

Implementation of the XGBoost Algorithm for Predicting Entrepreneurial Interest Categories of Accounting Department Students at SMK Islam Al Fadhila Demak

Ridwan Budi Satria, Endang Supriyati, Tri Listyorini

Universitas Muria Kudus, Indonesia

202351137@std.umk.ac.id

Submitted:	Revised:	Accepted:	Published:
Sep 16, 2026	Oct 15, 2026	Oct 26, 2026	Oct 31, 2026

Abstract

Digital entrepreneurship offers vocational high school graduates a potential career pathway, yet schools have limited approaches for identifying individual students' entrepreneurial interests. This study aimed to classify entrepreneurial interest as a proxy for digital entrepreneurial potential among accounting students at SMK Islam Al Fadhila Demak, identify influential predictors, and integrate the classifier into a prototype web system. A quantitative survey covered all 64 twelfth-grade students, assessing ten indicators through 30 Likert-scale items. Entrepreneurial interest scores of at least 13 defined the high-interest category ($n = 20$), while lower scores defined the low-interest category ($n = 44$). The remaining nine indicators served as predictive features. Extreme Gradient Boosting (XGBoost) was employed, with the synthetic minority oversampling technique (SMOTE) restricted to training data. Model development and evaluation involved stratified five-fold cross-validation and an 80:20 data split. The model achieved 84.49% accuracy, 84.79% macro precision, 80.56% macro recall, and an 80.92% macro F1-score. Accuracy exceeded the majority-class baseline of 68.75%, although recall for the high-interest category was 70.00%, indicating limitations in identifying students within this group. Risk-taking and digital marketing received the highest feature importance scores, at 52.25% and 17.15%, respectively. These importance values describe their contributions to model predictions and do not establish causal relationships. The results suggest that XGBoost can support exploratory

classification of entrepreneurial interest within the studied cohort. However, the small sample, self-reported measurements, and use of shared folds for tuning and evaluation limit confidence in generalizability. The prototype has also not undergone user testing. Further validation with larger, independent samples and evaluation by intended users are needed before broader implementation. The proposed system offers a preliminary basis for student screening and prioritizing entrepreneurship coaching, while entrepreneurial interest remains distinct from demonstrated entrepreneurial capability.

Keywords: Digital Entrepreneurship; XGBoost; Vocational Education; Classification; Feature Importance

INTRODUCTION

The rapid growth of digital technology has changed how businesses are started and run. In a systematic review of digital entrepreneurship research published in twenty-five indexed journals, Paul et al. (2023) showed that digital technology no longer acts merely as a supporting tool but also shapes entrepreneurial opportunities themselves, allowing conventional businesses to transform into digital ones. In Vietnam, digital platform ecosystems are likewise described as drivers of innovation, economic growth, and job creation (Tran, 2023). This shift opens wide opportunities for young people to start ventures with relatively affordable capital, while demanding competencies that differ from those of conventional entrepreneurship. Vocational high schools (Sekolah Menengah Kejuruan, SMK) are designed to prepare graduates either for employment or for starting their own businesses, and entrepreneurship education in Indonesian SMKs emphasizes practical competence and partnerships with industry (Widodo et al., 2025).

Evidence indicates that the entrepreneurial intention of vocational students is not innate but is formed by measurable factors. The teaching factory model influences intention both directly and through attitudes and perceived behavioral control (Sutiadiningsih & Mahfud, 2023), and personality dimensions and entrepreneurial self-efficacy predict intention among vocational students (Juhari et al., 2023). Structural equation modeling with five hundred vocational respondents showed that creativity influences intention through self-efficacy and attitude, with risk propensity acting as a moderator (Niu et al., 2022), and creativity and entrepreneurship education also show positive effects on intention (Atrup et al., 2023). In the digital context, digital literacy positively affects the entrepreneurial intention

of vocational students in Indonesia (Suryani & Chaniago, 2023), and digital competence influences intention through opportunity alertness, an effect strengthened by motivation (Zhao et al., 2025). Opportunity-related competencies are linked to business performance (Sakib et al., 2022); digital marketing is related to innovation capability and the performance of micro, small, and medium enterprises (Padmawidjaja et al., 2025); financial literacy raises entrepreneurial intention (Anggraini & Sakti, 2025); and digital accounting and digital platform capability contribute to financial performance (Maryanti et al., 2026).

Accounting students therefore hold a distinctive position, because the competencies in recording transactions and managing finances that they learn are direct assets for running a business. Nevertheless, the coaching of accounting students is still largely directed toward work readiness (Puspita & Nurkhin, 2026), so systematic mapping of entrepreneurial potential in this group is rare. Most research on student entrepreneurship uses inferential approaches such as regression or structural equation modeling. These approaches are well suited to testing relationships among variables, but they yield conclusions at the group level rather than classifications of individual students, whereas schools need per-student information in order to target coaching. Predictive machine learning can address this need. A systematic review of seventy articles shows that machine learning in education has grown rapidly, although the predicted outcomes remain largely academic performance and graduation (Pelima et al., 2024).

Extreme Gradient Boosting (XGBoost) builds an ensemble of decision trees sequentially, with each new tree correcting the prediction errors of the previous ones and with regularization that prevents the model from becoming overly complex (Chen & Guestrin, 2016). On student data, XGBoost was among the best of six algorithms compared (Kumar et al., 2024), and it outperformed a single decision tree in sentiment classification (Sinaga & Agustian, 2022). For imbalanced data, combining XGBoost with the Synthetic Minority Over-sampling Technique (SMOTE) (Elreedy et al., 2024) and Bayesian hyperparameter search raised the area under the curve from 0.618 to 0.767 (Ubaidillah et al., 2022), and the algorithm also allows the contribution of each feature to the prediction to be traced (Latifah et al., 2026). The last property matters for decision-oriented research, because results go beyond a predicted label to explain which aspects matter most. Explainable output is also increasingly expected of applied intelligent systems; Supriyati et al. (2026), for example, paired classification with SHAP-based explanations in a user-facing verification system for halal cosmetic products and reported an accuracy of 94.5%.

Two gaps follow from the literature. Research on the entrepreneurship of vocational students usually stops at testing relationships among variables, while predictive research in education mostly forecasts academic achievement. Entrepreneurial indicators that are widely measured among vocational students have not been used to classify students individually in a way that can be traced back to the indicators, although the coaching needs of SMKs arise at the individual level. The novelty of this study is applied and methodological rather than algorithmic: (1) a design that separates the target indicator from the nine feature indicators so that the label is not recomputed by the model, (2) an evaluation protocol for small questionnaire datasets that uses stratified cross-validation and applies SMOTE only to training data, and (3) a traceable ranking of influential indicators that schools can read. The study does not seek to develop a new algorithm, and its results are positioned as a preliminary case study of one school. The study therefore aims to design a preliminary model that predicts the category of entrepreneurial interest, used as a proxy for digital entrepreneurial potential, of accounting students at SMK Islam Al Fadhila Demak using XGBoost on ten questionnaire indicators, to measure the performance of the model, to identify the most influential indicators through feature importance, and to deploy the model in a prototype web system.

METHODS

This study used a quantitative approach with a predictive modeling design, in which questionnaire responses were converted into numerical features to train and test a classification model (Figure 1).

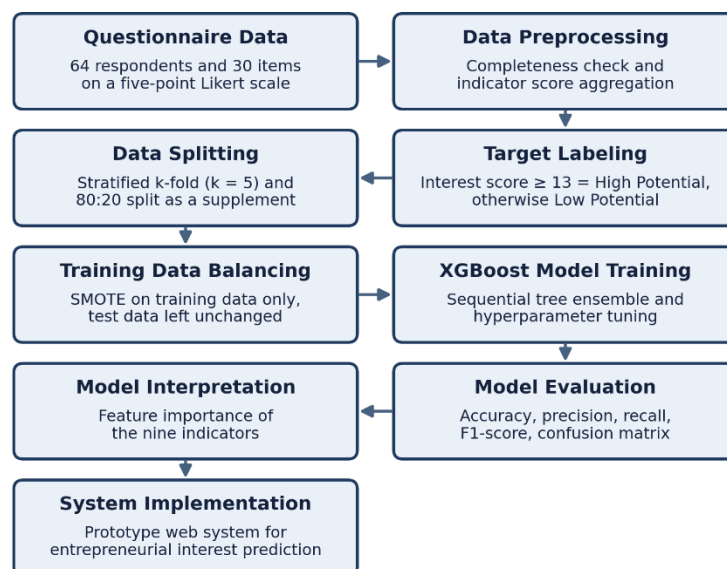


Figure 1. Research Stages

Figure 1. Research Stages explains that the study moved from questionnaire data collection to preprocessing and labeling, data splitting and class balancing, XGBoost training, evaluation, interpretation of feature importance, and implementation in a prototype web system.

The population was all 64 twelfth-grade students of the Accounting Department at SMK Islam Al Fadhila Demak, spread across two classes (XII Accounting 1 and XII Accounting 2); because every member completed the questionnaire, total sampling was applied and the sample equals the population.

Data were collected with a closed questionnaire containing ten indicators, each measured by three five-point Likert items (30 items in total; Table 1), which was distributed online at school starting on 31 August 2026 without a time limit so that students could finish it at home, and closed on 1 October 2026 at 10:00 p.m. (31 days in total); documentation and a literature study supplemented the questionnaire.

Table 1. Research Indicators and Their Sources

No	Indikator	Code	Items	Sources
1	Entrepreneurial Interest	Y	3	(Sutiadiningsih & Mahfud, 2023; Juhari et al., 2023)
2	Creativity	X2	3	(Niu et al., 2022; Atrup et al., 2023)
3	Accounting Competence	X5	3	(Maryanti et al., 2026; Puspita & Nurkhin, 2026)
4	Digital Literacy	X6	3	(Suryani & Chaniago, 2023; Zhao et al., 2025)
5	Entrepreneurial Experience	X8	3	(Puspita & Nurkhin, 2026)
6	Opportunity Identification Ability	X3	3	(Zhao et al., 2025; Sakib et al., 2022)
7	Risk-Taking	X9	3	(Niu et al., 2022)
8	Digital Marketing	X7	3	(Paul et al., 2023; Padmawidjaja et al., 2025)
9	Financial Literacy	X4	3	(Anggraini & Sakti, 2025; Maryanti et al., 2026)
10	Entrepreneurial Motivation	X1	3	(Zhao et al., 2025)

Source: Authors' compilation based on the cited literature (2026).

Table 1. Research Indicators and Their Sources explains that entrepreneurial interest (Y) serves as the source of the target label, while the nine other indicators (X1 to X9) serve as model features, each grounded in earlier studies.

Item validity was tested by item–total correlation and reliability by Cronbach's alpha with a minimum of 0.60 (Zakariya, 2022), the items of each indicator were summed, and the

entrepreneurial interest score was labeled high when it was at least 13 and low otherwise, as a proxy of digital entrepreneurial potential (Juhari et al., 2023; Sutiadiningsih & Mahfud, 2023). An XGBoost classifier (Chen & Guestrin, 2016) was then trained on the nine other indicators with SMOTE applied only to training data (Elreedy et al., 2024), hyperparameter tuning, and stratified five-fold cross-validation with an 80:20 split as a supplement, and it was evaluated with accuracy, precision, recall, macro F1-score, a confusion matrix, and feature importance (Rainio et al., 2024).

The study was conducted at SMK Islam Al Fadhila Demak, Central Java, Indonesia, with a data collection period of 31 days (31 August to 1 October 2026). Permission was requested directly at the school and approved on 31 August 2026, and the questionnaire was administered in classes XII Accounting 1 and XII Accounting 2. The students of both classes also agreed to participate.

RESULTS

Main Findings

All 64 students completed the questionnaire, all 30 items were valid, and the instrument was highly reliable (Cronbach's $\alpha = 0.948$). Labeling produced 20 high-potential and 44 low-potential students, a ratio of about 31 to 69.

The model was trained on the nine feature indicators and evaluated mainly with stratified five-fold cross-validation, with an 80:20 train–test split as a supplement (Table 2). The metrics in Table 2 are those reported by the system in Figure 4.

Table 2. Model Evaluation Results

Metric	Five-fold cross-validation	80:20 split
Accuracy	84.49%	84.62%
Precision (macro)	84.79%	90.91%
Recall (macro)	80.56%	75.00%
F1-score (macro)	80.92%	78.33%

Source: Authors' analysis of questionnaire data (2026).

Table 2. Model Evaluation Results explains that the model reached an accuracy of 84.49% in cross-validation, 15.74 percentage points above the 68.75% of a predictor that always guesses the largest class. The 80:20 split contained only 13 test cases, so a single error shifts accuracy by about 7.7 points; its accuracy was similar (84.62%), but its precision (90.91%) and recall (75.00%) diverged.

The confusion matrix in Table 3 shows that errors were not evenly distributed between the two classes.

Table 3. Confusion Matrix of Five-Fold Cross-Validation (n = 64)

Actual / Predicted	Low Potential	High Potential
Low Potential	40	4
High Potential	6	14

Source: Authors' analysis of questionnaire data (2026).

Table 3. Confusion Matrix of Five-Fold Cross-Validation (n = 64) explains that 40 of 44 low-potential students were classified correctly (recall 90.91%), whereas 14 of 20 high-potential students were classified correctly (recall 70.00%); 6 high-potential students were classified as low potential and 4 low-potential students as high potential.

Feature importance in Table 4 shows that two indicators dominate the model.

Table 4. Ranking of Indicator Importance

Rank	Indicator	Importance (%)
1	Risk-Taking (X9)	52.25
2	Digital Marketing (X7)	17.15
3	Digital Literacy (X6)	7.20
4	Opportunity Identification Ability (X3)	6.77
5	Creativity (X2)	6.66
6	Entrepreneurial Motivation (X1)	3.77
7	Accounting Competence (X5)	3.12
8	Financial Literacy (X4)	3.08
9	Entrepreneurial Experience (X8)	0.00

Source: Authors' analysis of questionnaire data (2026).

Table 4. Ranking of Indicator Importance explains that Risk-Taking (X9) contributed 52.25%, more than three times Digital Marketing (X7) at 17.15%, whereas Accounting Competence (X5) contributed only 3.12% and Entrepreneurial Experience (X8) 0.00%.

Data Visualization

The results were also visualized through the prototype web system and a feature importance chart (Figures 2 to 5). The prototype has five menus (home, student prediction, student data, model evaluation, and research information) and keeps the trained model as a file so that predictions do not require retraining; it has not yet been tested by users at the school.

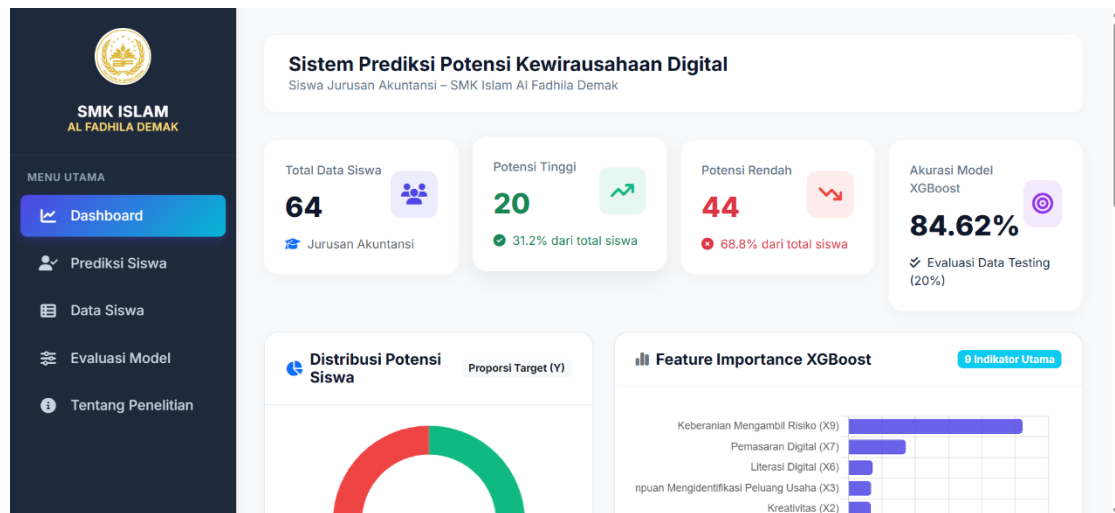


Figure 2. Home Page of the Prediction System (Interface in Indonesian)

Figure 2. Home Page of the Prediction System (Interface in Indonesian) shows that the system summarizes the number of respondents (64), the distribution of potential categories (20 high and 44 low), the model accuracy on the test split (84.62%), and indicator importance on a single page.

Indicator	Score (0-100)
1. Motivasi Berwirausaha (X1)	75
2. Kreativitas (X2)	75
3. Inovasi (X3)	75
4. Literasi Keuangan (X4)	75
5. Kompetensi Akuntansi (X5)	75
6. Literasi Digital (X6)	75
7. Pemasaran Digital (X7)	75
8. Pengalaman Usaha (X8)	0 = Tidak Memiliki Penge
9. Kemampuan Business Plan (X9)	75

Figure 3. Student Prediction Page (Interface in Indonesian)

Figure 3. Student Prediction Page (Interface in Indonesian) shows that a user enters the scores (0–100) of the nine indicators of one student with sliders and then requests a prediction of the potential category.

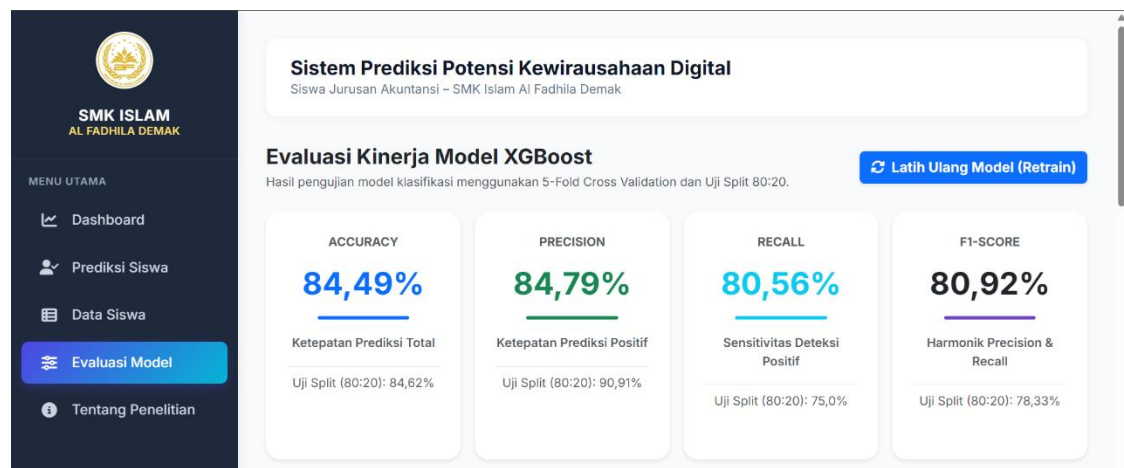


Figure 4. Model Evaluation Page of the System (Interface in Indonesian)

Figure 4. Model Evaluation Page of the System (Interface in Indonesian) shows that the system reports accuracy, precision, recall, and F1-score from five-fold cross-validation (84.49%, 84.79%, 80.56%, and 80.92%) together with the split results, and offers a button for retraining the model.

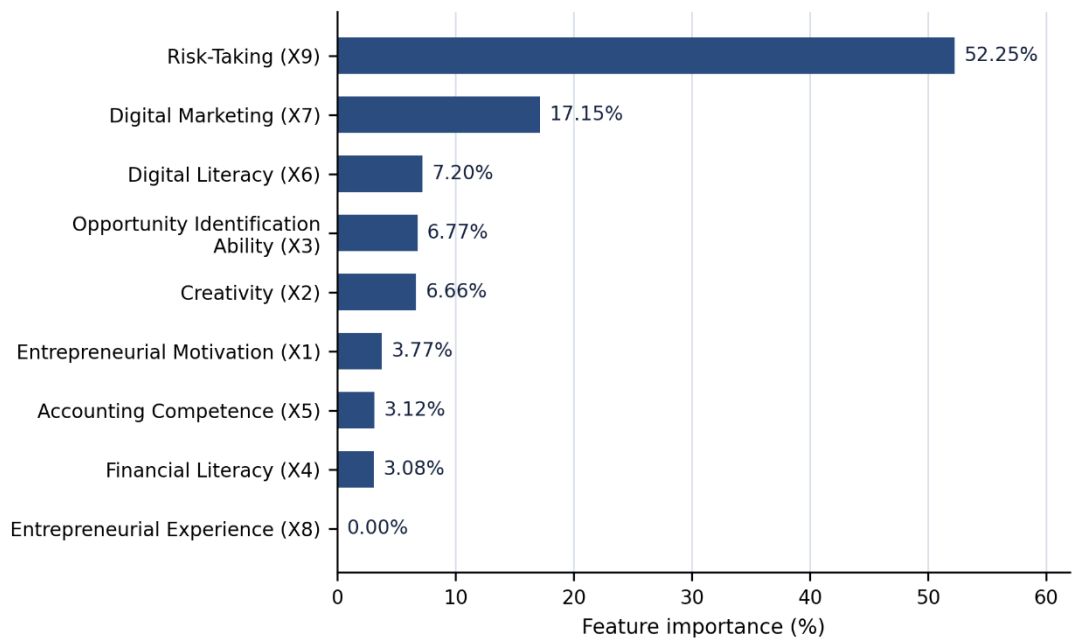


Figure 5. Indicator Importance from the XGBoost Model

Figure 5. Indicator Importance from the XGBoost Model shows that importance is concentrated in one indicator, Risk-Taking, followed by Digital Marketing, while each of the remaining seven indicators contributes less than 8%.

Negative and Divergent Findings

Several findings did not follow the general pattern. First, the recall of the high-potential class (70.00%) was clearly lower than that of the low-potential class (90.91%). Second, Entrepreneurial Experience (X8) received an importance of 0.00%, and Accounting Competence (X5) received only 3.12%, even though all respondents were accounting students. Third, the accuracy computed from the pooled predictions in Table 3 is 84.38%, slightly different from the 84.49% in Table 2, and precision and recall under the 80:20 split diverged (90.91% versus 75.00%). Finally, the benefit of SMOTE was not tested by comparing models with and without it, and XGBoost was not compared with other algorithms.

DISCUSSION

Analysis and Interpretation of Findings

The results indicate that the entrepreneurial interest category of students can be predicted from the other nine indicators in this dataset. The model exceeded the majority-class baseline by 15.74 percentage points, which suggests that it captured patterns beyond a tendency to favor the majority class, although this conclusion is limited by the small sample.

The lower recall of the high-potential class reflects the imbalance between the classes. It matters in practice because the high-potential group is the main target of entrepreneurship coaching, so some high-potential students would be overlooked. The system should therefore serve as an initial mapping aid rather than as a sole decision-maker. In contrast, classifying a low-potential student as high potential has a lighter impact, because it only adds one more student to the coaching group. The divergence between precision and recall under the 80:20 split, together with the 7.7-point effect of a single error, shows that a single split on such a small dataset is a weak basis for the main evaluation, which supports using cross-validation as the main reference.

Risk-taking was the indicator most often used by the model to split the data, which can be read as a preliminary indication that willingness to bear uncertainty distinguishes the two categories in this dataset. Digital marketing ranked second. The zero importance of X8 and the low importance of X5 do not mean that these aspects are unimportant theoretically; they more likely indicate that, among these 64 respondents, the indicators separated the two

categories less than risk-taking and digital marketing did. One possible explanation is the small variation in answers within a single program, but this has not been tested, for example through the standard deviation of the score of each indicator. Because the importance values come from one model fitted on 64 cases and reflect association rather than causation, they should be read as indications of direction and not as firm rankings.

Comparison with Theories and Previous Literature

The finding that entrepreneurial interest can be predicted from measurable indicators is consistent with studies showing that the entrepreneurial intention of vocational students is formed by observable factors (Juhari et al., 2023; Sutiadiningsih & Mahfud, 2023). The prominence of risk-taking agrees with Niu et al. (2022), who identified risk propensity as a moderator in the formation of intention, and the prominence of digital marketing and digital literacy is not contradicted by studies that place these factors among the important ones for digital ventures (Padmawidjaja et al., 2025; Suryani & Chaniago, 2023). A difference is that earlier work stops at testing relationships among variables, whereas this study produces an individual-level classification that schools can use for initial mapping.

Kumar et al. (2024) reported a test accuracy of 84.38% for XGBoost on student data, but that figure cannot be compared directly with the present results because the prediction target, sample size, and evaluation scheme differ; it is cited only to show that XGBoost is commonly applied to student data. Applying SMOTE only to the training data of each fold is in line with the practice for imbalanced data reported by Ubaidillah et al. (2022), so that the test data contain no synthetic samples. Separating the interest indicator from the feature list was a decisive design choice: without it, the model would merely recompute the labeling rule and yield a high but meaningless accuracy. Building a model from data collected in one specific local setting follows the approach of Permata et al. (2026), who trained a translation model for the Kudus regional language and noted that non-standardized local variation in digital data is a challenge; a similar concern applies here, because a model trained on the responses of one school may not transfer to other schools. Limited annotated data is likewise acknowledged as a constraint on model strength in a classification study of four Nigerian languages (Abdulyekeen & Abdullahi, 2026).

Research Implications

Practically, schools can use the model and the prototype system as an initial mapping aid to identify students whose entrepreneurial interest is high or low and to decide which

aspects, such as risk-taking and digital marketing, deserve more emphasis in coaching, provided that the ranking is interpreted with caution. Methodologically, the study offers a protocol for small questionnaire datasets: separating the target indicator from the features, using stratified cross-validation, and applying SMOTE to training data only. Theoretically, the study adds a predictive, individual-level perspective to a literature dominated by group-level inferential analysis, without claiming a new algorithm.

Research Limitations and Directions for Future Research

This study has several limitations. First, 64 respondents with 20 students in the high-potential class is small for machine learning modeling, so the evaluation results may change if more data are added. Second, at this sample size the ranking of feature importance is not fully stable, so the order in Table 4 should be read as an indication of direction rather than a definite ranking. Third, the recall of the high-potential class was only 70.00%, so some high-potential students are still missed. Fourth, all indicators, both features and target, come from the self-assessment of respondents in a single questionnaire at a single time, which carries a risk of common method bias and lacks data on actual business behavior as a comparison. Fifth, the target label was formed from the entrepreneurial interest score with a threshold of 13 that has not been tested for sensitivity to other values. Sixth, hyperparameter tuning and reporting of results used the same cross-validation scheme, so the performance estimates tend to be optimistic. Seventh, the study was conducted in one school and one program, so its results are not intended to be generalized to other schools or programs. In addition, XGBoost was not compared with other algorithms such as logistic regression or random forest, so its superiority over them cannot be concluded. Future studies are encouraged to enlarge the sample and the number of schools, compare algorithms with nested cross-validation that separates parameter tuning from performance reporting, test the sensitivity of the labeling threshold, add indicators of actual business behavior, and evaluate the usability of the system with school users. Future studies could also complement the gain-based importance used here with SHAP-based explanations, as applied by Supriyati et al. (2026).

CONCLUSION

This study produced a preliminary model that predicts the entrepreneurial interest category of accounting students at SMK Islam Al Fadhila Demak using XGBoost on nine questionnaire indicators, with entrepreneurial interest as the source of a two-category target.

The model obtained an accuracy of 84.49% and a macro F1-score of 80.92% in five-fold cross-validation, above the 68.75% of a majority-class predictor, with a recall of 70.00% for the high-potential class. Risk-taking (52.25%) and digital marketing (17.15%) were the indicators most often used by the model to distinguish the two categories. Given the small data, these results are indicative and do not show superiority over other algorithms.

The study contributes an applied and methodological example of individual-level prediction of entrepreneurial interest in vocational education, consisting of a design that separates target and features, an evaluation protocol for small questionnaire data, and a prototype web system, while making no claim of algorithmic novelty.

Future research is recommended to enlarge the number of respondents and schools, compare several algorithms using nested cross-validation, test the sensitivity of the labeling threshold, add indicators of actual business behavior, and evaluate the usability of the system with users at schools.

REFERENCES

- Abdulyekeen, R., & Abdullahi, B. A. (2026). Cross-lingual sentiment analysis with natural language processing: Insights from selected Nigeria languages (Yoruba, Hausa, Igbo, and Nigerian Pidgin). *Journal of Multidisciplinary Science: MIKAILALSYS*, 4(1), 60–86. <https://doi.org/10.58578/mikailalsys.v4i1.8596>
- Anggraini, A. N., & Sakti, N. C. (2025). Pengaruh Literasi Keuangan terhadap Intensi Berwirausaha dengan Pengetahuan Technopreneurship sebagai Variabel Moderasi [The effect of financial literacy on entrepreneurial intention with technopreneurship knowledge as a moderating variable]. *Jurnal Ilmiah Manajemen dan Kewirausahaan*, 4(3), 704–718. <https://doi.org/10.55606/jimak.v4i3.4972>
- Atrup, Diawati, P., Syamsuri, Pramono, S. A., & Ausat, A. M. A. (2023). The effect of entrepreneurship education and creativity on students' entrepreneurial intention: The perspective of effectuation and cognitive flexibility theory. *Jurnal Kependidikan: Jurnal Hasil Penelitian dan Kajian Kepustakaan di Bidang Pendidikan, Pengajaran dan Pembelajaran*, 9(2), 555–569. <https://doi.org/10.33394/jk.v9i2.7822>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM. <https://doi.org/10.1145/2939672.2939785>
- Elreedy, D., Atiya, A. F., & Kamalov, F. (2024). A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning. *Machine Learning*, 113(7), 4903–4923. <https://doi.org/10.1007/s10994-022-06296-4>
- Juhari, H. S., Ismail, N., Roslan, S., & Zaremohzzabieh, Z. (2023). Personality traits, entrepreneurial self-efficacy, and entrepreneurial intention among vocational students. *International Journal of Evaluation and Research in Education (IJERE)*, 12(1), 9–14. <https://doi.org/10.11591/ijere.v12i1.23472>

- Kumar, M., Singh, N., Wadhwa, J., Singh, P., Kumar, G., & Qtaishat, A. (2024). Utilizing random forest and XGBoost data mining algorithms for anticipating students' academic performance. *International Journal of Modern Education and Computer Science*, 16(2), 29–44. <https://doi.org/10.5815/ijmeecs.2024.02.03>
- Latifah, S., Martanto, Dana, R. D., Dikananda, F., & Hayati, U. (2026). Predicting student academic performance based on learning habits using XGBoost and SHAP. *Journal of Artificial Intelligence and Engineering Applications (JAIEA)*, 5(2), 2329–2333. <https://doi.org/10.59934/jaiea.v5i2.1860>
- Maryanti, Mediaty, Arifin, A. H., & Mas'ud, A. A. (2026). Digital accounting and financial performance of MSMEs in Indonesia: The mediating role of digital innovation. *International Journal of Financial Studies*, 14(3), Article 66. <https://doi.org/10.3390/ijfs14030066>
- Niu, X., Niu, Z., Wang, M., & Wu, X. (2022). What are the key drivers to promote entrepreneurial intention of vocational college students? An empirical study based on structural equation modeling. *Frontiers in Psychology*, 13, Article 1021969. <https://doi.org/10.3389/fpsyg.2022.1021969>
- Padmawidjaja, L., Rokhman, M. T. N., & Mitasari, A. R. (2025). Bridging entrepreneurial and digital marketing to enhance business performance: The mediating role of innovation capability and the moderating role of gender. *Jurnal Aplikasi Manajemen*, 23(3), 985–1004. <https://doi.org/10.21776/ub.jam.2025.023.3.12>
- Paul, J., Alhassan, I., Binsaif, N., & Singh, P. (2023). Digital entrepreneurship research: A systematic review. *Journal of Business Research*, 156, Article 113507. <https://doi.org/10.1016/j.jbusres.2022.113507>
- Pelima, L. R., Sukmana, Y., & Rosmansyah, Y. (2024). Predicting university student graduation using academic performance and machine learning: A systematic literature review. *IEEE Access*, 12, 23451–23465. <https://doi.org/10.1109/ACCESS.2024.3361479>
- Permata, A. B., Listyorini, T., & Supriyati, E. (2026). Pengembangan Mesin Penerjemah Bahasa Indonesia ke Bahasa Daerah Kudus Menggunakan NMT Berbasis RNN-GRU [Development of an Indonesian to Kudus regional language machine translator using RNN-GRU based NMT]. *JITEKH (Jurnal Ilmiah Teknologi Harapan)*, 14(1), 15–27.
- Puspita, M. P., & Nurkhin, A. (2026). Pengaruh Teaching Factory dan Praktik Kerja Industri terhadap Kesiapan Kerja Siswa Akuntansi Dimoderasi Motivasi Kerja [The influence of teaching factory and industrial work practices on work readiness of accounting students moderated by work motivation]. *Jurnal Pendidikan Akuntansi Indonesia*, 24(1), 97–110. <https://doi.org/10.21831/jpai.v24i1.83607>
- Rainio, O., Teuho, J., & Klén, R. (2024). Evaluation metrics and statistical tests for machine learning. *Scientific Reports*, 14, Article 6086. <https://doi.org/10.1038/s41598-024-56706-x>
- Sakib, M. N., Rabbani, M. R., Hawaldar, I. T., Jabber, M. A., Hossain, J., & Sahabuddin, M. (2022). Entrepreneurial competencies and SMEs' performance in a developing economy. *Sustainability*, 14(20), Article 13643. <https://doi.org/10.3390/su142013643>
- Sinaga, H. H., & Agustian, S. (2022). Pebandingan Metode Decision Tree dan XGBoost untuk Klasifikasi Sentimen Vaksin COVID-19 di Twitter [Comparison of decision tree and XGBoost methods for sentiment classification of COVID-19 vaccines on

- Twitter]. *Jurnal Nasional Teknologi dan Sistem Informasi*, 8(3), 107–114. <https://doi.org/10.25077/TEKNOSI.v8i3.2022.107-114>
- Supriyati, E., Iqbal, M., & Listyorini, T. (2026). Multimodal explainable AI and blockchain integration for automated halal verification of cosmetic products. *IIUM Engineering Journal*, 27(2), 175–188. <https://doi.org/10.31436/iiumej.v27i2.4059>
- Suryani, S., & Chaniago, H. (2023). Digital literacy and its impact on entrepreneurial intentions: Studies on vocational students. *International Journal Administration, Business & Organization*, 4(2), 16–22. <https://doi.org/10.61242/ijabo.23.261>
- Sutiadiningsih, A., & Mahfud, T. (2023). Can the teaching factory model improve the entrepreneurial intentions of vocational high school students? *International Journal of Evaluation and Research in Education (IJERE)*, 12(3), 1654–1662. <https://doi.org/10.11591/ijere.v12i3.25652>
- Tran, T. M. (2023). The digital economic ecosystem in Vietnam. *Journal of Multidisciplinary Science: MIKAILALSYS*, 1(3), 260–272. <https://doi.org/10.58578/mikailalsys.v1i3.1895>
- Ubaidillah, R., Muliadi, M., Nugrahadhi, D. T., Faisal, M. R., & Herteno, R. (2022). Implementasi XGBoost pada Keseimbangan Liver Patient Dataset dengan SMOTE dan Hyperparameter Tuning Bayesian Search [Implementation of XGBoost on the balancing of a liver patient dataset with SMOTE and Bayesian search hyperparameter tuning]. *Jurnal Media Informatika Budidarma*, 6(3), 1723–1729. <https://doi.org/10.30865/mib.v6i3.4146>
- Widodo, Baswedan, A. R., Suyata, P., & Saputra, W. N. E. (2025). Entrepreneurship education in vocational schools: An Indonesian model. *International Journal of Evaluation and Research in Education (IJERE)*, 14(1), 373–381. <https://doi.org/10.11591/ijere.v14i1.32317>
- Zakariya, Y. F. (2022). Cronbach's alpha in mathematics education research: Its appropriateness, overuse, and alternatives in estimating scale reliability. *Frontiers in Psychology*, 13, Article 1074430. <https://doi.org/10.3389/fpsyg.2022.1074430>
- Zhao, J., Zhao, L., Chong, K. M., Hamid, M. B. B. Y. B., & Osman, M. Z. (2025). Entrepreneurial intention among university students in the digital economy: The mediating role of alertness and the moderating effect of motivation. *Humanities and Social Sciences Communications*, 12, Article 1861. <https://doi.org/10.1057/s41599-025-06136-5>