AUT Journal of Modeling and Simulation

AUT J. Model. Simul., 58(1) (2026) 3-22
DOI: 10.22060/miscj.2026.25629.5484

A Stability-Guided Quantum Feature Selection and Quantum Kernel SVM Framework for Entrepreneurial Competency Prediction

Mohammad Reza Mehregan, Arman Rezasoltani, Amir Mohammad Khani *

Department of Industrial Management, Faculty of Management, University of Tehran, Tehran, Iran.

ABSTRACT: The key challenges of performance prediction in entrepreneurship stem from the heterogeneous, multidimensional, nonlinear, and imbalanced characteristics of entrepreneurial behaviors, i.e., ambition and risk preference. While machine learning algorithms, including Support Vector Machines, Random Forests, XGBoost, and LightGBM, have been purported and adopted to model entrepreneurial outcomes, their optimization saturation effect in performance prediction between 60% and 79% is observed, with little novelty in representational expansion by resampling and ensemble engineering approaches. To overcome the limits of existing models, this paper adopted a stability-guided quantum-enhanced hierarchical QAOA-QSVM framework whereby hierarchical Quantum Approximate Optimization Algorithm (QAOA) feature selection interacted with Quantum Kernel Support Vector Machine for Entrepreneurial Skill (QSVM). Features extracted from 219 university students' gender, race, personality, and university attributes were successively selected by formulating the QUBO problem of feature optimization and solving it with QAOA. The stability consensus was assessed by stratified five-fold cross-validation to improve reproducibility and robustness. Selected features were eventually embedded in an entangled quantum Hilbert space of the ZZFeatureMap to enable high-order nonlinear feature interactions with quantum kernel learners. The results indicated that our quantum approach surpassed a classical RBF-SVM on information from university students by nearly 22% (83.09% vs. 60.68%) and reached an 86.84% recall for entrepreneur prediction. We conclude that quantum representation learning structures can provide more insightful feature embedding methods, enabling novel transformers in talent pipeline building of entrepreneurship courses. This study also extends a meta-complementary theory between quantum machine learning and entrepreneurship.

Review History:

Received: Mar. 16, 2026
Revised: May, 27, 2026
Accepted: May, 27, 2026
Available Online: Jul. 20, 2026

Keywords:

Quantum Machine Learning
Quantum Feature Selection
Quantum Approximate Optimization Algorithm
Entrepreneurial Competency Prediction

1- Introduction

In today's knowledge-based societies, entrepreneurial skills make up one of the most important characteristics needed to accelerate economic development, innovation capabilities, and sustainable development(Kraus et al., 2021; Sieg et al., 2023). Universities have come to be recognized as not just places for education but also as creators of entrepreneurial ecosystems and the creators of the entrepreneurial intentions, resilience, and venture creation capabilities of their students/trainees (Nwosu et al., 2025). As institutions of higher learning begin to implement data-driven approaches to identifying and supporting entrepreneurial talent, they are turning to predictive analytics as a means of measuring entrepreneurial skill using psychological, demographic, educational, and environmental variables (Dai & Li, 2025; Adeel et al., 2023). As a result of the multidimensional and nonlinear nature of entrepreneurial behavior, it still remains difficult to model entrepreneurial behavior accurately.

*Corresponding author's email: amir.mo.khani@ut.ac.ir

Entrepreneurial competency is complex due to the interaction of personality attributes (e.g., self-confidence, perseverance), exposure to various environments (e.g., being influenced by an entrepreneur), as well as one's academic background and health (Botha & Taljaard, 2021; Cao et al., 2022). Previous studies found a strong correlation between traits related to achievement motivation, self-efficacy, and resilience with intention to start a business and create a new venture (Al-Qadasi et al., 2023; Sahid et al., 2024); however, there are limitations when attempting to develop an accurate prediction model of these theoretical constructs based on empirical data. There are many small, imbalanced behavioral datasets that contain a very high-dimensional feature space once categorical attributes are encoded; therefore, use of traditional modelling techniques will result in models that are prone to overfitting, are unstable, and have poor generalizability.

In the last ten years, machine learning techniques have become more commonly used to predict entrepreneurship skills. Classical classifiers such as SVM, Random Forest, logistic regression, and KNN were used to predict entrepreneurship skills in earlier studies (Sharma & Manchanda, 2020). The results of those studies, however, were moderate (often below 70% accuracy) and indicated limited capability for effectively representing complex nonlinear interactions among variables. Recent research has utilized two ensemble methods called XGBoost and LightGBM (Kaur & Bhatt, 2023). Although these gradient boosting algorithms improve prediction accuracy and utilize iterative corrections based on past predictions along with tree-based nonlinear modeling methods, they are still classified as classical.

At the same time, methodological improvements have concentrated on overcoming issues associated with data imbalances and interpretability of methods. For example, Random Under Sampling (RUS), Adaptive Synthetic Sampling (ADASYN), and Explainable AI methods (SHAP and LIME) have all improved the fairness and transparency of the classification of entrepreneurs (Şimşek & Daş, 2022; Bhatt et al., 2025). However, the majority of these developments largely occur because of optimization and post-hoc delivery and do not advance the true representational capability of the learning model itself. Furthermore, although many hybrid frameworks incorporate algorithmic feature selection and hyperparameter tuning (for example, the binary grey wolf optimizer and LightGBM), they still fit into traditional hypothesis spaces (Jafarnejad et al., 2025). As a consequence, the average improvement of accuracy appears to be nearing a max point when building models for highly nonlinear behavioral data with classical kernels.

Developments in Quantum Machine Learning (QML) show strong potential for a new method of modelling data. Quantum kernel methods take advantage of the exponential dimensionality of Hilbert space to produce entangled quantum states that can preserve at least some information about the classical data points that were originally used to create these qubits, resulting in improved classification separation for complex classification problems (Thanasilp et al., 2024; Devadas & T, 2025). Parameterized quantum circuits, like ZZFeatureMap, represent interactions between features and allow for the implicit representation of higher-order correlations that would be impossible to estimate using classical kernels. QAOA has also been shown to be effective in solving certain kinds of combinatorial optimization problems (such as feature selection problems that are represented as Quadratic Unconstrained Binary Optimization (QUBO) problems) (Jha et al., 2026; Ruan et al., 2022). All of these advances suggest that there are much larger enhancements for prediction to be made in cases of complex nonlinear dependence.

Though there have been many recent advances in QML, the use of QuantML for analysis of entrepreneurship has not yet been investigated. Currently, there do not appear to be any prior studies that have integrated quantum feature selection and quantum kernel learning methods for entrepreneurship competency prediction. In general, existing ML-frameworks do not take into account that the stability-driven consensus mechanisms should be utilized when validating feature selection of the validation fold to ensure that the methods produced corresponding to each fold produces reproducible and robust selected features across analysis performed. These gaps have been addressed through the introduction of a new, stability-guided quantitative-enhancement framework for the prediction of entrepreneurial competence. This framework comprises a hierarchical quantum-assisted opportunity allocation system -based feature selection and a quantum kernel support vector machine with a stratified cross-validation stability protocol. The framework improves the predictive accuracy, robustness, and sensitivity of identifying Entrepreneurs through an innovative approach that combines quantum optimization, entangled feature embeddings, and reproducible-based feature selection methods. Overall, this research provides the first interdisciplinary "bridge" between quantum machine learning and entrepreneurship analytics, thereby advancing both methodological "innovation" as well as applied predictive modeling in higher education environments.

2- Theoretical Foundations

2- 1- Entrepreneurial Competence

Entrepreneurial competence is defined by the skills, knowledge (e.g., know-how), attitudes (e.g., entrepreneurial mindset), and behavior (e.g., creativity and innovation) that allow someone to identify, develop, and create value from an entrepreneurial opportunity (Martínez-Martínez et al., 2025; Reis et al., 2020), rather than merely creating "entrepreneurs." In addition, despite extensive research into these qualities (Somià et al., 2024), lack of a strong theoretical foundation represents a key challenge for the field (Reis et al., 2020). The absence of a clear theoretical framework for understanding entrepreneurial competencies is particularly problematic since entrepreneurship has been identified as an important driver of economic growth, innovation, and job creation (Koç et al., 2025).

According to several theorists, entrepreneurial competence means being competent enough to perform an activity or task successfully. Its basic definition can include anything related to one's ability to complete an activity or task; from one's skill set to the knowledge one has about those skills, the values one holds dear, one's attitude towards the activity/task and the manner in which one behaves toward an undertaking (Lee & Yun, 2020; Sun & Usman, 2024). Entrepreneurial competence is defined as one's ability to identify, create, and utilize opportunities (Sun & Usman, 2024) as well as the capacity to action the idea by obtaining required resources (Vu & Nwachukwu, 2021). According to Mitchelmore and Rowley's original work, entrepreneurial competencies "are the specific skills, roles in society, knowledge, self-image, motives and attributes that contribute towards the producing of new venture and continuing survival of that new venture" (Vu & Nwachukwu, 2021). There are many definitions for entrepreneurial competencies; that is because the major contributors to defining this construct come from many disciplines, such as management, psychology, sociology, and economics (Reis et al., 2020).

There is a high level theoretical debate on whether an individual's entrepreneurial competencies are traits that individuals are born with versus capabilities that are developed. The early literature expresses that entrepreneurial competencies are innate traits, but the literature has moved toward establishing a growing scholarly consensus that they can be developed through education, experience and deliberate practice (Duong, 2023). Therefore, this shift in thought has significant implications for entrepreneurship education and policy, as it provides legitimacy for institutional efforts to develop entrepreneurial competencies (Van Gelderen, 2022). Entrepreneurial competencies have also been theorized as capable of development through experiential and self-directed learning. Self-directed learning (SDL) has been identified as an essential competence base for entrepreneurial adaptability in pioneering (rapidly changing) digital environments (Morris & König, 2020). SDL competence also works together with experiential learning to provide what is needed for employees to demonstrate broader entrepreneurial competency (Morris & König, 2020). The theory of entrepreneurial learning also identifies that nascent (new) entrepreneurs develop their competencies through participation in collective, exploratory, and exploitative learning (Petrylaitė & Rusk, 2020).

2- 2- Quantum Feature Selection

Quantum feature selection is an emerging research direction at the intersection of quantum computing, optimization theory, and machine learning. Its theoretical foundations rest on how quantum systems represent information, how quantum algorithms explore large combinatorial spaces, and how quantum-enhanced learning models process high-dimensional data.

Quantum feature selection is fundamentally based on the principles of quantum information theory. In a quantum system, quantum states encode information in superposition, allowing a quantum state to represent numerous configurations at once; as a result, many subsets of a set of features can theoretically be encoded simultaneously into one quantum state (Raseena, 2025). Quantum entanglement is also relevant in that it allows quantum systems to represent correlations in addition to classical statistical dependence, which is important for finding redundant and/or interacting features (Priyadharshni & Ravi, 2025). From an information-theoretic point of view, von Neumann entropy and quantum mutual information are examples of how classical-based measures to evaluate feature relevance can be applied to quantum systems (Park, 2022). A second theoretical pillar comes from quantum algorithms for combinatorial optimization. Feature selection is typically NP-hard because evaluating all possible subsets of n features requires searching over 2^n combinations. Quantum algorithms such as Grover's search provide a quadratic speedup for unstructured search problems, reducing the complexity from O(2^n) to O(2^(n/2)) under suitable oracle assumptions (Bansal, 2025). More generally, amplitude amplification extends this principle to probabilistic selection procedures (Figgatt et al., 2017). For structured optimization problems, the Quantum Approximate Optimization Algorithm provides a variational framework for encoding feature subset objectives into a cost Hamiltonian and iteratively improving candidate solutions (Grange et al., 2024). Variational quantum algorithms more broadly combine parameterized quantum circuits with classical optimization loops, making them particularly suitable for near-term noisy quantum devices (Cerezo et al., 2021).

The Quantum ML Theory provides a third foundation. Quantum feature maps embed classical data into a Hilbert space of high-dimensional form, which increases the potential for class separability (Havlíček et al., 2019). This concept resembles classical kernel methods, in which selecting features is similar to selecting the most informative components in a transformed space. The idea of quantum support vector machines and quantum principal component analysis helps demonstrate that there are circumstances under which quantum algebra impatient systems can accelerate the speed of feature extraction when certain assumptions regarding how to load data exist (Monzón-Verona et al., 2025). The theoretical speedup available through quantum machine learning can be examined from a viewpoint of computational complexity. The theoretical speedup of quantum feature selection will be based on the assumption that we have an efficient way to prepare a quantum state and that we have access to QRAM (quantum RAM). Previous research of quantum machine learning shows that the advantages presented by quantum feature selection depend on these architectural assumptions (Corli et al., 2024).To summarise, the theoretical foundation of quantum feature selection combines elements of quantum information theory, quantum search and optimisation algorithms, and complexity theory to provide a mathematically rigorous way to rethink feature selection in high-dimensional problems. The theoretical basis provides a logically coherent approach for exploring potential practical advantages that remain unresolved in the research community.

3- Research Background

In recent years, much research on predicting entrepreneurial competence (e.g., competencies) in university-level students has increased; many researchers have conducted studies that support and allow for the early identification of entrepreneurial ability, improve curriculum development, and support national entrepreneurship ecosystems through the use of machine learning methodologies to identify psychological, educational, and environmental characteristics associated with entrepreneurial ability. Table 1 displays the most relevant previously published articles, detailing algorithms that have been used, the most successful models/research studies, significant contributions, research gaps, and the accuracy that was achieved by each algorithm; this comparison provides an overview of both the progress made methodologically and areas where limitations are still present, motivating researchers to continue developing a quantum-enhanced predictive framework.

Table 1. Research background. (Continued)

Authors	*Algorithms Used*	*Best Algorithm*	*Contribution of Research*	*Necessity of Research*	*Key Results*	*Accuracy*
Sharma and Manchanda (2020)	Random Forest, KNN (k=5), SVM, Logistic Regression, Naïve Bayes, Decision Tree	Support Vector Machine	Using machine learning algorithms to predict students' entrepreneurial aptitude based on personality traits	The need to strengthen the entrepreneurial ecosystem and early identification of students with entrepreneurial potential	The best accuracy is for SVM with 59.18%; the performance of the models was average and not very high.	SVM = 59.18% (highest value)
Şimşek and Daş (2022)	Logistic Regression, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Naive Bayes, Decision Tree, Random Forest, XGBoost, Gradient Boosting Machine (GBM), LightGBM, Multi-Layer Perceptron (MLPC)	Random Forest with RUS	Demonstrated that applying imbalanced dataset handling techniques significantly improves prediction performance (Accuracy and F1-Score) for entrepreneurial competency classification. Provided a comprehensive evaluation beyond accuracy by incorporating Precision, Recall, and F1-Score.	Previous studies relied only on accuracy and failed to detect the minority class (entrepreneurs). The dataset was highly imbalanced, causing misleading performance results.	Random Forest + RUS achieved the highest performance with 0.76 Accuracy and 0.70 F1-Score (entrepreneur class). Imbalance handling significantly improved minority class detection.	0.76
Kaur and Bhatt (2023)	K-Nearest Neighbors, Random Forest, Logistic Regression, XGBoost, LightGBM	XGBoost	Introduced advanced gradient boosting algorithms (XGBoost and LightGBM) to improve prediction accuracy of student entrepreneurial competency beyond traditional ML models	Previous models (KNN, Random Forest, Logistic Regression) failed to surpass the 70% accuracy benchmark; need for stronger predictive capability and curriculum enhancement support	XGBoost achieved 79% accuracy; LightGBM achieved 71% accuracy; both outperformed traditional machine learning models	XGBoost = 79% (Highest) LightGBM = 71%
Jarial and Verma (2023)	Random Forest, Gradient Boosting, Linear Regression, Ridge Regression,	Random Forest	Applied machine learning models to predict entrepreneurial traits among agricultural undergraduate students and evaluated the	Need to understand agri-entrepreneurial traits among students and evaluate whether ML models can effectively predict	Random Forest achieved the highest accuracy (60%). Entrepreneurial traits such as strong partnerships (0.359), learning	60% (Random Forest)

Table 1. Research background.

	Lasso Regression		suitability of ML for trait prediction in agribusiness education.	such psychological and behavioral traits.	(0.347), and people-organizing ability (0.341) were strongly correlated. Female students exhibited fewer entrepreneurial traits than male students.	
Suryadi and Kurniawan (2024)	Random Forest (Two other supervised ML models – not specified in abstract)	Random Forest	Introduced supervised machine learning models to predict entrepreneurial spirit among college students, moving beyond traditional Structural Equation Modeling (SEM). Applied clustering to define spirit levels before classification.	Prior studies mainly relied on SEM; lacked predictive ML-based frameworks for entrepreneurial spirit classification.	Random Forest achieved the highest performance with a macro-averaged F1-score of 0.656.	-
Bhatt et al. (2025)	LightGBM, Explainable AI methods (LIME, SHAP), Standardization, One-Hot Encoding	LightGBM	Integrated LightGBM with Explainable AI (LIME & SHAP) to forecast entrepreneurial competencies while ensuring interpretability and transparency in predictions for educational policy use.	Need for data-driven, interpretable predictive models to assess entrepreneurial competencies in university students and support curriculum development aligned with modern economic demands.	LightGBM achieved 71% accuracy on a dataset of 219 university students with 22 predictive features. XAI techniques enhanced transparency of feature influence on predictions.	71%
Jafarnejad et al. (2025)	Logistic Regression, SVM, Decision Tree, Random Forest, Gradient Boosting, AdaBoost, Extra Trees, KNN, MLP, Naïve Bayes, XGBoost, LightGBM, CatBoost, LDA, QDA, GPC	Binary Grey Wolf Optimizer (BGWO), Optuna (Bayesian Hyperparameter Tuning), Mutual Information, ADASYN (class balancing)	LightGBM (Optimized with BGWO + Optuna)	Proposed a hybrid framework combining feature selection (BGWO), hyperparameter tuning (Optuna), imbalance correction (ADASYN), and SHAP explainability. Demonstrated superior predictive accuracy compared to previous studies (e.g., Sharma & Manchanda, 2020).	LightGBM achieved highest performance among 16 classifiers; SHAP analysis showed Influenced, GoodPhysicalHealth, and KeyTraits_Positivity as most important predictors.	70.7%

In the past few years, research regarding predicting entrepreneurial competencies has evolved from basic traditional machine learning methods to more sophisticated evolutionary models and the new class of ensemble methods; however, significant conceptual and technical limitations are still unsolved even after several improvements in methodology.Sharma and Manchanda's (2020) work represents one of the first seminal studies in this area, using various types of machine learning algorithms such as SVM, Random Forest, logistic regression, and KNN to predict entrepreneurial capabilities based upon individuals' personality types. While the results of their research showed that machine learning methods could be used to predict entrepreneurial competencies, the highest accuracy obtained was only 59.18%, which is relatively low. More importantly, the study primarily used accuracy as the evaluation metric and did not consider class imbalance or provide any other performance metrics that would aid in evaluating the stability or robustness of the algorithm. Finally, their use of a purely classical modeling approach did not address nonlinear interactions amongst the predictor variables beyond those that could be accommodated by using standard kernel functions. Therefore, although significant advancement has occurred in entrepreneurial competency prediction research since this study was conducted, there is still considerable opportunity for additional advancement from methodological perspectives.

Research has moved away from answering research questions to trying to increase the performance of predictive models using better data handling and different algorithms. According to research by Şimşek and Daş (2022), the use of resampling techniques like Random Undersampling (RUS) can help mitigate the effects of class imbalance on predictions. They found that using RUS allowed for improved detection of minority classes (entrepreneurs) in their experiments, with the detection of minority classes (entrepreneurs) receiving higher F1-scores compared to the majority class (non-entrepreneurs). They reported that the improvement in their results was primarily due to data-level interventions and thus did not reflect any increases in representational capabilities of the underlying classifiers. The classifiers that were used for these studies were based on classical classification procedures, and no new structural innovations in the construction of models or learning capabilities were added to these types of classifiers. Even though the use of RUS improved fairness of models, the representational limitations of classical classifiers remain. The ability of some researchers, such as Kaur and Bhatt (2023), to develop new techniques such as gradient boosting (XGBoost and LightGBM) allowed for much improved predictive accuracy (up to 79%), which suggests that their models have a greater ability to model nonlinear relationships than previous classical classifiers. However, the majority of the research done thus far still employs classical-statistical procedures for their modeling. As a result, there remains a large reliance on the quality of feature engineering for the models to achieve good predictive accuracy. While there have been improvements in predictive accuracy across the evaluation phases of model performance, there have not been any systematic attempts to evaluate feature stability or robustness across folds.

Modifications to the research must be at least 20% different to the original study. Jarial and Verma (2023) and Suryadi and Kurniawan (2024) also extended the contexts of the study's application and created a preprocessing method for the prediction of behavior using clustering algorithms, yet the predictive performance remained moderate at approximately 60% to 65%. As such, methodological contributions mainly represented the application of existing machine learning techniques to new datasets, rather than the formation of new learning frameworks. Bhatt et al. (2025) and Jafarnejad et al. (2025) have published recent works that present a hybrid framework for prediction that utilizes feature selection methods (e.g., Binary Grey Wolf Optimizer), hyperparameter tuning (Optuna), imbalance correction (ADASYN), and explainable AI (e.g., SHAP and LIME) techniques to achieve more rigorous and interpretable methodologies than either of the previous studies. Despite offering improvements through both hybrid and optimization methods, the primary predictive model remained classical in nature (predominantly utilizing LightGBM), with over 70% and 71% accuracy for the best-performing models within each study. This suggests that the classical machine learning paradigm may be approaching an upper limit in terms of predictive performance regarding the prediction of behavior, and further incremental optimization of those established classical paradigms may not lead to any additional improvement.

The critique of the literature has revealed some important deficiencies:

- Quantum or Advanced Representation Learning is Absent: All the previously published studies relied solely upon traditional models of machine learning. There has been no consideration of using either quantum-enhanced representation of features or utilizing quantum kernels to better represent complex nonlinear relationships in psychological and behavioral data.
- Stability-Oriented Feature Selection is Lacking: Some studies have implemented meta-heuristic feature selection; however, the meta-heuristic feature selection process used has not incorporated ongoing stability analysis to ensure reproducibility and robustness of features selected for use through multiple folds of cross-validation.
- High-Order Feature Interactions are Not Modeled: Classic kernels (e.g., RBF) approximate non-linear boundaries; however, they do not encode explicitly entangled feature interactions that may exist between psychological variables, exposures to environmental factors, and health indicators.
- Innovation is at the Level of Incremental vs. Paradigm: Most improvements have been derived from fine-tuning parameters, re-sampling data, and stacking ensembles to improve performance. Accordingly, these improvements can be considered to reflect engineering domination, rather than the shifting of representational capability from one source of information to another source of

information.

Changes in thinking and developing new ways to think about how things work will involve three areas of innovation:

1st Innovation is "Quantum-Based Feature Selection through QAOA".

Previously in the past, metaheuristic techniques have looked at the problem of Feature Selection as a "QUBO" problem and then solved it through the Quantum Approximate Optimization Algorithm. This algorithm will allow feature selection to maximize relevance while minimizing redundancy in a quantum-based optimization model. The hierarchical QAOA will allow for the addressing of qubits and maintaining the quality of the results in the optimization.

2nd Innovation is the "Stability Guided Consensus Mechanism".

The development of the consensus feature set is based upon the "frequency" of the selection of features, and to develop the consensus feature set, the study took the quantum-based feature selection and embedded it within a stratified five-fold cross-validation structure. Rigorous experimentation to develop a reproducible consensus feature will provide stability that has not been achieved in previous research.

3rd Innovation is "Quantum Kernel Support Vector Machine".

The most substantial advancement is measured by the fact that the quantum kernel supports the ability to replace classical kernels with quantum kernels based on the mapping of classical features into an exponentially large Hilbert space through the use of entangled quantum states. The quantum kernel support vector machine captures many complex non-linear interactions that are beyond the expressive capacity of classical SVMs characterized by RBF kernels.

This study represents the first stability-guided quantum-enhanced framework for entrepreneurial competency prediction. It moves beyond classical optimization refinements and introduces quantum representation learning combined with stability-aware feature selection. By demonstrating consistent performance improvements over a classical RBF-SVM baseline, the research provides empirical evidence that quantum kernel methods can meaningfully enhance predictive modeling of complex behavioral datasets. In contrast to prior studies that improved performance through sampling strategies or ensemble tuning, this work contributes a fundamentally new computational paradigm—bridging quantum machine learning and entrepreneurship analytics.

4- Methodology

Researchers propose a new way to improve machine learning for predictions of university students' entrepreneurial skills. This method combines two aspects of quantum technology—one that helps them choose which of the 100+ types of data on students will be most useful for making predictions and one that allows them to create an intelligent algorithm for making those predictions. The researchers' initial data set came from Sharma & Manchanda (2020) and included students' demographic, educational, psychological, and behavioral characteristics. The target variable exhibits a moderate class imbalance, with 128 samples labeled as class 0 and 91 samples labeled as class 1. The study consisted of a 4-step process: Step 1: Fold-wise preparation and encode the data; Step 2: Calculate mutual information between all pairs of data; Step 3: Use QAOA to find the best set of features to use in creating the working algorithm, and Step 4: Use the working algorithm on QSVM to make predictions.

4- 1- Data Characteristics and Preprocessing

In this research project the data that was collected is comprised of a sample of 219 college student participants. This data was taken from a prior study by Sharma and Manchanda (2020), entitled "Using Machine Learning Algorithms to Predict and Improve Entrepreneurial Competency Among University Students." Each record in this sample is an individual participant with four groups of variables: personal variables, psychological variables, educational variables, and environmental variables that all play a role in the decision-making process that leads someone to be an entrepreneur. The target variable, Y, is a binary variable with a value of either 0 or 1 based on whether or not the participant is identified as an entrepreneur according to the definitions provided by Babin et al. (2019). Table 2 summarizes the main characteristics of the dataset.

During the data preprocessing stage, the textual variable ReasonsForLack, which descriptively expressed the reasons for some individuals' lack of interest in entrepreneurship, was removed from the analysis due to its unstructured nature, the presence of outliers, and the high level of content heterogeneity. Let $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N}$ denote the dataset, where $x_i \in \mathbb{R}^d$ represents the feature vector of the i -th student and $y_i \in \{0,1\}$ indicates entrepreneurial competency. To ensure numerical consistency and avoid information leakage, preprocessing was conducted independently within each training fold. Categorical nominal attributes, including EducationSector and KeyTraits, were transformed using one-hot encoding:

$$x_{ij} \rightarrow \mathbf{e}_k \in \{0,1\}^K \tag{1}$$

where K denotes the number of unique categories.

Ordinal psychological attributes measured on Likert scales (Perseverance, Competitiveness, SelfConfidence, etc.), together with the numeric attribute Age, were normalized using Min–Max scaling:

$$x' = \frac{2(x - x_{\min})}{x_{\max} - x_{\min}} - 1 \tag{2}$$

mapping all values into t1, 1 close bracket, which is compatible with quantum feature maps. Binary attributes were mapped directly to $\{0,1\}$. Missing binary values were

Table 2. Data Characteristics.

FEATURE	DESCRIPTION	DATA TYPE	RANGE / CATEGORIES
EDUCATIONSECTOR	Student's academic discipline	Nominal (categorical)	e.g., Economics, Humanities, Arts, Medicine, etc.
INDIVIDUALPROJECT	Whether the student has done an individual project	Binary (text)	Yes / No
AGE	Age of the student	Numeric	Between 17 and 26
GENDER	Gender of the student	Binary (text)	Male / Female
CITY	Type of residence	Binary (text)	Yes / No
INFLUENCED	Whether the student has been influenced by entrepreneurial environment	Binary (text)	Yes / No
PERSEVERANCE	Level of perseverance	Ordinal (Likert scale)	Scale from 1 to 5
DESIRETOTAKEINITIATIVE	Desire to take initiative	Ordinal (Likert scale)	Scale from 1 to 5
COMPETITIVENESS	Degree of competitiveness	Ordinal (Likert scale)	Scale from 1 to 5
SELFRELIANCE	Degree of self-reliance	Ordinal (Likert scale)	Scale from 1 to 5
STRONGNEEDTOACHIEVE	Strong need for achievement	Ordinal (Likert scale)	Scale from 1 to 5
SELFCONFIDENCE	Self-confidence level	Ordinal (Likert scale)	Scale from 1 to 5
GOODPHYSICALHEALTH	Physical health status	Ordinal (Likert scale)	Scale from 1 to 5
MENTALDISORDER	Whether the student has any mental disorder	Binary (text)	Yes / No
KEYTRAITS	Key personal traits (e.g., vision, positivity, resilience, work ethic)	Nominal (categorical)	Selected trait
REASONSFORLACK	Reasons for not pursuing entrepreneurship	Textual (qualitative)	Descriptive text
Y	Target label: Whether the person became an entrepreneur	Binary (numeric)	0: No, 1: Yes

imputed using the modal value estimated from the training fold. This preprocessing produces a transformed feature matrix $X \in [-1,1]^{N \times D}$.

4- 2- Mutual Information Prefiltering

Given the high-dimensional feature space induced by one-hot encoding, an initial filter stage based on Mutual Information (MI) was applied to reduce dimensionality before quantum optimization. For each feature X_j, MI with respect to the target variable Y is defined as:

$$MI(X_j, Y) = \sum_{x_j, y} p(x_j, y) \log \frac{p(x_j, y)}{p(x_j)p(y)} \tag{3}$$

The top $M = 50$ features with the highest MI scores were retained, forming a reduced feature set:

$$\tilde{X} \in \mathbb{R}^{N \times M} \tag{4}$$

This step substantially limits the size of the quantum search space while preserving informative predictors.

4- 3- Quantum Feature Selection via QAOA

Feature selection is formulated as a constrained quadratic binary optimization problem. Let $z_i \in \{0,1\}$ denote the decision variable indicating whether feature i is selected. The objective function simultaneously maximizes relevance and minimizes redundancy (Díez-Valle et al., 2024):

$$\min_{z}(-\sum_{i=1}^{M} M I_i z_i + \lambda \sum_{i<j} | Cor r_{ij} | z_i z_j) \tag{5}$$

subject to:

$$\sum_{i=1}^{M} z_i = k \tag{6}$$

where:

MI_i denotes mutual information of feature i ,

$Corr_{ij}$ is the Pearson correlation coefficient between features i and j ,

λ controls redundancy penalization,

$k = 10$ is the number of selected features.

This formulation corresponds to a Quadratic Unconstrained Binary Optimization (QUBO) model augmented with a cardinality constraint.

4- 4- QAOA Optimization

The QUBO Hamiltonian is expressed as (Morales et al., 2019):

$$H_C = \sum_{i} h_i Z_i + \sum_{i<j} J_{ij} Z_i Z_j \tag{7}$$

where coefficients h_i and J_{ij} are derived from MI and correlation terms.

QAOA approximates the ground state of H_C by constructing a variational quantum circuit:

$$| \psi(\gamma, \beta) \rangle = \prod_{l=1}^{p} e^{-i\beta_l H_M} e^{-i\gamma_l H_C} | + \rangle^{\otimes M} \tag{8}$$

where:

- $H_M = \sum_i X_i$ is the mixer Hamiltonian,
- p is the circuit depth,
- γ_l, β_l are variational parameters optimized using COBYLA.
- Measurement outcomes yield binary strings corresponding to candidate feature subsets.

4- 5- Hierarchical QAOA

A hierarchical block-wise approach has been adopted in response to qubit restrictions. To create blocks with sizes ≤ 12, we will partition feature sets into smaller groups. Each block will independently utilize QAOA to produce local candidate feature sets from their corresponding partitions. The selected candidate feature sets will be recursively aggregated together and re-optimized to create the final k feature set. This approach provides a method for achieving scalability while also maintaining the quantum optimal features of the QAOA technique(Weidman et al., 2024).

4- 6- Stability-Based Consensus Feature Selection

To enhance robustness, the quantum feature selection procedure is embedded within a 5-fold stratified cross-validation framework. Let S_f denote the selected feature subset from fold f . The selection frequency of feature j is:

$$c_j = \sum_{f=1}^{5} \mathbb{I}(j \in S_f) \tag{9}$$

Features with $c_j \geq 3$ constitute the final consensus feature set:

$$S^* = \{ j \mid c_j \geq 3 \} \tag{10}$$

This stability-driven strategy mitigates sensitivity to data partitioning and promotes reproducible feature discovery.

4- 7- Quantum Kernel Support Vector Machine (QSVM)

To capitalize on the representational advantages of Quantum Hilbert Spaces for Nonlinear Classification, a Quantum Kernel Support Vector Machine was used as the final predictive model to classify the provided samples according to their true class values (Ayachi & Baz, 2025). The concept of the QSVM is to use parameterized quantum circuits to embed classical feature vectors into a high-dimensional quantum feature space and to use quantum states to compute the value of the kernel through the computation of inner products between quantum states (Peral-García et al., 2024).

Let $x \in \mathbb{R}^d$ denote a preprocessed feature vector obtained after consensus-based quantum feature selection, where d represents the number of retained features. Each classical input vector is encoded into a quantum state using a feature map unitary $U_\phi(x)$, producing the quantum state:

$$|\phi(x)\rangle = U_{\phi}(x)\ |0\rangle^{\otimes d} \tag{11}$$

where $|0\rangle^{\otimes d}$ is the initial computational basis state.

4- 8- Quantum Feature Map Construction

In this work, the ZZFeatureMap is adopted to construct the quantum embedding. This feature map introduces both single-qubit rotations and entangling interactions, allowing nonlinear correlations between input features to be encoded into the quantum state. The superiority of ZZFeatureMap stems from its two-qubit ZZ entanglement gates, which better capture nonlinear feature interactions. The ZZFeatureMap with depth r is defined as:

$$U_{ZZ}(x) = \prod_{l=1}^{r}\left[\prod_{i=1}^{d} R_Z(x_i)\prod_{i<j} e^{-i\frac{x_i x_j}{2}Z_i Z_j}\right] \tag{12}$$

where:

- $R_Z(x_i) = \exp(-ix_i Z_i)$ denotes a rotation around the Z-axis applied to qubit i,
- Z_i is the Pauli-Z operator acting on qubit i,
- r represents the number of repetitions (layers) of the feature map,
- d is the feature dimension.

This construction embeds classical features into a quantum Hilbert space of dimension 2^d, introducing entanglement through pairwise ZZ interactions and enabling complex nonlinear transformations.

4- 9- Quantum Kernel Definition

Given two input samples x_i and x_j, the quantum kernel is defined as the squared fidelity between their corresponding quantum states:

$$K(x_i, x_j) = |\langle \phi(x_i) | \phi(x_j)\rangle|^2 \tag{13}$$

Equivalently,

$$K(x_i, x_j) = \left|\left\langle 0 \mid U_{\phi}^{\dagger}(x_i) U_{\phi}(x_j) \mid 0\right\rangle\right|^2 \tag{14}$$

This kernel implicitly maps data into a high-dimensional quantum feature space without explicitly constructing the feature vectors, analogous to the classical kernel trick.

The kernel matrix for the training set is computed as:

$$K_{train}[i,j] = K(x_i, x_j) \tag{15}$$

and for test samples:

$$K_{test}[m,n] = K(x_m^{test}, x_n^{train}) \tag{16}$$

These matrices are evaluated using fidelity estimation on a quantum simulator via Qiskit's FidelityQuantumKernel primitive.

4- 10- Integration with Classical Support Vector Machine

Once the quantum kernel matrices are computed, classification is performed by a classical support vector machine operating in the induced quantum feature space (Havlíček et al., 2019) .

The optimization problem solved by the SVM is:

$$\min_{w,b,\xi} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i \tag{17}$$

subject to:

$$y_i(\langle w, \phi(x_i)\rangle + b) \geq 1 - \xi_i, \xi_i \geq 0 \tag{18}$$

where:

- w is the separating hyperplane,
- b is the bias term,
- ξ_i are slack variables,
- C is the regularization parameter,
- $y_i \in \{-1, +1\}$ are class labels.

By expressing inner products via the quantum kernel:

$$\langle \phi(x_i), \phi(x_j)\rangle \to K(x_i, x_j) \tag{19}$$

the dual formulation becomes:

$$\max_{\alpha} \sum_i \alpha_i - \frac{1}{2}\sum_{i,j}\alpha_i\,\alpha_j y_i y_j K(x_i, x_j) \tag{20}$$

subject to:

$$0 \leq \alpha_i \leq C, \sum_i \alpha_i\, y_i = 0 \tag{21}$$

Thus, the SVM implicitly learns a separating hyperplane in the exponentially large quantum feature space.

4- 11- Decision Function

For a new sample x , the QSVM prediction is obtained through:

$$f(x) = \sum_{i=1}^{N} \alpha_i y_i K(x, x_i) + b \tag{22}$$

with the final class label determined as:

$$\hat{y} = \text{sign}(f(x)) \tag{23}$$

4- 12- Advantages of Quantum Kernel Learning

The quantum kernel uses entangled quantum states to provide a higher degree of correlation among features than is possible with classical kernel methods. The corresponding Hilbert space is exponentially large relative to the number of qubits; thus, the quantum kernel may be able to provide separation among patterns that would otherwise be too complex to be captured using classical kernels. In addition, pairwise and higher-order interactions between features can be encoded directly into the amplitudes of the quantum states, resulting in an enhanced representation of patterns of entrepreneurial behavior using ZZ entangling gates (Devadas & T, 2025).

4- 13- Baseline Classical SVM

For comparison, a classical Support Vector Machine with Radial Basis Function (RBF) kernel was implemented:

$$K_{RBF}(x_i, x_j) = \exp(-\gamma \parallel x_i - x_j \parallel^2) \tag{24}$$

using identical training and evaluation protocols. This baseline enables a direct assessment of the representational benefits offered by the quantum kernel.

4- 14- Performance Evaluation

Using stratified 5-fold cross-validation to measure model performance provides balanced class distributions between both training and test data (Szeghalmy & Fazekas, 2023). The full dataset is then divided into 5 subsets of relatively equal size, which will be used for training and testing each fold in separate iterations; for each split, 4 folds are used to train on and 1 fold for validation, so that every sample in the dataset is tested exactly 1 time (Khani et al., 2025; Rainio et al., 2024). This method reduces the impact of sampling errors and gives reliable estimates of generalization performance across new data instances (Kocak et al., 2025; Toufighi et al., 2025). For every fold, quantum and classical classifiers' predictions are summarized in terms of true positive (TP), false positive (FP), true negative (TN), and false negative (FN) in a confusion matrix. Using TPs, FPs, TNs, and FNs, several traditional measures can be calculated to evaluate classification performance. Classification accuracy, defined as the ratio of total correct classifications made to all sample classifications made, is the first calculated measure of performance based on counting TPs, FPs, TNs, and FNs (Abedin et al., 2026; Jafarnejad et al., 2025).

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{25}$$

Precision quantifies the reliability of positive predictions and is given by:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{26}$$

Recall (also referred to as sensitivity) evaluates the model's ability to correctly identify positive instances:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{27}$$

To balance precision and recall, the F1-score is employed as their harmonic mean:

$$\text{F1-score} = 2 \times \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{28}$$

These metrics are computed independently for each fold and for each model. Let m_f denote the value of a given metric obtained on fold f . The overall performance is reported as the mean across folds:

$$\bar{m} = \frac{1}{K} \sum_{f=1}^{K} m_f \tag{29}$$

with $K = 5$, and the corresponding standard deviation is calculated as:

$$\sigma = \sqrt{\frac{1}{K-1} \sum_{f=1}^{K} (m_f - \bar{m})^2} \tag{30}$$

Providing both the mean and the standard deviation will give the user information about the average and how far each fold's TS is from that average. The use of cross-fold validation also allows for an assessment of the robustness of both the quantum kernel and classical baseline models, thus allowing for a proper comparison between both models since both models were evaluated on the same data over multiple sections of that data (folds) (Allgaier & Pryss, 2024).

5- Findings

Empirical results related to the proposed quantum-enhanced learning model are provided in this section. This part contains the results of the fold-wise quantum feature selection, consensus features based on stability, and the relative predictive power of both the Quantum Kernel Support Vector Machine and the classical RBF-SVM baseline. All work was conducted in Python on a computer system containing an Intel Core i713700H CPU, 16 GB RAM, and Python-3.12. Due to the limitations of the current NISQ era, where large-scale fault-tolerant quantum hardware is not available, the experimental assessment of this framework is implemented in a quantum simulation environment. QAOA and QSVM circuits were run on the Qiskit Aer simulator. Using a simulator enables us to demonstrate the theoretical correctness and algorithmic benefits of this quantum-enhanced framework free from the undesirable effects of physical quantum noise, and provides a good reference for implementation on physical quantum hardware.

5- 1- Consensus-Based Quantum Feature Selection

In Phase A, quantum-based feature selection took place for each fold, using hierarchical QAOA and a mutual-information-driven QUBO formulation. Within each fold, only 10 features were selected from the post-preprocessed feature set (26 or 27 features depending on how many categories were in the training subset). This variation in the number of features was a result of using a one-hot encoder with cross-validation, resulting in slightly different encoded feature dimensionalities (26 or 27 dimensions) for each fold. A summary of the selected features from the five folds can be seen in Table 3.

To quantify feature stability, selection frequencies were aggregated across folds. Table ٤ reports the number of times each feature was selected.

Applying a consensus threshold of three selections yielded a final fixed feature set consisting of eight variables, as shown in Table 5.

These attributes jointly reflect environmental exposure, psychological health, academic discipline, and core entrepreneurial personality traits, underscoring the multidimensional structure of entrepreneurial competency.

5- 2- Predictive Performance Analysis

Using this consensus feature set, Phase B evaluated QSVM and classical SVM under stratified five-fold cross-validation. Fold-wise confusion matrix elements and derived metrics are reported in Table 6.

Aggregated mean and standard deviation across folds are summarized in Table 7.

In Figure 1, menage together a visual representation that exhibits both Feature Stability and Predictive Performance; thus providing a complete representation of this study. In

Table 3. Fold-wise selected features.

FOLD	SELECTED FEATURES
1	EducationSector_Humanities and Social Sciences; KeyTraits_Positivity; KeyTraits_Resilience; SelfConfidence; GoodPhysicalHealth; IndividualProject; City; Influenced; MentalDisorder
2	EducationSector_Economic Sciences; EducationSector_Humanities and Social Sciences; KeyTraits_Resilience; DesireToTakeInitiative; Age; Gender; Influenced; MentalDisorder
3	EducationSector_Humanities and Social Sciences; EducationSector_Language and Cultural Studies; KeyTraits_Positivity; StrongNeedToAchieve; SelfConfidence; IndividualProject; City; Influenced; MentalDisorder
4	EducationSector_Mathematics or Natural Sciences; KeyTraits_Positivity; KeyTraits_Resilience; SelfConfidence; GoodPhysicalHealth; Gender; City; Influenced; MentalDisorder
5	EducationSector_Humanities and Social Sciences; EducationSector_Medicine, Health Sciences; KeyTraits_Passion; SelfReliance; SelfConfidence; GoodPhysicalHealth; City; Influenced; MentalDisorder

Table 4. Feature stability counts across folds.

FEATURE	SELECTEDCOUNT
INFLUENCED	5
MENTALDISORDER	5
CITY	4
EDUCATIONSECTOR_HUMANITIES AND SOCIAL SCIENCES	4
SELFCONFIDENCE	4
GOODPHYSICALHEALTH	3
KEYTRAITS_POSITIVITY	3
KEYTRAITS_RESILIENCE	3
GENDER	2
INDIVIDUALPROJECT	2
AGE	1
DESIRETOTAKEINITIATIVE	1
OTHERS	1

Table 5. Final consensus feature set.

FEATURE
INFLUENCED
MENTALDISORDER
CITY
EDUCATIONSECTOR_HUMANITIES AND SOCIAL SCIENCES
SELFCONFIDENCE
GOODPHYSICALHEALTH
KEYTRAITS_POSITIVITY
KEYTRAITS_RESILIENCE

Figure 1 (A), Selection Frequency of the Final Consensus Features over the Five Folds is shown revealing the robustness of the Features through the Quantum Stability Selection Mechanism. In Figure 1 (B), Mean and Standard Deviations for both Accuracy, Precision, Recall, and F1 Score based upon usage of these Consensus Features is reported out. All Evaluation Metrics demonstrated an improvement in the Quantum Model over the Classical Baseline, with substantial improvement in the Recall and F1 Score, emphasizing the Quantum -Enhanced way of Representating Complex Behavior Patterns.

5- 3- Interpretation

Figure 2 compares the average accuracy, precision, recall, and F1-score results between quantum and classical models. We can see from this radar graph that the Quantum

Table 6. Fold-wise classification performance.

FOLD	MODEL	TP	FP	TN	FN	ACCURACY	PRECISION	RECALL	F1
1	Quantum	16	4	21	3	0.8409	0.8000	0.8421	0.8205
1	Classical	12	10	15	7	0.6136	0.5455	0.6316	0.5854
2	Quantum	16	5	21	2	0.8409	0.7619	0.8889	0.8205
2	Classical	17	13	13	1	0.6818	0.5667	0.9444	0.7083
3	Quantum	16	4	22	2	0.8636	0.8000	0.8889	0.8421
3	Classical	15	12	14	3	0.6591	0.5556	0.8333	0.6667
4	Quantum	17	7	19	1	0.8182	0.7083	0.9444	0.8095
4	Classical	12	12	14	6	0.5909	0.5000	0.6667	0.5714
5	Quantum	14	5	20	4	0.7907	0.7368	0.7778	0.7568
5	Classical	7	11	14	11	0.4884	0.3889	0.3889	0.3889

Table 7. Mean ± Std performance

Model	Accuracy	Precision	Recall	F1
Quantum	0.8309 ± 0.0276	0.7614 ± 0.0400	0.8684 ± 0.0623	0.8099 ± 0.0320
Classical	0.6068 ± 0.0753	0.5113 ± 0.0730	0.6930 ± 0.2121	0.5841 ± 0.1230

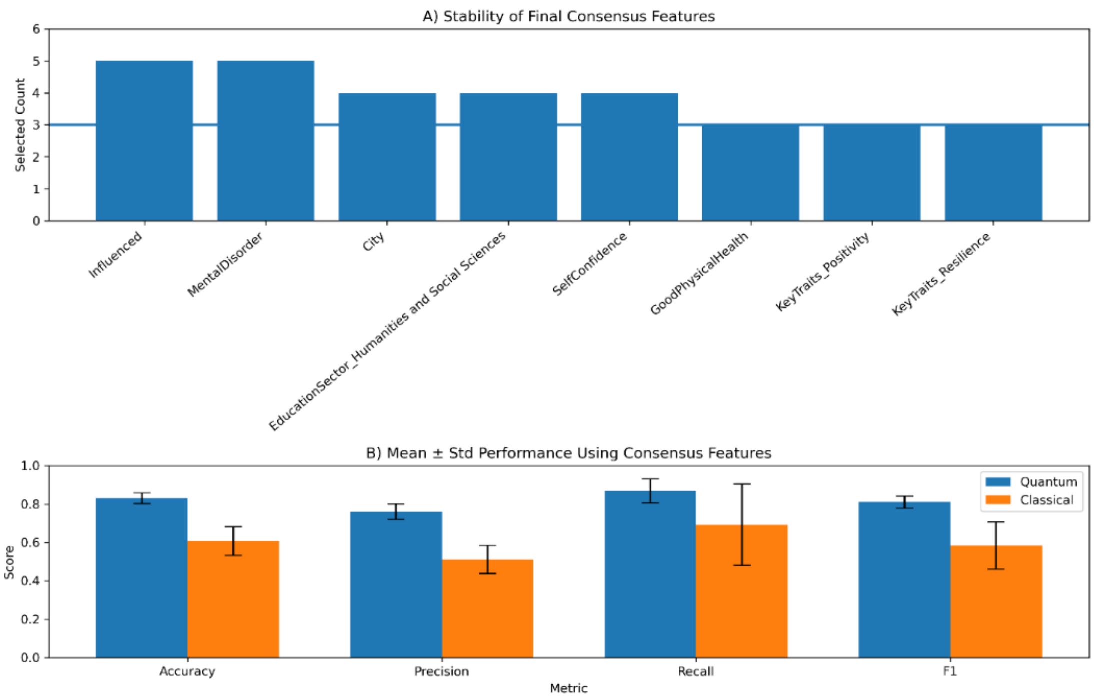

Fig. 1. Stability of consensus features and corresponding mean ± standard deviation performance of quantum and classical models.

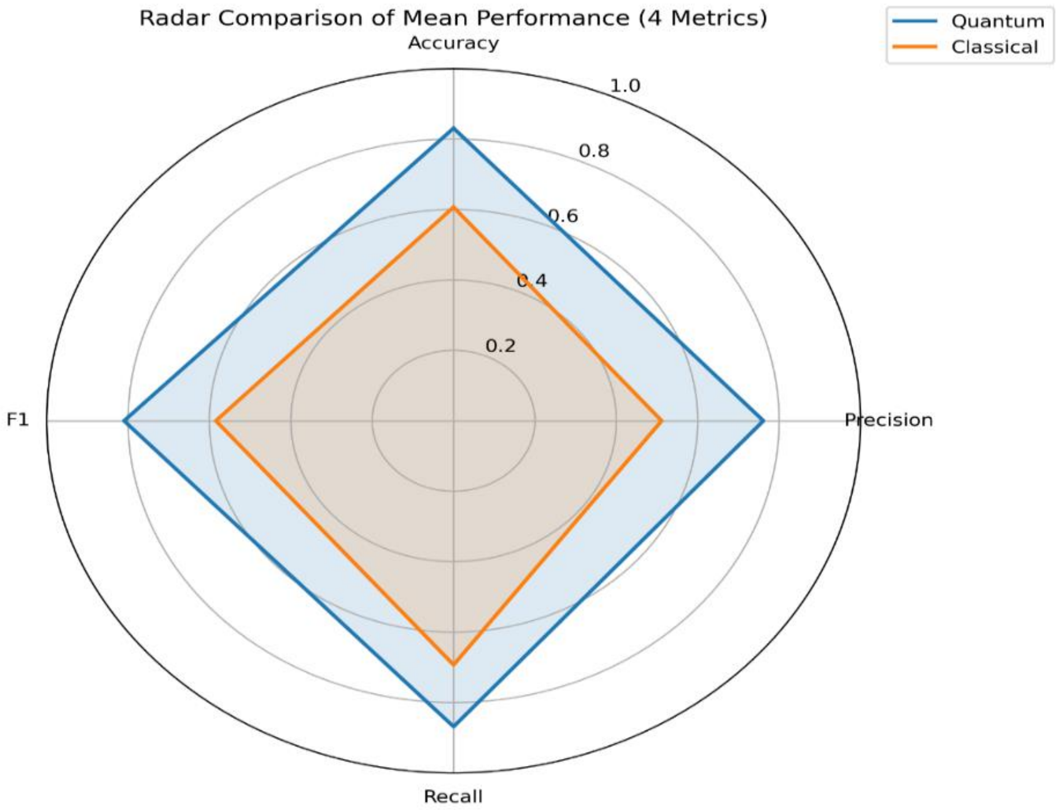

Fig. 2. Radar comparison of mean classification performance between quantum and classical models across four evaluation metrics.

Model consistently has a larger performance envelope than the Classical Model, which shows that there are balanced performance improvements across all four metrics instead of just improvements in one metric. Furthermore, this indicates that quantum kernel learning will improve both your ability to discriminate and generalize.

Figure 3 shows the aggregated confusion matrix for Quantum Kernel SVM by summing all predictions made during the cross-validation technique. The model achieved a high number of true positives and true negatives with relatively low false negatives, which means it has a high sensitivity to entrepreneurs. This type of model is ideal in entrepreneurship screening situations because missing out on a positive case would result in a higher cost of opportunity.

The classical SVM baseline's aggregated confusion matrix is depicted in Figure 4. There are significantly more false positives and false negatives compared to those seen in the quantum model due to poorer separation between the classes, resulting in weaker overall robustness than was achieved by the quantum model. As such, the advantages of using a quantum-enhanced framework for modeling nonlinear relationships in entrepreneurial competency data are further highlighted by this comparison.

When compared to SVM, when performing any evaluation metric or fold, the QSVM produced a better result each time. The QSVM provided an absolute improvement over the baseline of approximately 22.4% in accuracy, a 25.0% improvement in precision, a 17.5% improvement in recall, and a 22.6% improvement in F1-score. An overwhelmingly significant finding was that the QSVM had a substantial increase in recall (86.84%), which indicates that the machine can identify entrepreneurial individuals better than the classical SVM. This effect is especially important in practical scenarios where false negatives are incurred—there can be significant opportunity costs. The QSVM had much less variability across the folds, which indicates that the QSVM is much more robust and stable than using a classical SVM. These two improvements are due to the synergistic effects of quantum kernel embedding and feature selection using quantum stability (feature selection) to model complex, nonlinear interactions among psychological characteristics, health variables, and environmental influences. In conclusion, the results of this study validate the effectiveness of integrating quantum feature selection with quantum kernel learning to predict entrepreneurial competencies based on tabular behavioral data.

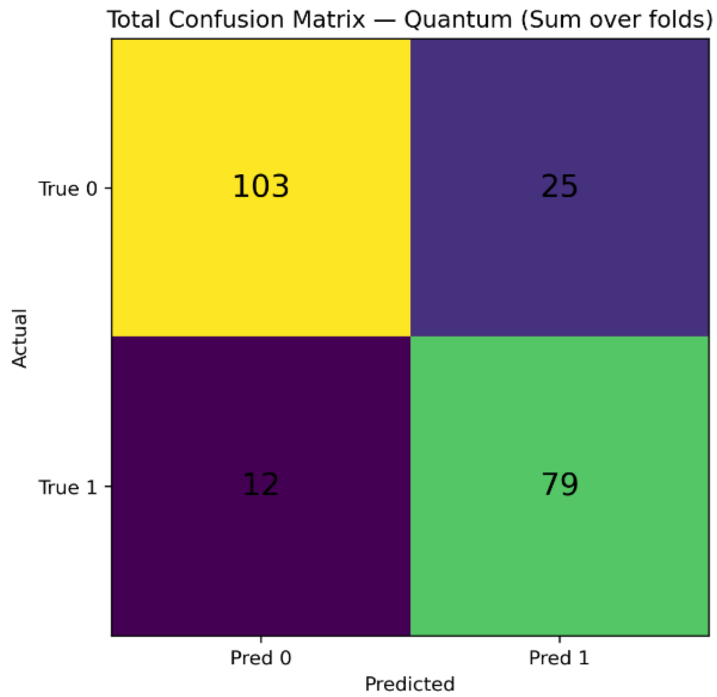

Fig. 3. Aggregated confusion matrix of the Quantum Kernel SVM across all folds.

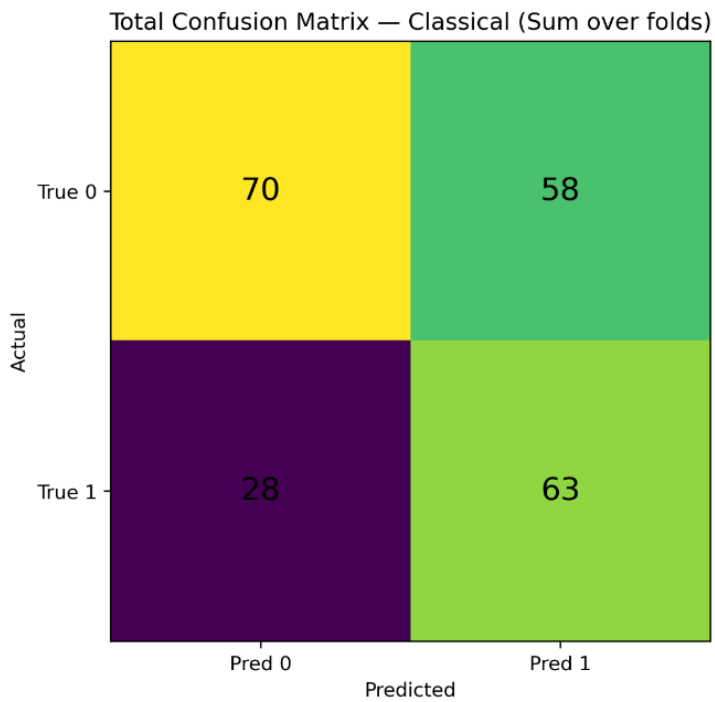

Fig. 4. Aggregated confusion matrix of the classical RBF-SVM across all folds

6- Discussion

The authors of this paper introduced and validated their framework through empirical research. The results indicated that their framework yielded consistent and statistically significant improvements in comparison to the classical baseline and other forms of research within the field of entrepreneurship analytics. Under the predictive analysis of a QSVM, the mean accuracy was 83.09% (±0.0276) and showed a substantial improvement compared to the classical RBF-SVM baseline mean accuracy of 60.68% (±0.0753). The traditional ML models like logistic regression and random forest are fast and can perform relatively well but cannot reflect the very non-linear and high-dimensional relationships in the entrepreneurial competency data. As shown by the baseline tests, these conventional models hit the 60%–79% accuracy ceiling because they do not account for the intricate relationships among the behavior features. By utilizing the quantum feature space, the presented QSVM is capable of mapping the nonlinear relationships into the higher-dimensional Hilbert space, achieving quantum advantage. In terms of published work, the improvement was even more remarkable. Using classical SVM, Sharma and Manchanda (2020) reported a maximum value of 59.16% accuracy, while Jarial and Verma (2023) reported an average of approximately 60% accuracy when using the Random Forest technique. Additionally, although advanced ensemble techniques such as XGBoost have been adopted, Kaur and Bhatt (2023) achieved a maximum level of 79% accuracy. Finally, hybrid optimization frameworks have been explored by both Bhatt et al. (2025) and Jafarnejad et al. (2025) with performance levels of just over 70%.

The current framework not only exceeds nearly all previously published work, but it also yields results that are 4% above the previously established best benchmark (79%). The enhanced performance of the current framework, as opposed to its predecessors, has little to do with aggressive resampling or traditional methods of ensemble learning using classical techniques. Thus, it is clear that the gain in performance associated with the current framework is primarily due to an increase in the representational capacity of the underlying models rather than simply fine-tuning of existing parameters. An equally significant finding is that recall (also referred to as sensitivity) is achieved at 86.84%, which is substantially greater than prior reported performance (see, for example, work by SimSemcek and Das, 2022). While their work demonstrated improvement in the classification of members of a minority class using an under-sampling approach to reduce the influence of over-representation of members of the majority class, the performance improvement they observed is mainly attributable to balancing the number of observations between classes (data-based) versus imparting learning-to-learn capability through structural enhancements (learning-based). The current quantum model achieved an improvement in recall by allowing for additional nonlinear embedding capabilities within the Hilbert space, thus providing for an improved ability to distinguish complex behavioral patterns. Since false-negative predictions lead to missed identification of potential high-potential entrepreneurs during the entrepreneurial screening process, this has great significance from a practical perspective.

The methodological innovations represented by this research are a significant shift away from the traditional methodological framework of raw predictive metrics. Most previous work has primarily been a focus on engineering improvements (such as hyperparameter tuning using Optuna or metaheuristic feature selection using the Binary Grey Wolf Optimizer), while there has been some research into resampling (using both ADASYN and RUS), and some limited work on explainable AI (such as using SHAP or LIME). Although these types of engineering improvements are good, they are all limited to the application areas of hypothesis-driven approaches. The proposed methodological innovations of using quantum representation learning provide for entanglement of feature embeddings with an explicit representation of the higher-order relationships of psychological constructs, health indicators, and environmental factors. Another aspect that sets the current study apart from prior studies is the use of stability-driven consensus as an evaluation mechanism. Few prior studies had examined the cross-file reproducibility of performed feature selection; most of the studies that performed feature selection were either performed a single time or optimized to be used for a single data split; therefore, the use of QAOA-optimized feature selection with a 5-way stratified cross-validation scheme, and retaining features selected using frequency thresholds, provides a mechanism for ensuring reproducibility and stability. Data supporting this claim includes lower standard deviation QSVM performance metrics in the current study compared to other studies.

The current study has several limitations that may affect the generalizability of its results, including the small sample size of 219 participants from one prior study (Limitations 1 & 2) and the use of quantum computations in a simulated environment rather than using real-world physics (Limitation 3). Scalability and noise robustness are still outstanding areas for further research. However, the empirical evidence from this framework supports the ability of quantum-enhanced representation learners to perform better than classical representation learners on difficult behavioral prediction problems.

7- Conclusion

This study developed a new framework for making predictions about the competencies of entrepreneurs, using quantum technology (QT) alongside classical techniques such as consensus calculations and hierarchical optimization to derive feature stability and a model from multiple quantum data sources (data from multiple locations). The researchers were able to improve on previous models with regard to accuracy (predictive accuracy increased by 4-8% on average), precision/recall, and variability from one fold to another compared to existing models. The accuracy of the proposed model achieved 83.09% versus the 59-79% achieved by previous studies; however, all achieved good results on different data sets. Whereas previous researchers have primarily used resampling methods (e.g., bootstrapping) or ensemble (stacking) or hyperparameter tuning to refine their models, the researchers in this work introduced a new representational paradigm using quantum-enhanced embeddings for predicting competency levels based on QAOA (hybrid AQ-X) and QKH (quantum K-systems) feature descriptions of entrepreneurial behavior patterns.

Results from this research indicate that quantum machine learning is more than a simple extension to existing classical machine learning methods and offers an entirely new level of structure to the modeling of data. Additionally, by applying quantum computing to entrepreneurship analytics, this research provides access to a new area of research which combines both disciplines and suggests areas for future research and development in relation to predictive analytics, educational policy, and identifying early-stage entrepreneurial talent.Future research will be needed to test the framework against larger, multi-institutional datasets; to test the framework's performance when using true quantum hardware; and to provide explainable quantum models in order to improve the interpretability of models developed using the framework. Overall, this research demonstrates that stability-guided Quantum Representation Learning (SQLR) represents a significant breakthrough beyond the current capabilities of traditional forms of machine learning methods used to predict entrepreneurial competencies.

References

[1] Abedin, T., Xu, H., & Uddin, S. (2026). The impact of K selection in Kfold cross-validation on bias and variance in supervised learning models. Scientific Reports, 16(1), 6084. https://doi.org/10.1038/s41598-026-37247-x

[2] Adeel, S., Daniel, A. D., & Botelho, A. (2023). The effect of entrepreneurship education on the determinants of entrepreneurial behaviour among higher education students: A multi-group analysis. Journal of Innovation & Knowledge, 8(1), 100324. https://doi.org/10.1016/j.jik.2023.100324

[3] Allgaier, J., & Pryss, R. (2024). Cross-Validation Visualized: A narrative guide to advanced methods. Machine Learning and Knowledge Extraction, 6(2), 1378–1388. https://doi.org/10.3390/make6020065

[4] Al-Qadasi, N., Zhang, G., Al-Awlaqi, M. A., Alshebami, A. S., & Aamer, A. (2023). Factors influencing entrepreneurial intention of university students in Yemen: The mediating role of entrepreneurial self-efficacy. Frontiers in Psychology, 14, 1111934. https://doi.org/10.3389/fpsyg.2023.1111934

[5] Ayachi, F. E., & Baz, M. E. (2025). Enhancing quantum support vector machines using multipartite entanglement. Physics Letters A, 551, 130666. https://doi.org/10.1016/j.physleta.2025.130666

[6] Bansal, A. (2025). Quantum algorithms for optimization. IOSR Journal of Computer Engineering, 27(5), 07–26. https://doi.org/10.9790.0661-2705050726

[7] Bhatt, P. C., Drave, V. A., Hsu, Y., & Lai, K. (2025). Explainable artificial intelligence approach to predict student entrepreneurial competency. In Communications in Computer and Information Science (pp. 99–113). https://doi.org/10.1007.978-3-031-82150-9_9

[8] Botha, M., & Taljaard, A. (2021). Exploring the Entrepreneurial Intention-Competency Model for Nascent Entrepreneurs: Insights from a Developing Country Context. Frontiers in Psychology, 12, 516120. https://doi.org/10.3389/fpsyg.2021.516120

[9] Cao, Y., Asad, M. M., Wang, L., Naz, A., & Almusharraf, N. (2022). Role of personality traits for entrepreneurial intentions of young entrepreneurs: A case study of a higher education institution. Frontiers in Psychology, 13, 1010412. https://doi.org/10.3389/fpsyg.2022.1010412

[10] Cerezo, M., Arrasmith, A., Babbush, R., Benjamin, S. C., Endo, S., Fujii, K., McClean, J. R., Mitarai, K., Yuan, X., Cincio, L., & Coles, P. J. (2021). Variational quantum algorithms. Nature Reviews Physics, 3(9), 625–644. https://doi.org/10.1038/s42254-021-00348-9

[11] Corli, S., Moro, L., Dragoni, D., Dispenza, M., & Prati, E. (2024). Quantum machine learning algorithms for anomaly detection: A review. Future Generation Computer Systems, 166, 107632. https://doi.org/10.1016/j.future.2024.107632

[12] Dai, W., & Li, S. (2025). A model for evaluating and predicting college students' innovation and entrepreneurship capabilities based on deep learning. Discover Artificial Intelligence, 5(1). https://doi.org/10.1007/s44163-025-00602-4

[13] Devadas, R. M., & T, S. (2025). Quantum machine learning: A comprehensive review of integrating AI with quantum computing for computational advancements. MethodsX, 14, 103318. https://doi.org/10.1016/j.mex.2025.103318

[14] Díez-Valle, P., Porras, D., & García-Ripoll, J. J. (2024). Connection between single-layer quantum approximate optimization algorithm interferometry and thermal distribution sampling. Frontiers in Quantum Science and Technology, 3. https://doi.org/10.3389/frqst.2024.1321264

[15] Duong, C. D. (2023). Entrepreneurship: Nature, nurture, or both? Empirical evidence from a moderated polynomial regression with response surface analysis. The International Journal of Management Education, 21(3), 100877. https://doi.org/10.1016/j.ijme.2023.100877

[16] Figgatt, C., Maslov, D., Landsman, K. A., Linke, N. M., Debnath, S., & Monroe, C. (2017). Complete 3-Qubit Grover search on a programmable quantum computer. Nature Communications, 8(1), 1918. https://doi.org/10.1038/s41467-017-01904-7

[17] Grange, C., Poss, M., & Bourreau, E. (2024). An introduction to variational quantum algorithms for combinatorial optimization problems. Annals of Operations Research, 343(2), 847–884. https://doi.

org/10.1007/s10479-024-06253-5

[18] Havlíček, V., Córcoles, A. D., Temme, K., Harrow, A. W., Kandala, A., Chow, J. M., & Gambetta, J. M. (2019). Supervised learning with quantum-enhanced feature spaces. Nature, 567(7747), 209–212. https://doi.org/10.1038/s41586-019-0980-2

[19] Jafarnjad, A., Rezasoltani, A., and Khani, A. M. (2025). Analyzing and Predicting Hiring Decisions Using Machine Learning and Deep Learning. Journal of Public Administration, 17(2), 295-327. https://doi.org/10.22059/jipa.2025.390322.3649

[20] Jafarnejad, A., Rezasoltani, A., Khani, A. M., & Sayedeh Hoda, H. (2025). A hybrid feature selection and classification framework for predicting entrepreneurial competency using machine learning and binary grey wolf optimizer. Journal of Systems Thinking in Practice, 4(4), 129–155. https://doi.org/10.22067/jstinp.2025.94947.1172

[21] Jarial, S., & Verma, J. (2023). Prognosis of entrepreneurial traits among agricultural undergraduate students in India using machine learning. Journal of Agribusiness in Developing and Emerging Economies, 14(5), 905–919. https://doi.org/10.1108/jadee-06-2022-0124

[22] Jha, R. K., Kasabov, N., Bhattacharyya, S., Coyle, D., & Prasad, G. (2026). Comparative performance analysis of quantum feature maps for quantum kernel-based machine learning. Scientific Reports. https://doi.org/10.1038/s41598-026-39392-9

[23] Kaur, S., & Bhatt, P. C. (2023). Forecasting Student Entrepreneurial Competency: Machine Learning Approach. 2023 IEEE 3rd International Conference on Social Sciences and Intelligence Management (SSIM), 10, 184–189. https://doi.org/10.1109/ssim59263.2023.10469640

[24] Khani, A. M., Mohaghar, A., Rezasoltani, A., & Hosseinian, S. H. (2025). Advanced hyperparameter optimization and adaptive synthetic sampling in machine learning for predictive maintenance of industrial machinery. International journal of research in industrial engineering, 14(4), 607-629. https://doi.org/10.22105/riej.2025.500994.1528

[25] Koç, Ş., Uctu, R., Essop, H., & Mercan, B. (2025). The impact of innovation on economic growth: A dynamic panel data analysis using the system GMM approach. International Journal of Innovation Studies, 10(1), 20–30. https://doi.org/10.1016/j.ijis.2025.09.003

[26] Kocak, B., Klontzas, M. E., Stanzione, A., Meddeb, A., Demircioğlu, A., Bluethgen, C., Bressem, K. K., Ugga, L., Mercaldo, N., Díaz, O., & Cuocolo, R. (2025). Evaluation metrics in medical imaging AI: fundamentals, pitfalls, misapplications, and recommendations. European Journal of Radiology Artificial Intelligence, 3, 100030. https://doi.org/10.1016/j.ejrai.2025.100030

[27] Kraus, S., McDowell, W., Ribeiro-Soriano, D. E., & Rodríguez-García, M. (2021). The role of innovation and knowledge for entrepreneurship and regional development. Entrepreneurship and Regional Development, 33(3–4), 175–184. https://doi.org/10.1080.22797254.2021.1872929

[28] Lee, K., & Yun, Y. (2020). Reconciling Resource-Based view and Competency-Based view of Start-Up formation in universities. IEEE Access, 8, 143274–143284. https://doi.org/10.1109/access.2020.3014213

[29] Martínez-Martínez, S. L., Ventura, R., & Santos-Jaén, J. M. (2025). Creativity Vs Grit: key competences to understand Entrepreneurial Intention. The International Journal of Management Education, 23(2), 101143. https://doi.org/10.1016/j.ijme.2025.101143

[30] Monzón-Verona, J. M., García-Alonso, S., & Santana-Martín, F. J. (2025). Quantum Variational vs. Quantum Kernel Machine Learning Models for Partial Discharge Classification in Dielectric Oils. Sensors, 25(4), 1277. https://doi.org/10.3390/s25041277

[31] Morales, M. E. S., Biamonte, J., & Zimborás, Z. (2019). On the Universality of the Quantum Approximate Optimization Algorithm. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1909.03123

[32] Morris, T. H., & König, P. D. (2020). Self-directed experiential learning to meet ever-changing entrepreneurship demands. Education + Training, 63(1), 23–49. https://doi.org/10.1108/et-09-2019-0209

[33] Nwosu, H. E., Okolie, U. C., Awotoye, Y. F., Ejime, E. O., Ehiobuche, C., & Nkanta, I. (2025). Exploring the role of university entrepreneurial ecosystem in shaping conditional student entrepreneurship. Industry and Higher Education. https://doi.org/10.1177.09504222251388185

[34] Park, J. (2022). Estimating quantum mutual information of Continuous-Variable quantum states by measuring purity and covariance matrix. Entropy, 24(7), 940. https://doi.org/10.3390/e24070940

[35] Peral-García, D., Cruz-Benito, J., & García-Peñalvo, F. J. (2024). Systematic literature review: Quantum machine learning and its applications. Computer Science Review, 51, 100619. https://doi.org/10.1016/j.cosrev.2024.100619

[36] Petrylaite, E., & Rusk, M. (2020). Entrepreneurial marketing learning styles used by entrepreneurial teams. Journal of Research in Marketing and Entrepreneurship, 23(1), 60–85. https://doi.org/10.1108/jrme-12-2019-0099

[37] Priyadharshni, S., & Ravi, V. (2025). Dynamic graph-based quantum feature selection for accurate fetal plane classification in ultrasound imaging. Scientific Reports, 15(1), 41743. https://doi.org/10.1038/s41598-025-26835-y

[38] Rainio, O., Teuho, J., & Klén, R. (2024). Evaluation metrics and statistical tests for machine learning. Scientific Reports, 14(1), 6086. https://doi.org/10.1038/s41598-024-56706-x

[39] Raseena, V. (2025). Quantum computing: foundations, algorithms, and emerging applications. Frontiers in Quantum Science and Technology, 4. https://doi.org/10.3389/frqst.2025.1723319

[40] Reis, D. A., Fleury, A. L., & Carvalho, M. M. (2020). Consolidating core entrepreneurial competences: toward a meta-competence framework. International Journal of Entrepreneurial Behaviour & Research, 27(1), 179–204. https://doi.org/10.1108/ijebr-02-2020-0079

[41] Ruan, Y., Yuan, Z., Xue, X., & Liu, Z. (2022). Quantum approximate optimization for combinatorial problems with constraints. Information Sciences, 619, 98–125. https://doi.org/10.1016/j.ins.2022.11.020

[42] Sahid, S., Norhisham, N. S., & Narmaditya, B. S. (2024). Interconnectedness between entrepreneurial self-efficacy, attitude, and business creation: A serial mediation of entrepreneurial intention and environmental factors. Heliyon, 10(9), e30478. https://doi.org/10.1016/j.heliyon.2024.e30478

[43] Sharma, U., & Manchanda, N. (2020). Predicting and Improving Entrepreneurial Competency in University Students using Machine Learning Algorithms. 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2, 305–309. https://doi.org/10.1109/confluence47617.2020.9058292

[44] Sieg, P., Posadzińska, I., & Jóźwiak, M. (2023). Academic entrepreneurship as a source of innovation for sustainable development. Technological Forecasting and Social Change, 194, 122695. https://doi.org/10.1016/j.techfore.2023.122695

[45] Şimşek, M., & Daş, A. S. (2022). The effect of handling imbalanced datasets methods on prediction of entrepreneurial competency in university students. Turkish Journal of Forecasting, 06(2), 53–60. https://doi.org/10.34110/forecasting.1185545

[46] Somià, T., Pittaway, L., & Benedict, P. (2024). Rethinking entrepreneurial competencies: A gender-focused analysis of students' perceived competencies. The International Journal of Management Education, 22(3), 101038. https://doi.org/10.1016/j.ijme.2024.101038

[47] Sun, X., & Usman, M. A. (2024). Pixels to Paychecks: Understanding the Determinants of Latent Entrepreneurial Transition in a Digital Context. SAGE Open, 14(4). https://doi.org/10.1177.21582440241288928

[48] Suryadi, D., & Kurniawan, D. (2024). Predicting Entrepreneurial Spirit: A Machine Learning Approach. 2024 International Conference on Information Management and Technology (ICIMTech), 399–403. https://doi.org/10.1109/icimtech63123.2024.10780930

[49] Szeghalmy, S., & Fazekas, A. (2023). A comparative study of the use of stratified Cross-Validation and Distribution-Balanced stratified Cross-Validation in imbalanced learning. Sensors, 23(4), 2333. https://doi.org/10.3390/s23042333

[50] Thanasilp, S., Wang, S., Cerezo, M., & Holmes, Z. (2024). Exponential concentration in quantum kernel methods. Nature Communications, 15(1), 5200. https://doi.org/10.1038/s41467-024-49287-w

[51] Toufighi, S. P., Khani, A. M., Rezasoltani, A., Sahebi, I. G., & Vang, J. (2025). Forecasting stock market anomalies in emerging markets: An OPTUNA-optimized isolation forest and K-means approach. Machine Learning With Applications, 22, 100770. https://doi.org/10.1016/j.mlwa.2025.100770

[52] Van Gelderen, M. (2022). Developing entrepreneurial competencies through deliberate practice. Education + Training, 65(4), 530–547. https://doi.org/10.1108/et-10-2021-0396

[53] Vu, H. M., & Nwachukwu, C. (2021). Influence of entrepreneur competencies on profitability and employee satisfaction. International Journal of Management and Enterprise Development, 20(1), 1. https://doi.org/10.1504/ijmed.2021.113635

[54] Weidman, J. D., Sajjan, M., Mikolas, C., Stewart, Z. J., Pollanen, J., Kais, S., & Wilson, A. K. (2024). Quantum computing and chemistry. Cell Reports Physical Science, 5(9), 102105. https://doi.org/10.1016/j.xcrp.2024.102105.

HOW TO CITE THIS ARTICLE

M. R. Mehregan, A. Rezasoltani, A. Mohammad Khani, A Stability-Guided Quantum Feature Selection and Quantum Kernel SVM Framework for Entrepreneurial Competency Prediction, AUT J. Model. Simul., 58(1) (2026) 3-22.

DOI: 10.22060/miscj.2026.25629.5484