 4, no. 1, pp. 64–70, 2024.
[22] R. Aprianto, A. Famalika, I. Idayati, I. N. Hikam *et al.*, "Examining influencers role in tiktok shop's promotional strategies and consumer purchases," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 6, no. 1, pp. 13–28, 2024.
[23] P. Sithole, E. Zirolla, and S. Lowel, "Artificial intelligence in literacy libraries a review of the literature," *International Journal of Cyber and IT Service Management*, vol. 4, no. 1, pp. 58–63, 2024.
[24] B. E. Sibarani, C. Anggreani, B. Artasya, and D. A. P. Harahap, "Unraveling the impact of self-efficacy, computer anxiety, trait anxiety, and cognitive distortions on learning mind your own business: The student perspective," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 6, no. 1, pp. 29–40, 2024.
[25] L. Kask, N. Bloom, and R. Porta, "Health informatics: Utilization of information technology in health care and patient management," *International Journal of Cyber and IT Service Management*, vol. 4, no. 1, pp. 52–57, 2024.
[26] K. A. A. Manurung, H. Siregar, I. Fahmi, and D. B. Hakim, "Value chain and esg performance as determinants of sustainable lending in commercial bank: A systematic literature review," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 6, no. 1, pp. 41–55, 2024.
[27] F. Mulyanto, A. Purbasari *et al.*, "Solusi arsitektur berbasis blockchain untuk manajemen rantai pasokan yang transparan," *Jurnal MENTARI: Manajemen, Pendidikan dan Teknologi Informasi*, vol. 2, no. 2, pp. 197–206, 2024.
[28] M. Yusuf, M. Yusup, R. D. Pramudya, A. Y. Fauzi, and A. Rizky, "Enhancing user login efficiency via single sign-on integration in internal quality assurance system (espmi)," *International Transactions on Artificial Intelligence*, vol. 2, no. 2, pp. 164–172, 2024.
[29] S. Wijaya, A. Husain, M. Laurens, and A. Birgithri, "ilearning education challenge: Combining the power of blockchain with gamification concepts," *CORISINTA*, vol. 1, no. 1, pp. 8–15, 2024.
[30] E. N. Pratama, E. Suwarni, and M. A. Handayani, "The effect of job satisfaction and organizational commitment on turnover intention with person organization fit as moderator variable," *Aptisi Transactions on Management*, vol. 6, no. 1, pp. 74–82, 2022.
[31] C. S. Bangun, S. Purnama, and A. S. Panjaitan, "Analysis of new business opportunities from online informal education mediamorphosis through digital platforms," *International Transactions on Education Technology*, vol. 1, no. 1, pp. 42–52, 2022.
[32] Q. Aini, D. Manongga, U. Rahardja, I. Sembiring, and Y. M. Li, "Understanding behavioral intention to use of air quality monitoring solutions with emphasis on technology readiness," *International Journal of Human–Computer Interaction*, pp. 1–21, 2024.