International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

A Comparative Review of Large Language Model Approaches for Startup Success Prediction and Market Research

Noel T Jomon, Raihan Tom, Sion Jose, Vishnu Lal T,
Athira Prakash, Dr. Sabeena K, Chinchu M Pillai
Department of Computer Science
College of Engineering Chengannur
Kerala, India

***Abstract* - The evaluation and planning of new business ventures remain a complex and high-risk endeavor, traditionally dependent on human expertise. The recent proliferation of Large Language Models (LLMs) has catalyzed a paradigm shift, introducing data-driven methodologies for startup analysis, market research, and strategic planning. This paper presents a comparative review of the state-of-the-art in LLM applications for the entrepreneurial domain. We synthesize findings from recent studies, categorizing them into three primary areas: (1) predictive modeling for startup success and funding, (2) simulation of market dynamics and customer behavior, and (3) the development of advanced architectures like Retrieval-Augmented Generation (RAG) and multi-agent systems to enhance reliability and capability. Our review identifies a clear trajectory from single-model analytical tools toward sophisticated, generative, and agentic frameworks. While significant progress has been made in prediction and guidance, we identify a critical research gap: the absence of systems capable of autonomously generating comprehensive, actionable business plans from a nascent idea. We conclude by positioning the development of such generative, multi-agent systems as the next logical frontier in applying agentic AI to the real-world challenge of business creation.**

Index Terms - Large Language Models, Startup, Business Plan, Agentic AI, Retrieval-Augmented Generation (RAG), Market Research, Multi-Agent Systems

I. INTRODUCTION

Entrepreneurial ventures are crucial engines of economic growth and innovation, yet they face notoriously high failure rates. A primary challenge in the early stages of a startup is the immense uncertainty surrounding its business model, market viability, and potential for success [1], [3]. Traditionally, navigating this uncertainty has relied on the intuition, experience, and pattern-matching abilities of seasoned entrepreneurs and venture capital (VC) investors. This human-centric approach, however, can be time-consuming, costly, and susceptible to cognitive biases such as overconfidence and anchoring, which can skew judgment and lead to suboptimal investment decisions [8]. In 2021 alone, angel investors in the United States invested approximately $29.5 billion across over 69,000 ventures, highlighting the scale of capital allocated based on these high-stakes evaluations [8].

The emergence of Large Language Models (LLMs) and generative Artificial Intelligence (AI) has introduced a transformative potential to this domain [10]. By their ability to process vast amounts of unstructured data, recognize complex patterns, and generate human-like text, LLMs offer a powerful toolkit for augmenting and automating many aspects of startup analysis and business planning [4]. Researchers and practitioners are rapidly exploring the application of these models to create more objective, scalable, and data-driven methods for evaluating and supporting new ventures. The goal is to move beyond subjective "gut feelings" and toward structured, evidence-based decision-making frameworks that can be deployed at scale.

This paper provides a comprehensive literature review of the recent advancements in applying LLMs to the entrepreneurial landscape. The structure is as follows: Section II outlines the methodology used for this review. Section III presents a taxonomy of the analyzed literature. Section IV delves into a thematic analysis of the literature, segmented by application area. Section V synthesizes these findings to identify a key research gap. Section VI discusses the broader implications, challenges, and future directions for the field. Finally, Section VII concludes the review. Through this synthesis, we chart a clear progression from simple analytical tools to complex, autonomous agentic systems, highlighting an opportunity for a new class of generative tools designed not just to evaluate ideas, but to help build them.

II. METHODOLOGY

This paper conducts a systematic literature review to synthesize the state-of-the-art in LLM applications within the entrepreneurial and startup context. The research methodology involved a multi-stage process to identify, select, and analyze relevant scholarly work. An initial search was conducted on academic databases and preprint archives (e.g., arXiv, IEEE Xplore) using keywords such as "Large Language Model," "startup," "venture capital," "RAG," "agentic AI," and "market research."

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

From the initial pool of literature, ten representative papers published or set to be published in 2024 and 2025 were selected for in-depth analysis. This selection was based on their relevance to the core topics of startup prediction, market simulation, and advanced AI architectures, as well as their methodological novelty. A thematic analysis approach was employed to categorize the selected papers. This involved identifying recurring concepts, methodologies, and objectives across the literature, leading to the formation of the three primary themes discussed in Section IV: (1) Prediction and Evaluation, (2) Simulation and Ideation, and (3) Advanced System Architectures. Each paper was analyzed for its core contribution, methodology, data sources, and identified limitations to build a comprehensive and critical overview of the field.

III. THEMATIC ANALYSIS OF THE LITERATURE

The application of LLMs in the startup ecosystem has evolved rapidly. This section provides a detailed analysis of the literature, organized by the primary themes identified during our review. A summary of these approaches is presented in Table I.

A. *Predicting Startup Success and Funding*

A significant body of research has focused on leveraging LLMs for the predictive task of identifying which startups will succeed or secure funding.

1) Data Sources and Feature Types: Researchers have demonstrated the versatility of LLMs by applying them to various data types. Maarouf et al. (2024) utilized data from Crunchbase, a comprehensive VC platform, combining structured fundamental variables (e.g., funding history, founder count) with unstructured textual self-descriptions [1]. Katcharovski and Maxwell (2024) focused on verbal communication, analyzing full transcripts of startup pitches from the television show “Shark Tank” to assess persuasive cues and strategic content [8]. Meanwhile, Wang et al. (2025) curated a dataset from founder LinkedIn profiles and Crunchbase to build a detailed picture of the founding team’s experience and background, which they found to be a highly predictive element [3].

2) Modeling Techniques and Key Findings: The modeling techniques employed are as varied as the data sources. Maarouf et al. (2024) developed a “fused” architecture where BERT embeddings of the textual descriptions were concatenated with fundamental variables and fed into a neural network. This approach yielded a significant increase in predictive power (AUROC of 82.78%) compared to using fundamentals alone (80.60%) [1].

In contrast, Katcharovski and Maxwell (2024) used an LLM not for direct prediction, but for automated *feature synthesis*. Their model operationalized the established “Critical Factor Assessment” (CFA) framework by scoring verbal pitches across eight dimensions. These generated scores were then used as input for traditional machine learning models like Naive Bayes, which achieved a predictive accuracy of 79% [8].

Pushing the architectural envelope, Wang et al. (2025) created the Startup Success Forecasting Framework (SSFF), a multi-agent system that combines LLMs with traditional models like Random Forests within a Retrieval-Augmented Generation (RAG) framework. Their work revealed that baseline LLMs often suffer from an “over-prediction bias,” and that a multi-agent “divide and conquer” approach significantly improves reliability, achieving a 30.8% relative improvement over GPT-4o [3].

3) Critique and Limitations: While powerful, these predictive models have limitations. They rely on historical data that may not capture future market shifts. Models trained on public data, such as “Shark Tank” pitches, are acknowledged by their authors to be potentially unrepresentative of real-world, confidential investor interactions due to editing for entertainment [8]. Furthermore, many studies, including that of Wang et al. (2025), are based on relatively small evaluation samples (e.g., 50 startups), indicating that further validation is necessary to confirm their robustness [3].

B. *Simulating Market Dynamics and Customer Behavior*

A newer but equally promising application of LLMs is in simulating market dynamics to aid in early-stage ideation and customer discovery.

1) Estimating Economic Metrics: Brand et al. (2024) conducted extensive experiments showing that LLMs can generate realistic estimates for consumer Willingness-To-Pay (WTP) using simulated conjoint analysis [4]. A key finding was that fine-tuning an LLM with existing human survey data improved its WTP estimates for *new features* on existing products. This suggests LLMs can learn underlying preference structures and extrapolate to novel-but-related concepts.

2) Validating Business Hypotheses: Ilagan et al. (2024) frame this simulation capability as an “exploratory customer discovery” tool, aligning LLM use with Lean Startup principles [2]. They used ChatGPT to create synthetic customer personas and simulated their reactions to a hypothetical business idea. This positions the LLM as a “co-pilot” for entrepreneurs, enabling rapid iteration on business model hypotheses before engaging in costly real-world experiments. The vision is for LLMs to become integral assistants in specialized domains, such as international business, where they can be trained on country-specific data to help navigate local cultures and regulations [10].

3) Critique and Limitations: The validity of LLMs as “simulated humans” is still an open question. Brand et al. (2024) themselves found that while GPT could simulate an average consumer, it failed to meaningfully capture the nuanced preferences of different demographic segments, a significant limitation for targeted market research [4]. Similarly, Ilagan et al. (2024) are careful to position their tool as an *exploratory* co-pilot, not a replacement for real customer interaction, acknowledging the simulation’s limits [2]. The risk is that

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

TABLE I
TAXONOMY OF RECENT LLM APPROACHES IN ENTREPRENEURSHIP AND MARKET RESEARCH

Paper	Primary Task	Core Methodology	Data Source	Key Contribution
Maarouf et al. [1]	Prediction	Fused LLM (BERT) with Neural Network	Crunchbase data	Demonstrates that textual self-descriptions significantly improve startup success prediction accuracy over fundamental data alone.
Katcharovski & Maxwell [8]	Prediction	LLM for Feature Synthesis + ML	"Shark Tank" transcripts	Automates the Critical Factor Assessment (CFA) framework, providing a scalable method to predict funding outcomes from verbal pitches.
Wang et al. [3]	Prediction	Multi-Agent RAG Framework	LinkedIn profiles	Introduces the SSFF framework that mimics VC analysis, mitigates LLM over-prediction bias, and shows a 30.8% improvement over GPT-4o.
Brand et al. [4]	Simulation	Conjoint Analysis Simulation	Simulated survey data	Shows LLMs can realistically estimate consumer Willingness-To-Pay (WTP) and that fine-tuning improves predictions for new features.
Ilagan et al. [2]	Ideation	Prompt Engineering for Persona Simulation	Simulated interactions	Proposes LLMs as a low-cost "co-pilot" for entrepreneurs to perform customer discovery and test business hypotheses by simulating personas.
Sunde et al. [5]	Mentorship	RAG	Academic research	Develops "StartupGPT," a RAG chatbot for advice. Finds users perceive it as reliable but too theoretical and lacking concrete examples.
Shenoy [10]	Market Research	Fine-Tuning / Specialization	Country-specific data	Proposes creating specialized LLM assistants for international business to help navigate foreign market entry and marketing challenges.
Xu et al. [9]	Methodological Paradigm	Agentic RAG with Human-in-the-Loop	Conceptual	Introduces the "Human-in-the-Retrieval" concept, where an AI agent can actively query human experts for tacit knowledge.
Kenneweg et al. [7]	Evaluation	Automated Dataset Generation	Knowledge base	Presents RAGVAL, a framework to automatically create datasets and quantitatively evaluate RAG systems without human annotation.
Pushpalatha & Aravindan [6]	Technical Comparison	RAG vs Web Scraping Tools	Web-scraped data	Compares different AI web scraping methods, finding that graph-based logic (ScrapeGraphAI) provides superior accuracy over simpler RAG.

these simulations reflect biases in the training data rather than true population characteristics.

C. The Rise of Advanced Architectures for Reliability

The effectiveness of any LLM-based system depends on the factual accuracy and relevance of its outputs. To overcome the known limitations of standalone LLMs, such as hallucination and outdated knowledge, the field has rapidly adopted more sophisticated architectures.

1) Overcoming LLM Limitations with RAG: Retrieval-Augmented Generation (RAG) grounds an LLM's response in real-time, external information, ensuring outputs are up- to-date and factually accurate. The system "StartupGPT" by Sunde et al. (2025) is a direct application of this principle, using RAG to provide startup advice from a knowledge base of academic papers [5]. However, user feedback on this system produced a critical insight: while reliable, the advice was

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

often perceived as “too theoretical” and lacking in “practical, applicable advice” [5]. This highlights that factual grounding is necessary but not sufficient; the system must also synthesize knowledge into actionable outputs.

2) The Technical Ecosystem of RAG: The implementation and evaluation of RAG systems are themselves active areas of research. The performance of a RAG pipeline depends heavily on its components. Pushpalatha and Aravindan (2025) analyzed different AI-powered web scraping and RAG tech- niques, finding that methods incorporating graph logic pro- vided superior accuracy [6]. Concurrently, the challenge of objectively measuring RAG performance led Kenneweg et al. (2024) to develop RAGVAL, an automated workflow for creating evaluation datasets. This framework allows developers to systematically test and optimize RAG parameters for a specific domain without costly human labeling [7].

3) The Emergence of Agentic AI: “Agentic AI” represents the next step, where a complex problem is decomposed into tasks and autonomous “agents” (specialized LLMs) collabo- rate to solve them. Wang et al.’s SSFF for startup evaluation is a prime example of an agentic system applied to analysis [3]. Visionary work by Xu et al. (2025) proposes an “Agentic RAG with Human-in-the-Retrieval” paradigm, where an AI agent can autonomously decide to query a human expert for “tacit knowledge” that is not documented [9]. This concept blurs the line between tool and collaborator, pointing toward a future of deeply integrated human-AI systems.

IV. SYNTHESIS AND RESEARCH GAP

The reviewed literature reveals a clear trajectory in the appli- cation of LLMs to the startup domain. The field is progressing from single, monolithic models performing narrow predictive tasks toward decentralized, multi-agent systems capable of complex reasoning and generation.

A synthesis of the key findings highlights three major trends:

- **From Prediction to Generation:** While foundational work has focused on prediction (i.e., will this startup succeed?), there is a growing interest in using LLMs for generative tasks that support the entrepreneur directly, such as market simulation and strategic ideation [2], [4].
- **From Monolithic Models to Multi-Agent Systems:** The limitations of single LLMs, such as bias and lack of specialized depth, are being overcome by multi-agent frameworks that decompose complex problems and as- sign tasks to specialized agents, mirroring human expert teams [3].
- **From General Advice to Actionable Output:** Systems designed to provide guidance, like StartupGPT, have demonstrated the value of RAG for reliability but also revealed a strong user demand for outputs that are more concrete, personalized, and less theoretical [5].

This trajectory leads to a significant and well-defined research gap. While existing systems can **predict** success, **simulate** customer responses, or offer **general advice**, there is a lack of a unified system that can autonomously **generate** a comprehensive, structured, and actionable business plan from a simple idea. Current tools assist with pieces of the puzzle, but the entrepreneur is still left with the task of integrating the outputs into a coherent strategy. The “Start-up Business Planner Agent” is designed to directly address this gap by using a multi-agent framework to orchestrate the generation of a complete business plan, from market research to a final pitch deck.

V. DISCUSSION AND IMPLICATIONS

The development of autonomous, generative AI systems for business planning carries significant implications for both industry practice and academic research.

A. Democratization of Entrepreneurial Strategy

For aspiring entrepreneurs, a tool that can transform a raw idea into a structured business plan represents a massive reduction in the initial barriers to entry. It democratizes access to strategic frameworks (like the Business Model Canvas or SWOT analysis) and market data that were previously the domain of experienced consultants or MBA graduates. For startup incubators and accelerators, such a tool could function as a powerful screening and support mechanism, allowing program managers to rapidly assess the potential of a large volume of applications and provide incoming ventures with a solid, data-driven foundation.

B. Remaining Challenges and Ethical Considerations

Despite the promise, several challenges must be addressed. The primary risk is the generation of generic or “templated” business plans that lack true strategic insight. While RAG can ensure factual accuracy, the system’s ability to synthesize that data into a unique and compelling strategy remains a key hurdle. Furthermore, ensuring the ethical use of such a system is paramount. The potential for biases inherited from the LLM’s training data could lead to skewed market analysis, as LLMs still struggle to model the preferences of diverse demographic groups accurately [4]. Another consideration is intellectual property: who owns the output of an AI-generated business plan, and how can entrepreneurs ensure their novel ideas are protected?

C. Future Research Directions

This research area is rich with opportunities for future work. A key direction is exploring hybrid, interactive systems. Rather than a single-shot generation, the system could engage in a di- alogue with the entrepreneur, asking clarifying questions. This aligns with the “human-in-the-retrieval” paradigm proposed by Xu et al. (2025), where the agent can consult the founder for their unique vision and tacit knowledge [9]. Another vital area is the development of robust evaluation metrics for *generative* business tools.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

While metrics for predictive accuracy are well-established, assessing the quality, creativity, and viability of a generated business plan is a more subjective and complex task. Future research could focus on creating frameworks for this evaluation, potentially involving panels of expert investors and entrepreneurs.

VI. CONCLUSION

This review has charted the rapid integration of Large Language Models into the fields of startup evaluation and market research. The literature demonstrates a clear and accelerating trend away from using single, general-purpose models for simple analytical tasks and toward the development of sophisticated, multi-agent systems. These advanced frameworks leverage Retrieval-Augmented Generation to ensure factual grounding and employ specialized agents to conduct deep, nuanced analysis.

Our synthesis of the current state-of-the-art reveals that while LLM-based tools for predicting startup success and simulating market conditions are becoming increasingly robust, a significant opportunity remains for generative applications. The demand is not just for tools that can evaluate an existing idea, but for tools that can help build and articulate that idea from its inception. The conceptualization of a multi-agent system like the “Start-up Business Planner Agent” represents a timely and logical progression in the field. By automating the creation of a comprehensive business plan, such a system would synthesize the predictive, analytical, and simulative capabilities demonstrated across the literature into a single, powerful generative tool. It addresses a clear need for practical, actionable, and data-driven support for entrepreneurs, promising to lower the barrier to entry for innovators and foster a more dynamic startup ecosystem.

REFERENCES

[1] A. Maarouf, S. Feuerriegel, and N. Pro¨llochs, “A Fused Large Language Model for Predicting Startup Success,” arXiv preprint arXiv:2409.03668, 2024.
[2] J. B. Ilagan, Z. M. Alabastro, C. L. Basallo, and J. R. Ilagan, “Exploratory customer discovery through simulation using ChatGPT and prompt engineering,” in *2024 ICICT*, 2024.
[3] X. Wang, Y. Ihlamur, and F. Alican, “SSFF: Investigating LLM Predictive Capabilities for Startup Success through a Multi-Agent Framework with Enhanced Explainability and Performance,” arXiv preprint arXiv:2405.19456v2, 2025.
[4] J. Brand, A. Israeli, and D. Ngwe, “Using LLMs for Market Research,” Harvard Business School Working Paper, No. 23-062, July 2024.
[5] H. F. Sunde, A. Nguyen-Duc, T. L. Ahlgren, and L. Jaccheri, “On the evaluation of StartupGPT: A Retrieval-Augmented AI Chatbot for Delivering Research-Driven Guidance to Startups,” in *2025 IEEE/ACM IWSiB*, 2025.
[6] M. Pushpalatha and M. S. Aravindan, “Comparative Analysis of Web Scraping Methodologies Using Generative AI,” in *2025 6th RAIT*, 2025.
[7] T. Kenneweg, P. Kenneweg, and B. Hammer, “RAGVAL: Automatic Dataset Creation and Evaluation for RAG Systems,” in *2024 2nd FLLM*, 2024.
[8] Y. Katcharovski and A. L. Maxwell, “Predicting Startup Funding using AI and the CFA,” Conference Paper, November 2024.
[9] X. Xu, D. Zhang, Q. Liu, Q. Lu, and L. Zhu, “Agentic RAG with Human-in-the-Retrieval,” in *2025 IEEE ICSA-C*, 2025.
[10] P. P. Shenoy, “Large Language Model Virtual Assistants for International Business and Product Marketing,” in *2024 13th SMART*, 2024.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)