Human-in-the-Loop Multi-Agent AI for Startup Decision Support: A Review of Orchestration, Evidence-Grounded Advice, and Adaptive Monitoring

Dr. R. S. Meshram[1], Vipul Barmukh[2], Smriti Dharmadhikari[3], Sanjana Godse[4]

[1]Mrs. R. S. Meshram (Project Guide / Mentor), Dept. of Artificial Intelligence & Data Science, AISSMS Institute of Information Technology, Pune, India

[2]Vipul Barmukh, Dept. of Artificial Intelligence & Data Science, AISSMS Institute of Information Technology, Pune, India | [vipulbarmukh@gmail.com]

[3]Smriti Dharmadhikari, Dept. of Artificial Intelligence & Data Science, AISSMS Institute of Information Technology, Pune, India |smriti291105@gmail.com]

[4]Sanjana Godse, Dept. of Artificial Intelligence & Data Science, AISSMS Institute of Information Technology, Pune, India | [sanjanagodse06@gmail.comsmi]

Abstract – Starting and growing a startup is a continuous chain of decisions across market validation, business modelling, finance, product, growth and operations, yet the Artificial Intelligence (AI) tools available to founders mostly address one of these tasks at a time, give generic advice and do not compare the plan with what actually happens. This paper reviews fifteen recent IEEE papers, supported by foundation work on Large Language Model (LLM) agents, Retrieval-Augmented Generation (RAG), graph-based orchestration and tree-based machine learning, along three technical themes: orchestration of specialised agents, evidence-grounded and explainable advice, and adaptive monitoring of planned against actual performance, with human oversight as a cross-cutting requirement. The review finds that machine-learning studies predict startup success or lifespan but stop at a prediction; early-stage multi-agent systems coordinate idea generation, market research and Minimum Viable Product (MVP) planning but give limited support after launch; the one reviewed marketing system adds founder approval before publishing but covers content only; and general agentic AI studies are not tied to the startup lifecycle. None of the reviewed papers combines dynamic agent selection by startup type and stage, financial and valuation analysis, explanation with confidence and alternatives under founder control, and continuous planned-versus-actual monitoring that re-triggers planning. The paper maps these gaps to a reference architecture built on a LangGraph orchestrator with seven specialised agents, discusses open challenges in data quality, LLM reliability, valuation and privacy, and outlines directions for future work. No new experimental results are reported; findings of the reviewed papers are stated as reported by their authors.

Key Words: Startup Decision Support, Agentic AI, Multi-Agent Systems, Human-in-the-Loop, Large Language Models, LangGraph, Retrieval-Augmented Generation, Explainable AI, Adaptive Monitoring, Startup Success Prediction.

1. INTRODUCTION

Startups drive innovation and employment, yet a founder faces a continuous chain of decisions. The founder validates a market, identifies customers, studies competitors, designs a business model, plans finances, builds a product, markets it and runs operations, and a weak decision at any step shows up later as lost time, money or growth. Founders with limited business experience or resources must analyse information from many sources and decide across these areas at the same time. Needs also differ by startup type: a technology startup concentrates on product development, technical architecture and software implementation, whereas a non-technology startup depends more on suppliers, inventory, operations, distribution and customer acquisition. A generalised advisory approach therefore seldom gives

sufficiently relevant guidance.

Existing startup-support platforms and AI tools usually handle one task at a time, such as market research, business-plan generation, financial analysis or content generation. They work independently, do not keep a continuous picture of the startup's situation, and do not compare planned objectives with actual performance. The founder is left to connect the insights, set priorities, build execution plans and decide what to change as the startup evolves.

Three lines of work make a more coordinated approach possible. The first is agentic AI, in which several specialised LLM agents collaborate under a central orchestrator [14], [15], [17], supported by graph-based orchestration of stateful workflows [20]. The second is grounding: RAG supplies retrieved text that an answer can be traced to [19], and tree-based models remain established methods for tabular business data [21], [22]. The third is human–AI collaboration, which treats AI as decision support while keeping the human involved in important decisions [18]. For advice about a founder's money and company this last point is more than a design preference, because LLM output can be fluent yet wrong, valuations depend on assumptions, and external services can fail.

The literature on AI for startups is, however, scattered across prediction, estimation, idea validation, marketing and general agentic frameworks. A reader who wants to know what is established about a coordinated, grounded, monitored and human-supervised decision-support system has to assemble the answer from papers that each address one piece. This paper performs that assembly. The main contributions of this review are:

- A structured review of fifteen recent IEEE papers [1]–[15] and seven foundation references [16]–[22], each summarised with its strengths and limitations as reported by the original authors.
- A four-dimension analytical lens – orchestration, evidence-grounded advice, adaptive monitoring and human oversight – and a comparison of the reviewed work against it.
- A research-gap statement showing that no reviewed paper combines dynamic orchestration, financial and valuation analysis, explainable advice under founder control and continuous monitoring in one platform.
- A reference architecture in which each reviewed limitation is addressed by a named component, together with open challenges and future directions.

Better decision support for founders also bears on SDG 8 (target 8.3, entrepreneurship and growth of micro-, small- and medium-sized enterprises) and SDG 9 (target 9.5, technological capabilities and innovation) [24]. The rest of the paper is organised as follows. Section 2 states the scope and method of the review. Section 3 gives the background. Section 4 reviews existing work by category, and Section 5 synthesises it along the four dimensions. Section 6 compares the work and states the gap. Section 7 presents the reference architecture, Section 8 discusses open challenges, Section 9 gives future directions and Section 10 concludes.

2. REVIEW SCOPE AND METHOD

2.1 Sources and Scope

The review covers fifteen IEEE papers. Five primary papers published in 2026 [1]–[5] cover startup success prediction, multi-agent startup support, strategic planning for technology startups, agentic marketing and agentic decision support. Ten further IEEE papers [6]–[15], published between 2022 and 2026, cover machine-learning prediction of startup success and lifespan, capital and profitability estimation, AI-based idea validation, and agentic and multi-agent LLM foundations. Seven foundation references complete the base: Zhang et al. [16] on LLM-based business ideation and validation, Yang et al. [17] on LLM-based autonomous agents and their coordination, Park et al. [18] on human–AI collaboration in decision support, RAG [19], LangGraph [20], random forests [21] and gradient boosting [22].

2.2 Analytical Lens

Each paper was summarised by its relevance, technique, strengths and limitations (Table 1) and then read against four dimensions that follow from the problem described in Section 1, as shown in Fig. 1. **Orchestration** asks whether several specialised agents are coordinated, and how. **Evidence-grounded advice** asks whether outputs are tied to data, retrieved sources or explanation. **Adaptive monitoring** asks whether planned and actual performance are compared and the plan updated. **Human oversight** asks whether the founder reviews, approves or can reject what the system proposes. Startup-specific scope is recorded as a fifth, descriptive column in the comparison of Section 6.

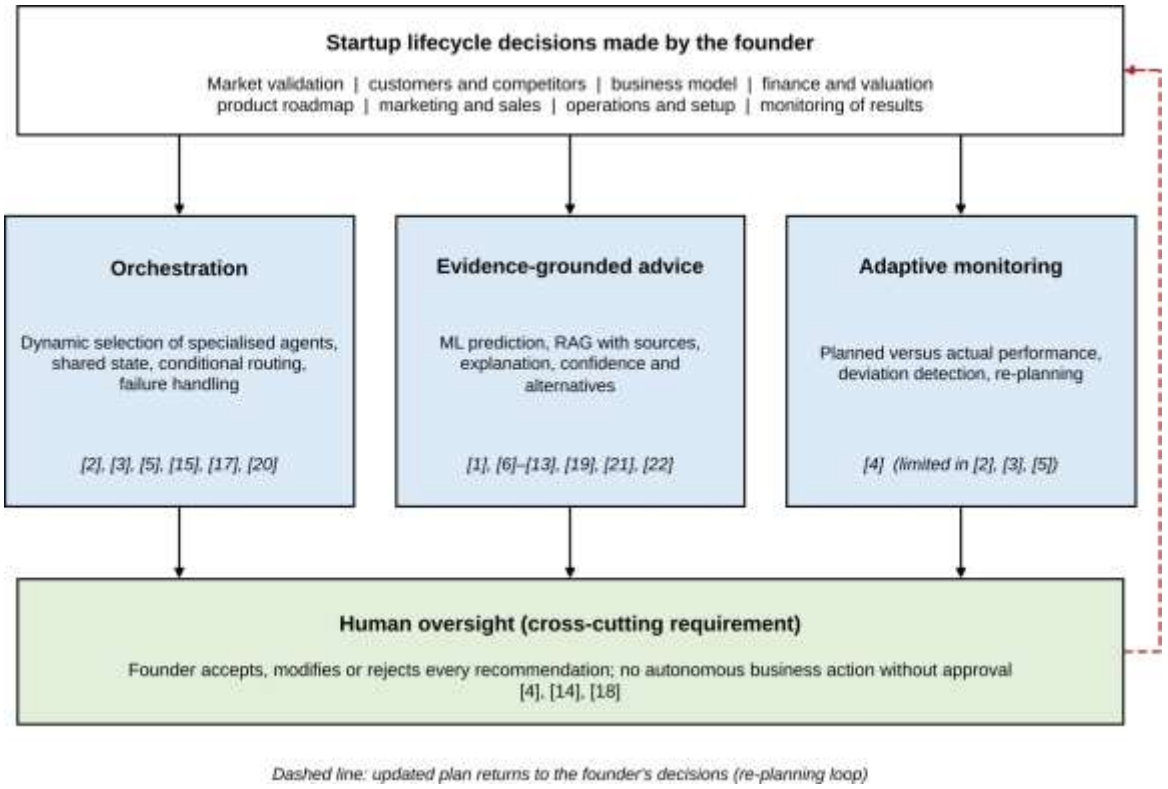


Fig. 1: Analytical framework of the review and the papers mapped to each dimension

2.3 Limits of the Review

Results and claims are those reported by the original authors; they are not results of this work, and the review does not re-implement any of the systems. Datasets and metrics differ between papers, so reported accuracies are not directly comparable. The ratings in Table 3 are the authors' reading of the strengths and limitations reported for each paper; a cross (✗) means that a capability is not reported in the reviewed summary, not that it has been shown to be absent. Bibliographic details for several entries are still being completed and are highlighted in the reference list.

3. BACKGROUND AND FOUNDATIONS

3.1 Decisions Across the Startup Lifecycle

A founder-facing decision-support system has to cover validation of the idea and market, customer and competitor analysis, the business model, the financial plan with funding requirements and valuation, product roadmaps and milestones, marketing, sales and customer retention, legal requirements, hiring and operational setup, and the monitoring of results against the plan. Technology and non-technology startups weight these areas differently, which is why coverage of both types matters. The founder remains the accountable decision maker throughout.

3.2 Agentic AI and LLM-Based Multi-Agent Systems

Acharya et al. [14] survey the concepts, methods and applications of agentic AI, including healthcare, finance and adaptive software, and discuss ethical challenges together with a framework for safe integration. The study in [15] uses LLMs as high-level reasoning engines that coordinate domain-specific agents, applied to process automation, data-driven decision-making and incident response. Yang et al. [17] survey LLM-based autonomous agents and their coordination in complex tasks. Together these works establish the coordination pattern that the startup-specific systems reviewed in Section 4 build on, but none of them is tied to the startup lifecycle.

3.3 Grounding, Prediction and Orchestration Building Blocks

RAG [19] retrieves relevant text and supplies it to the language model, so that an answer can be grounded in sources. LangGraph [20] provides graph-based orchestration of stateful workflows: nodes hold the steps, edges define their order, conditional edges route work, and a shared state carries the startup context between nodes. Random forests [21]

and gradient boosting [22] are established methods for tabular business data and underlie most of the prediction studies reviewed below.

4. REVIEW OF EXISTING WORK

Table 1 summarises the fifteen reviewed papers. Strengths and limitations are as reported for each paper; years marked TBC are still to be confirmed from IEEE Xplore.

4.1 Startup Success and Risk Prediction

Seven of the fifteen papers [1], [6]–[9], [11], [12] predict how a startup will fare. Start-Wise AI [1] trains a Random Forest survival predictor on a synthetic set of over 2,000 startup records and exposes survival scores, risk indicators and recommendations through Streamlit and FastAPI. The study in [6] combines Random Forest, AdaBoost and XGBoost in a hybrid ensemble on founder-background and funding features and reports accuracy above 90% for most results, while
[7] compares Decision Tree, Random Forest, Gradient Boost, Logistic Regression and a multilayer perceptron on valuation and funding features and reports about 92%. The work in [8] applies XGBoost with SHAP values to the IT Orange dataset to find determinants of start-up lifespan in the Chinese market, Abhinand and Poonam [9] propose a stacking ensemble for Indian ventures, [11] studies entrepreneurial success in Bahrain, and [12] compares several techniques for predicting success rate. These studies perform well on their own datasets and [8] alone adds explanation. All of them stop at a prediction: none plans, acts or monitors, and their datasets are synthetic, regional or single-source, so the reported accuracies do not transfer to other settings.

4.2 Financial Estimation

The study in [10] applies linear regression to collected business metrics to predict the start-up capital requirement and future profitability of e-commerce start-ups. It addresses funding requirement and profitability, which a Strategy and Finance function needs, but it is limited to one domain and to simple models, and it offers estimation without execution support.

4.3 Idea Validation and Pitching

PitchPal-AI [13] combines an Idea Intelligence Engine for semantic analysis and market relevance, pitch-deck generation and a Validation Scoring System for innovation, market fit and feasibility, using NLP and transformer models. It is useful for the idea stage but does not cover financial planning, operations or continuous monitoring. Zhang et al. [16] provide a foundation for LLM-based business ideation and validation.

4.4 Early-Stage Multi-Agent Support

Two 2026 papers coordinate specialised agents for early-stage support. The Agentic AI Startup Cofounder [2] uses specialised agents with open-source LLMs, vector memory, FastAPI, React and ChromaDB across an eight-stage lifecycle from idea generation to MVP. The strategic-planning system for tech startups [3] uses agents for validation, market research, business modelling, MVP planning and pitch, with vector memory, external APIs, workflow orchestration, dashboards and feedback refinement. Both show that agents sharing memory can coordinate several business tasks, yet both concentrate on the early stage and on producing deliverables, with limited monitoring, valuation and adaptation to actual performance.

4.5 Marketing Automation

The system in [4] manages startup social-media content with a Planner–Actor design built on n8n, LLM APIs, Buffer and VADER sentiment analysis. It adapts content from engagement and sentiment and requires human approval before publishing, which makes it the clearest example of a human-approval step among the startup-specific papers. It is limited to social-media marketing, with no finance, valuation, product or operations support, and it names multi-agent design and ROI tracking as future work.

4.6 Decision-Support Frameworks and Agentic Foundations

The study in [5] confirms the value of collaborating LLM-based agents for business analysis, strategic planning and growth decisions, without covering continuous monitoring, valuation, execution management or adaptive decision-making. The general works [14], [15] and [17] give foundations for agentic AI and LLM-guided agent coordination but are not tailored to startups, finance, valuation or business monitoring, while Park et al. [18] ground the use of AI as decision support in which the human stays involved in important decisions. Table 2 groups the reviewed work by category.

Table 2: Technical Comparison by Category

Category	Papers	Typical technique	Common limitation
Startup success prediction	[1], [6]–[9], [11], [12]	Random Forest, XGBoost, AdaBoost, stacking and hybrid ensembles, MLP, feature importance, SHAP	Prediction only; no agents, monitoring or adaptation; often synthetic, regional or single-dataset
Financial estimation	[10]	Linear regression for capital and profitability	Narrow domain (e-commerce); estimation only
Idea validation and pitching	[13]	NLP and transformer models, scoring	No finance, operations or monitoring
Early-stage multi-agent support	[2], [3], [16]	LLM agents, vector memory, orchestration layer, external APIs	Early-stage or MVP focus; limited valuation and post-launch adaptation
Marketing automation	[4]	Planner–Actor, n8n, LLM APIs, sentiment analysis	Marketing only; no wider business functions
Decision-support frameworks	[5], [18]	Agentic LLM framework; human–AI collaboration	Limited monitoring, valuation and execution management
LLM and agentic AI foundations	[14], [15], [17], [19], [20]	Autonomous LLM agents, RAG, graph-based orchestration	General frameworks; not tied to the startup lifecycle

5. THEMATIC SYNTHESIS

5.1 Orchestration

The startup-specific multi-agent systems [2]–[5] coordinate agents through shared memory, APIs and workflow tools, for example vector memory in [2] and [3] and n8n in [4], but they are described as a fixed pipeline or agent set, such as the eight-stage lifecycle of [2]. Selection of agents by startup type, stage and goals is not reported. General work points to how this could be done: the LLM acts as a high-level reasoning engine that decides which domain agents to call [15], and graph-based orchestration provides a shared state with conditional routing [20]. With conditional routing, a Developer agent can be activated only for technology startups, while a non-technology startup is routed to operations, suppliers and distribution. A second orchestration concern is failure: language-model APIs and external services can be unavailable, changed or costly, so an orchestrator should return an explicit “unavailable” result for the affected agent instead of a guess, and keep the other results.

5.2 Evidence-Grounded Advice

Three kinds of evidence appear in the reviewed work. The first is statistical: models trained on historical data produce survival scores, risk indicators or success labels [1], [6]–[12]. The second is retrieval: vector memory in [2] and [3] and, in general, RAG [19] supply text that outputs can be tied to. The third is explanation: feature importance in [1] and SHAP values in [8] show which inputs drive a prediction. Two gaps remain. A prediction is not yet advice: the prediction papers end with a score or a label, and even [1], which adds risk indicators and recommendations, does not act on its outcome. And explanation is rarely given to the founder: none of the reviewed papers is reported to attach assumptions, a confidence level and alternatives to a recommendation, and per-answer source citation is not reported for any of the startup-specific systems. The evidence base is also thin. The prediction in [1] comes from a synthetic dataset and should not be read as real-world performance, and [8], [9] and [11] are regional. Valuations in particular are estimates that

depend on assumptions and on the accuracy of founder-entered data, so they should be shown as ranges with their assumptions.

5.3 Adaptive Monitoring

Monitoring is the least developed dimension. The prediction papers produce a single outcome and do not follow the startup after the prediction [1], [6]–[12]. The multi-agent systems [2], [3] and [5] report limited monitoring, and the only reported adaptation loop is the content adaptation of [4], which uses engagement and sentiment for social-media posts. A general formulation of planned-versus-actual monitoring is to track indicators such as revenue, expenses, customers and milestone status, compute the relative deviation of each from plan, and re-trigger planning when a deviation exceeds a threshold:

$$d_i = (a_i - p_i) / p_i, \quad \text{re-plan if } |d_i| > \tau_i$$

where a is the actual value, p the planned value and τ the threshold for indicator i. The deviating indicator decides which agents are re-triggered: a revenue shortfall points to growth and finance, a milestone delay to product and operations. This formulation is a generic illustration, not a result from the reviewed papers. The thresholds have to be fixed during design, and founder-entered figures are a noisy source, so how thresholds are chosen and how often they trigger re-planning remain open questions.

5.4 Human Oversight

Human oversight is reported explicitly in only one startup-specific paper, [4], where a human approves content before publishing. Acharya et al. [14] treat human–AI collaboration and safe integration at the level of agentic AI in general, and Park et al. [18] ground AI as decision support in which people remain involved in important decisions. The system in [3] reports feedback refinement, which is an indirect form of human input. For startup advice the practical design is: the founder accepts, modifies or rejects every recommendation; a rejected result returns through the workflow for revision; and no business action is taken autonomously without founder approval. Because valuations, legal requirements and investment choices carry real consequences, the output is decision support and not certified legal, financial or investment advice, and the founder is told to consult qualified professionals. A review step does cost founder effort, and where to place such steps so that they protect decisions without becoming a bottleneck is an open design question.

6. COMPARATIVE ANALYSIS AND RESEARCH GAP

Table 3 rates the reviewed work on startup-specific scope and the four dimensions of Section 2.2, using the strengths and limitations reported for each paper. Partial means that a capability is reported in a limited or indirect form.

Table 3: Capability Comparison of Reviewed Work and the Reference Architecture

Reference	Startup-specific	Multi-agent orchestration	Evidence and explanation	Human oversight	Monitoring and adaptation
[1] Start-Wise AI (2026)	✓	✗	Partial	✗	✗
[2] Agentic AI Startup Cofounder (2026)	✓	✓	Partial	✗	Partial
[3] Strategic planning for tech startups (2026)	✓	✓	Partial	Partial	Partial
[4] Startup social-media marketing (2026)	✓	Partial	✗	✓	Partial
[5] Agentic AI for startup growth (2026)	✓	✓	✗	✗	Partial
[6], [7], [9], [11], [12] Success prediction	✓	✗	✗	✗	✗
[8] Lifespan prediction with SHAP (2023)	✓	✗	Partial	✗	✗

[10] Capital and profitability estimation (2024)	✓	✗	✗	✗	✗
[13] PitchPal-AI idea validator (2026)	✓	✗	✗	✗	✗
[14] Agentic AI survey (2025)	✗	Partial	✗	Partial	✗
[15] LLM-guided multi-agent collaboration (2025)	✗	✓	✗	✗	✗
Reference architecture (Section 7, planned)	✓	✓	✓	✓	✓

✓ *reported; Partial: reported in a limited or indirect form;* ✗ *not reported in the reviewed summary.*

Four findings follow. First, startup-specific prediction is the best-covered area, but it ends at a prediction. Second, multi-agent support for startups exists, but it concentrates on the early stage and on deliverables. Third, adaptive monitoring and human oversight are the least covered dimensions: only [4] reports an approval step, and no startup-specific paper reports continuous planned-versus-actual monitoring. Fourth, general agentic frameworks [14], [15], [17] cover coordination but are not tied to the startup lifecycle. Table 4 contrasts what the review found with what a coordinated system needs.

Table 4: Gap Between Reviewed Work and the Reference Architecture

Aspect	Existing work (from the review)	Reference architecture
Scope	Single function or early-stage support	Whole startup lifecycle on one platform
Architecture	Single model, fixed pipeline or fixed agent set	LangGraph orchestrator that selects among seven agents by startup type and stage
Finance and valuation	Survival prediction or limited financial output	Financial planning, funding requirement and valuation estimate
Explanation	Rarely given to the founder	Factors, assumptions, confidence and alternatives; founder approves
Monitoring	Rare; marketing feedback only	Planned-versus-actual comparison that re-triggers agents
Startup types	Mostly technology startups	Technology and non-technology startups

Research gap. None of the reviewed papers combines (i) an orchestrator that selects agents dynamically by startup type, stage and goals, (ii) financial and valuation analysis, (iii) product, growth and operational planning for both technology and non-technology startups, (iv) explainable recommendations with confidence and alternatives under founder control, and (v) continuous monitoring that compares planned and actual performance and triggers re-planning. Combining these in one founder-oriented platform is the open problem this review identifies.

7. REFERENCE ARCHITECTURE

To show how the gaps can be addressed together, this section outlines the reference architecture followed in the authors' ongoing project, a Multi-Agent AI Platform for Startup Ideation and Execution (Group A3, AISSMS Institute of Information Technology, Pune). It is presented as a design. No performance has been measured at this stage, and any target values are design goals, not achieved results.

7.1 Architecture Overview

The presentation layer is one Angular founder app. The application layer is one FastAPI backend that validates input, exposes REST APIs, reads and writes PostgreSQL, runs the LangGraph Startup Orchestrator and calls the Grok API; LangGraph runs inside the FastAPI backend and is not a second backend. The data layer holds PostgreSQL, which stores the founder profile, onboarding answers and saved agent results. The Angular app talks to the backend over HTTPS using JSON, and secrets such as API keys are kept in environment variables. A RAG knowledge base and

machine-learning models for startup risk and financial estimation are planned components (dashed in Fig. 2).

7.2 Orchestrator, Agents and Explanation

The orchestrator holds the startup state and routes work to seven specialised agents, each implemented as a LangGraph node with a defined role and input and output schema. Table 5 lists the components and the reviewed limitation each addresses, and Fig. 2 shows how they connect.

Table 5: Components of the Reference Architecture and the Limitations They Address

Component	Responsibility	Reviewed limitation addressed
Startup Intelligence Agent	Market, customer, competitor and feasibility analysis	Coverage limited to idea validation [13]
Component	**Responsibility**	**Reviewed limitation addressed**
Strategy & Finance Agent	Business model, financial plan, funding requirements, key metrics and valuation estimate	Prediction only [1], [6]–[12]; narrow estimation [10]
Product & Roadmap Agent	Product or service roadmaps, milestones and actionable tasks	Early-stage, deliverable focus [2], [3]
Developer Agent	Technical planning and implementation support, activated for technology startups only	Technology and non-technology needs differ
Growth & Customer Agent	Marketing, sales, customer acquisition and retention	Marketing-only coverage [4]
Operations & Setup Agent	Legal requirements, hiring, resources and operational setup	Operations rarely covered [4], [13]
Monitoring & Adaptation Agent	Compares planned and actual performance, identifies deviations and triggers agents to update strategy	Limited or no monitoring [1], [2], [3], [5]
Explainable AI module	Factors, supporting data, assumptions, expected impact, confidence and alternatives; founder accepts, modifies or rejects	Explanation rarely given [1], [8]; no oversight [5]
RAG knowledge base (planned)	Chunking, embeddings and retrieval of external and founder-provided information, with sources	Ungrounded LLM output [19]

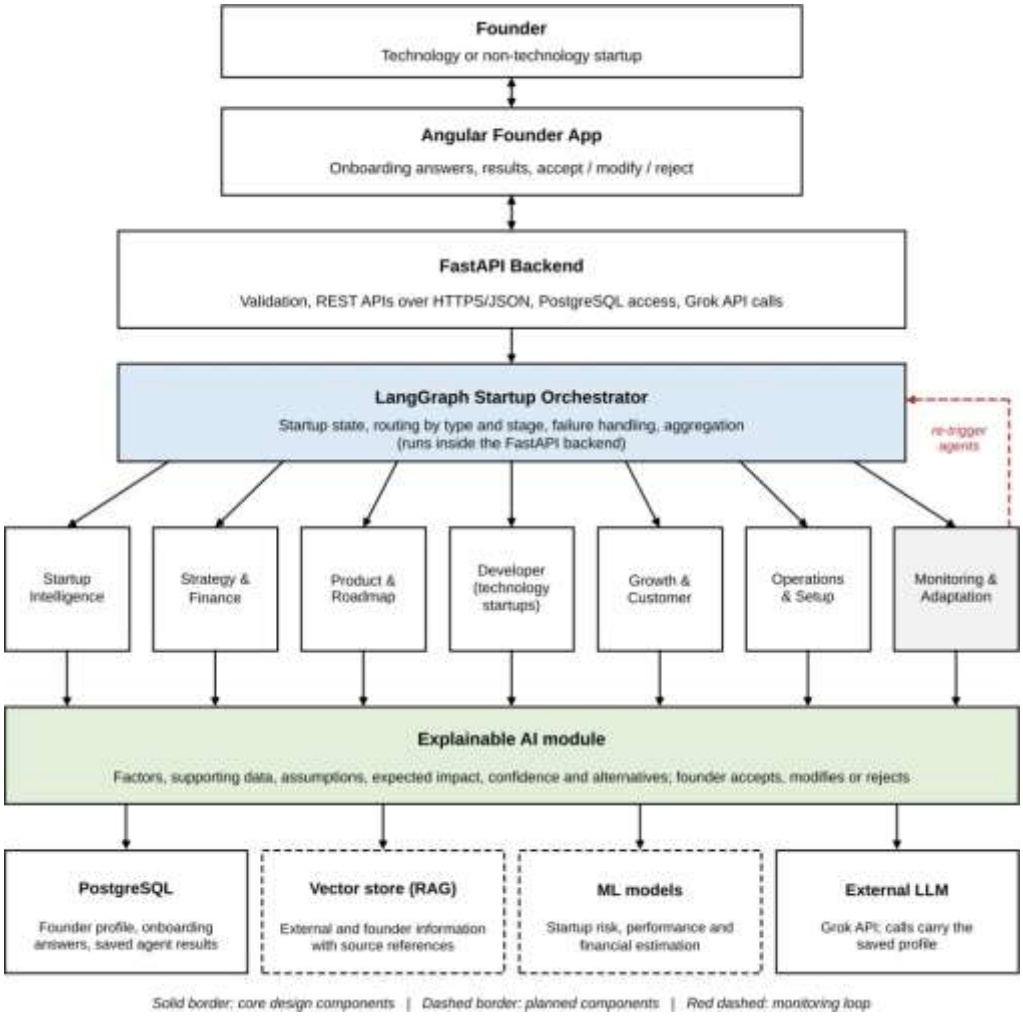


Fig. 2: Reference architecture of the multi-agent startup decision-support platform

7.3 Workflow

The founder enters startup information covering domain, stage, problem, customers, business model, financial position, resources and goals, and the system structures it into a startup profile. The orchestrator analyses the profile, selects the agents the startup needs, collects their results and returns one explained set of recommendations. The initial LangGraph graph routes idea refinement, market analysis, a revision step, founder review and an "unavailable" branch. Each node calls the Grok API with the saved profile, and the market-analysis node must attach sources. Tasks, outputs and progress are stored so that the Monitoring & Adaptation Agent can compare them with the plan and trigger the relevant agents when the market, customers, finances, operations or goals change. A feature is accepted only after it works from the dashboard to the database.

7.4 Proposed Validation Approach

For systems of this kind the validation approach should match each claim to a test:

- Machine-learning models are validated with train/test splits and cross-validation, and report measured metrics such as accuracy, F1, MAE and RMSE on held-out data.
- Agents are unit-tested with mocked LLM and API responses, and the orchestrator with a labelled set of sample startup scenarios covering technology and non-technology startups at different stages.
- RAG answers are checked manually against reference questions, and every market answer is checked for at least one cited source.
- The Monitoring & Adaptation Agent is tested on simulated planned-versus-actual data to confirm that deviations are detected and the right agents are triggered.
- Injected service failures confirm that the system returns an "unavailable" result without crashing, and the full founder journey is covered by end-to-end tests.

8. OPEN CHALLENGES AND LIMITATIONS

Data. Public startup and financial datasets are limited, synthetic, regional or single-source [1], [8], [9], [11], so reported accuracies do not transfer. Studies should use several datasets, report their limits openly and avoid real-world accuracy claims beyond the tested categories.

LLM reliability. Language-model output can be fluent yet wrong. Grounding through RAG with source citation, explanation with confidence, and founder approval reduce this risk but do not remove it.

Valuation. Valuations are estimates that depend on assumptions and on founder-entered data. They should be shown as ranges with assumptions, and guaranteed valuations are out of scope.

External services. LLM and data services can be unavailable, changed or costly. Mocked tests, caching, usage limits and an explicit "unavailable" result limit the damage, and API cost constrains how often models can be called.

Privacy, ethics and law. Business plans and financial information are confidential. Only necessary data should be stored, with consent, encrypted in transit and protected by access control, and only the minimum context should be sent to the language-model API. Practice should follow the Digital Personal Data Protection Act, 2023 [23] and the Information Technology Act, 2000, respect dataset, model and library licences, and document possible bias in datasets and language models.

Evaluation. There is no common benchmark for startup-advice systems, and the reviewed papers use different datasets and metrics. Comparison with reported results must therefore carry the caveat that datasets differ.

Limits of this review. The review covers fifteen IEEE papers and seven foundation references, relies on the strengths and limitations reported for each, and does not re-implement any system. The capability ratings are interpretive, and several bibliographic details are still to be completed.

9. FUTURE DIRECTIONS

The review points to the following directions:

- Integration with real-time market, financial, competitor and industry APIs for continuously updated insights.
- Integration with CRM, accounting, project-management, e-commerce and analytics systems for automated performance monitoring.
- More advanced multi-agent collaboration and autonomous task execution, with agents performing selected business activities only after founder approval.
- Predictive analytics for revenue forecasting, customer churn, funding requirements, business risks and growth opportunities.
- Cloud deployment and scalability for many startups, with persistent startup memory and long-term performance tracking.
- Integration with investor and funding platforms for investor matching, pitch evaluation and fundraising support.
- A shared set of labelled startup scenarios and metrics, so that orchestration, grounding and monitoring can be compared across systems.

10. CONCLUSION

This paper reviewed fifteen recent IEEE papers and seven foundation references on AI-based startup decision support along four dimensions: orchestration, evidence-grounded advice, adaptive monitoring and human oversight. The reviewed work supports individual functions or the early stage and leaves the post-launch loop of monitoring and adaptation open. Prediction studies stop at a prediction, early-stage multi-agent systems concentrate on deliverables, the one marketing system with founder approval covers content only, and general agentic frameworks are not tied to the startup lifecycle. No reviewed paper combines dynamic orchestration, financial and valuation analysis, explainable advice under founder control and continuous planned-versus-actual monitoring. A single explainable, lifecycle-oriented platform built as specialised agents under one orchestrator, with the founder approving every recommendation, is the combination that remains to be built and measured, for both technology and non-technology

startups.

ACKNOWLEDGEMENT

The authors sincerely thank their project guide, Mrs. R. S. Meshram, for her guidance, mentorship and encouragement throughout this work. The authors also thank the Department of Artificial Intelligence and Data Science, AISSMS Institute of Information Technology, Pune, for providing the facilities and academic environment that made this work possible.

REFERENCES

[1] "Start-Wise AI: Predicting Startup Success or Failure Using Data-Driven Insights," in Proc. IEEE ICETEMS, 2026,

doi: 10.1109/ICETEMS66917.2026.11469447. (Author names and pages to be added from IEEE Xplore.)

[2] "Agentic AI Startup Cofounder," in Proc. IEEE CI2A, 2026, doi: 10.1109/CI2A69097.2026.11576988. (Author names and pages to be added; DOI to be verified, as it is identical to the DOI recorded for [3].)

[3] "AI-Empowered System for Personalized Strategic Planning for Tech Startups," in Proc. IEEE CI2A, 2026. (Author

names, pages and DOI to be added from IEEE Xplore.)

[4] "Autonomous Content Generation and Management for Startup Social Media Marketing using Agentic AI," in Proc. 2026 2nd Int. Conf. Advances in Intelligent Computing and Applications (AICAPS), 2026, doi: 10.1109/AICAPS68631.2026.11453213.

[5] "Agentic AI for Startup Growth and Business Decision Support," IEEE, 2026.

[6] "Machine Learning-Based Prediction of Startup Success and Using Founder's Background and Funding Features,"

IEEE, IEEE Xplore document 11325516. (Authors, venue, year, pages and DOI to be added from IEEE Xplore.)

[7] "Predicting Success Rate of Startups using Machine Learning Algorithms," IEEE, 2022, IEEE Xplore document 9908921. (Authors, venue, pages and DOI to be added from IEEE Xplore.)

[8] "Machine Learning-based Start-up Company Lifespan Prediction: the Chinese Market as an Example," IEEE, 2023,

IEEE Xplore document 10185504. (Authors, venue, pages and DOI to be added from IEEE Xplore.)

[9] G. Abhinand and B. N. Poonam, "An efficient stacking ensemble technique for success prediction of Indian ventures," in Proc. 2022 Int. Conf. Data Science, Agents & Artificial Intelligence (ICDSAAI), IEEE, 2022, doi: 10.1109/ICDSAAI55433.2022.10028895.

[10] "Start-Up Capital Estimation and Profitability Prediction for E-Commerce Start-Ups Using Machine Learning," IEEE, 2024, doi: 10.1109/SEB4SDG60871.2024.10630338. (Authors, venue and pages to be added.)

[11] "Integrating Machine Learning to Predict Entrepreneurial Success in Bahrain," IEEE, 2024, doi: 10.1109/SIBF63788.2024.10883881. (Authors, venue and pages to be added.)

[12] "Comparative Analysis of Start-up Success Rate Prediction Using Machine Learning Techniques," IEEE, 2024, doi: 10.1109/ICSTSN61422.2024.10671362. (Authors, venue and pages to be added.)

[13] "PitchPal-AI Startup Pitch Deck & Idea Validator," IEEE, 2026, IEEE Xplore document 11507555. (Authors, venue, pages and DOI to be added from IEEE Xplore.)

[14] D. B. Acharya, K. Kuppan and B. Divya, "Agentic AI: Autonomous Intelligence for Complex Goals—A Comprehensive Survey," IEEE Access, vol. 13, pp. 18912–18936, 2025, doi: 10.1109/ACCESS.2025.3532853.

[15] "LLM-Guided Multi-Agent Collaboration for Complex Task Automation," in Proc. IEEE ICERECT, 2025, doi: 10.1109/ICERECT65215.2025.11377466. (Authors and pages to be added.)

[16] Zhang et al., "StartupGPT: Automated Business Ideation and Validation Using Large Language Models," in Proc.

IEEE Int. Conf. Big Data, 2024. (Full author list and details to be verified.)

[17] Yang et al., "Large Language Models as Autonomous Agents: A Survey of Capabilities, Challenges, and Applications," IEEE Trans. Artif. Intell., 2024. (Full author list and details to be verified.)

[18] Park et al., "Human–AI Collaboration in Decision Support Systems: A Survey and Research Directions," Information Systems Frontiers, 2024. (Full author list and details to be verified.)

[19] P. Lewis et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in Proc. Adv. Neural Inf.
Process. Syst. (NeurIPS), 2020, arXiv:2005.11401.

[20] LangChain Inc., "LangGraph documentation." [Online]. Available: LangChain official documentation website (accessed 2026).

[21] L. Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.

[22] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in Proc. 22nd ACM SIGKDD Int. Conf.
Knowledge Discovery and Data Mining, 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[23] Government of India, "The Digital Personal Data Protection Act, 2023," Gazette of India, Aug. 2023.

[24] United Nations, "Transforming our world: the 2030 Agenda for Sustainable Development, Sustainable Development Goals 8 and 9." [Online]. Available: sdgs.un.org (accessed 2026).