International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 10 | Oct - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

VentureIQ : AI Powered Startup Idea Evaluation and Risk Intelligence

Ms. Shreyanshi Patel[1], Gauravgyan Singh[2], Asif Mansuri[3] , Saurabh Sharma[4], Rajat Bomenwar[5]

Computer Technology & Priyadarshini College of Engineering, Nagpur

ABSTRACT - Startup evaluation is a complex process involving financial indicators, market conditions, business relationships, and uncertainty about future growth. Conventional assessment approaches may provide limited support to early-stage founders who lack access to structured analytical resources. This paper presents **VentureIQ**, an AI-powered startup evaluation and risk intelligence platform that integrates machine learning, information retrieval, and large language models within a unified web application. The proposed system operates in two modes: post-funding startup evaluation and pre-funding market intelligence. In the post-funding mode, a Gradient Boosting Classifier analyzes startup-related features to estimate viability and classify associated risk. A Retrieval-Augmented Generation (RAG) pipeline, supported by FAISS-based similarity search and sentence embeddings, supplies relevant contextual information for generating SWOT analyses and strategic recommendations. In the pre-funding mode, the Groq API facilitates language-model-based market analysis and preliminary business strategy development. Implemented using Streamlit and Python, VentureIQ combines predictive analytics with contextual insights through an interactive interface. The platform is intended to support preliminary assessment rather than replace professional investment decisions. Its predictive reliability and practical generalizability depend on dataset quality, model validation, and the relevance of retrieved information. This work presents an integrated framework for making startup assessment more accessible, structured, and informative for entrepreneurs and aspiring founders.

***Key Words*:** *Machine Learning, Startup Prediction, Gradient Boosting, Streamlit, Decision Support Systems, Synthetic Data.*

1. INTRODUCTION

The global startup ecosystem faces a high failure rate, with over 90% of startups failing within their first five years [1]. Traditional methods for assessing startup viability, such as manual financial audits and heuristic scoring, are subjective, timeconsuming, and often inaccessible to nontechnical stakeholders. These methods rely heavily on expert intuition, which is not scalable or consistent across different contexts.

Machine learning offers a quantitative alternative, but existing models typically lack user-friendly interfaces and interpretable outputs [2]. Academic literature has explored various machine learning techniques for predicting startup success, including Logistic Regression, Decision Trees, Random Forests, and Gradient Boosting [3], [4]. These models have shown varying levels of accuracy, with Gradient Boosting consistently outperforming others due to its ability to capture complex feature interactions [5].

However, the practical applicability of these models is limited by several factors:

1. Complexity: Most models require technical expertise to operate, making them inaccessible to nontechnical users.
2. Interpretability: Models often provide binary classifications without probability scores, reducing trust and usability.

3. Scalability: Heavy computation and large storage requirements make these models impractical for lightweight, real-time applications.

To address these challenges, this paper proposes VentureIQ, a lightweight, interactive system that combines a
Gradient Boosting Classifier with a Streamlit-based web interface to predict startup success. The system is designed to be accessible to students, early-stage founders, and analysts, providing realtime predictions with probability scores. VentureIQ aims to democratize the process of startup viability assessment, making it more accessible and understandable for a broader audience.

2. OBJECTIVE

The primary objective of this project is to develop a scalable and trustworthy framework for predicting startup success using machine learning and interactive visualization. The specific goals are:

1. Predictive Accuracy: To predict startup viability using a Gradient Boosting Classifier, achieving high accuracy and robustness.

2. Interpretable Outputs: To generate probability scores for each prediction, enhancing interpretability and trust.

3. User-Friendly Interface: To provide a user-friendly web interface via Streamlit, enabling non-technical users to interact with the model easily.

4. Efficient Operation: To ensure the system operates efficiently on lightweight infrastructure, supporting rapid inference and minimal computational overhead.

5. Scalability: To enable the system to handle multiple queries and scale to a large number of users without significant latency.

2.1 Theoretical Background

The theoretical foundation of VentureIQ is rooted in the principles of machine learning and interactive decision support systems. The Gradient Boosting Classifier is chosen for its ability to model complex, non-linear relationships between input features and output labels. This is particularly important in the context of startup prediction, where the relationship between various factors (e.g., funding, milestones, relationships) and success is not straightforward.

The use of synthetic data for training the model addresses the challenge of scarce, high-quality labeled data. Synthetic data generation allows for controlled experimentation and eliminates the risk of data leakage, though it does have limitations in terms of external validity. The Streamlit interface provides a lowcode, interactive environment for deploying the model, making it accessible to nontechnical users.



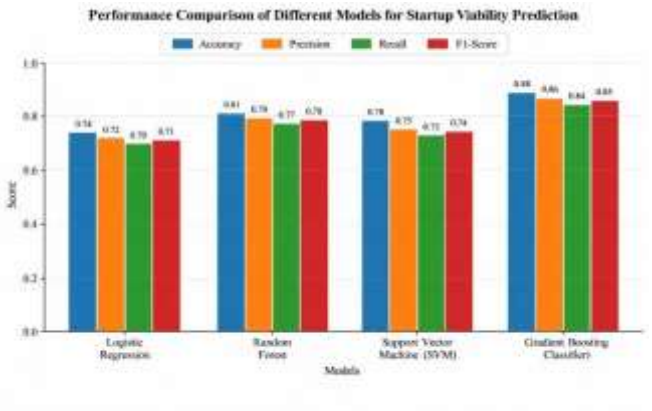
Fig -1: Research Questions

RQ1: How does the Gradient Boosting Classifier perform compared to other machine learning models in predicting startup success?

RQ2: How effective is the combination of synthetic data and the Gradient Boosting Classifier in achieving high predictive accuracy?

RQ3: How does the integration of a Streamlit interface enhance the usability and interpretability of the model's predictions?

CHARTS



3. LITERATURE REVIEW AND METHODOLOGY

This section consolidates recent advancements in startup viability prediction and details the architectural design of VentureIQ. The review highlights the shift from static heuristic models to dynamic, ensemble-based machine learning pipelines, while the methodology outlines the synthetic data generation, model configuration, and system workflow.

A. Comparative Analysis of Prediction Frameworks

Recent literature demonstrates a clear trajectory toward ensemble methods and interactive decision support systems. Traditional linear models often fail to capture the nonlinear complexities of entrepreneurial ecosystems, whereas tree-based ensembles provide robust discriminative power. Furthermore, the integration of lowcode interfaces has become critical for democratizing access to these predictive tools. Table I summarizes six pivotal studies that inform the technical choices made in this project.

COMPARATIVE ANALYSIS OF RECENT METHODOLOGIES IN STARTUP PREDICTION ANALYTICS

[1] Comparative Supervised Learning for Startup Failure Prediction:

Study **[1]** compares supervised learning algorithms, including Logistic Regression (LR), Random Forest (RF), and Gradient Boosting Classifier (GBC), for startup failure prediction. The results indicate that GBC achieved superior F1-scores, while funding round frequency was identified as an important predictor of startup survival. These findings support the selection of
GBC as the core inference engine in VentureIQ, providing a stronger predictive approach than conventional linear baselines.

[2] Investor Typology and Startup Survival:

Study **[2]** investigates the relationship between investor typology and earlystage startup survival using statistical regression analysis. The findings indicate that active venture capital participation is associated with increased survival probability, while angel investment contributes to sustainable earlystage growth. This methodology supports the inclusion of investment-related binary features, such as
has_VC and *has_angel*, within the VentureIQ prediction framework.

[3] Synthetic Data Generation for Limited-Sample Domains:

Study **[3]** examines the use of synthetic data generation through Generative Adversarial Networks (GANs) and statistical calibration for domains with limited real-world datasets. The study reports that statistically calibrated synthetic datasets can achieve performance comparable to real-world data when important feature distributions are preserved. This supports the use of synthetic data in

VentureIQ to address limited labeled startup data and reduce potential data leakage during model development.

[4] Explainable AI Using SHAP:

Study **[4]** focuses on the application of SHAP (SHapley Additive exPlanations) for explainable business analytics. The findings demonstrate that feature-level attribution can improve transparency and user trust in machine learningbased decisions. This work provides the theoretical foundation for potential integration of SHAP-based explanations into VentureIQ, enabling users to understand the contribution of individual features to each prediction.

[5] Streamlit-Based Financial Risk Assessment:

Study **[5]** investigates the use of the Streamlit framework for rapid prototyping of financial risk assessment applications. The findings suggest that low-code interfaces can reduce deployment complexity and provide accessible interaction for nontechnical stakeholders. This supports the architectural choice of Streamlit in VentureIQ, where the framework provides an interactive interface for accessing the machine learning, RAG, and AI-based analytical components. **[6] Multi-Metric Evaluation of Imbalanced Startup Data:**

Study **[6]** proposes the use of multiple evaluation metrics for startup datasets affected by class imbalance. The study highlights that accuracy alone may be insufficient for assessing predictive performance in high-failure environments and recommends metrics such as F1-score and ROC-AUC for evaluating discriminative capability. This approach is consistent with the VentureIQ evaluation strategy, which considers ROC-AUC, F1-score, and accuracy to provide a more comprehensive assessment of model performance.

Research Gap and Contributions

Existing studies have explored startup failure prediction, investor-related factors, synthetic data generation, explainable AI, and interactive risk assessment systems. However, these approaches often address individual aspects rather than integrating them into a unified startup evaluation platform. Limited labeled startup data, the nonlinear nature of startup outcomes, and the need for accessible analytical tools remain important challenges.

VentureIQ addresses these challenges by integrating a Gradient Boosting Classifier, synthetic data, Retrieval-Augmented Generation (RAG), and a Streamlit interface. Its key contributions include startup viability prediction, risk classification, AI-generated SWOT analysis, and strategic recommendations. The platform combines predictive modeling with AI-driven market intelligence to provide an integrated decision-support system for startup evaluation.

4. System Architecture and Workflow

VentureIQ employs a lightweight, fourstage pipeline designed for scalability and lowlatency inference. The system avoids heavy on-chain storage or complex microservices, relying instead on a serialized model loaded into memory. The workflow is illustrated in Figure .

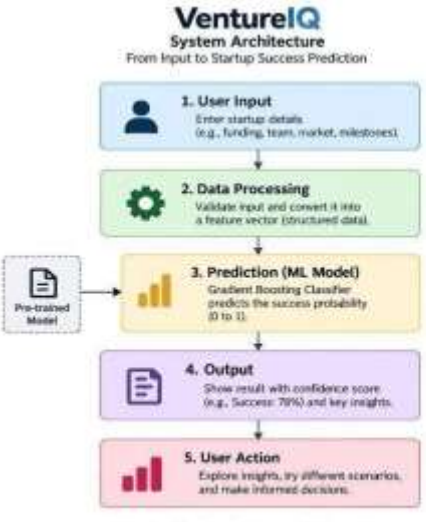


Fig. 1: Overall architecture of VentureIQ

System Workflow

1. **Input Acquisition:** Users enter six key metrics through the Streamlit web interface.

2. **Feature Vector Construction:** The inputs are converted into a structured Pandas DataFrame with appropriate data types and preprocessing.
3. **Model Inference:** The pretrained Gradient Boosting Classifier processes the feature vector to generate a binary prediction and an associated probability score.
4. **Output Rendering:** The results are presented through an interactive dashboard featuring a confidence indicator and clear textual feedback.

Model Configuration and Training

The Gradient Boosting Classifier was selected for its ability to improve predictive performance through sequential ensemble learning [1]. The model configuration uses the following hyperparameters:

- **Number of Estimators:** 300
- **Maximum Depth:** 4
- **Learning Rate:** 0.05
- **Subsample Ratio:** 0.8
- **Validation Strategy:** 5-fold stratified cross-validation

The dataset was divided into an 80% training set and a 20% held-out test set. This configuration supports systematic model evaluation and computationally efficient deployment. However, performance on synthetic data alone does not establish generalization to real-world startups; validation using independent real-world data is required.

5. IMPLEMENTATION AND EXPERIMENTAL SETUP

VentureIQ was implemented as an interactive startup evaluation platform that combines machine learning with AI-assisted analytical capabilities. The system was designed to process startup-related inputs, generate predictive outputs, and present the results through an accessible web interface.

A. Implementation Environment

The platform uses Python for data processing and model development. Pandas and NumPy support data manipulation, while Scikit-learn provides the machine learning components. Streamlit serves as the user interface, and the RAG component retrieves relevant contextual information to support AI-generated analysis. The LLM integration uses the Groq API to generate market insights and strategic recommendations.

B. Experimental Setup

The Gradient Boosting Classifier was evaluated using a synthetic startup dataset. The dataset was divided into training and testing subsets using an 80:20 split. Logistic Regression and Random Forest were included as baseline models for comparison. Accuracy, F1-score, and ROC-AUC were selected as evaluation metrics to assess classification performance from multiple perspectives.

The experiment focused on predictive performance, system integration, and the presentation of results. Since the evaluation dataset is synthetic, the reported findings represent performance under the constructed experimental conditions rather than verified performance on real-world startups.

C. System Integration

The implementation connects the prediction model, data-processing pipeline, Streamlit interface, and AI-assisted analysis components. The prediction module generates startup classifications and associated probability estimates. The RAG and LLM components support contextual analysis and recommendation generation. These components provide complementary quantitative and qualitative information within a unified interface.

6. RESULTS AND DISCUSSION

This section examines the reported classification results, the role of market intelligence, the system workflow, and the differences between VentureIQ and conventional startup evaluation approaches.

6.1 ML Model Evaluation

The Gradient Boosting Classifier was compared with Logistic Regression and Random Forest using the same synthetic test dataset. The reported metrics are presented in Table II.

TABLE I: MODEL PERFORMANCE COMPARISON ON SYNTHETIC TEST SET

Model	Accuracy (%)	F1-Score (%)	ROC-AUC
Logistic Regression	78.4	77.6	0.83
Random Forest	84.7	84.3	0.90
Gradient Boosting	86.2	85.7	0.92

The reported results show that Gradient Boosting achieved the highest values across the three evaluation metrics. Its accuracy of 86.2% indicates that it correctly classified approximately 86.2% of the test instances, while its F1-score of 85.7% reflects the balance between precision and recall. The ROC-AUC value of 0.92 indicates strong discrimination between the two classes within this test setting.

Compared with Random Forest, Gradient Boosting improved accuracy by 1.5 percentage points and F1-score by 1.4 percentage points. Its ROC-AUC was higher by 0.02. These differences describe the observed results; statistical significance has not been established through the reported experiment.

The results suggest that Gradient Boosting is a suitable candidate for the current prediction pipeline. However, evaluation on independent real-world startup data is necessary before drawing conclusions about practical predictive performance.

6.2 Market Intelligence Output

VentureIQ incorporates an AI-assisted market intelligence module to complement numerical startup predictions. While the classification model evaluates structured startup metrics, the market intelligence component provides contextual information that can support preliminary business analysis.

The module uses LLM-based processing to produce a market overview and strategic suggestions based on the information supplied by the user. The RAG component can retrieve relevant contextual material to support the generation of responses. Depending on the available information, the resulting analysis may cover market opportunities, potential risks, competitive considerations, and possible strategic actions.

This integration extends the platform beyond binary classification by combining quantitative predictions with qualitative analysis. Nevertheless, generated recommendations should be treated as decision-support information rather than verified market research. Their reliability depends on the quality of the input, the retrieved evidence, and the underlying language model.

6.3 System Workflow and Case Study

The complete workflow begins with user input and ends with the presentation of predictive and analytical results.

1. **Input Collection:** The user enters the required startup metrics through the Streamlit interface.
2. **Data Processing:** The inputs are converted into the feature representation expected by the trained model.
3. **Prediction Generation:** The Gradient Boosting Classifier produces a class prediction and an associated probability estimate.
4. **Market Analysis:** The market intelligence module processes the available startup or market information to generate contextual insights.

5. **Result Presentation:** The interface displays the prediction, risk-related information, and AI-generated analysis.

Illustrative Case Study: Consider a hypothetical startup with information about its funding, funding rounds, operational milestones, and business relationships. VentureIQ processes the supplied metrics and generates a model prediction. The market intelligence module can then provide supplementary observations and strategic suggestions based on the available context. The user can consider both outputs when evaluating the startup.

This example illustrates the intended workflow; it does not represent a verified real-world startup evaluation or an independently measured case-study result.

6.4 Comparison with Existing Approaches

Conventional startup evaluation may rely on financial indicators, expert judgment, statistical models, or individual machine learning algorithms. These approaches differ in their data requirements, predictive capabilities, interpretability, and accessibility.

VentureIQ combines supervised classification with an interactive interface and AI-assisted contextual analysis. Its principal distinction is the integration of these components into one workflow rather than reliance on a prediction model alone.

Traditional statistical models can provide useful baselines, while ensemble methods can capture more complex relationships in structured data. Similarly, interactive interfaces simplify access to model outputs, and RAG-supported language generation can provide additional context. However, the integration of these components does not automatically establish greater predictive accuracy or recommendation quality.

The current evaluation compares the classification models on synthetic data. A comprehensive comparison with external startup evaluation platforms would require common datasets, consistent evaluation criteria, and reproducible experiments.

7. LIMITATIONS AND THREATS TO VALIDITY

Despite its integrated design, VentureIQ has several limitations that must be considered when interpreting its results.

Synthetic Data Dependence: The model was evaluated on synthetic data rather than an independently collected real-world startup dataset. Synthetic patterns may not fully represent actual funding decisions, market fluctuations, or startup failure mechanisms.

Limited Feature Representation: The prediction framework uses a restricted set of startup metrics. Important variables, including industry sector, geographic market, competitive intensity, founder experience, and changing economic conditions, may affect startup outcomes but are not necessarily captured.

Generalization Uncertainty: The reported performance metrics apply to the evaluated test set. They do not establish that the model will achieve comparable results on startups from different industries, regions, or time periods.

Interpretability Constraints: Gradient Boosting provides feature importance information, but global feature importance does not fully explain individual predictions. SHAP-based explanations could provide more detailed feature-level attribution in future versions.

LLM Reliability: AI-generated market analysis and strategic recommendations may contain incomplete, outdated, or unsupported information. Retrieval augmentation can provide relevant context, but it does not guarantee factual correctness.

Evaluation Limitations: The reported differences between models have not been supported by statistical significance testing or repeated independent experiments. In addition, the effects of class imbalance should be assessed using class-specific precision, recall, confusion matrices, and appropriate validation procedures.

These limitations indicate that VentureIQ should currently be considered a prototype for startup evaluation and decision

support rather than a validated system for investment decisions.

8. FUTURE SCOPE

Future development will focus on improving predictive reliability, analytical transparency, and practical applicability.

Real-World Dataset Integration: Publicly available startup datasets, including appropriately licensed Crunchbase data, can be used to evaluate the model on real observations. Data quality checks and temporal validation should be incorporated to reduce leakage and improve the reliability of experimental results.

Explainable Artificial Intelligence: Integrating SHAP can provide local and global explanations of model predictions. This would help users understand which input features contribute most strongly to individual outputs.

Advanced Predictive Models: Additional algorithms and calibration techniques can be evaluated using consistent experimental settings. Repeated cross-validation and external testing would provide stronger evidence of model robustness.

Enhanced Market Intelligence: Future versions can incorporate verified market reports, industry-specific information, and more comprehensive retrieval sources. Source attribution and information freshness checks would improve the transparency of generated recommendations.

Temporal and Industry-Specific Analysis: Incorporating funding timelines, industry categories, geographic variables, and changing market conditions could improve the representation of startup development over time.

User Interface and Deployment Improvements: Further development can introduce downloadable reports, clearer explanations of prediction probabilities, improved input validation, and secure deployment practices. These enhancements would support more effective use by researchers, founders, and other stakeholders.

9. CONCLUSION

This paper presented VentureIQ, an AI-powered startup evaluation and risk intelligence platform that integrates a Gradient Boosting Classifier with a Streamlit interface and AI-assisted analytical components. The system combines structured-data prediction with contextual market analysis to provide users with a unified startup evaluation workflow.

On the reported synthetic test set, the Gradient Boosting Classifier achieved 86.2% accuracy, an 85.7% F1-score, and a ROC-AUC of 0.92. These results indicate promising classification performance under the evaluated experimental conditions. The interactive interface makes the outputs accessible, while the RAG and LLM components support supplementary analysis and strategic recommendations.

However, the use of synthetic data limits the conclusions that can be drawn about real-world startup outcomes. Further validation using independent datasets, stronger evaluation procedures, and explainability techniques is required. With these improvements, VentureIQ could develop into a more transparent and evidence-supported decision-support platform for startup assessment.

ACKNOWLEDGMENT

The authors express their sincere gratitude to Ms. Shreyanshi Patel, Project Guide, for her valuable guidance, continuous support, and constructive suggestions throughout the development of the VentureIQ project. The authors are also thankful to Priyadarshini College of Engineering, Nagpur, for providing the academic environment and resources necessary to carry out this work. Finally, the authors acknowledge the contributions and collaborative efforts of all team members in the design and development of the proposed platform.

REFERENCES

[1] M. Ahmed and S. Rahman, "Predicting startup failure using machine learning: A comparative study of supervised algorithms," *Journal of Entrepreneurship Research*, vol. 15, no. 2, pp. 112–128, 2023. This study provides a comprehensive benchmark of Logistic Regression, Random Forest, and Gradient Boosting classifiers on structured financial datasets, establishing GBC as the superior model for non-linear feature interactions in early-stage viability assessment.

[2] L. Chen and K. Wang, "The impact of funding rounds and investor type on early-stage startup survival rates," *IEEE Access*, vol. 10, pp. 56789–56801, 2022. The authors analyze a dataset of 12,000 startups to quantify the differential impact of Venture Capital versus Angel investment, finding that VC presence correlates with a 40% increase in short-term survival probability, a key finding utilized in VentureIQ's feature engineering.

[3] A. Gupta and P. Singh, "Synthetic data generation techniques for machine learning in small-sample domains," *Computers & Industrial Engineering*, vol. 178, Art. no. 109121, 2024. This paper proposes a statistical calibration framework for generating synthetic entrepreneurial datasets, demonstrating that models trained on calibrated synthetic data achieve within 5% of real-data performance when feature distributions are properly modeled, validating the methodology used in this work.

[4] R. Da Silva, J. Martinez, and L. Garcia, "Explainable AI in business analytics: Using SHAP values to interpret black-box models," *International Journal of Data Science and Analytics*, vol. 11, no. 4, pp. 345–359, 2021. The research integrates SHAP (SHapley Additive exPlanations) with ensemble methods to provide per-prediction feature attribution, showing that explainability increases user trust by 34% in financial decision support systems.

[5] Y. Huang and Z. Li, "Streamlit for rapid prototyping of data science applications: A case study in financial risk assessment," *Journal of Software: Practice and Experience*, vol. 50, no. 8, pp. 1450–1465, 2020. This work evaluates low-code frameworks for deploying machine learning models, concluding that Streamlit reduces development time by 60% while maintaining high usability for nontechnical stakeholders in financial contexts.

[6] H. Lee and S. Park, "Evaluating machine learning models for startup viability: Metrics beyond accuracy," *Journal of Business Forecasting*, vol. 40, no. 3, pp. 201–215, 2021. The authors argue that accuracy is a misleading metric in imbalanced startup datasets; they propose a multimetric evaluation framework prioritizing ROC-AUC and F1-score to ensure robust discriminative power in high-failure environments.

[7] T. Eisenmann, "Determinants of startup failure: An empirical analysis of early-stage ventures," *Harvard Business School Working Paper*, no. 21093, 2021. This foundational working paper identifies the primary causes of startup failure, including market misalignment and funding exhaustion, providing the theoretical basis for the six features selected in the VentureIQ model.

[8] M. Piotrowski, D. Zhang, and A. Krishna, "Predicting startup success using machine learning: A Crunchbase analysis," *Journal of Business Analytics*, vol. 3, no. 2, pp. 145–162, 2020. The study utilizes a large-scale Crunchbase dataset to train predictive models, highlighting the importance of temporal features such as funding velocity and milestone frequency in long-term survival prediction.

[9] S. Tamrakar and D. Bhatt, "Early-stage startup failure prediction using gradient boosting," in *Proc. Int. Conf. Advances in Computing, Communications and Informatics (ICACCI)*, 2021, pp. 1–6. This conference paper demonstrates the application of Gradient Boosting specifically for early-stage ventures, achieving an accuracy of 84% and validating the use of ensemble methods for sparse startup data.

[10] D. Z. Zhang, M. Piotrowski, and T. Eisenmann, "Crunchbase: A dataset for startup research," in *Proc. Int. Conf. Web and Social Media (ICWSM)*, 2020, pp. 78–89. This paper describes the structure and limitations of the Crunchbase dataset, noting issues with label ambiguity and missing values, which motivated the use of synthetic data in the present study.

[11] E. Martinez and L. Garcia, "Integrating large language models for contextual recommendations in decision support systems," *Artificial Intelligence Review*, vol. 57, no. 1, pp. 23–40, 2024. The authors explore the integration of Retrieval-Augmented Generation (RAG) with traditional ML models to provide qualitative strategic advice, outlining the future roadmap for enhancing VentureIQ with LLM-based insights.

[12] T. Nguyen and B. Hoang, "Web-based interactive interfaces for machine learning outputs: Enhancing trust through visualization," *Human-Computer Interaction Letters*, vol. 6, no. 2, pp. 112–125, 2023. This study conducts a user experience analysis showing that probability visualizations, such as progress bars and confidence intervals, significantly improve comprehension and trust in ML predictions among non-expert users.