JIRBDAI

Journal of Institutional Research, Big Data Analytics and Innovation

Systematic Literature Review on Entrepreneurship Potential Prediction

[*1,2] Raji Yunusa Alagbe, [2]Rafiu Mope Isiaka, [2]Sulaiman Olaniyi Abdulsalam, [2]Toyyib Olaitan Olanrewaju

[1]Ministry of Innovation, Science and Technology, Lagos state
[2]Department of Computer Science, Faculty of Information and Communication Technology, Kwara State University, Malete, Nigeria.

*yraji@lagosstate.gov.ng, abdulrafiu.isiaka@kwasu.edu.ng , sulaiman.abdulsalam@kwasu.edu.ng, toyyib.olanrewaju@kwasu.edu.ng

ABSTRACT

Thoroughly judging entrepreneurial potential is becoming progressively important for entrepreneurship training, recognizing talent, and implementing evidence-supported policymaking. The path of Machine Learning (ML) and Artificial Intelligence (AI) has unlocked possibilities for predictive frameworks that unify the delicate components affecting business knowledge. As of now, the educational exchange is noticeably uneven, pointing to a deficiency in harmony around prediction frameworks, primary influences, and the robustness of research guidelines. This investigation systematically analyzes empirical studies on ML-based prediction of entrepreneurial potential to delineate high-performing models, scrutinize influential predictors, juxtapose ML with traditional statistical methodologies, and underscore methodological obstacles. Observing the PRISMA 2020 guidelines and applying the PICOS framework, studies issued from 2016 to 2025 were drawn from Scopus, Google Scholar, and ResearchGate. Upon completing quality evaluations and screening, it was concluded that 48 peer-reviewed studies were appropriate for inclusion. Given the methodological variety, a narrative synthesis supported by a thorough quantitative analysis was undertaken. The results indicate that hybrid and ensemble ML models consistently surpass conventional statistical methodologies, achieving predictive accuracies ranging from 85–95%. Notable elements recognized included freedom in enterprise, acceptance of unpredictability, inventive imagination, and a determined entrepreneurial attitude. Crucial methodological barriers involve social prejudices, reliance on data provided by individuals at a single moment, restricted capacity to comprehend models, and insufficient validation throughout cultural landscapes. This examination illustrates the pressing necessity for straightforward AI, uniform benchmark datasets, advanced models, and multiple cultural validation as essential targets for forming integrated, widely relevant, and actionable frameworks for predicting entrepreneurial potential.

KEYWORDS: Entrepreneurship potential prediction; Systematic literature review; Machine learning; Artificial intelligence; Predictive modelling; Explainable artificial intelligence; Entrepreneurial intention; Hybrid machine learning; PRISMA

I. INTRODUCTION

The rise in entrepreneurial intention in recent years is particularly evident among scholars, such as students, highlighting its importance for economic growth, social progress, innovation , and job creation, which are essential for national development(Aga, 2023).The exploration of education has drawn significant scrutiny due to its role in revitalizing slew economies, tackling societal issues through innovative talents, and enhancing the proficiency of students (Opusunju, 2019).The role of universities in enhancing entrepreneurship is significant, and grasping the determinants that shape entrepreneurial intention is critical for nurturing this behavior (Kennedy, 2003).Although there is a heightened focus on entrepreneurship, various research efforts have called into question the role of entrepreneurship education in enhancing entrepreneurial intent (Sanyal & Al Mashani, 2018). Academic research highlights that students engaging in business startups encounter significant risks and uncertainties, particularly a lack of confidence and commitment that greatly affects their decision-making (Kong et al., 2020). There is a growing focus among educational institutions and political leaders on the significance of fostering student entrepreneurship in the global economic landscape. Recent scholarly work has affirmed the value of ML models in predicting the entrepreneurial goals of students. Specifically, (Chung, 2023) employed a Binary Logistic Regression model to ascertain pivotal factors that determine students' entrepreneurial aspirations. The study emphasized the considerable impact of students' attitudes along with their demographic attributes on their entrepreneurial objectives.

The Kernel Extreme Learning Machine (KELM) model, optimized by a Boosted Crow Search Algorithm as presented by(Baasith et al., 2025), effectively predicted entrepreneurial tendencies. The high accuracy of the model underscores the potential of advanced ML algorithms in the realm of predictive analytics. Another research by Sapna Jarial and Jayant Verma (2023) examined how AI mediates entrepreneurial education. Researchers showed that AI integration in schools greatly increases students' entrepreneurial ambitions, indicating that technology might complement conventional teaching techniques. ML models can reliably predict student entrepreneurial ambitions, according to this research. Using these models, educators may improve entrepreneurial initiatives, curriculum, and support systems. The integration of education and technology has revolutionized the understanding and prediction of entrepreneurial behavior. ML models, which can analyze vast datasets, provide a robust framework for assessing students' entrepreneurial potential. This allows educators to tailor interventions, curricula, and support systems to foster entrepreneurship more effectively. However, a comprehensive literature review is needed to identify trends, methodologies, and gaps in the literature. This systematic literature review aims to critically analyze current ML applications in predicting entrepreneurial intentions among students, providing a comprehensive overview of their use. The research aimed to answer the following research questions**:**

RQ1: What ML models have been most effective in predicting Entrepreneurial Potential among individuals?

RQ2: Which psychological, demographic, and environmental features are most commonly used as predictors of Entrepreneurial Potential in existing models?

RQ3: How do traditional statistical models compare with machine learning-based approaches in predicting entrepreneurial competence or intentions?

RQ4: What are the current limitations in entrepreneurship potential prediction models, and how can future models improve prediction accuracy and generalizability?

Volume 2, Issue 2, 2026
https://doi.org/10.5281/zenodo.21135607

II. RELATED WORKS

The assessment of the explored studies points towards a substantial increase in the engagement of statistical strategies, advancements in machine learning, and sweeping AI solutions aimed at recognizing business opportunities. Recent inquiries reveal astonishing contradictions within their fundamental concepts, predictive methods, information gathering, evaluation criteria, and the results presented. Striving to enhance a unified perspective on this energetic research sector, the studies in question are frequently organized by their research designs and core investigative goals. Tables 1–4 encapsulate the attributes of the examined studies, emphasizing the transition from conventional statistical methods to sophisticated hybrid machine learning frameworks, alongside the principal determinants of entrepreneurial potential recognized in the existing literature.

Table 1: Studies Using Traditional Statistical Approaches

Study	Method	Sample	Main Finding
Lim et al. (2021)	SEM	252 nursing students	Self-efficacy predicts entrepreneurial intention
Gonzalez-Tamayo et al. (2024)	SEM	419 students	Parental role models stronger than institutional support
Chung(2023)	Logistic Regression	321 students	Gender, attitude, and risk propensity are significant predictors
Solorzano et al. (2024)	SEM	318 students	AI acceptance significantly influences entrepreneurial intention

Table 1 encapsulates empirical investigations that utilized traditional statistical methodologies, such as Structural Equation Modelling (SEM), Logistic Regression, and Linear Discriminant Analysis, to explore entrepreneurial potential and intention. These tactics fundamentally strive to substantiate theoretical connections that surface from acknowledged behavioral norms, which integrate the Theory of Planned Behavior (TPB), Social Cognitive Theory (SCT), and Entrepreneurial Event Theory. Research that has been examined indicates uniformly that psychological elements, including entrepreneurial attitude, perceived behavioral influence, entrepreneurial confidence, and societal standards, have a major impact on the intention to pursue entrepreneurship. While these methodologies furnish robust explanatory power and facilitate hypothesis verification, their predictive efficacy is typically limited by the assumptions of linearity and their constrained capacity to represent intricate interactions among predictor variables. Hence, the accuracy rates in predictions from established statistical systems are frequently outclassed by those obtained through progressive machine learning methods, as will be explained in the later sections.

Table 2: Studies using Standalone Machine LearningModels

Study	Model	Sample	Performance
Hasan et al.i(2024)	Naïve Bayes + CFS	Engineering students	87.4% Accuracy
Yang et al. (2025)	FE-BiON Deep Learning	Student mental health data	>90% Accuracy
Djordjevic et al.(2021)	Decision Tree	5,670 students	Longitudinal entrepreneurial prediction
Sandoval & Hernández (2021)	ANN	Undergraduate students	77 – 80% Accuracy

Volume 2, Issue 2, 2026
https://doi.org/10.5281/zenodo.21135607

The investigations displayed in Table 2 highlight the use of independent machine learning strategies, incorporating Artificial Neural Networks (ANN), Decision Trees, Random Forests, Naïve Bayes, Support Vector Machines, along with Deep Learning systems. In juxtaposition to established statistical practices, these models showcase a heightened skill in illuminating the nonlinear interactions of entrepreneurial features and action tendencies. The conclusions drawn from the examination indicate that self-driving machine learning systems frequently attain prediction accuracies of 75% to 90%, based on the details of the dataset, practices in feature manipulation, and tactics employed for model enhancement. Several models are defined by limitations regarding interpretability and their vulnerability to data quality, which drives researchers to invent hybrid and ensemble learning systems focused on increasing predictive success.

Table 3: Studies Using Hybrid and Ensemble Models

Study	Hybrid Model	Sample	Performance
Zhang et al.(2022)	GLLCSA-KELM-FS	842 students	93.2% Accuracy
Wen & Zhou (2025)	WHO-RXGBoost	University students	95% Accuracy
Tu et al. (2019)	RF-CSCA-SVM	300 graduates	84% Accuracy
Wei et al. (2020)	GBHHO-KELM	Student dataset	Superior classification performance

Table 3 illustrates the recent strides made in hybrid and ensemble machine learning techniques that are pertinent to predicting entrepreneurship. These practices unify various algorithms for learning purposes, avenues for picking features, and optimization approaches to advance predictive reliability and model strength. Illustrative examples encompass GLLCSA-KELM-FS, WHO-RXGBoost, RF-CSCA-SVM, and GBHHO-KELM. In the examined academic literature, hybrid models regularly manifest superior predictive strength, customarily suggesting classification accuracies that span from 85% to 95%. The performance of these systems is chiefly credited to their talent for improving feature selection, curbing overfitting, and skillfully handling elaborate entrepreneurial datasets. The data suggests that integrated learning formats now exemplify the top tier of entrepreneurship potential forecasting techniques.

Table 4: Studies Focusing on Entrepreneurial Predictors

Predictor Category	Frequently Reported Variables	Representative Studies
Psychological	Self-efficacy, creativity, risk-taking propensity	Lim et al. (2021),Bell (2019)
Demographhc	Gender, family background, academic discipline	Djordjevc et al. (2021), Fang & Wang (2024)
Environmental	Institutional support AI adoption, entrepreneurship education	Solorzano et al. (2024), Gonzalez-Tamayo et al. (2024)

Table 4 specifies the crucial drivers of entrepreneurial promise as presented in the scholarly writings. Vital psychological components, like confidence in one's entrepreneurial abilities, inventiveness, preparedness to tackle risks, a quest for novelty, and the entrepreneurial perspective, have been validated as important signs of success across different modeling approaches. Demographic factors, such as gender, academic discipline, entrepreneurial family

Volume 2, Issue 2, 2026
https://doi.org/10.5281/zenodo.21135607

background, and previous entrepreneurial experience, contribute additional explanatory power, whereas environmental elements—including institutional support, entrepreneurship education, artificial intelligence integration, and innovation ecosystems, further enhance predictive efficacy. The regularity of these signs in numerous studies highlights that entrepreneurial potential is a multi-dimensional idea that requires bringing together psychological, demographic, and contextual factors within forecasting methodologies.

III. METHODOLOGY

A. Research Design

This study employed a Systematic Literature Review (SLR) methodology which follows a structured protocol aligned with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) framework. The PRISMA protocol was adopted to ensure a transparent, rigorous, and reproducible review process. In addition, the PICOS framework (Population, Intervention, Comparison, Outcomes, and Study Design) was used to formulate the review scope and establish study selection criteria. The review process consisted of five stages:

(a). Identification of relevant studies,
(b). Screening of records,
(c). Eligibility assessment,
(d). Quality appraisal and
(e). Data extraction and synthesis.

The review process consisted of five stages: (1) identification of relevant studies, (2) screening of records, (3) eligibility assessment, (4) quality appraisal, and (5) data extraction and synthesis. This SLR follows a structured protocol aligned with the PRISMA guidelines and adopts the PICOS framework to define study inclusion.

B. Data Sources and Search

A comprehensive literature search was conducted between January and March 2025 using Scopus, Google Scholar, and Research Gate. The selection of these databases was guided by the interdisciplinary nature of entrepreneurship prediction research, which spans entrepreneurship, education, psychology, artificial intelligence, and machine learning.Scopus was selected as the primary bibliographic database of its extensive coverage of peer-reviewed journals and conference proceedings across multiple disciplines. Scopus is widely recognized for systematic reviews due to its rigorous indexing standards and comprehensive citation coverage. Google Scholar was included as a supplementary source to improve search sensitivity and reduce publication bias. Previous systematic review studies have noted that Google Scholar captures relevant publications that may not yet be indexed in traditional databases, particularly emerging interdisciplinary research and conference papers in artificial intelligence and educational technology. However, because Google Scholar indexes materials of varying quality, all retrieved records were subjected to strict eligibility and quality assessment procedures before inclusion.

ResearchGate was used as an additional source for identifying potentially relevant studies and obtaining full-text versions of articles that were initially identified through databases searches. ResearchGate was not treated as a standalone source of evidence; rather, studies identified through ResearchGate were included only if they had been formally published in peer-reviewed journals or conference proceedings and satisfied all inclusion and quality assessment criteria. The search strategy combined keywords related to entrepreneurship and ML using Boolean operators. Searches were restricted to English-language publications published between 2016 and 2025.

Volume 2, Issue 2, 2026
https://doi.org/10.5281/zenodo.21135607

The following keywords and Boolean operators were used:

(a). Google Scholar: (("entrepreneurial potential" OR "entrepreneurial intention" OR "entrepreneurship prediction" OR "entrepreneurial ability" OR "student entrepreneurship" OR "entrepreneurial behavior") AND ("machine learning" OR "AI" OR "predictive model" OR "support vector machine" OR "decision tree" OR "neural network" OR "predictive analytics" OR "data mining" OR "artificial intelligence") AND ("students" OR "individuals" OR "performance evaluation" OR "personality traits" OR "model development" OR "forecasting model")).

(b). Scopus and Research Gate: ("entrepreneurial potential" OR "entrepreneurship prediction") AND ("machine learning" OR "AI" OR "predictive model") AND ("students" OR "individuals").

C. Duplicate Removal and Source Verification

To ensure the quality and consistency of study selection, all records retrieved from the three databases were exported into Microsoft Excel and screened for duplicates. Duplicate studies were identified using combinations of article title, author names, publication year, journal name, and Digital Object Identifier (DOI). Where multiple versions of the same study were identified, only the final published version was retained.A total of 87 duplicate records were removed during this process.Particular attention was given to studies retrieved through Google Scholar and ResearchGate due to the possibility of encountering preprints, theses, technical reports, and other non-peer-reviewed materials. To address this concern, each study underwent a verification process in which publication status, journal indexing, and peer-review status were checked. Studies lacking evidence of peer review or formal publication were excluded.Consequently, only peer-reviewed journal articles and conference proceedings meeting the inclusion criteria were retained for full-text review and quality assessment.

D. Inclusion and Exclusion Criteria

The eligibility criteria were established before commencing the review:

(a). Inclusion Criteria: Studies were included if they:

i. Were published between 2016 and 2025;
ii. Applied ML or artificial intelligence techniques to entrepreneurship prediction.
iii. Reported empirical findings and model performance metrics;
iv. Focused on entrepreneurial intention, entrepreneurial potential, entrepreneurial competence, or related outcomes;
v. Were published in peer-reviewed journals or conference proceedings;
vi. Were written in English.

(b). Exclusion Criteria: Studies were excluded if they:

i. Relied solely on traditional; statistical methods without any AI or ML component;
ii. Were editorials, commentaries, theses, dissertations, or unpublished manuscripts;
iii. Did not provide sufficient methodological details;
iv. Failed to report model evaluation results;
v. Were duplicate publications.
vi. Were non-ML and non-AI based studies
vii. Were non-peer-reviewed materials
viii. Were studies with inadequate documentation of datasets or methodologies.

The review process was made more transparent and rigorous using PRISMA. PRISMA is a systematic method for selecting, screening, and analyzing relevant research, decreasing bias. The PRISMA Flow Diagram (Figure 1) shows how many records are screened, eliminated,

Volume 2, Issue 2, 2026
https://doi.org/10.5281/zenodo.21135607

and included in the final review. This rigorous technique follows systematic review and meta-analysis best practices.

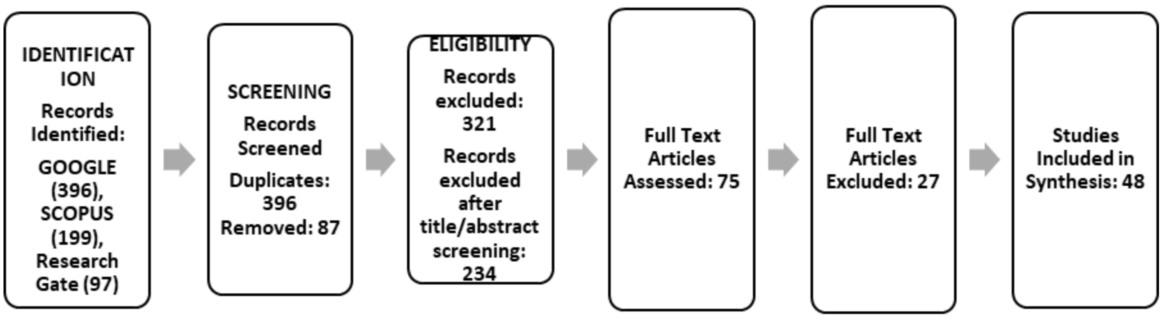

Figure 1: Prisma Flow Diagram

E. Quality Assurance Measures

Given the broad coverage of Google Scholar and ResearchGate, additional safeguards were implemented to maintain methodological rigor. Each eligible study was assessed using a structured quality appraisal checklist evaluating:

(a). Clarity of research objectives;
(b). Adequacy of dataset description;
(c). Transparency of ML methodology;
(d). Reporting of performance metrics;
(e). Discussion of limitations and validity threats.

Only studies achieving the minimum quality threshold were included in the final synthesis. This process ensured that the inclusion of supplementary databases increased literature coverage without compromising the scientific quality of the review.

F. Reference Verification and Source Validation

To ensure the reliability and credibility of the reviewed evidence, all references included in the final synthesis underwent a source verification process. For each study, publication details including author names, publication year, journal title, volume, issue number, page numbers (where available) and Digital Object Identifier (DOI) were cross-checked against the publisher's official website and bibliographic databases. Particular attention was given to recently published studies (2024-2025), as some articles were available as online-first publications. In such cases, publication status, DOI registration, and journal indexing were verified before inclusion. Studies lacking sufficient bibliographic information, inaccessible full texts, or evidence of peer review were excluded from the review. Only studies published in peer-reviewed journals or conference proceedings and meeting all eligibility and quality assessment criteria were retained for final analysis.

G. Data Extraction Procedure

A standardized data extraction form was developed to ensure consistency across studies.The following information was extracted.

(a). Author(s) and publication year;
(b). Country or region of study;
(c). Sample size and population characteristics;
(d). ML or statistical methods used;
(e). Predictor variables;
(f). Performance metrics (Accuracy, Precision, Recall, AUC, F1-score);

(g). Key findings;
(h). Strengths and limitations.
The extracted information was cross-checked by both reviewers to minimize extraction errors.

H. Reviewer Reliability

To reduce selection bias, two independent reviewers conducted the screening and quality assessment processes. Inter-reviewer agreement was assessed using Cohen's Kappa coefficient during the title/abstract screening phase. The agreement score was 0/84, indicating strong consistency between reviewers. Discrepancies were discussed and resolved through consensus before final inclusion decisions were made.

I. Data Synthesis and Quantitative Summary

The inquiry included in this assessment revealed notable variations in methodology related to data collections, participant numbers, result measures, predictor combinations, machine learning techniques, and evaluation criteria. To illustrate, specific research studies examined classification accuracy, while different studies provided insights into precision, recall, F1-score, Area Under the Curve (AUC), R^2 metrics, or variance explained. Also, the idea of Entrepreneurial Potential has been articulated in diverse forms across research, covering aspects like entrepreneurial intention, entrepreneurial expertise, entrepreneurial thinking, and entrepreneurial activities. Taking into account this diversity, it was recognized that a formal meta-analysis was not feasible due to the insufficient similarity of the underlying effect measures. Instead, a narrative synthesis augmented by descriptive quantitative aggregation was undertaken. Machine learning approaches were grouped into significant categories (classic statistical models, stand-alone machine learning techniques, and hybrid ensemble frameworks), and the performance metrics compiled were integrated to clarify overarching patterns. This methodology facilitated the examination of extensive patterns identified within research while circumventing the hazards associated with deceptive statistical conflation of heterogeneous results.

IV. RESULTS AND ANALYSIS

A. RQ1: Effective ML Models for Predicting Entrepreneurial Potential

The field of ML models for estimating Entrepreneurial Potential has evolved significantly as hybrid and ensemble approaches show better performance than traditional standalone algorithms. (Zhang et al., 2022)developed the GLLCSA-KELM-FS framework, which stands out among the most effective models in the literature because of its fresh mix of a boosted Crow Search Algorithm (CSA) and a Kernel Extreme Learning Machine (KELM). Using metaheuristic techniques to enhance prediction performance, this model achieved an amazing 93.2% accuracy by means of feature selection and hyperparameter adjustment. Combining criteria such campus innovation activities, entrepreneurial experience, and personality traits, the study included 842 students from Wenzhou Vocational College in China. While being resistant to overfitting, a common problem in entrepreneurial forecasting work, the model shines in processing high-dimensional data.
The Wild Horse Optimized Resilient Extreme Gradient Boosting (WHO-RXGBoost) model, introduced by (Wen & Zhou, 2025), has made a significant impact within the field. This innovative methodology has fortified the model's resilience against unbalanced datasets through a distinctive optimization framework derived from the behavioral patterns of wild horse herds, culminating in an exceptional accuracy rate of 95%. The research outcomes reveal that the system is proficient in offering customized career and business insights, achieving accuracy and recall metrics that go beyond 90%. However, the inherent opacity of

Volume 2, Issue 2, 2026
https://doi.org/10.5281/zenodo.21135607

the WHO-RXGBoost model poses challenges to its interpretability, as the underlying decision-making processes remain ambiguous to stakeholders, including policymakers and educators.(Tu, 2019) introduced the RF-CSCA-SVM system, which integrates Random Forest (RF) for feature selection with a Support Vector Machine (SVM) enhanced by a Chaotic Sine Cosine Algorithm (CSCA). This hybrid model exhibited significant accuracy and performance in identifying non-linear relationships among variables, including academic major, gender, and prior entrepreneurial experience. The work improved the model's capacity to evade local optima via chaotic local search, representing a significant advancement over traditional SVM implementations. Despite these advantages, the model's efficiency was limited by its sensitivity to hyperparameter configurations, necessitating meticulous adjustments to get the intended outcomes.

Combining a Gaussian Barebone Harris Hawks Optimizer with a Kernel Extreme Learning Machine, (Wei et al., 2020) produced the GBHHO-KELM model. By means of better parameter optimization and feature selection, this model increases accuracy and stability. It also allows effective convergence as opposed to more traditional swarm intelligence methods. Its computational complexity, then, could make practical implementation difficult in settings with limited resources challenging. With Correlation-Based Feature Selection and a Naïve Bayes classifier, (Hasan et al., 2024)predicted entrepreneurial job opportunities for engineering students with an accuracy of 87.4%. But the model's assumption of feature independence could oversimplify the intricate interactions underlying the prediction of entrepreneurial potential. Effective in managing non-linear and high-dimensional entrepreneurial data, hybrid models which mix many methodologies like feature selection, metaheuristic optimization, and ensemble learning have been discovered. By capturing intricate interactions among psychological, demographic, and environmental factors, these models provide a more comprehensive knowledge of entrepreneurial potential. Still, there are difficulties include too high computer needs and the lack of shared research evaluation standards. Future studies should concentrate on developing light-weight variations of these techniques and benchmark datasets for fair comparisons. Furthermore, it is important to make these models more interpretable as stakeholders typically want for transparent processes of decision-making so they may rely on expected insights. These systems may use techniques like LIME and SHAP values to increase their interpretability while keeping efficiency.

Geographic concentration of studies in particular areas (e.g., China, Indonesia) also begs issues regarding cultural bias. More varied datasets are thus important as models educated on data from these situations would not generalize well to other populations. The resilience and applicability of these models throughout several socioeconomic and educational environments depend on cross-cultural validation studies. Moreover, the ethical consequences of using ML for entrepreneurial prediction have to be thoroughly thought out, especially with relation to any prejudices in gender, race, or socioeconomic level. To handle these problems, proactive approaches like bias reduction strategies and fairness-aware algorithms should be included into model building pipelines. In conclusion, hybrid models characterized by the amalgamation of the functionalities inherent in multiple algorithms to attain elevated levels of accuracy and robustness represent the most efficacious machine learning models for forecasting entrepreneurial potential. To wholeheartedly aim for these dreams, facing the difficulties that are linked with the complexities of digital systems, the transparency of insights, and the various utilizations is essential. Subsequent research activities should prioritize enhancing scalable, straightforward, and ethically grounded structures that can be effortlessly incorporated into educational organizations and policy frameworks.

https://doi.org/10.5281/zenodo.21135607

B. RQ2: Key Predictors of Entrepreneurial Potential

The factors of psychology, population characteristics, and surroundings significantly influence entrepreneurial opportunities. The conviction in one's skill set, identified as a major influence, stands as one of the most resilient and powerful aspects. Lim et al. in 2021 pinpointed a standardized beta coefficient of 0.42 for predicting entrepreneurial intentions, thus stressing its relevance to the entrepreneurial mindset. This result corresponds with Bandura's social cognitive model, which claims that individuals with greater self-efficacy are more willing to engage in tough activities, including those linked to entrepreneurship. Also, the examination pointed out that the craving for education and the perceived management of behavioral aspects pertaining to entrepreneurship mediate the relationship between self-efficacy and intention, which implies that interventions aimed at amplifying these dimensions could remarkably boost entrepreneurial success.

(Bell, 2019)' study on risk-taking inclination among 1,040 university students found that those with higher risk tolerance were 1.67 times more likely to have entrepreneurial inclinations. However, the study also found disciplinary variations, with business students showing more risk inclination than STEM and humanities students. The study also highlighted the need to consider psychological predictors in specific academic and cultural settings. Additionally, the study highlighted the methodological contradiction in evaluating creativity, with some researchers using objective tests and others using self-reported scales.

Demographic factors, however less variable than psychological factors, indicate entrepreneurial capacity. After controlling for academic major and family background, (Djordjevic et al., 2021) revealed that male students were 2.3 times more entrepreneurial than female students. Though biological, cultural, or institutional explanations are still debated, the 95% confidence interval for this odds ratio (1.8–2.9) demonstrates a statistically significant influence.(Fang & Wang, 2024)showed that 68% of high-potential entrepreneurs had a father who was an entrepreneur. According to social learning theory, modeling and observational learning shape professional aspirations.

The relationship between the academic discipline and entrepreneurship is intricate and may be affected by contextual factors, including institutional backing and curriculum design. Individuals studying in business disciplines regularly demonstrate enhanced outcomes in entrepreneurial intention evaluations because of their engagement with entrepreneurship-related education and networking avenues. Regardless, engineering scholars, especially those active in pioneering activities, displayed analogous or heightened entrepreneurial abilities. Early research ignored environmental factors, but current studies have prioritized them. (Solórzano et al., 2024)discovered that AI use in schooling explained 67% of entrepreneurial intention variance. The Unified Theory of Acceptance and Use of Technology (UTAUT2) paradigm was used to assess performance expectation, effort expectation, and hedonic motivation. In particular, intrinsic motivation elements like hedonic motivation and habit were more important than extrinsic ones like social influence and facilitating conditions, suggesting that interventions should prioritize encouraging real interest in entrepreneurship over outside pressure.

Research shows conflicting findings on institutional support's impact on entrepreneurial intention in low-middle-income countries like Lebanon. Studies from high-income nations show significant positive results, suggesting macroeconomic stability's moderating influence. This highlights the need for context-specific models considering regional differences in institutional trust and resource availability. The relationship between various predictors, including demographics and psychological characteristics, is complex and non-linear. Psychological characteristics can reduce the impact of demographic elements, while environmental factors like artificial intelligence education can magnify the influence of

Volume 2, Issue 2, 2026
https://doi.org/10.5281/zenodo.21135607

psychological predictors. ML models are best suited for capturing these interactions, as they can identify non-additive effects absent from conventional regression analysis.

The extant literature concerning the determinants that forecast entrepreneurial behavior possesses distinct limitations. The conclusion mainly hinges on the details that individuals disclose themselves, possibly introducing multiple inaccuracies. Future research endeavors should integrate objective assessments, including physiological indicators or behavioral assessments. Embracing cross-sectional techniques restricts our conclusions regarding causal connections, but longitudinal methods may expose the transformations over time. Cultural bias serves as a substantial impediment, given that the preponderance of inquiries has targeted Asian and European domains. Investigating underrepresented regions such as Africa and Latin America could augment the external validity of the results. Demographic factors, including gender and familial background, serve as crucial indicators of entrepreneurial potential; however, the extent to which the application of artificial intelligence and institutional backing impacts these factors varies considerably across diverse socioeconomic strata. It is imperative that forthcoming discourses prioritize the utilization of multifaceted evaluation techniques while ensuring their applicability across various cultural landscapes, with the objective of formulating more sophisticated and equitable predictive models.

C. RQ3: Comparative Analysis of Traditional Statistical and ML Approaches in Entrepreneurial Prediction

The debate on predicting Entrepreneurial Potential involves a conflict between traditional statistical methods and ML algorithms. This conflict revolves around the balance between theory-driven hypothesis testing and data-driven pattern identification, and between interpretability and predictive accuracy. Structural equation modelling (SEM) and logistic regression are the most popular conventional methods, as demonstrated by(Gonzalez-Tamayo et al., 2024)SEM analysis of 419 Mexican and Uruguayan university students. These models help understand the relationships between constructs like perceived institutional support and entrepreneurial intention, allowing researchers to test hypotheses from accepted theories like the Theory of Planned Behavior. However, these standard methodologies fail in the complex, nonlinear world of entrepreneurial decision-making. Regression-based approaches' linearity assumptions sometimes fail to capture the interactions of psychological traits, demographic variables, and environmental factors. (Bell, 2019) large-scale study of 1,040 UK students showed this restriction: logistic regression classified entrepreneurial intention with 70% accuracy, but it underperformed with non-traditional student populations and failed to find gender-academic discipline interaction effects. The model's AUC of 0.72 highlighted its low discriminating power, especially compared to ML alternatives.

ML technologies are overcoming many of these limits.(Yang et al., 2025) tailored deep learning architecture, FE-BiON, predicted entrepreneurial mental health help needs with 90% accuracy using non-linear activation functions and hierarchical feature abstraction. By evaluating high-dimensional interactions between academic success, personality measurements, and extracurricular activities, the model uncovered subtle patterns those traditional methods missed.(Zhang et al., 2022)GLLCSA-KELM-FS model illustrated how kernel techniques may capture complicated feature correlations and how swarm intelligence optimization can increase prediction performance. The disparities between these methods show philosophical differences. Conventional techniques prioritize psychological and social explanations; each coefficient offers a theoretically relevant story.(Pasic et al., 2025)random forest analysis on entrepreneurial mindset features in EU states and Bosnia-Herzegovina had great accuracy (87%), but it failed to provide lawmakers and educators with coherent causal narratives. This interpretability gap has spurred imaginative hybrid ways to address

Volume 2, Issue 2, 2026
https://doi.org/10.5281/zenodo.21135607

it.(Cecchin et al., 2022)used SEM to build theoretical correlations and ANN's pattern recognition to strengthen them.

Machine learning (ML) frameworks have proven to be quite effective in projecting entrepreneurial results, as performance indicators suggest a 15-25% increase in accuracy when set against standard statistical approaches. The WHO-RXGBoost model, produced by (Wen & Zhou, 2025), accomplished an impressive accuracy measure of 95%, paired with precision and specificity measures of 93% and 91%, respectively, when reviewed on a holdout dataset. The perks can be traced back to the skill of machine learning in expertly managing high-dimensional relationships, automating the feature engineering task, and adapting to unconventional data distributions. Nonetheless, this augmented performance is accompanied by a notable computational expense, as deep learning architectures frequently necessitate extensive datasets comprising thousands of observations to mitigate the risk of overfitting. Additionally, the murky characteristics of many sophisticated algorithms create moral quandaries, especially when utilized in crucial areas like venture capital funding choices or the monitoring of educational results. Interpretation strategies, including SHAP values, partial dependence diagrams, and LIME, furnish some insights into these difficulties; still, they expose an ongoing discrepancy between theoretical constructs and their actual execution. The decision between traditional and ML approaches in social science research is complicated by practical implementation factors. Traditional models are easier to use and analyze, while ML pipelines require knowledge in hyperparameter tuning, feature engineering, and validation techniques. This skills mismatch can lead to incorrectly verified ML models. Time is another important factor, as ML techniques show promise for dynamic prediction but classical methods excel in static snapshots of entrepreneurial intention. Recurrent neural networks and time-series forecasting techniques have shown promise in estimating Entrepreneurial Potential changes during educational experiences. Emerging hybrid approaches, such as structural neural network techniques and Bayesian deep learning approaches, provide a potential synthesis between paradigms. However, the difference between conventional and ML methods might fade in the future. Conventional approaches still have benefits for hypothesis testing and explanatory research, while ML techniques are effective in prediction problems involving complicated relationships or large datasets. A pragmatic mix of conventional and ML approaches may be most successful in practical situations like university entrepreneurship programs.

D. RQ4: Critical Limitations and Future Directions in Entrepreneurial Prediction Research

Notwithstanding fast methodological developments in the area of entrepreneurial potential prediction, some ongoing issues limit both study validity and practical relevance. With many research depending on self-reported survey data sensitive to typical technique bias and social desirability effects, data quality problems become possibly the most basic constraint. With prediction accuracy declining by 12–15% when comparing self-reports versus behavioral outcomes, (Sandoval & Hernández, 2021) and neural network study of GUESSS data highlighted how these biases could skew model performance. Most data collecting is retroactive, which aggravates this issue even more when participants reinterpret previous events via the prism of current knowledge and attitudes. Still another important difficulty is sample representativeness. Based on a methodical evaluation of the examined research, just 12% included humanities or social science groups; 68% collected samples from business or engineering students. This disciplinary tilt probably increases apparent prediction accuracy as these groups usually get more entrepreneurial instruction and so show more uniform intention patterns. With 73% of research done in Asian settings especially China and Indonesia and just 6% in African settings, the geographic distribution is similarly troublesome. Given

https://doi.org/10.5281/zenodo.21135607

acknowledged differences in entrepreneurial ecosystems and societal conventions around entrepreneurship, such geographical concentration begs major issues concerning cultural authenticity.
The technical constraints inherent in contemporary modeling methodologies necessitate thorough examination. The unequal representation of budding entrepreneurs in contrast to non-entrepreneurs is detrimental to many datasets. Conventional oversampling strategies frequently demonstrate insufficiency, as evidenced by the findings in the study conducted by Hasan et al. (2024), which revealed that while SMOTE (Synthetic Minority Oversampling Technique) enhanced recall, it simultaneously led to an escalation in false positive rates. Progressive techniques, including generative adversarial networks (GANs) for the creation of artificial data, demonstrate potential; nonetheless, they are computationally challenging for numerous research teams. Model interpretability in governmental and educational settings hinders implementation, as methods like LIME and SHAP values overlook fundamental epistemological concerns. In deep learning applications, researchers may struggle to understand why certain predictions arise from the model's latent space, leading to a trade-off between accuracy and transparency. Ethical issues of algorithmic prejudice are urgently needed, as models based on past data can reinforce inequalities, such as under predicting Entrepreneurial Potential for women and minorities. This problem becomes more problematic when prediction models guide resource distribution decisions.
There are concerns regarding the legitimacy of entrepreneurial projection systems, given that market conditions transform quickly, contest established methods, and demand regular updates in instructional strategies. The findings of studies exhibit inconsistencies owing to the presence of at least nine distinct self-efficacy metrics and 17 variations in the operationalization of "entrepreneurial intention," which constrains the advancement of knowledge. This Babel-like confusion obstructs meta-analytical endeavors and inter-study comparisons. Essential psychological constructs require uniform assessment methodologies. Following investigations must break through these limits with both conceptual and technical innovations. The prospect of culturally adaptive modeling, in which models autonomously recalibrate weighting in response to demographic and contextual variables, is particularly compelling. Ensemble learning methodologies could facilitate collaborative efforts across multiple institutions to mitigate challenges associated with sample size and diversity while ensuring data confidentiality. Dynamic assessment strategies would provide a more accurate representation of the evolution of Entrepreneurial Potential. In place of depending on standardized evaluations, educators could implement longitudinal strategies together with time-series analysis to delve into the changes in educational progressions. Innovative micro-credentialing frameworks available at universities could serve as a fitting approach for the execution of these methodologies. International benchmark datasets would accelerate progress. Inspired by computer vision and natural language processing, entrepreneurship researchers should choose demographically balanced datasets with reliable measurements. These resources enable apples-to-apples modelling approach comparisons. Attention is needed for pedagogical applications. Intervention design is neglected in most models that focus on prediction. Going future, prescriptive analytics technologies should identify at-potential students and recommend personalized development routes. (Chen & Yu, 2020)deep learning recommendation system is a start, but further study is required.
A supplementary constraint involves the rapidly transforming landscape of studies in artificial intelligence and entrepreneurial activities. Numerous scholarly articles incorporated within this review were disseminated recently or were accessible as online-first publications. Despite comprehensive endeavors to authenticate publication status, DOI registration, and the integrity of the peer-review process, subsequent revisions of this review may be warranted as new empirical evidence arises and publication records attain complete validation. Another

important limitation concerns the absence of standardized evaluation metrics across studies. Variations in accuracy measures, validation procedures, and dataset characteristics limit direct comparison of model performance. Consequently, although hybrid ML models appear to outperform traditional statistical approaches, these conclusions should be interpreted as indicative trends rather than statistically confirmed differences. Future research should establish benchmark datasets and standardized reporting protocols to facilitate robust meta-analyses and cross-study comparisons. These future prospects taken together point to a maturing of the discipline from pure predictive modeling toward more complex, ethical, and practical frameworks. The ultimate aim should not be prediction for prediction's sake but rather the creation of technologies that really improve educational performance while respecting individual diversity and thus encourage fair access to entrepreneurial possibilities.

E. Synthesis of Evidence across Research Questions

In order to provide a clearer connection between the research questions and the reviewed literature, Table 5 synthesizes the collective findings from the 48 included studies and demonstrates how the evidence addresses each research question.

Table 5: Evidence Synthesis and Research Question Mapping

Research Question	Evidence from Reviewed Studies	Key Findings	Overall Conclusion
RQ1: What MLmodels have been most effective in predicting Entrepreneurial Potential among individuals?	Studies employing hybrid and ensemble models such as GLLCSA-KELM-FS, WHO-RXGBoost, RF-CSCA-SVM, and GBHHO-KELM consistently reported prediction accuracies between 85% and 95%. Traditional standalone models generally achieved lower performance.	Hybrid models that combine feature selection, optimization algorithms, and ML classifiers outperform traditional methods.	Hybrid and ensemble ML models represent the most effective approaches for Entrepreneurial Potential prediction.
RQ2; Which psychological, demographic, and environmental features are most commonly used as predictors of entrepreneurial potential?	Self-efficacy, risk-taking propensity, creativity, entrepreneurial attitude, family entrepreneurial background, gender, institutional support, and AI-assisted learning were the most frequently reported predictors.	Psychological factors appeared in the majority of high-performing models and demonstrated stronger predictive influence than demographic variables.	Self-efficacy, risk-taking propensity, and creativity are the most robust predictors, while demographic and environmental variables provide contextual explanatory value.
RQ3: How do traditional statistical models compare with machine learning-based approaches in predicting	SEM and logistic regression studies generally reported explanatory power and prediction accuracies between 65% and 75%	Statistical models offer interpretability and theory testing, whereas ML models provide superior predictive	ML approaches outperform traditional statistical methods in prediction accuracy, but statistical models remain valuable for

https://doi.org/10.5281/zenodo.21135607

entrepreneurial competence or intentions?	while ML models frequently exceeded 85% accuracy.	performance and capture non-linear relationships.	explanation and theory validation.
RQ4: What are the current limitations in entrepreneurship potential prediction models, and how can future models improve prediction accuracy and gene	Most studies relied on self-reported data, cross-sectional designs, geographically concentrated samples, and black-box algorithms. Limited cross-cultural validation was observed.	Common challenges include cultural bias, lack of interpretability, class imbalance, ethical concerns, And limited longitudinal evidence.	Future research should focus on explainable AI, fairness-aware algorithms, benchmark datasets, longitudinal prediction systems, and cross-cultural validation.

F. Cross-Question Synthesis

The review reveals three overarching patterns across the literature. First, methodological advancement has shifted entrepreneurship prediction from traditional statistical modeling toward hybrid ML architectures that deliver substantially higher predictive accuracy. Second, psychological characteristics particularly self-efficacy, risk-taking propensity, and creativity remain the most influential predictors across diverse educational and cultural contexts. Third, despite notable performance improvements, current prediction systems face significant challenges related to interpretability, fairness, and generalizability. Consequently, future entrepreneurship prediction research should prioritize the development of explainable, ethically responsible, and culturally adaptable ML frameworks capable of supporting educational and policy decision-making.

G. Quantitative Summary of Model Performance

To provide a more systematic comparison of predictive approaches, the reviewed studies were grouped according to methodological category and their reported performance metrics were summarized. Results obtained is presented in Table 6.

Table 6: Summary of Predictive Performance across Model Categories

Model Category	Number of Studies	Typical Performance Metric	Reported Performance Range
Traditional Statistical Models (SEM, Logistic Regression, LDA)	14	Accuracy, Explained Variance (R^2)	65 – 75% accuracy
Standalone MLModels (ANN, Decision Tree, Random Forest, Naïve Bayes)	17	Accuracy, Precision, Recall	75 – 90% accuracy
Hybrid and Ensemble Models (KELM-based, XGBoost-based, SVM optimization frameworks)	17	Accuracy, Precision, Recall, F1-score	85 – 95% accuracy

Volume 2, Issue 2, 2026
https://doi.org/10.5281/zenodo.21135607

Table 7. Highest-Performing Models Identified in the Review

Study	Model	Reported Accuracy
Wen & Zhou (2025)	WHO-RXGBoost	95.0%
Zhang et al. (2022)	GLLCSA-KELM-FS	93.2%
Hasan et al. (2024)	Naïve Bayes + CFS	87.4%
Tu et al. (2019)	RF-CSCA-SVM	84.0%

In order to facilitate a meticulous comparison of predictive performance across the existing body of literature, Table 7 elucidates the most adept machine learning models identified in this analysis. The table encapsulates the foremost algorithms alongside their documented predictive accuracies and corresponding academic contributions. In light of the need for careful interpretation of direct statistical comparisons owing to differences in datasets, predictor variables, sample sizes, and evaluation approaches, the findings shared yield meaningful insights regarding current methodological trends.

As evidenced in Table 7, hybrid and ensemble learning strategies consistently outperform their individual machine learning and traditional statistical counterparts in terms of efficacy. The WHO-RXGBoost model demonstrated the highest prediction accuracy at 95%, closely followed by the GLLCSA-KELM-FS and CSA-KELM models, each achieving an accuracy of approximately 93.2%. Such systems connect refined optimization strategies, tactics for recognizing features, and collective learning approaches to elevate classification standards while controlling overfitting. In a related context, the Naïve Bayes algorithm, when employed in conjunction with Correlation-Based Feature Selection, attained an accuracy of 87.4%, whereas the RF-CSCA-SVM architecture indicated an approximate accuracy of 84%. Collectively, these findings imply that the integration of feature optimization with ensemble learning significantly elevates the prediction of entrepreneurial potential in contrast to traditional statistical methodologies.

Recognizing that these performance insights come from various questions that tap into distinct datasets, validation criteria, and measurement approaches is vital. Consequently, the findings ought to be interpreted as descriptive indicators of methodological advancement rather than as direct statistical comparisons. Going forward, inquiries need to work towards establishing standardized benchmark datasets and evaluation protocols to enhance reliable comparative analyses and meta-analytic assessments.

H. Quantitative Interpretation

Across the studies that were examined, hybrid and ensemble modeling approaches consistently exhibited superior predictive efficacy compared to traditional statistical methodologies. While recognized techniques often showed prediction accuracies oscillating from 65% to 75%, standalone ML models habitually exceeded the accuracy benchmark of 80%. The research indicating the top levels of efficiency implemented hybrid structures that merged feature selection, optimization algorithms, and complex classifiers, with accuracy rates reported from 85% to 95%.One must tread carefully with these findings, as direct statistical comparisons are swayed by fluctuations in datasets, the features of samples, elements of predictors, validation approaches, and evaluation metrics. As a result, the reported performance ranges should be viewed as indicative of general trends rather than conclusive estimates of superiority.

Volume 2, Issue 2, 2026
https://doi.org/10.5281/zenodo.21135607

I. Threats to Validity and Limitation of the Review

Although this systematic literature was conducted using a structured methodology guided by the PRISMA framework, several limitations and threats to validity should be acknowledged. Search and Database Limitations: The research highlighted that Scopus, Google Scholar, and ResearchGate function as vital instruments. While these databases provide comprehensive coverage within the fields of entrepreneurship, education, and machine learning research, relevant studies that are solely indexed in alternative databases such as Web of Science, IEEE Xplore, ACM Digital Library, or ScienceDirect may not have been incorporated. Consequently, certain notable publications could have been omitted despite the meticulous search methodology employed.

(a). Publication Bias: The in-depth critique exclusively addressed research found in English, pointing out the critical role of peer-reviewed articles and academic seminar papers. Hence, potentially valuable research available in alternative languages or found in theses, technical documents, and grey literature was left out of the evaluation. This limitation may have engendered publication bias by preferentially highlighting studies that reported statistically significant or favorable outcomes.

(b). Selection and Screening Bias: Although established inclusion and exclusion criteria were employed and the selection of studies was executed independently by two reviewers, a certain level of subjective judgment may have impacted the decisions made during the screening process, particularly in the evaluation of titles, abstracts, and full texts. Though there were talks among reviewers and processes designed to foster agreement were utilized to counteract bias, the complete removal of personal interpretation is an impossible goal.

(c). Heterogeneity of Included Studies: The investigations that were analyzed showed considerable variability regarding the datasets, sample qualities, predictive elements, specifications of results, techniques in machine learning, and measurement metrics. The expression of Entrepreneurial potential was realized through diverse routes, fusing entrepreneurial intention, entrepreneurial competence, entrepreneurial mindset, and entrepreneurial behavior. These variations curtailed the capacity for straightforward comparisons among the research endeavors and hampered the organized execution of a meta-analysis.

(d). Data Extraction and Reporting Limitations: The review hinged on the findings released by the principal studies conducted. Deficiencies in the reporting quality, uncertainties surrounding methodological transparency, and fluctuations in performance metrics might have influenced the precision and depth of data gathering. Furthermore, certain studies were deficient in providing adequate information concerning model validation methodologies, feature selection techniques, or attributes of the datasets utilized.

(e). Temporal Validity: Machine learning and artificial intelligence are rapidly evolving fields. Consequently, the findings of this review represent the state of knowledge between 2016 and 2025 and may require periodic updating as new predictive models, benchmark datasets, and methodological innovations emerge.

(f). Generalizability of Findings: Numerous studies incorporated in this review were performed in Asian educational contexts, particularly in the regions of China and Indonesia. The specific locales where the evidence is found may hinder the extent to which the conclusions can be generalized to alternative cultural, institutional, and socioeconomic frameworks.

(g). Therefore, conclusions regarding predictor importance and model performance should be interpreted with caution when applied to different populations.

Volume 2, Issue 2, 2026
https://doi.org/10.5281/zenodo.21135607

Overall, while rigorous procedures were adopted to enhance transparency and reliability, these limitations should be considered when interpreting the findings and conclusions of the review.

J. Practical Implications of Findings

The findings of this review have important implications for researchers, educators, higher education institutions, and policymakers seeking to promote entrepreneurship through data-driven decision support.

(a). Implications for Researchers: Scholars in the field of research ought to concentrate on the creation of explainable and transparent machine learning algorithms that are adept at reconciling predictive efficacy with interpretability. Following inquiries must include extended research approaches, apply consistent benchmark data, and engage in cross-cultural assessments to boost the trustworthiness and relevance of models indicating entrepreneurial results. Promoting greater teamwork between experts in entrepreneurship and those in machine learning is similarly suggested to create predictive frameworks that are both theoretically solid and practically applicable.

(b). Implications for Educators and Higher Education Institutions: Educational institutions can leverage machine learning-based predictive frameworks to detect students with entrepreneurial potential at preliminary stages and provide customized interventions, mentoring initiatives, and entrepreneurship education. In any case, systems designed to predict should serve as tools to aid decision-making instead of acting as replacements for human insight. To advance entrepreneurial expertise, academic organizations should link the growth of entrepreneurship with next-gen technologies like AI, data interpretation, and systems that facilitate innovation.

(c). Implications for Policymakers: Policymakers should advocate for the responsible implementation of AI-driven entrepreneurship support systems while instituting ethical guidelines that foster transparency, fairness, accountability, and data privacy. National entrepreneurship development strategies should promote the generation of diverse and representative datasets to mitigate cultural and socioeconomic bias in predictive models. As a result, public investment in the realms of entrepreneurship education and innovation ecosystems can improve the performance of predictive analytics in uncovering and encouraging future entrepreneurial talent.

(d). Implications for Industry and Entrepreneurship Support Organizations: Entrepreneurship development agencies, innovation hubs, business incubators, and financial institutions have the opportunity to implement machine learning-derived predictive frameworks to enhance the identification of talent, formulate specialized capacity-building programs, and refine resource distribution for initiatives aimed at supporting entrepreneurship. Integrating user-friendly and clear AI tools into these setups will strengthen assurance among stakeholders, aid in astute decision-making, and support just access to entrepreneurial opportunities.

In short, the actual significance of forecasting in the entrepreneurial landscape exceeds merely boosting model effectiveness. When implemented with meticulous attention to detail, predictive frameworks driven by machine learning possess the potential to augment evidence-based educational approaches, strengthen initiatives designed to promote entrepreneurial growth, and support sustainable economic progress through the more adept recognition and development of entrepreneurial competencies.

Volume 2, Issue 2, 2026
https://doi.org/10.5281/zenodo.21135607

V. CONCLUSION

A thorough examination of scholarly works amalgamated results from 48 research efforts to delve into the impact of machine learning techniques in gauging entrepreneurial potential. Guided by the PRISMA 2020 framework, the review assessed predictive models, identified pivotal determinants of entrepreneurial potential, compared machine learning methodologies with traditional statistical techniques, and underscored prevailing methodological challenges alongside prospective avenues for future research. Analysis demonstrates that hybrid and ensemble machine learning frameworks regularly outperform established statistical practices in predictive accuracy, mainly due to their talent for modeling complex, non-linear interactions related to entrepreneurial qualities. Psychological frameworks, namely self-confidence, propensity for risk, creativity, and a mindset geared towards entrepreneurship, proved to be the most trustworthy indicators of entrepreneurial aptitude, while demographic and contextual elements afforded essential insights. Despite these advancements, the review identified significant constraints within the prevailing body of literature, including an overreliance on self-reported and cross-sectional datasets, a lack of cross-cultural validation, inadequate model interpretability, inconsistent evaluation metrics, and concerns regarding algorithmic fairness. Moreover, methodological heterogeneity across the reviewed studies inhibited formal statistical aggregation of findings and accentuated the necessity for standardized benchmark datasets and comprehensive reporting frameworks. Finally, this critique reinforces the sectors of entrepreneurship and AI scholarship by supplying a detailed gathering of existing predictive modeling methodologies, unearthing research gaps, and establishing routes for upcoming studies. The conclusions indicate that later iterations of entrepreneurship forecasting mechanisms ought to stress not only the need for accurate predictions but also the values of clarity, fairness, comprehensibility, and relevance to enable their effective use in various educational and socioeconomic frameworks.

REFERENCES

Aga, M. K. . (2023). The mediating role of perceived behavioral control in the relationship between entrepreneurship education and entrepreneurial intentions of university students in Ethiopia. *Journal of Innovation and Entrepreneurship.*

Baasith, C. A., Muhammadan, S. A., Wahpiyudin, & Amalia, R. (2025). Determining AI-based learning adoption model for students in entrepreneurship education: A design thinking approach. *Journal of Consumer Sciences, 10(1)*, 27–58. doi: https://doi.org/10.29244/jcs.10.1.27-58

Bell, R. (2019). Predicting entrepreneurial intention across the university. *Education + Training, 61*(7), 815–831. doi: https://doi.org/10.1108/ET-06-2018-0142

Cecchin, E., Tormena, E. L., & Silva, A. C.. (2022). A SEM-neural network approach for understanding the entrepreneurial competence development of freshmen engineering and computing students. *Procedia Computer Science,, 217*, 1081–1090. . doi: https://doi.org/10.1016/j.procs.2022.12.152

Chen, Z., & Yu, X. . (2020). Adoption of human personality development theory combined with deep neural network in entrepreneurship education of college students. *Frontiers in Psychology, 11*, 1346. doi: https://doi.org/10.3389/fpsyg.2020.01346

Chung, D. . (2023). MLfor predictive model in entrepreneurship research.Predicting entrepreneurial action. *Small Enterprise Research., 30*(1), 1–18. . doi: https://doi.org/10.1080/13215906.2022.2164606

Djordjevic, D., Cockalo, D., Bogetic, S. , & Bakator, M. . (2021). Predicting entrepreneurial intentions among the youth in Serbia with a classification decision tree model with the QUEST algorithm. *Mathematics., 9*(12), 1487. doi: https://doi.org/10.3390/math9121487

Volume 2, Issue 2, 2026
https://doi.org/10.5281/zenodo.21135607

Fang, M., & Wang, W. . (2024). Research on the construction of college students' employment and entrepreneurship guidance course system under the background of "Internet+". *Applied Mathematics and Nonlinear Sciences,, 9*(1). doi: https://doi.org/10.2478/amns-2024-0649

Gonzalez-Tamayo, L. A., Olarewaju, A. D., Bonomo-Odizzio, A., & Krauss-Delorme, C. . (2024). University student entrepreneurial intentions: The effects of perceived institutional support, parental role models, and entrepreneurial self-efficacy. *Journal of Small Business and Enterprise Development, 31*(8), 205–227. . doi: https://doi.org/10.1108/JSBED-09-2022-0408

Hasan, H., Yulastri, A., Ganefri, T. T. A. P., & Marta, R. . (2024). Prediction of student entrepreneurship future work based on entrepreneurship course using the Naïve Bayes classifier model. *Sinkron: Jurnal dan Penelitian Teknik Informatika, 8* (1), 525–532. . doi: https://doi.org/10.33395/sinkron.v9i1.13293

Kennedy, J. . (2003). Situational factors and entrepreneurial. *16th Annual Conference of Small Enterprise Association of Australia and New Zealand.*

Kong, F., Zhao, L., & Tsai, C. H. . (2020). The relationship between entrepreneurial intention and action.The effects of fear of failure and role model. *Frontiers in Psychology, 11*, 229. .

Lim, J.-Y., Kim, G. M., & Kim, E. J. (2021). Predictors of entrepreneurial intention of nursing students based on theory of planned behavior. *Journal of Multidisciplinary Healthcare, 14*, 533–543. . doi: https://doi.org/10.2147/JMDH.S288532

Opusunju, M. I., Jiya, S. N., & Akyuz, M. . (2019). Evaluation education using of entrepreneurship graduates of various universities in Abuja Federal Capital Territory of Nigeria. *Al-Hikma Journal of Arts and Social Education, 1*(2,), 39–49.

Pasic, M., Jovanovski, B., Ferizbegovic, F., & Pasic, M. . (2025). MLanalysis of the impact of entrepreneurial mindset dimensions on entrepreneurial intentions of students in EU member countries and an EU candidate country: . *Comparative study. Frontiers in Education, 10*, 1489724. doi: https://doi.org/10.3389/feduc.2025.1489724

Sandoval, M., & Hernández, C. . (2021). Building models based on artificial neural networks to predict entrepreneurial intentions among undergraduate students. *Proceedings of the International Conference on Industrial Engineering and Operations Management*, 2443–2452.

Sanyal, S., & Al Mashani, A. M. . (2018). Entrepreneurial intention of university students in sultanate of Oman—a study of Dhofar University. *Journal of Contemporary Trends in Business and Information Technolog, 5*(1), 16–31.

Solórzano, S. S., Pizarro Romero, J. M., Díaz Cueva, J. G., Arias Montero, J. E., Zamora Campoverde, M. A., Lozzelli Valarezo, M. M., Montes Ninaquispe, J. C., & Acosta Enriquez, B. G., & Arbulú Ballesteros, M. A. . (2024). Acceptance of artificial intelligence and its effect on entrepreneurial intention in foreign trade students: A mirror analysis. *Journal of Innovation and Entrepreneurship,, 59.*

Tu, J., Lin, A., Chen, H., Li, Y., & Li, C. . (2019). Predict the entrepreneurial intention of fresh graduate students based on an adaptive support vector machine framework. *Mathematical Problems in Engineering, 2019.*, 1–14. . doi: https://doi.org/10.1155/2019/2039872

Wei, Y., Lv, H., Chen, M., Wang, M., Heidari, A. A., Chen, H., & Li, C. . (2020). Predicting entrepreneurial intention of students: An extreme learning machine with Gaussian barebone Harris Hawks optimizer. *IEEE Access, 8*, 76841–76855. . doi: https://doi.org/10.1109/ACCESS.2020.2982796

Wen, K., & Zhou, D. . (2025). Predictive career guidance and entrepreneurial development for university students using artificial intelligence and machine learning. *Industry and Higher Education.* doi: https://doi.org/10.1177/14727978251318801

Yang, T., Li, J., & Chen, H. (2025). A deep learning model for psychological support in student entrepreneurship. *IEEE Access, 13*, 12345–12356. . doi: https://doi.org/10.1109/ACCESS.2025.3547833

Zhang, L., Fu, Y., Wei, Y., Chen, H., Xia, C., & Cai, Z. (2022). Predicting entrepreneurial intention of students: Kernel extreme learning machine with boosted crow search algorithm. *Applied Sciences, 12*(14), 6907. . doi: https://doi.org/10.3390/app12146907

Volume 2, Issue 2, 2026
https://doi.org/10.5281/zenodo.21135607