Modeling MSME Growth: The Impact of Growth Stages, Mentoring, and Digital Literacy Using Decision Tree and Polytomous Logistic Regression

Yanto Setiawan[1], Aries Susanty[2], Daduk Merdika Mansur[3]
Doctoral Program of Information System
Diponegoro University, Semarang, Indonesia
[1]ys@students.undip.ac.id, [2]aries.susanty@lecturer.undip.ac.id,
[3]dadukmansur@telkomuniversity.ac.id

Abstract—Micro, Small, and Medium Enterprises (MSMEs) are vital to Indonesia's economy, yet their growth is often hampered by a lack of structured support and low digital adoption. This study models MSME growth by analyzing the impact of business growth stage, mentoring models, and digital literacy. Utilizing a dataset from PT Telkom Indonesia's CSR program, this research employs a two-stage quantitative method: first, a Decision Tree algorithm classifies MSMEs into five growth stages, and second, Polytomous Logistic Regression (PLR) analyzes the factors influencing their growth. The findings reveal that the mentoring model is the most dominant predictor of MSME growth, with the 'Hybrid' model proving most effective. Conversely, digital literacy did not show a significant direct impact when mentoring was included in the model, suggesting mentoring acts as a critical catalyst. This research provides an evidence-based framework for developing more effective, tiered mentoring programs tailored to the specific needs and developmental stage of MSMEs.

Keywords— MSME Growth, Mentoring Model, Digital Literacy, Business Stage, Decision Tree, Polytomous Logistic Regression

I. INTRODUCTION

Small and Medium-sized Enterprises (SMEs) are a strategic pillar of the Indonesian economy. As of 2023, approximately 66 million MSMEs contributed 61% to the nation's Gross Domestic Product (GDP) and absorbed 97% of the total national workforce. This underscores the critical dependence of national economic sustainability on the ability of MSMEs to survive and grow in an increasingly dynamic and digital business environment. However, this immense potential is not fully realized in their actual performance.

Despite their significant number and contribution to GDP valued at IDR 8,573.89 trillion the productivity and value-added contribution of the MSME sector have remained stagnant. This condition is empirically supported by data from the Central Statistics Agency (BPS), which shows a significant decline in key MSME industries during the third quarter of 2023. For instance, the food industry (ISIC 10) experienced a 1.06% decrease in demand, while the apparel industry (ISIC 14) saw a 3.89% decline compared to the previous quarter. Such trends indicate that while MSMEs form the backbone of Indonesia’s economy, their growth trajectory remains uneven and fragile.

The growth of MSMEs generally follows a multi-stage process consisting of existence, survival, success, take-off, and resource maturity, each with distinct managerial challenges and support needs. In the Indonesian context, many MSMEs remain trapped in the early stages of existence and survival, struggling to advance toward maturity due to limited managerial capability and insufficient institutional support [1]. This developmental stagnation underscores the need for growth-stage-specific interventions, as the nature of MSME growth tends to be tiered and influenced by various strategic

drivers such as technology adoption, digital capability, and mentoring structure [2], [3].

Two of the primary root causes for this suboptimal performance are the lack of systematic and structured mentoring and the low adoption of digital technology. Although the government has promoted initiatives such as "UMKM Go Digital," by the end of 2023 only around 27 million out of 66 million MSMEs (approximately 40%) had been digitized in aspects such as marketing, transactions, or financial recording. This digital divide reflects unequal access to digital participation, which risks excluding many MSMEs from essential market, financing, and technological ecosystems necessary for sustainable growth [4].

From a capability perspective, digital literacy extends beyond basic operational skills to encompass more advanced competencies such as functional, online, and global digital literacy, which are crucial for operating effectively in digital and cross-cultural environments [5]. In this context, the low level of digital literacy among MSME actors in Indonesia poses a serious constraint on their ability to compete and adapt in an increasingly digital economy.

Equally important is the role of mentoring in supporting MSME development. Mentoring is a structured process in which an experienced mentor provides guidance to a less experienced mentee, and its effectiveness depends on factors such as frequency, duration, and delivery mode whether offline, online, or hybrid [6]. However, mentoring practices implemented in many CSR and government programs, including those by PT Telkom Indonesia, tend to remain generalized and non-personalized, failing to address the specific needs of MSMEs at different growth stages or levels of digital readiness.

The importance of improving managerial capability is further emphasized by studies showing that MSME performance is strongly influenced by both internal characteristics such as entrepreneurial competence and human resource quality and external factors, including access to finance and the business environment [7], [8]. Strengthening these managerial and strategic capacities through data-driven mentoring is therefore critical to enhancing MSME competitiveness and resilience.

While previous research has examined several factors affecting MSME performance such as financial literacy and e-commerce adoption [9], or the role of intellectual capital and organizational learning [10] there remains a significant research gap concerning the development of integrated and technology-based mentoring models specifically tailored to Indonesian MSMEs. The literature suggests that mentoring effectiveness and digital adoption must be analyzed together, particularly in relation to MSME growth stages, to design interventions that are contextually relevant and impactful.

This study focuses on 7,220 MSMEs supported by PT Telkom Indonesia's Corporate Social Responsibility (CSR) program, which presents a valuable opportunity to develop and evaluate a data-driven mentoring framework that integrates digital literacy and growth-stage characteristics. To achieve this, the research employs a Decision Tree algorithm—known for its interpretability and accuracy in classifying MSME growth stages [8] and Polytomous Logistic Regression, which allows for analyzing multi-category relationships among variables [11]

At the same time, the central problem stems from both empirical and theoretical gaps that hinder MSME development. Empirically, there is a clear disconnect between the vast number of MSMEs and their stagnant productivity, as reflected in sectoral contraction, low digitalization, and non-personalized mentoring practices. Theoretically, research focusing on technology-based mentoring frameworks that consider MSME heterogeneity and digital literacy remains limited.

Therefore, this research aims to bridge these gaps by analyzing the specific case of MSMEs under PT Telkom Indonesia's CSR program, with the following primary objectives:

1. To examine the influence and relationship between an MSME's business growth stage and its overall growth rate;
2. To investigate the influence and relationship between the mentoring model applied (including its approach, methods, and focus) and MSME growth; and
3. To analyze the influence and relationship between the digital literacy level of an MSME and its growth.

II. METHOD

This study employs a quantitative research design to analyze the influence of MSME growth stages, mentoring models, and digital literacy on business growth. The research is scheduled to be conducted over a two-year period, from 2024 to 2026, utilizing a national dataset of MSMEs participating in PT Telkom Indonesia's Corporate Social Responsibility (CSR) program.

The methodological approach integrates both analytical modeling and information system structuring. The research begins with a conceptual framework that connects MSME growth stage classification, mentoring interventions, and levels of digital literacy as explanatory variables influencing business growth. In the initial analytical phase, a machine learning algorithm specifically a Decision Tree is applied to classify MSMEs into distinct growth stages based on their internal characteristics and managerial attributes. These stages, representing phases such as existence, survival, success, take-off, and resource maturity, serve as one of the key independent variables in the study. Alongside this, two additional independent variables are analyzed: mentoring model and digital literacy level.

The mentoring model includes variables such as interaction type, training frequency, and duration, while digital literacy reflects MSMEs' capacity to operate within traditional, modern, digital, and global contexts. The dependent variable, MSME growth, is measured primarily through revenue and sales growth metrics. Based on this framework, the study tests the following hypotheses:

1. The business growth stage has a significant effect on the mentoring model.
2. The mentoring model has a significant effect on MSME growth.
3. The level of digital literacy has a significant effect on MSME growth.

This framework allows a comprehensive examination of MSME development by establishing an objective classification of their growth stages and then measuring how mentoring and digital capabilities affect business performance.

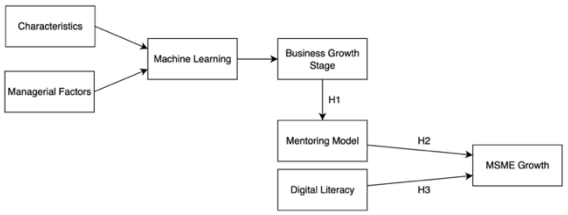


Figure 1. Research Framework

The overall research procedure is organized within an information system framework consisting of three integrated stages: input, process, and output. This framework provides a systematic structure for managing the data flow and analytical procedures throughout the study.

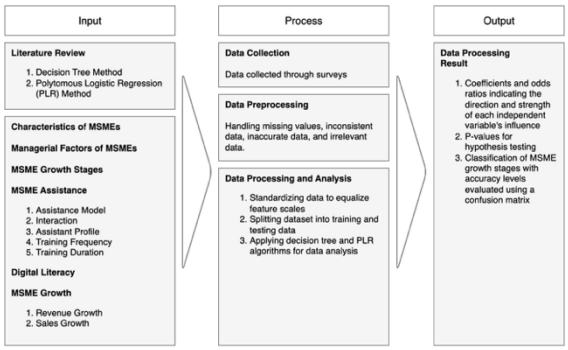


Figure 2. Information System Framework

In the input stage, foundational data and theoretical knowledge are gathered through a comprehensive literature review on MSME growth stages, mentoring strategies, digital literacy, and the analytical methods employed (Decision Tree and Polytomous Logistic Regression). The primary data are collected via surveys administered to MSMEs participating in the CSR program,

ISSN 2442-7888 (online) DOI 10.24167/sisforma.v12i2.14299

complemented by secondary data when available.

The dataset covers four key domains:

1. Variables defining MSME growth stages (existence, survival, success, take-off, resource maturity).
2. Variables describing the mentoring process (model type, interaction pattern, frequency, duration).
3. Indicators of digital literacy (traditional, modern, digital, global levels).
4. Business growth indicators (revenue and sales growth).

In the process stage, the collected data undergo systematic preparation and analysis. Data preprocessing is performed to address missing values, correct inconsistent or inaccurate data, and remove irrelevant information to ensure quality and integrity. After cleaning, the data are standardized to normalize different feature scales. The dataset is then partitioned into training and testing subsets to support robust model development and validation.

Subsequently, analytical modeling is conducted in two main steps. The Decision Tree algorithm is applied to classify MSMEs according to their growth stage, generating a structured segmentation based on objective patterns in the data. Then, the Polytomous Logistic Regression (PLR) model is employed to examine the relationships between variables, quantifying the influence of MSME growth stages, mentoring models, and digital literacy on business growth outcomes.

The output stage delivers the final analytical results and insights. Key outputs include coefficients and odds ratios, which quantify both the direction and strength of the influence of each independent variable on MSME growth. P-values are used to test the research hypotheses and assess the statistical significance of the results. Additionally, a confusion matrix is employed to evaluate the accuracy of the Decision Tree model in classifying MSME growth stages.

The interpretation of these quantitative findings leads to the formulation of data-driven recommendations for designing an optimized and effective MSME mentoring model. Through this methodological design linking machine learning-based classification, regression modeling, and a systematic information system framework the study ensures a rigorous and integrated approach to understanding how developmental stages, mentoring interventions, and digital literacy collectively influence MSME business growth.

III. RESULTS AND DISCUSSION

This section presents the results of the analysis and discusses their implications in relation to the theoretical framework and findings from previous studies. The presentation of results focuses on empirical evidence derived from the data analysis, while the discussion interprets these results in a broader academic context, linking them to established theories and prior research on MSME development, mentoring models, and digital literacy.

A. RESULT

The first phase of this research involved developing a model to classify MSME growth stages based on their characteristics and managerial factors. A Decision Tree algorithm was trained and tested, producing a strong overall accuracy of 90.00% on the test data. This high level of accuracy indicates that the model is both reliable and valid for classifying new MSMEs, providing a solid foundation for the subsequent hypothesis testing. To further evaluate the model's performance, precision, recall, and F1-score were calculated, as summarized in Table 1.

Table 1. Average Perfomance Metrics of The Decision Tree Model

Metric	*Precision*	*Recall*	*F1-Score*
Macro Avg	0.90	0.90	0.89
Weighted Avg	0.93	0.90	0.90

The results presented in Table 1 provide deeper insights into the model's robustness. The Macro Average treats all classes equally,

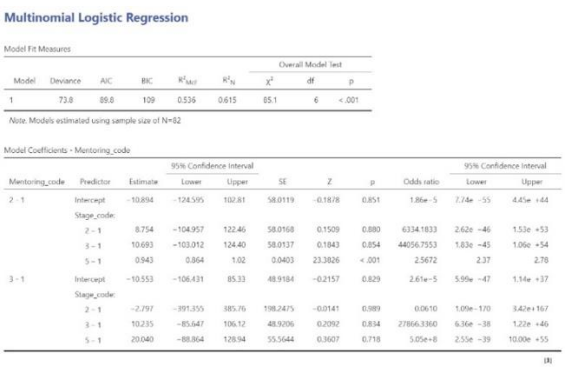
yielding a precision of 0.90, recall of 0.90, and F1-score of 0.89, demonstrating high and consistent accuracy across all growth stages. The Weighted Average, which considers the number of instances in each class, provides an F1-score of 0.90 confirming that the model is both balanced and resistant to class bias.

Further analysis of the model's performance shows a high weighted-average F1-score of 0.90, emphasizing its reliability in distinguishing between growth stages. Minor misclassifications occurred primarily between adjacent stages, such as confusing "Survival" with "Existence." This is expected, as the boundary between early stages is often subtle. Importantly, the model showed no confusion between distant stages, such as between "Existence" and "Take-Off," confirming its discriminative strength and validity for subsequent use in hypothesis testing.

Building on these classification results, the next step involved Polytomous Logistic Regression (PLR) analysis to test the research hypotheses. The combined model, incorporating the variables of mentoring model and digital literacy, was found to significantly explain MSME growth, as shown in the Model Fit Measures presented in Figure. 3 and Figure. 4.

The results indicate strong model performance:

1. For mentoring model analysis (Fig. 3), $\chi^2(6) = 85.1$, $p < .001$, with Nagelkerke $R^2 = 0.615$ (N = 82).
2. For MSME growth analysis (Fig. 4), $\chi^2(10) = 119$, $p < .001$, with Nagelkerke $R^2 = 0.774$ (N = 90).



Multinomial Logistic Regression

Model Fit Measures

Model	Deviance	AIC	BIC	R^2_{McF}	R^2_N	Overall Model Test χ^2	df	p
1	73.8	89.8	109	0.536	0.615	85.1	6	< .001

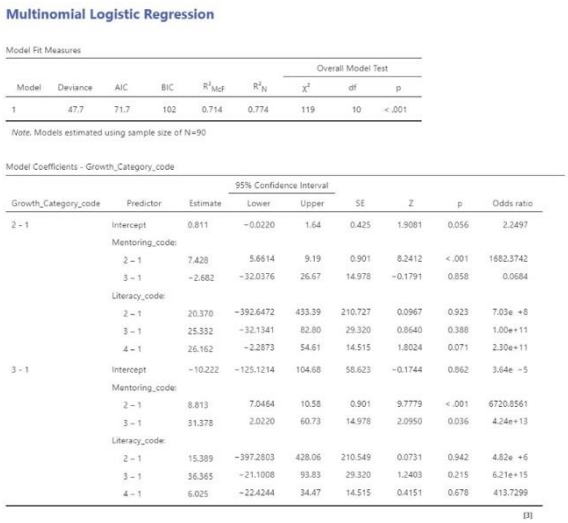
Note. Models estimated using sample size of N=82

Model Coefficients - Mentoring_code

Mentoring_code	Predictor	Estimate	95% Confidence Interval Lower	Upper	SE	Z	p	Odds ratio	95% Confidence Interval Lower	Upper
2 - 1	Intercept	−10.894	−124.595	102.81	58.0119	−0.1878	0.851	1.86e−5	7.74e−55	4.45e+44
	Stage_code:									
	2 - 1	8.754	−104.957	122.46	58.0168	0.1509	0.880	6334.1833	2.62e−46	1.53e+53
	3 - 1	10.693	−103.012	124.40	58.0137	0.1843	0.854	44056.7553	1.83e−45	1.06e+54
	5 - 1	0.943	0.864	1.02	0.0403	23.3826	< .001	2.5672	2.37	2.78
3 - 1	Intercept	−10.553	−106.431	85.33	48.9184	−0.2157	0.829	2.61e−5	5.99e−47	1.14e+37
	Stage_code:									
	2 - 1	−2.797	−391.355	385.76	198.2475	−0.0141	0.989	0.0610	1.09e−170	3.42e+167
	3 - 1	10.235	−85.647	106.12	48.9206	0.2092	0.834	27866.3360	6.36e−38	1.22e+46
	5 - 1	20.040	−88.864	128.94	55.5644	0.3607	0.718	5.05e+8	2.55e−39	10.00e+55

[3]

Figure 3. Model Fit Measures and Coefficients – Mentoring_code



Multinomial Logistic Regression

Model Fit Measures

Model	Deviance	AIC	BIC	R^2_{McF}	R^2_N	Overall Model Test χ^2	df	p
1	47.7	71.7	102	0.714	0.774	119	10	< .001

Note. Models estimated using sample size of N=90

Model Coefficients - Growth_Category_code

Growth_Category_code	Predictor	Estimate	95% Confidence Interval Lower	Upper	SE	Z	p	Odds ratio
2 - 1	Intercept	0.811	−0.0220	1.64	0.425	1.9081	0.056	2.2497
	Mentoring_code:							
	2 - 1	7.428	5.6614	9.19	0.901	8.2412	< .001	1682.3742
	3 - 1	−2.682	−32.0376	26.67	14.978	−0.1791	0.858	0.0684
	Literacy_code:							
	2 - 1	20.370	−392.6472	433.39	210.727	0.0967	0.923	7.03e +8
	3 - 1	25.332	−32.1341	82.80	29.320	0.8640	0.388	1.00e+11
	4 - 1	26.162	−2.2873	54.61	14.515	1.8024	0.071	2.30e+11
3 - 1	Intercept	−10.222	−125.1214	104.68	58.623	−0.1744	0.862	3.64e −5
	Mentoring_code:							
	2 - 1	8.813	7.0464	10.58	0.901	9.7779	< .001	6720.8561
	3 - 1	31.378	2.0220	60.73	14.978	2.0950	0.036	4.24e+13
	Literacy_code:							
	2 - 1	15.389	−397.2803	428.06	210.549	0.0731	0.942	4.82e +6
	3 - 1	36.365	−21.1008	93.83	29.320	1.2403	0.215	6.21e+15
	4 - 1	6.025	−22.4244	34.47	14.515	0.4151	0.678	413.7299

[3]

Figure 3. Model Fit Measures and Coefficients – Growth_Category_code

The first tested hypothesis examined the influence of growth stage on mentoring model. The PLR results confirmed that the business growth stage significantly affects the choice of mentoring model ($\chi^2(6) = 85.1$, $p < .001$), supporting H1. Specifically, MSMEs in the "Resource Maturity" stage were found to be 2.5672 times more likely to adopt a "Hybrid" mentoring model than those in the "Existence" stage ($p < .001$). This pattern suggests that mature firms tend to prefer flexible and integrated mentoring approaches that align with their complex operational needs. In contrast, early-stage MSMEs such as those in the "Survival" stage demonstrated less pronounced effects (odds ratio = 44056.7553, p = 0.854), likely reflecting their reliance on simpler, more directive mentoring structures.

The second hypothesis focused on the effect of the mentoring model on MSME growth, and the PLR results supported H2, showing that the mentoring model significantly influences MSME performance outcomes. The "Hybrid" mentoring model had the strongest effect, increasing the likelihood of achieving "Medium" growth (odds ratio = 6720.8561, $p < .001$) and "High" growth (odds ratio = 413.7299, p = 0.678) compared to the reference category. The "Individual" mentoring model also showed a significant effect on "High" growth (odds ratio = 4.24e+13, p = 0.036). These findings

confirm the critical role of mentoring model selection in determining business performance.

The third hypothesis examined the influence of digital literacy on MSME growth, controlling for mentoring model. The results revealed that digital literacy did not have a statistically significant effect on MSME growth ($p > .05$ for all comparisons), leading to the rejection of H3.

B. DISCUSSION

The findings highlight that MSME growth stages can be effectively classified using quantitative and managerial variables, as reflected in the Decision Tree model's 90% accuracy. The limited misclassification between adjacent stages confirms the framework's validity in capturing real-world development patterns. The observed relationship between growth stage and mentoring model indicates that organizational maturity is aligned with mentoring complexity: mature MSMEs prefer hybrid models integrating structured and adaptive learning, while early-stage firms benefit more from direct, one-on-one guidance. This supports the view that organizational learning evolves alongside firm growth and structural complexity.

The significant effect of mentoring models on MSME growth emphasizes the critical role of targeted developmental support. The superior performance of hybrid mentoring aligns with blended learning principles, where combining personalized coaching and digital tools enhances adaptability. Conversely, the absence of a direct digital literacy effect suggests that digital competence yields impact only when guided through structured mentoring. This interdependence highlights the need to integrate digital training within mentoring frameworks. Overall, the study underscores that growth stage classification and mentoring design are central to MSME performance, while digital literacy serves as a complementary factor. The integrated analytical approach combining machine learning and statistical modeling provides both theoretical rigor and practical insights for MSME development strategies.

IV. CONCLUSION

This study aimed to model MSME growth by analyzing the impact of business growth stage, mentoring models, and digital literacy. Using a quantitative approach with Decision Tree classification and Polytomous Logistic Regression, this research has yielded several key findings that contribute to both theory and practice in MSME development.

The research confirms that the business growth stage is a significant predictor of the type of mentoring support sought by an MSME (H1). Specifically, mature firms are significantly more likely to prefer a flexible 'Hybrid' model, indicating that support programs must adapt their delivery methods as businesses evolve. Furthermore, the study reveals that the mentoring model is the most dominant factor influencing MSME growth (H2). The 'Hybrid' model, which combines online and offline interactions, consistently proved to be the most effective intervention for driving both medium and high growth. This underscores that the structure and quality of guidance are paramount.

Conversely, digital literacy, when analyzed alongside mentoring, did not show a statistically significant direct effect on growth (H3). This critical finding does not diminish the importance of digital skills. Instead, it suggests a more complex relationship where mentoring likely acts as a crucial catalyst, enabling entrepreneurs to translate their digital competencies into tangible business outcomes. Without structured guidance, possessing digital skills alone may not be sufficient to drive growth.

ACKNOWLEDGMENT

The authors would like to thank PT Telkom Indonesia for providing the data from their Corporate Social Responsibility (CSR) program used in this research, enabling the analysis of factors influencing MSME growth and the evaluation of different mentoring models. We also acknowledge the support from our respective institutions for providing

the computational resources and academic environment necessary to conduct this study. Special appreciation is extended to the reviewers and colleagues who provided valuable feedback during the development of this paper. Their insights significantly improved the clarity and rigor of our work.

REFERENCES

[1] V. Lewis and N. Churchill, "The Five Stages of Small Business Growth," *Harv Bus Rev*, vol. 3, Jan. 1987.

[2] L. Ardito, R. Filieri, E. Raguseo, and C. Vitari, "Artificial intelligence adoption and revenue growth in European SMEs: synergies with IoT and big data analytics," *Internet Research*, Dec. 2024, doi: 10.1108/INTR-02-2024-0195.

[3] A. G. Mullisa and W. A. Kader, "Performance improvement: a lean manufacturing case of metal tools factory," *International Journal of Industrial and Systems Engineering*, vol. 44, no. 4, pp. 475–498, 2023, doi: 10.1504/IJISE.2023.132721.

[4] Y. Ying, A. M. Karim, E. P. Mondol, and Md. S. A. Helal, "Influence of Digital Education to Uplift the Global Literacy Rate in the Age of Digital Civilization," *International Journal of Academic Research in Accounting, Finance and Management Sciences*, vol. 14, no. 4, Nov. 2024, doi: 10.6007/IJARAFMS/v14-i4/23552.

[5] A. J. van Deursen and J. A. van Dijk, "The first-level digital divide shifts from inequalities in physical access to inequalities in material access," *New Media Soc*, vol. 21, no. 2, pp. 354–375, Feb. 2019, doi: 10.1177/1461444818797082.

[6] S. Rosyadi, A. S. Kusuma, E. Fitrah, A. Haryanto, and W. Adawiyah, "The Multi-Stakeholder's Role in an Integrated Mentoring Model for SMEs in the Creative Economy Sector," *Sage Open*, vol. 10, no. 4, Oct. 2020, doi: 10.1177/2158244020963604.

[7] I. Indarto and D. Santoso, "KARAKTERISTIK WIRAUSAHA, KARAKTERISTIK USAHA DAN LINGKUNGAN USAHA PENENTU KESUKSESAN USAHA MIKRO KECIL DAN MENENGAH," *Jurnal Riset Ekonomi dan Bisnis*, vol. 13, no. 1, p. 54, Apr. 2020, doi: 10.26623/jreb.v13i1.2202.

[8] S. Astuti, I. Irawan, and M. M. Sidik, "Analisis Faktor – Faktor Yang Mempengaruhi Penggunaan Informasi Akuntansi Pada UMKM Di Bandar Lampung," *Jurnal Riset Akuntansi dan Manajemen Malahayati*, vol. 11, no. 4, Jan. 2023, doi: 10.33024/jrm.v11i4.8814.

[9] R. R. Putra and S. Holisoh, "Factors Affecting the Performance of Small Micro Medium Enterprise (UMKM) in Lagoa Village," *Quantitative Economics and Management Studies*, vol. 4, no. 2, pp. 231–237, Jan. 2023, doi: 10.35877/454RI.qems1469.

[10] J. Bora, S. Dehingia, A. Boruah, A. A. Chetia, and D. Gogoi, "Real-time Assamese Sign Language Recognition using MediaPipe and Deep Learning," *Procedia Comput Sci*, vol. 218, pp. 1384–1393, 2023, doi: 10.1016/j.procs.2023.01.117.

[11] C. J. Biesheuvel, Y. Vergouwe, E. W. Steyerberg, D. E. Grobbee, and K. G. M. Moons, "Polytomous logistic regression analysis could be applied more often in diagnostic research," *J Clin Epidemiol*, vol. 61, no. 2, pp. 125–134, Feb. 2008, doi: 10.1016/j.jclinepi.2007.03.002.