International Journal of Interdisciplinary Innovative Research & Development (IJIIRD)
ISSN: 2456-236X
Vol. 11 Special Issue 03 | 2026

Idea Feasibility Checker and Market Predictor

Miss. Sonali J. Kalyankar[1], Miss. Prachi G. Parab[2], Mr. Nikhil P. Patil[3], Miss. Sanika S. Sawant[4], Sagar V. Harne[5]
Department of Computer Engineering, Sindhudurg Shikshan Prasarak Mandal's C.O.E. Kankavli, Maharashtra, India 416602 University Of Mumbai

DOI: 10.5281/zenodo.20631132

ABSTRACT

Startup businesses experience a large number of failures due to uncertain market conditions, competitive forces, and a lack of strategic planning. Conventional approaches for predicting the success of startups involve machine learning models trained on historical datasets that are not adaptable in real-time. This paper presents a Real-Time AI-Based Startup Evaluation Framework that combines Google Gemini API and Cloud Firestore. The proposed framework extracts real-time competitor and market information from a cloud-based NoSQL database and uses structured prompt engineering to produce context-specific analytical reports. Unlike conventional machine learning models, the proposed framework focuses on dynamic data integration and generative AI reasoning for evaluating startups.

Keyword: - Startup Evaluation, Generative AI, Gemini API, Firestore, Real-Time Data, Decision Support System.

1. INTRODUCTION

Startup ecosystems are important for innovation, technology development, and economic development around the globe. Startups bring innovative solutions, create job opportunities, and help with the digitalization of various sectors like fintech, health-tech, software, and e-commerce. However, despite their potential, startups face highly uncertain and competitive environments, and even survival is not guaranteed. Industry studies show that a large number of startups fail in their initial operational phases because of a lack of market demand, financial instability, poor strategic planning, and high competition. The conventional approach to evaluating startups involves financial analysis, human expert opinion, and statistical forecasting techniques. However, with the advent of data science, various machine learning algorithms like Logistic Regression, Decision Trees, Random Forest, Support Vector Machines, and Neural Networks have been extensively investigated for predicting the success of startups. These algorithms process structured data sets with financial information, entrepreneur information, funding phases, and growth factors to predict which startups are successful and which are not. Although these approaches provide precise predictive results, they are heavily reliant on static structured data sets and need to be retrained periodically to adjust to the dynamic market environment.

In actual startup ecosystems, market trends, competitor valuations, funding scenarios, and industry growth rates keep fluctuating. Static ML-based systems face challenges in adapting to real-time contextual changes without complex retraining processes. Additionally, conventional classification models can predict probabilities but fail to offer in-depth contextual reasoning and interpretability necessary for strategic decisionmaking. Recent breakthroughs in Large Language Models (LLMs) have opened up new avenues in contextual reasoning, knowledge compilation, and analytical report writing. Generative AI models such as Google Gemini are equipped with the ability to process structured inputs and generate in-depth insights based on domain knowledge. Unlike conventional ML classifiers, LLMs focus on reasoning-driven output generation, allowing for qualitative analysis in addition to quantitative analysis. With the above limitations and opportunities in mind, this research work aims to develop a Real-Time AI-Based Startup Evaluation Framework by combining Google Gemini API with Cloud Firestore. The aim is to design a scalable, adaptive, and interpretable decision-support system that can analyze startup concepts based on dynamic competitor and market information.

2. RELATED WORK

In this section, we concentrate on the discussion of various Artificial Intelligence and Machine Learning techniques used for the prediction of success in startups and the forecasting of business performance. The existing literature on the topic can be broadly classified into supervised learning, deep learning, natural language processing (NLP), graph-based learning, and AI-based decision support systems. Based on the availability of labeled outcome variables like "success," "failure," "bankruptcy," or "growth," the models are generally classified into classification and regression models. Supervised learning models can be further categorized into binary classification, multi-class classification, and continuous outcome prediction models. Besides, unsupervised and hybrid models of AI have been extensively used to discover hidden patterns and enhance business intelligence systems.

2.1 Supervised Learning For Startup Success Prediction

Supervised learning models are commonly employed for binary classification tasks in startup environments, aiming to separate successful startups from unsuccessful ones. Unal and ¨ Ceasu (2019) [1] created Random Forest and Logistic Regression models to predict startup success, with 85% accuracy. The authors concluded that financial variables and founder related variables are robust predictors. Nevertheless, the authors stressed the need for integrating real-time financial and investor sentiment variables to improve prediction adaptability.

Likewise, Kumar et al. (2022) [8] used Logistic Regression and Random Forest models to predict startup performance, with about 85% accuracy. The authors' results underscored the need to combine internal startup variables with external market information for better predictive results. Choi (2024) [2] proposed the use of Principal Component Analysis (PCA) together with machine learning classifiers to improve model interpretability and eliminate feature redundancy in startup data. PCA helped maintain predictive accuracy while improving model interpretability. Future studies should explore the application of PCA-improved models in large-scale, multi-national startup ecosystems. Regression-based forecasting has also been used in business analysis. Patel et al. (2021) [10] used Linear Regression and ARIMA models for financial forecasting and early risk identification. Their study showed that financial time-series variables play a significant role in the prediction of the sustainability of start-ups. But they recommended the use of non-financial variables for a holistic feasibility analysis. Ensemble learning models have been found to provide better predictions. Garcia and Huang (2019) [12] designed an ensemble model based on XGBoost and stacking techniques for market demand forecasting, which outperformed statistical models. Their study recommended the use of transfer learning for model adaptation in new markets.

2.2 Deep Learning And Neural Network Approaches

Deep learning models have emerged as popular tools for modeling complex nonlinear relationships in business ecosystems. Potanin et al. (2023) [5] used Deep Neural Networks to optimize venture capital portfolios, achieving better risk adjusted returns than conventional financial models. The results indicate that deep learning models are able to model complex investment and performance dynamics. Kaminski et al. (2020) [3] used neural networks integrated with Natural Language Processing to forecast the success of crowdfunding campaigns. The results showed that sentiment and tone of text have a substantial impact on funding success, establishing the importance of linguistic variables in startup analysis. The authors proposed transformer models for NLP to improve contextual understanding in future studies. Recent transformer models for NLP (2024) [6] have achieved 90% accuracy on company classification problems. These results confirm the power of large language models in understanding business descriptions and contextual scenarios. Nevertheless, explainability, bias, and fairness are still pending issues

2.3 Natural Language Processing In Startup Evaluation

NLP tools are being increasingly employed for the analysis of startup descriptions, investor reports, and market sentiments. Lee and Cho (2023) [9] used NLP techniques to identify the business domain and possible market segments from the descriptions of startup concepts. The study proved the feasibility of automated text analysis for validating ideas at the nascent stage. Rao et al. (2020) [15] conducted sentiment analysis of social media and news articles to determine market receptiveness. The study concluded that positive market sentiments are a precursor to short-term startup success. Future studies can explore the causal link between media trends and the long-term success of startups

2.4 Graph-Based And Relationship-Oriented Models

The relationships in the network between founders, advisors, and investors play an important role in the success of startups. Nguyen (2022) [14] introduced graph-based neural networks for the representation of entrepreneurial ecosystems. The work showed that there is a strong relationship between network centrality and fundraising success. Modeling temporal graphs for dynamic representation of startup ecosystems was proposed as future work.

2.5 Hybrid And AI-Driven Decision Support Systems

In addition to standalone ML models, hybrid AI models that leverage expert knowledge and predictions from machine learning models have been suggested. Muller et al. (2021) [13] ¨ proposed a hybrid expert-AI evaluation model that combined expert evaluation and machine learning predictions, enhancing the accuracy of early-stage evaluations. Silva and Torres (2022) [7] performed a comparative study of AI-driven business intelligence systems and found enhancements in decision accuracy and forecasting. The study underscored the increasing role of AI-driven decision support systems in strategic planning. Sharma and Mehta (2020) [11] investigated data mining approaches for entrepreneurial decision-making and investment forecasting, uncovering hidden patterns that inform startup evaluation. The authors focused on developing dynamic systems that adapt to real-time business dynamics. Pasayat (2022) [4] suggested a systematic feature engineering approach, underlining the role of team experience and investment ratios as key predictors of startup success. The study suggested the incorporation of qualitative innovation metrics and market sentiment features to enhance robustness in predictions.

2.6 Research Gap Identification

Despite the presence of extensive research work on the high accuracy of predictions using supervised learning, deep learning, and NLP models, certain gaps in the research work are apparent:

- Heavy dependence on static historical data corpora
- Less emphasis on real-time competitor intelligence incorporation
- Limited contextual reasoning in prediction results
- Less focus on integrated real-time AI-driven decision support systems

2.7 Positioning of the Proposed Work

To overcome the above-mentioned drawbacks, the proposed work incorporates:

- Real-time Cloud Firestore database for dynamic competitor intelligence
- Structured prompt engineering
- Google Gemini Large Language Model for contextual reasoning
- Analytical report generation for decision support

Unlike conventional supervised classification models, the proposed work focuses on real-time adaptability, contextual understanding, and explainable AI-driven analysis for startup assessment.

3. PROPOSED METHODOLOGY

The proposed system will be based on a multi-layered architecture that integrates cloud-based data storage with generative AI analysis. The methodology will be broken down into the following steps:

3.1 Data Collection

The user-inputted startup parameters, including industry type, funding level, target market, and product description, will be collected using an input interface. At the same time, real-time competitor and market data will be obtained from Cloud Firestore.

3.2 Real-Time Database Integration

Cloud Fire store will be used as a NoSQL real-time database to store competitor valuations, funding rounds, industry growth data, and market data. Firestore will enable dynamic querying and real-time updates, ensuring that the analysis is based on up-to-date market information.

3.3 Data Retrieval and Preprocessing

The collected data will be structured into input parameters suitable for prompt construction. Unlike traditional ML preprocessing, which includes normalization and feature encoding, this step will concentrate on contextual structuring.

3.4 Prompt Engineering

Layer Structured prompts are developed by integrating startup parameters with competitor insights in real-time. The prompt engineering layer ensures that the Gemini model is provided with all the necessary contextual information required for analytical reasoning.

3.5 AI Analysis Using Gemini API

The structured prompt is passed to the Gemini API, which analyzes the context and produces the following:

- Risk Assessment
- Competitive Positioning Analysis
- Success Likelihood Insight
- Strategic Recommendations

3.6 Output and Decision Support

The result is provided to entrepreneurs and investors as a decision-support document.

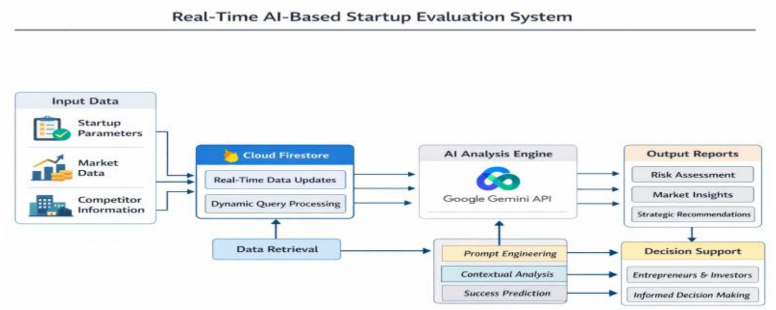

Fig.: Proposed System

4. SYSTEM ACHITECTURE

The proposed system architecture is based on a vertical data flow paradigm that includes the following layers:
• Input Layer
• Real-Time Database Layer (Firestore)
• Data Retrieval Module
• Prompt Engineering Layer
• Gemini AI Analysis Engine
• Output & Decision Support Layer
The combination of real-time databases and generative AI supports dynamic contextual reasoning, making the proposed system different from traditional ML-based systems that lack dynamic contextual reasoning capabilities.

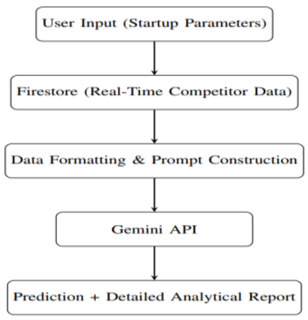

Fig.2 . System Architecture Flowchart

5. IMPLEMENTATION

The system was implemented using:
• Backend: Python-based server integration
• Database: Cloud Firestore (NoSQL, real-time updates)
• AI Engine: Google Gemini API
• Prompt Strategy: Structured contextual prompt engineering.
Firestore enables real-time synchronization and scalable data storage. The Gemini API processes structured prompts and generates analytical reports in natural language format.

6. RESULTS AND DISCUSSION

The system successfully generated contextual evaluation reports based on dynamic competitor data. Compared to traditional ML classifiers, the proposed framework demonstrated superior interpretability and adaptability. The qualitative nature of generative outputs enhances usability for strategic planning purposes.
Although quantitative accuracy benchmarking was not performed in the current phase, the system lays the foundation for hybrid ML-LLM architectures in future research.

7. CONCLUSION

This research presents a real-time AI-driven startup evaluation framework integrating Cloud Firestore with Google Gemini API. By combining dynamic database retrieval with contextual generative reasoning, the system addresses limitations of traditional static machine learning approaches. The framework enhances adaptability, interpretability, and decision-support capability for startup ecosystems.Future work includes quantitative benchmarking against supervised ML models and exploring hybrid predictive frameworks.

8. REFERENCES

[1] Unal and O. Ceasu. Machine learning classification models for startup ¨ success predict- Tion. International Review of Technology Management, 2019. Achieved 85% accuracy differentiating successful startups.
[2] Choi. Pca-enhanced machine learning models for startup success analysis. Journal of Emerging Technology in Business, 2024. Applied PCA to reduce feature redundancy.
[3] Kaminski and Others. Neural network and nlp-based analysis of crowdfunding campaigns. Entrepreneurship Research Journal, 2020.
[4] K. Pasayat. Feature engineering and predictive modeling for startup evaluation. Applied Intelligence Review, 2022.
[5] Potanin and Colleagues. Deep neural networks for venture capital portfolio optimiza- tion. AI in Finance Journal, 2023.
[6] Rizinski and Team. Transformer-based nlp for company classification. Information Journal, 15(2):77–88, 2024. [7] L. Silva and M. Torres. Comparative analysis of ai-driven business intelligence systems. In International Conference on Data Analytics, 2022.
[8] N. Kumar and Coauthors. ML classification for startup success prediction. Applied AI Journal, 2022.
[9] J. Lee and K. Cho. Nlp for startup idea analysis. Computational Entrepreneurship Review, 2023.
[10] S. Patel and Team. Ai-driven financial forecasting using arima models. Finance and AI Journal, 2021.
[11] R. Sharma and D. Mehta. Data mining techniques for entrepreneurial decision-support. Technical report, Institute of Business Analytics, 2020.
[12] L. Garcia and S. Huang. Ensemble ml for market demand forecasting. International Journal of Forecasting, 2019.
[13] R. Muller and Others. Hybrid expert–ml models for early-stage evaluation. Decision Support Systems, 2021. [14] T. Nguyen. Graph neural network models for founder–investor networks. Network Science in Business, 2022. [15] P. Rao et al. Sentiment analysis of media for startup traction. Journal of Media Economics, 2020.
[16] H. Kim and J. Park. Automl pipelines for startup prediction. Automated Intelligence Journal, 2023.
[17] M. Oliveira et al. Explainable ai for investor decision models. AI and Society, 2024.
[18] G. Smith et al. Survival analysis for startup longevity. Journal of Business Analytics, 2018.