





Article

Startup Survival Forecasting: A Multivariate AI Approach Based on Empirical Knowledge

Francesc Font-Cot ¹, Pablo Lara-Navarra ^{1,*}, Claudia Sánchez-Arnau ² and Enrique A. Sánchez-Pérez ³

¹ Information & Communication Faculty, Open University of Catalonia, Rambla Poblenou, 08018 Barcelona, Spain; ffontco@uoc.edu

² School of Engineering, Universitat València, Avinguda de l'Universitat, 46100 Burjassot, Spain; sanar4@alumni.uv.es

³ Applied Mathematics Department, Universitat Politècnica València, Camino de Vera, 46022 Valencia, Spain; easancpe@upv.es

* Correspondence: plara@uoc.edu

Abstract: Predicting the survival of startups is a complex challenge due to the multifaceted nature of entrepreneurial ecosystems and the dynamic interplay of internal and external factors. Despite advances in empirical research, existing models often lack integration with robust conceptual frameworks. This study addresses these gaps by developing a multivariate AI-driven model for predicting startup survival, leveraging Lipschitz extensions, neural networks, and linear regression. Using a dataset of 20 startups, selected across diverse industries and evaluated on attributes such as team dynamics, market conditions, and financial metrics, the model demonstrated high accuracy and clustering capabilities. Key findings highlight the pivotal role of team dynamics and product differentiation in determining survival probabilities. By integrating conceptual insights with empirical data, the study bridges gaps in existing literature and offers a practical decision-making tool for entrepreneurs, investors, and policymakers. These findings underscore the importance of fostering collaborative, innovative ecosystems to enhance entrepreneurial success and societal well-being.



Academic Editor: Martin Wynn and Jose Irizar

Received: 2 December 2024

Revised: 4 January 2025

Accepted: 9 January 2025

Published: 16 January 2025

Citation: Font-Cot, F.; Lara-Navarra, P.; Sánchez-Arnau, C.; Sánchez-Pérez, E.A. Startup Survival Forecasting: A Multivariate AI Approach Based on Empirical Knowledge. *Information* **2025**, *16*, 61. <https://doi.org/10.3390/info16010061>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: startup survival; artificial intelligence; Lipschitz extensions; neural networks; entrepreneurial ecosystems

1. Introduction

Predicting startup survival remains a significant challenge in entrepreneurship research, as approximately 90% of startups fail within their first few years (Cantamessa et al., 2018) [1]. These high failure rates underscore the inherent volatility of entrepreneurial ecosystems, driven by factors such as resource constraints, intense market competition, and scaling complexities. Traditional forecasting models, which focus on static indicators like financial performance, team composition, or market fit (Cooper, 1990; Wernerfelt, 1984) [2,3], often fall short of capturing the dynamic and nonlinear processes that shape startup trajectories.

Recent advancements in artificial intelligence (AI) have opened new avenues for addressing these limitations. AI techniques, including neural networks and clustering algorithms, are well-suited for uncovering complex data relationships and latent variables that traditional methods often overlook (Nambisan, 2017; Sun et al., 2024) [4,5]. These models enable more precise survival predictions and offer actionable insights for entrepreneurs, investors, and policymakers. However, while AI-based approaches have

demonstrated potential, their effectiveness hinges on robust datasets and the integration of conceptual frameworks that account for the multidimensional and dynamic nature of entrepreneurial ecosystems.

This study seeks to bridge the gap between empirical data and conceptual understanding in startup survival analysis. We develop a multivariate AI model that integrates Lipschitz extensions, neural networks, and linear regression within a conceptual framework. The framework, inspired by theoretical foundations such as the Resource-Based View (Wernerfelt, 1984) [3] and Open Innovation (Chesbrough, 2003) [6], emphasizes the role of dynamic team alignment, market conditions, and innovation in predicting success. This approach provides a structured method for analyzing startup scalability, addressing the gaps in existing literature by combining conceptual insights with empirical validation.

The dataset used in this study includes evaluations of 20 startups across diverse industries, gathered through structured surveys, interviews, and publicly available data. Key variables include team dynamics, market conditions, financial metrics, and strategic vision. These variables were chosen based on their relevance to startup survival and were iteratively refined through industry expert feedback to ensure their practical applicability.

This paper is structured as follows: Section 2 provides a detailed review of existing literature, focusing on the integration of conceptual and empirical approaches. Section 3 describes the methodology, including data collection and AI modeling techniques. Section 4 presents the results of the analysis, highlighting critical factors influencing survival probabilities. Section 5 discusses the implications of these findings in the context of existing literature, and Section 6 concludes with practical, societal, and research implications.

By combining conceptual rigor with AI-driven empirical analysis, this study contributes to the evolving understanding of startup survival, offering practical tools for stakeholders and advancing the integration of AI in entrepreneurial research.

2. Literature Review

Startup survival has been a central theme in entrepreneurship research due to its pivotal role in driving innovation and economic growth. Despite this focus, startups continue to face high failure rates, with approximately 90% ceasing operations within the first few years (Cantamessa et al., 2018) [1]. This persistent challenge underscores the need for more comprehensive methodologies to identify key factors that predict long-term success.

Traditional models for evaluating startup success have predominantly relied on static indicators such as financial ratios, team composition, and market fit (Cooper, 1990; Wernerfelt, 1984) [2,3]. Operational frameworks like the Stage-Gate System emphasize structured product development but often lack the flexibility required to adapt to dynamic market environments (Cooper, 1990) [2]. While these approaches provide valuable insights, their static nature limits their effectiveness in addressing the complexities of modern entrepreneurial ecosystems, where adaptability and innovation are crucial for survival.

Dynamic frameworks have emerged to address these limitations. The Resource-Based View (RBV) emphasizes internal capabilities as sources of competitive advantage but may overlook the significance of external collaboration and market volatility (Wernerfelt, 1984; Granstrand & Holgersson, 2020) [3,7]. Chesbrough's (2003) [6] Open Innovation model highlights the value of integrating external ideas and technologies to enhance scalability through collaboration. Additionally, Taleb's (2012) [8] concept of antifragility advocates for startups to not only withstand uncertainty but also thrive under adverse conditions. However, these models often lack practical frameworks for real-world application.

Advancements in artificial intelligence (AI) have introduced transformative tools for startup evaluation. Machine learning techniques, including neural networks and clustering

algorithms, facilitate the analysis of complex, nonlinear relationships, offering nuanced insights into startup trajectories (Nambisan, 2017; Sun et al., 2024) [4,5]. Mathematical tools like Lipschitz extensions further enhance predictive accuracy by ensuring stability and continuity across multidimensional datasets (Arnau et al., 2023; Blom & Mooij, 2020) [9,10]. These AI-driven methodologies address gaps in traditional models, enabling a more holistic evaluation of startup scalability.

Integrative frameworks such as the Business Model Navigator encourage adaptability by identifying patterns and recombining business models to meet evolving market demands (Gassmann et al., 2014) [11]. Similarly, the Platform Ecosystem Model explores how network effects can drive scalability in digital platforms (Eisenmann et al., 2006) [12]. While these approaches offer valuable insights, they often lack empirical validation in the context of startup survival, highlighting the necessity for robust datasets and structured evaluation frameworks.

Recent studies have sought to bridge these gaps by applying advanced predictive techniques to model startup survival. For instance, a study examined the effectiveness of Random Survival Forests, Cox proportional hazards models, and Gradient Boosting in predicting the duration of business activities among startups, finding that Gradient Boosting provided superior predictive capability (Fuentes-Callés, 2022) [13]. Another systematic review identified emerging success factors for startups, emphasizing the importance of adaptability and innovation in dynamic ecosystems (Levie & Lichtenstein, 2010) [14].

Despite these advancements, significant gaps remain in understanding startup scalability. Traditional models often fail to account for the dynamic interplay of internal and external factors, while emerging AI-driven techniques require robust datasets to ensure reliability. This study aims to bridge these gaps by integrating theoretical insights with empirical evidence, offering a comprehensive framework for evaluating startup survival and scalability.

Our goal is to provide an empirical approach to modeling the evolution of startups, focusing on five operational aspects that can be represented within a real-world observational framework. Synthesizing the concepts outlined in the theoretical discussion above, we introduce our Scaling Wheel Model, which is grounded in our professional experience supporting startups and assessing their survival probabilities. This assessment considers their capabilities, contextual variables, and other critical factors, processed through a conceptual analysis performed by an analyst from our team, who acts as an advisor to the startup.

A comprehensive explanation of the model is available in Font-Cot et al. (2023) [15]. For the purposes of this study, it is sufficient to note that the analysis revolves around the following dimensions: Team Dynamics, concerning the structure and interactions within the working team; Market Conditions, related to the specific circumstances of the market and its interaction with the new enterprise; Financial Metrics, focusing on economic data regarding the company and the market; Product and Service Differentiation, addressing the uniqueness and creativity of the startup's foundational idea; and Strategic Vision and Timing, which considers how the company interprets and navigates the market context.

In the remainder of this paper, we demonstrate how these theoretical ideas are translated into data and numerical insights. We also propose various mathematical and AI-based tools to process, analyze, and interpret the current status of the startup, as well as to estimate its likelihood of survival.

3. Methodology

3.1. Experimental Data Acquisition

The dataset used in this study comprises 20 startups, selected to represent a broad spectrum of industries, sizes, and market dynamics. These startups, founded between

2010 and 2024, span 14 years of technological and market evolution. This period captures significant changes in entrepreneurial ecosystems, offering a rich basis for analyzing survival trajectories. The data acquisition process was informed by the authors' decade-long engagement with startups, providing privileged access to context-rich information from both successful and failed ventures.

Data were collected using a multi-pronged approach. Structured surveys consisting of 81 items were administered to startup managers, focusing on key attributes critical to survival. These were supplemented by semi-structured interviews with founders, investors, and other stakeholders, offering qualitative depth. Observations during startup activities provided additional insights, while secondary data, including financial statements, funding histories, and market reports, ensured a comprehensive perspective. This triangulation of data sources aligns with best practices in startup research, enhancing reliability and contextual relevance.

The attributes collected were structured around the five dimensions of the Scaling Wheel Framework, encompassing team dynamics, market conditions, financial metrics, product and service differentiation, and strategic vision. Team dynamics, for example, were measured through indicators such as team size, leadership experience, goal alignment, and relational capital, reflecting their pivotal role in fostering scalability (Brinckmann & Kim, 2015; Banerji & Reimer, 2019) [16,17]. Market conditions captured external variables like market size, competitive intensity, and growth potential, while financial metrics focused on funding history, revenue streams, and resource accessibility (Davila & Foster, 2007) [18]. Product differentiation was assessed through innovation levels and adaptability, while strategic vision emphasized long-term planning and milestone achievements (Chesbrough, 2003; Porter, 1980) [6,19].

To ensure the data's reproducibility and integrity, qualitative responses were converted into numerical scores following a standardized coding protocol. Logs documenting the origin and nature of each attribute were maintained throughout the process. Validation of the dataset involved both qualitative and quantitative methods. Qualitative validation contextualized the metrics through expert feedback, while quantitative validation cross-referenced financial and operational data with publicly available records (Altman, 1968; Amat, 1990) [20,21]. These steps ensured that the dataset was both reliable and reflective of real-world startup dynamics.

Survival outcomes were expressed on a scale from 0.0 to 1.0, with 1.0 indicating fully successful startups that demonstrated sustained growth and market presence. Startups with partial success, characterized by challenges in financial stability or market adaptation, received values below 1.0, while failed startups that ceased operations entirely were assigned a value of 0.0.

3.2. *Mathematical Tools and Model Explanation*

The analysis utilized advanced mathematical tools to organize and interpret the dataset, enabling a nuanced understanding of startup survival. Principal Component Analysis (PCA) was employed as a preprocessing step to reduce the dimensionality of the dataset, identifying key features and facilitating clustering. This step ensured that the clustering process minimized intra-cluster variance while maximizing inter-cluster differences, providing clear groupings of startups based on their survival probabilities.

Three complementary predictive models were used to estimate survival probabilities. Lipschitz regression, a mathematically rigorous approach, operated within a metric space using McShane and Whitney formulas. This method ensured that predictions aligned with dataset continuity, making it particularly effective for sparse or uneven datasets (Ferrer-Sapena et al., 2020; Erdoğan et al., 2022) [22,23]. Neural networks captured complex,

nonlinear relationships, leveraging a simplified architecture to prevent overfitting given the dataset's size (Di Franco & Santurro, 2021) [24]. Linear regression provided a baseline, offering transparency in linking variables to survival probabilities while serving as a comparative benchmark (Dobson & Barnett, 2018) [25].

The three models were integrated into a machine learning framework: linear regression, Lipschitz regression, and neural networks are integrated into an adaptive machine learning framework. Initially, the survival score is computed as a convex combination of the three models.

$$g(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x) + \alpha_3 f_3(x),$$

where $\alpha_1 + \alpha_2 + \alpha_3 = 1$. As new data are added, the dataset is enriched with updated evaluations and survival outcomes. The functions f_1, f_2, f_3 are refined accordingly, and the coefficients $\alpha_1, \alpha_2, \alpha_3$ are optimized to minimize the overall quadratic error. This adaptive approach evolves with the dataset, enabling continuous improvement and serving as a foundation for reinforcement learning. By integrating these models, the methodology balances interpretability, flexibility, and predictive power, offering a comprehensive toolkit for understanding and simulating startup survival.

4. Results

This section presents the empirical findings of the study, organized according to the theoretical dimensions of the Scaling Wheel Framework: team dynamics, market conditions, financial metrics, product differentiation, and strategic vision. The results incorporate computational insights, including clustering, machine learning models, and innovation analysis, highlighting the contributions of these advanced methodologies to understanding startup survival. Our approach provides a unified framework for addressing the problem of startup survival from a methodological perspective, refining and concretizing the existing methodologies while also introducing new tools, such as Lipschitz regression, to enhance predictive accuracy (see Gangwani and Zhu, 2024) [26]. The main objective of the section is to present these results, as well as to show how the tools we have designed (the Scaling Wheel, but also the combined use of clustering techniques and Lipschitz extensions, and the other AI procedures explained here) can be adapted to different contexts. The idea is that these formal tools accompany the analysis of the results of the Scaling Wheel when used by the analyst in the elaboration of technical reports, adapting them then and choosing the most appropriate ones. That is why we put emphasis on the methodology, as well as on the results themselves.

4.1. Overview of the Analytical Tool and Its Results

The dataset comprises evaluations of 20 startups, collected through structured surveys, interviews, and secondary data. Each startup's attributes were translated into numerical vectors with 81 components, normalized on a scale from 0 to 1 to ensure consistency across variables. The fundamental premise is that the problem of startup survival can be approached as an experimental challenge (Blank and Dorf, 2020; Kerr et al., 2014) [27,28]. Principal Component Analysis (PCA) was applied to reduce dimensionality, facilitating the visualization of survival probabilities and enabling clustering analysis. Figure A1, which can be found in Appendix A, presents a two-dimensional representation of the dataset, showcasing survival probabilities along the principal components. This visualization captures the diversity in team dynamics, financial stability, and market adaptability, providing the foundation for further computational analysis, including feature extraction and predictive modeling.

Various classification methods previously tested for startup survival were considered as the starting point of the analysis, including the k-nearest neighbors algorithm, random

forest, extreme gradient boosting, support vector machine, and artificial neural networks (Koumbarakis and Volery, 2023) [29]. Finally, we opted to use standard clustering techniques, specifically k-means, to categorize startups into three groups: high, medium, and low survival probabilities. This process identifies patterns linking key attributes, such as team cohesion, innovation, and financial stability, to survival outcomes. The result can be seen in Figure 1. As can be seen in Figure 2, startups with high survival probabilities consistently demonstrated robust team dynamics and strong strategic visions, while medium-probability startups exhibited moderate strengths but often lacked innovation or differentiation. Low-probability startups faced challenges in financial stability and adaptability to market conditions, underscoring the critical role of resource allocation and strategic focus. These findings align with the IBM Triangle Framework, where investment, business features, and market adaptability converge to determine success (see, for example, Section 2 of Gangwani and Zhu, 2024 [30]).

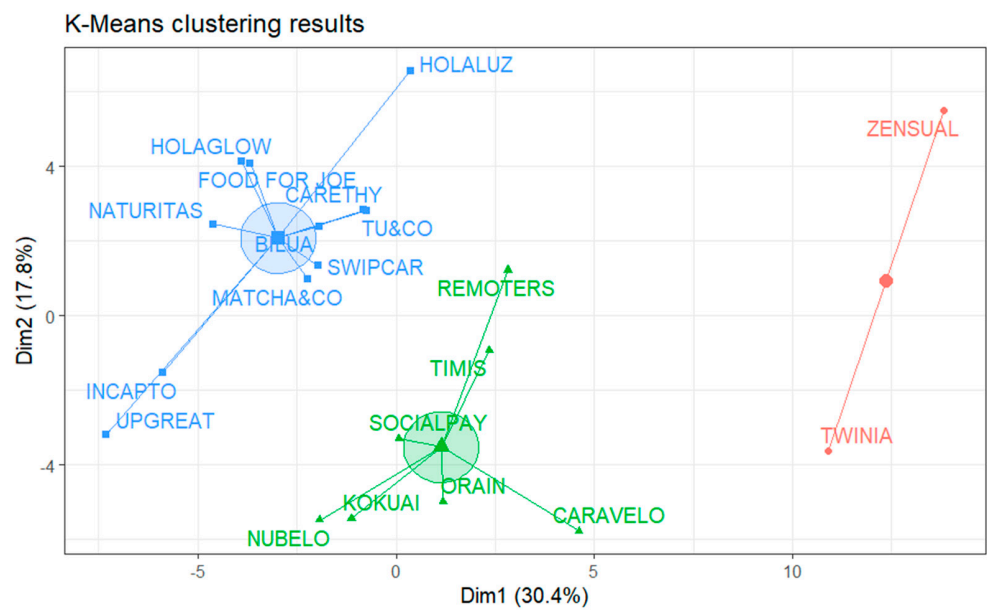


Figure 1. Clustering of startups into survival probability categories. K-means has been used, showing that the three groups appearing respond to different general marks on all the variables, as explained in the text.

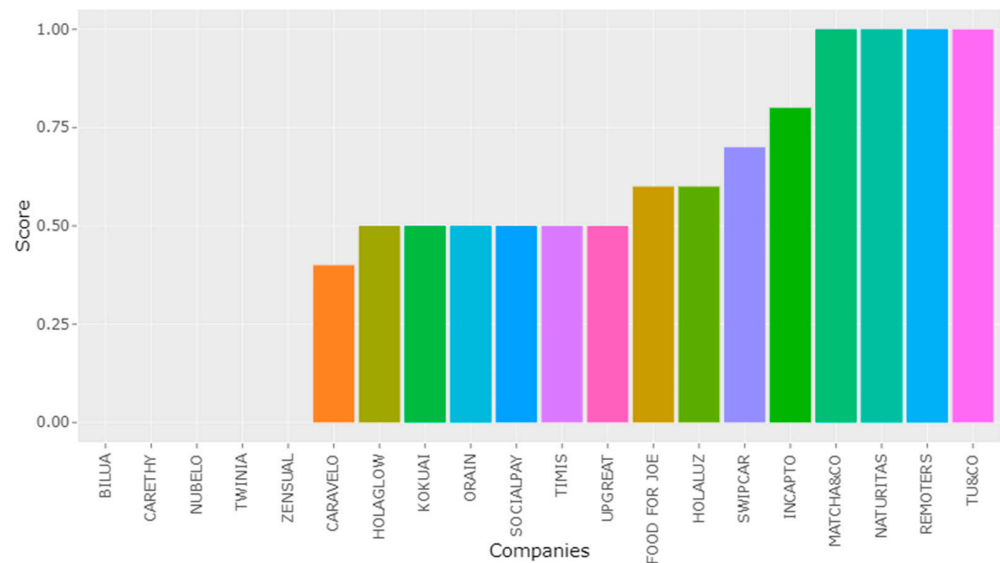


Figure 2. Histogram representation of the success ratio of the companies in the dataset.

The results of the clustering analysis provide positive information about the usefulness of the model. The cluster with a success estimate of less than 0.5 has two elements from each group, but it should be noted that the blue set is considerably larger. The vast majority of the firms in the blue set are in good health after the time interval considered and show survival probabilities above 0.5. The group with a score equal to 0.5 is split between the green (4) and blue (2) groups, but the proportion benefits the blue group. Finally, startups with ratings strictly above 0.5 up to 1 are almost all in the blue group. These results demonstrate that the clustering method offers a valuable first tool for the analysis of success. For the classification of any other company, we only have to introduce the results of the questionnaire in the database and see which group is automatically classified.

Three predictive models—Lipschitz regression, neural networks, and linear regression—were applied to estimate survival probabilities. Each model provided unique insights into the startup dataset. Lipschitz regression delivered stable and precise predictions, making it ideal for sparse datasets. Neural networks, which have already demonstrated success in this domain (Huang et al., 2024) [31], effectively captured complex and nonlinear relationships within the dataset. However, their performance exhibited variability, likely attributed to the limited sample size. Linear regression offered transparency and interpretability, although it struggled to model variable interactions effectively. The combination of these methods highlighted the multidimensional nature of startup survival analysis, as can be seen in Figure 3. Lipschitz regression emerged as the most consistent, while neural networks identified intricate patterns often missed by simpler models. The integration of these tools underscores the potential of ensemble approaches in predictive modeling.

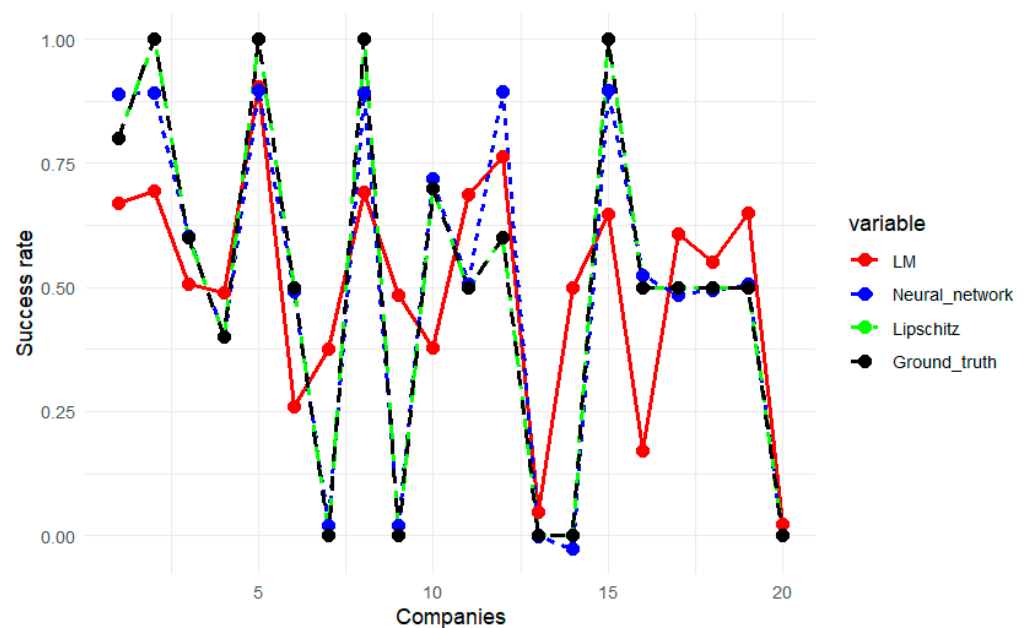


Figure 3. Representation of the scores of all elements of the dataset together with the estimates of the three referenced models. The ground truth points (black) cannot be seen because they are covered by the Lipschitz model (green), which fully matches them.

The Scaling Wheel Framework's dimensions were analyzed to determine their influence on survival probabilities. The findings emphasized the unique contribution of each dimension. Thus, the influence of survey components on survival predictions was analyzed by isolating responses for each block of questions (team dynamics, market conditions, product differentiation, financial metrics, and strategic vision). Figure 4 highlights the relative impact of these factors, which will be separately analyzed in the rest of this

section. Our conceptualization, although different from other comprehensive purposes, is essentially compatible with the mainstream of the valuation methods for startups. All of them coincide in recalling that Startup valuation is a crucial aspect of the entrepreneurial journey, requiring a deep understanding of the relevance of the startup team, dynamic ecosystem, and various valuation methods (e.g., DCF, Venture Capital, Scorecard), and their impact on fundraising and business growth while addressing challenges such as uncertainty and data limitations (Köseoğlu and Patterson, 2023) [32].

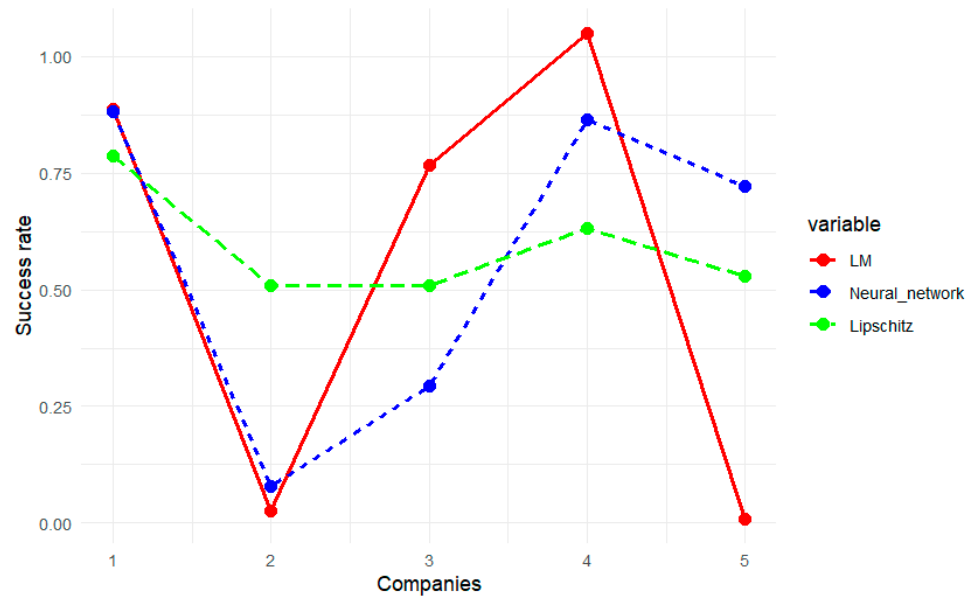


Figure 4. Estimates of the relevance of each block of questions in the final success index.

4.2. Team Dynamics (Block 1)

The sets of variables related to team dynamics emerged from the data analysis as the most significant factor influencing startup survival. This dimension emphasizes the importance of a well-prepared, cohesive team with effective leadership and strong relational capital. Studies such as McCarthy et al. (2023) [33] have shown that the personality traits of founders and diversity within teams are critical determinants of startup success. Their findings align closely with the role of team dynamics as prioritized by the Scaling Wheel Framework, reinforcing the notion that internal cohesion is fundamental for scalability. Other studies reveal that variables such as educational background (STEM or arts), entrepreneurial experience, and diversity within the team play a significant role. For example, the number of organizations founded by individuals within a team is a key predictor of success (Thirupathi et al., 2021) [34]. These and other factors can be detected by the analyst when the Scaling Wheel interviews are applied; the results obtained reveal the importance of this information, becoming the most important variable to take into account.

Another factor that directly affects the survival of a startup is the relationship between the founders and obtaining financial resources. Although our model has a specific point that relates the success of the startup to financial issues, it is clear that good initial funding could be a consequence of the previous experience of the partners, which is reflected in their ability to find good investors. A comprehensive discussion on the topic, related to the financial opportunities opened up by the precedence and previous experience of team members (which we refer to as Team Dynamics), can be found in Dworak, 2022 [35] (see also the references therein). We will give more clues on this topic in the next section dedicated to Financial Metrics.

4.3. Market Conditions (Block 2)

Market conditions, while influential, were less critical compared to internal factors such as team and product differentiation in our analysis. Although this is a controversial issue, as far as startup survival is concerned, we emphasize that although external market forces, such as competition and market size, affect startups, intrinsic characteristics, such as innovativeness and team resilience, often carry more weight than these external variables (Risku, 2021) [36]. This finding aligns with the relatively lower influence of market conditions observed in the Scaling Wheel analysis.

In addition, startups with resilient teams capable of managing change were more likely to endure in exceptional circumstances (e.g., during the COVID-19 pandemic), underscoring the importance of team resilience in overcoming external challenges, in the face of variables representing market conditions (Polese et al., 2022) [37].

4.4. Product and Service Differentiation (Block 3)

Most of the documents that can be found in the Scientific literature reveal that this block of variables plays a vital role in determining survival probabilities, highlighting the need for market-relevant innovation. Research by Sun et al. (2024) [5] demonstrates that innovative and differentiated offerings increase a startup's ability to capture market share and adapt to competitive pressures. Their use of machine learning models to predict startup success mirrors the Scaling Wheel Framework's emphasis on product differentiation.

However, although in the primary information we obtain from our model, this block is moderately relevant, the mathematical models give less importance to its influence on the final survival of the startup. It should be noted that although our method is intended to be globally applicable, the testing procedure we present in this paper is restricted to the local ecosystem of companies that have been under the surveillance of our analysts for several years. This block of variables may become principal if a broader analytical project is designed, but, as we have explained in the paper, our methodology has been created as a tool for local analysis of startup projects and not as a global analytical instrument for scientific, strategic, or foresight research.

4.5. Financial Metrics (Block 4)

Financial metrics related to the economic situation of the startup and the chances of getting good financing for the project are, of course, fundamental in the design of a new startup. As stated by Fuertes-Callén et al. (2022) [13], startups with healthier early financial indicators, such as profitability, liquidity, and manageable debt levels, exhibit significantly higher survival rates, with these metrics continuing to influence their viability for up to eight years. This study and other related analyses demonstrate the importance of financial stability to the survival of new ventures, and resource allocation and funding are critical to long-term viability. This research supports the inclusion of financial metrics as a key dimension of the Scaling Wheel Framework. Moreover, from the standpoint of technical mathematical modeling of probabilistic prediction of startup success (Gujarathi et al. 2024) [30], this is a critical variable.

Also, as explained in Section 4.2, the startup team is a critical factor for survival, particularly in securing financial funding at an early stage. Readers interested in this topic can find extensive information on the role of team members in acquiring the necessary financial resources to initiate startup activities in McCarthy et al., 2023 [33], and the references therein.

4.6. Strategic Vision and Timing (Block 5)

The analysis provided by the Scaling Wheel revealed high variability in the influence of strategic factors and timing, indicating a complex and context-dependent relationship with survival outcomes. As a result, strategic matters do not appear to constitute a fundamental variable in the framework, likely due to their relative weight when compared to other factors.

Granstrand and Holgersson (2020) [7] highlight that a startup's strategic alignment with market opportunities and the timing of its market entry can substantially influence its success. However, they note that this impact depends on industry-specific dynamics and external conditions. These findings are consistent with the Scaling Wheel Framework, which also emphasizes the situational and context-driven nature of strategic vision.

These results align with computational analyses, like machine learning models, which rank team dynamics and financial metrics (closely tied to team properties) as more important than external factors. The Scaling Wheel Framework brings these elements together in a unified approach, offering valuable insights into the complex nature of startup survival. This blend of theory and real-world data highlights the importance of using different analytical methods to better understand what drives entrepreneurial success.

5. Discussion

The findings of this study underscore the strengths and versatility of the proposed multivariate AI-based model in evaluating startup survival. By combining empirical insights with conceptual frameworks, the model addresses key gaps in the literature and offers actionable tools for stakeholders in entrepreneurial ecosystems. This discussion elaborates on the model's strengths, theoretical and empirical contributions, and its relevance to existing literature, presenting a holistic view of its impact.

5.1. Strengths of the Proposed Model

The proposed model demonstrates exceptional predictive capabilities, particularly through its integration of Lipschitz regression, linear models, and neural networks. Lipschitz regression stands out for its accuracy and stability, offering consistent predictions even in datasets with sparse or uneven distributions. In contrast, neural networks excel at capturing nonlinear relationships, albeit with variability that necessitates careful calibration. The inclusion of linear regression adds interpretability, serving as a transparent baseline for comparison. Together, these elements achieve a balance between simplicity, flexibility, and precision.

The adaptive nature of the framework further enhances its utility. Stakeholders can derive actionable insights through interpretable outputs, scale the model to accommodate datasets of varying sizes and complexities, and leverage clustering techniques to effectively categorize startups based on survival probabilities. Unlike traditional approaches, such as financial ratio analyses (Altman, 1968) [20] or static market assessments (Porter, 1980) [19], this framework captures dynamic and nonlinear relationships, offering a nuanced understanding of survival determinants. The stability of the Lipschitz model, in particular, provides a reliable foundation for decision-making in environments characterized by data scarcity or high variability.

5.2. Theoretical Contributions

This study bridges mathematical rigor with conceptual insights, advancing the theoretical discourse on startup survival. By incorporating Lipschitz extensions and neural networks, the framework aligns with calls for AI-driven methodologies that enhance strate-

gic foresight in innovation ecosystems (Granstrand and Holgersson, 2020) [7]. The study contributes to and extends several prominent theoretical perspectives.

The Resource-Based View (RBV) is reinforced through the emphasis on team dynamics and internal capabilities as critical factors for survival, consistent with Wernerfelt's (1984) [3] seminal work. Open Innovation, as articulated by Chesbrough (2003) [6], finds support in the model's validation of adaptability and external collaboration as key drivers of scalability. Furthermore, the study echoes Taleb's (2012) [8] concept of antifragility, advocating for startups to cultivate resilience and adaptability in the face of market uncertainties. By empirically validating these frameworks, the model bridges theoretical constructs with practical applications, demonstrating their relevance to real-world entrepreneurial challenges.

5.3. Empirical Contributions

The empirical findings of this study validate the practical relevance of the Scaling Wheel Framework and provide critical insights into the factors influencing startup survival. Team dynamics emerged as the most significant predictor, emphasizing the importance of leadership alignment, relational capital, and cohesive team structures. These findings align with prior research highlighting the role of team cohesion in fostering scalability (Brinckmann and Kim, 2015; Banerji and Reimer, 2019) [16,17].

Product differentiation also played a pivotal role, underscoring the necessity of innovation and adaptability in competitive ecosystems. This result reinforces the perspective that startups must establish unique market positions to sustain their growth (Zhang et al., 2021) [38]. Market conditions and strategic vision exhibited context-dependent impacts, reflecting the variability of external dynamics and the criticality of timing in entrepreneurial success. The robustness of the proposed model, particularly its ability to integrate multidimensional datasets and preserve continuity across metrics, further supports its empirical contributions. Lipschitz regression demonstrated exceptional performance in maintaining stability and coherence, even when handling complex data relationships (Arnau et al., 2023) [9].

5.4. Contributions to the Literature

This study makes significant contributions to the literature by advancing AI-driven methodologies for evaluating startups. Recent research has highlighted the need for dynamic, data-driven frameworks that go beyond traditional methods (Park et al., 2024) [39], which emphasize the utility of predictive techniques like neural networks in modeling survival probabilities. Building on this, our work incorporates clustering and Lipschitz regression to enhance stability and interpretability, offering a more comprehensive analytical toolkit. Techniques such as random forest, XGBoost, and support vector machines could help with this task (see Shi et al., 2024 [40], and the references therein).

Similarly, several investigations have identified adaptability as a core success factor in startup ecosystems (Sevilla-Bernardo et al., 2022) [41]. Our findings reinforce this perspective, particularly through the lens of team alignment and innovation. By combining empirical data with conceptual insights, this study extends existing knowledge, addressing complex entrepreneurial challenges with a multifaceted approach. The integration of adaptive learning mechanisms and the emphasis on practical implications further position this research as a valuable contribution to the evolving discourse on startup survival.

5.5. Practical Implications

The findings of this study have significant practical implications for entrepreneurs, investors, and policymakers, addressing key challenges in evaluating and fostering startup survival. For entrepreneurs, the framework provides actionable insights into the most critical dimensions of success, particularly team dynamics and product differentiation. By

focusing on leadership alignment, team cohesion, and innovation, entrepreneurs can make strategic decisions to mitigate risks and enhance their scalability potential. For example, startups can use the model to identify weaknesses in team capabilities or gaps in market differentiation, allowing them to address these areas proactively.

Investors can benefit from the model's predictive capabilities by incorporating its outputs into their decision-making processes. The multivariate AI approach allows investors to assess survival probabilities with greater nuance, moving beyond static indicators like financial ratios. This deeper understanding enables better portfolio diversification, risk assessment, and investment targeting, particularly in identifying early-stage startups with strong long-term potential. Additionally, the framework's transparency and adaptability ensure that investors can apply it across various industries and market conditions.

Policymakers can leverage the insights provided by the study to design targeted interventions and support mechanisms for startups. By understanding the role of team dynamics, innovation, and strategic timing in driving success, policymakers can develop programs that address specific ecosystem weaknesses. For instance, initiatives to foster leadership training, improve access to funding for innovative projects, or support collaborative networks could significantly enhance startup viability. Furthermore, the framework's scalability makes it suitable for application at regional or national levels, enabling the creation of policies tailored to the unique characteristics of local entrepreneurial ecosystems.

5.6. Limitations and Future Research

While this study provides valuable contributions to understanding startup survival, several limitations must be acknowledged, paving the way for future research. One major limitation is the dataset size, which consists of 20 startups. Although these startups were selected to represent diverse industries and stages of development, the sample size constrains the generalizability of the findings. Future studies should incorporate larger, more diverse datasets, including startups from different regions, industries, and time frames, to validate and extend the applicability of the model.

Another limitation is the exclusion of external variables, such as macroeconomic conditions, cultural influences, and regulatory environments. These factors often play a significant role in shaping entrepreneurial success but were beyond the scope of the current study. Future research should integrate such variables to provide a more comprehensive analysis, enabling models to account for external dynamics and their interactions with internal startup attributes.

The current framework also lacks real-time adaptability. While the adaptive learning mechanism allows for continuous improvement as new data are added, the model does not yet incorporate real-time data streams. Future work could focus on developing interactive, real-time applications of the framework, potentially through online platforms. Such platforms could allow startups, investors, and policymakers to input live data and receive dynamic survival predictions, further bridging the gap between theoretical insights and practical decision-making.

Additionally, exploring advanced AI techniques, such as hybrid models or ensemble learning, could enhance the model's predictive accuracy and robustness. These methods could combine the strengths of existing predictive tools while minimizing their weaknesses. For instance, reinforcement learning could be applied to continuously optimize the model as new data becomes available, ensuring that it evolves in response to changes in entrepreneurial ecosystems.

Lastly, future research should explore the societal implications of startup survival, considering how successful startups contribute to broader economic and social outcomes. This could include analyzing the impact of startups on job creation, innovation diffusion,

and regional development. By linking survival predictions to societal benefits, the framework could provide even greater value to stakeholders and contribute to a more holistic understanding of entrepreneurship.

These expansions provide a deeper exploration of practical applications and future research opportunities, addressing potential gaps while setting the stage for continued advancements in the field. Let me know if you'd like further refinements or additional areas of focus!

6. Conclusions

This study demonstrates the application of multivariate models to analyze and predict the survival probabilities of startups using structured datasets. By employing clustering techniques and predictive models, including linear regression, neural networks, and Lipschitz functions, the research evaluates the feasibility and effectiveness of diverse methodologies for estimating survival outcomes. The findings highlight the clustering method's ability to categorize startups into groups with distinct probabilities of success, offering a practical approach for initial evaluations and enabling the seamless integration of additional startups into the framework. This approach ensures scalability and adaptability, making it a valuable tool for dynamic startup ecosystems.

The predictive models used in the study each bring unique strengths to the analysis. Lipschitz regression demonstrates its value by providing stable and conservative predictions, minimizing variability and the risk of overfitting. Neural networks, while capable of capturing complex, nonlinear relationships, show greater variability, particularly with sparse datasets. Linear regression serves as an interpretable baseline model, offering simplicity without compromising the ability to assess key trends. The integration of these models through a convex approach enhances overall predictive reliability and accuracy, illustrating the value of combining complementary methodologies to address the complexities of startup survival.

The analysis of the survey's conceptual blocks further underscores the importance of team dynamics as the most significant predictor of survival. The findings align with previous research emphasizing the critical role of leadership, goal alignment, and relational capital in fostering scalability. Product and service differentiation also emerges as a key factor, highlighting the necessity of innovation and adaptability to maintain competitiveness in dynamic markets. In contrast, market conditions appear to play a less pivotal role, while strategic vision and timing exhibit context-dependent variability, reflecting the nuanced nature of their impact on survival outcomes.

The findings of this study carry practical implications for multiple stakeholders. Entrepreneurs can leverage the framework to identify and strengthen critical success factors, particularly those related to team alignment and innovation. Investors can use the model to evaluate potential investments, gaining insights into multidimensional survival probabilities and associated risks. Policymakers, too, can benefit by employing data-driven recommendations to design targeted interventions that support startup ecosystems and enhance overall entrepreneurial success rates.

Despite its contributions, this study is not without limitations. The dataset, limited to 20 startups, constrains the generalizability of the findings. Future research should address this by expanding the dataset to include more diverse industries and geographical contexts, thereby increasing the robustness of the analysis. Additionally, external variables such as macroeconomic conditions, regulatory environments, and cultural influences, which were not included in the current framework, should be explored to provide a more comprehensive understanding of the factors affecting startup survival. Incorporating real-

time data and adaptive learning mechanisms could further refine the model’s scalability and enhance its predictive accuracy, enabling a more dynamic and responsive framework.

By integrating structured evaluation tools with advanced AI methodologies, this study contributes to the growing body of research on startup survival. The framework presented is scalable, adaptable, and practical, addressing the complexities of entrepreneurial ecosystems and offering a foundation for more accurate and nuanced decision-making. The results not only enhance the understanding of startup survival dynamics but also provide a pathway for future studies to explore innovative, data-driven approaches to addressing the challenges faced by startups in rapidly evolving markets.

Author Contributions: Conceptualization, F.F.-C., P.L.-N. and E.A.S.-P.; Methodology, F.F.-C., P.L.-N. and C.S.-A.; Validation, F.F.-C.; Formal analysis, C.S.-A.; Investigation, F.F.-C., P.L.-N., C.S.-A. and E.A.S.-P.; Resources, F.F.-C. and P.L.-N.; Writing—original draft, F.F.-C.; Writing—review & editing, E.A.S.-P.; Supervision, P.L.-N. and E.A.S.-P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by University Grants Management Agency of Generalitat de Catalunya grant number 2022DI086.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Some figures on how the mathematical model works. Figures A1–A3 represent different features of the model. The corresponding explanation can be found below each figure.

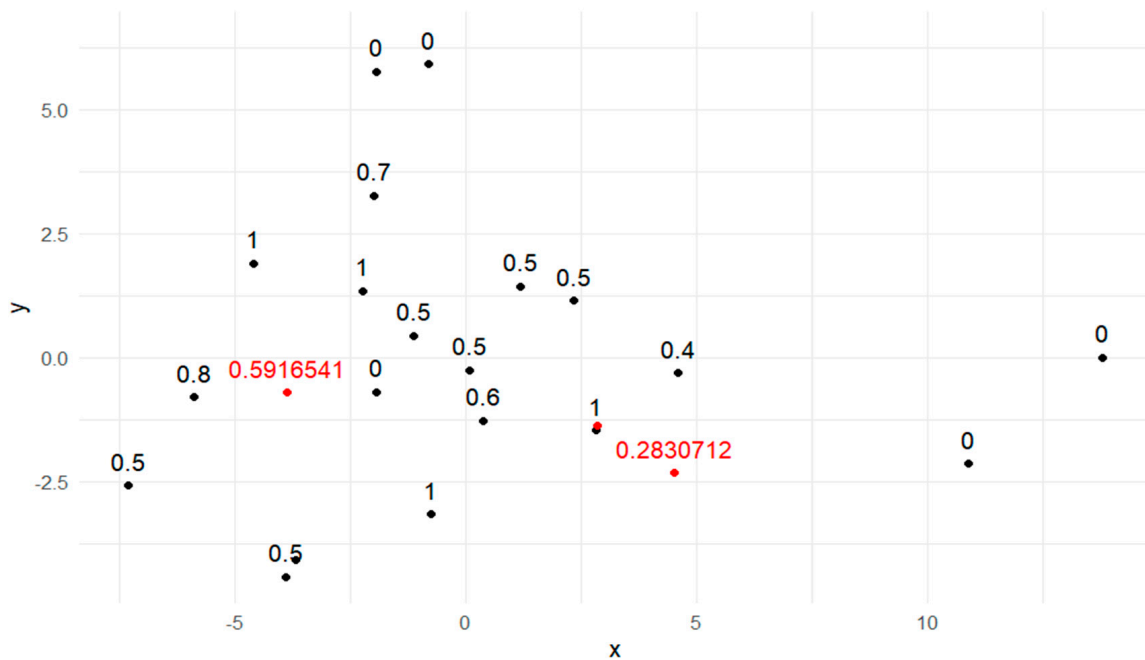


Figure A1. Representation of an example of the classification of three random vectors (red) together with the initial database using PCA. These vectors represent imaginary companies and can be seen together with the original points of the data set (20 evaluated startups). The label shows the final estimated rate obtained with the explained Lipschitz model.

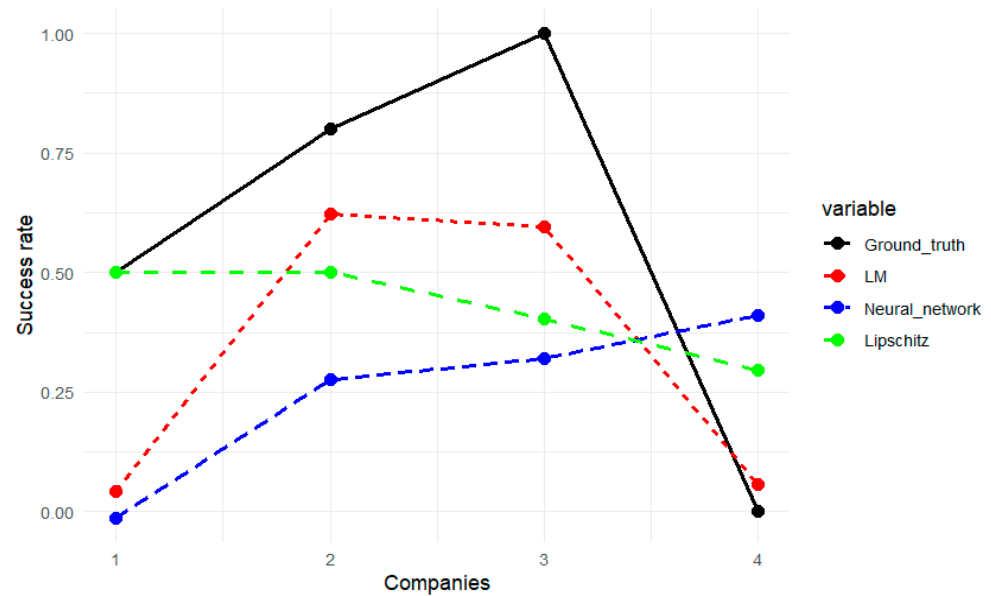


Figure A2. Model testing. Estimated values of 4 real companies provided by each of the three models considered (Linear LM, Neutral Network, and Lipschitz Regression) when we train them with the rest of the 16 benchmark start-ups.

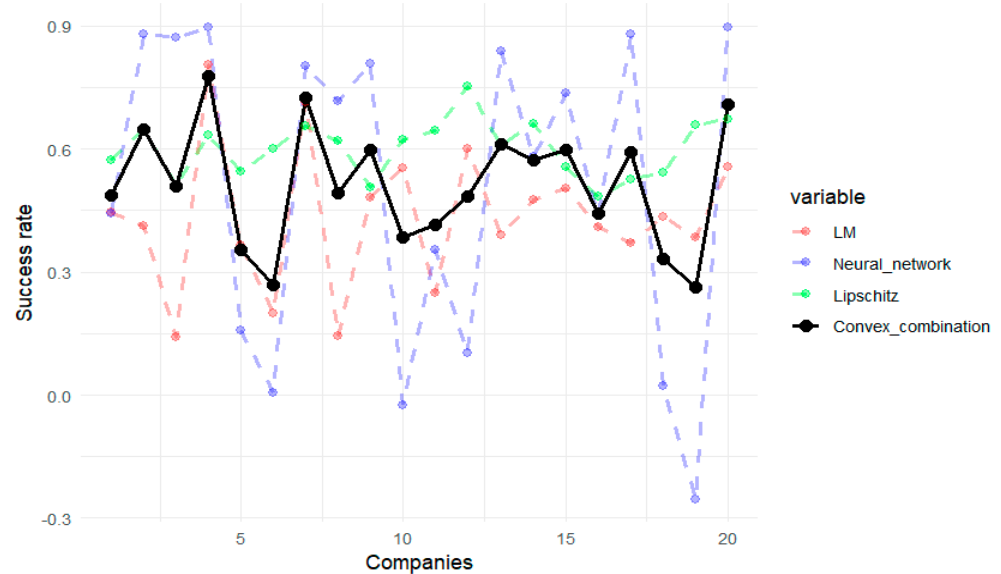


Figure A3. Convex combination of the three models with equally distributed coefficients (black). It can be used as an integrated model for forecasting to estimate the survival probability when applied to a new startup.

References

1. Cantamessa, M.; Gatteschi, V.; Perboli, G.; Rosano, M. Startups' roads to failure. *Sustainability* **2018**, *10*, 2346. [[CrossRef](#)]
2. Cooper, R.G. Stage-gate systems: A new tool for managing new products. *Bus. Horiz.* **1990**, *33*, 44–54. [[CrossRef](#)]
3. Wernerfelt, B. A resource-based view of the firm. *Strateg. Manag. J.* **1984**, *5*, 171–180. [[CrossRef](#)]
4. Nambisan, S. Digital entrepreneurship: Toward a digital technology perspective of entrepreneurship. *Entrep. Theory Pract.* **2017**, *41*, 1029–1055. [[CrossRef](#)]
5. Sun, X.; Abdullahi Usman, M. Drivers of platform ecosystem adoption: Does innovation capability translate these drivers into improved firm performance. *Bus. Process Manag. J.* **2025**, *31*, 118–145. [[CrossRef](#)]
6. Chesbrough, H.W. *Open Innovation: The New Imperative for Creating and Profiting from Technology*; Harvard Business Press: Brighton, MA, USA, 2003.
7. Granstrand, O.; Holgersson, M. Innovation ecosystems: A conceptual review and a new definition. *Technovation* **2020**, *90*, 102098. [[CrossRef](#)]

8. Taleb, N.N. *Antifragile: Things That Gain from Disorder*; Random House: New York, NY, USA, 2012.
9. Arnau, R.; Calabuig, J.M.; Erdogan, E.; Sánchez Pérez, E.A. Extension procedures for lattice Lipschitz operators on Euclidean spaces. *Rev. Real Acad. Cienc. Exactas Fis. Nat. Ser. A-Mat.* **2023**, *117*, 76. [[CrossRef](#)]
10. Blom, T.; Mooij, J.M. Robustness of model predictions under extension. *arXiv* **2020**, arXiv:2012.04723.
11. Gassmann, O.; Frankenberger, K.; Csik, M. *The Business Model Navigator: 55 Models That Will Revolutionise Your Business*; Pearson: London, UK, 2014.
12. Eisenmann, T.; Parker, G.; Van Alstyne, M. Strategies for two-sided markets. *Harv. Bus. Rev.* **2006**, *84*, 92–101.
13. Fuertes-Callén, Y.; Cuellar-Fernández, B.; Serrano-Cinca, C. Predicting startup survival using first years financial statements. *J. Small Bus. Manag.* **2022**, *60*, 1314–1350. [[CrossRef](#)]
14. Levie, J.; Lichtenstein, B.B. A terminal assessment of stages theory: Introducing a dynamic states approach to entrepreneurship. *Entrep. Theory Pract.* **2010**, *34*, 317–350. [[CrossRef](#)]
15. Font-Cot, F.; Lara-Navarra, P.; Serradell-Lopez, E. Digital transformation policies to develop an effective startup ecosystem: The case of Barcelona. *Transform. Gov. People Process Policy* **2023**, *17*, 344–355. [[CrossRef](#)]
16. Banerji, D.; Reimer, T. Startup founders and their LinkedIn connections: Are well-connected entrepreneurs more successful? *Comput. Hum. Behav.* **2019**, *90*, 46–52. [[CrossRef](#)]
17. Brinckmann JA, N.; Kim, S.M. Why we plan: The impact of nascent entrepreneurs' cognitive characteristics and human capital on business planning. *Strateg. Entrep. J.* **2015**, *9*, 153–166. [[CrossRef](#)]
18. Davila, A.; Foster, G. Management control systems in early-stage startup companies. *Account. Rev.* **2007**, *82*, 907–937. [[CrossRef](#)]
19. Porter, M.E. *Competitive Strategy: Techniques for Analyzing Industries and Competitors*; Free Press: Los Angeles, CA, USA, 1980.
20. Altman, E.I. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *J. Financ.* **1968**, *23*, 589–609. [[CrossRef](#)]
21. Amat, O. *Análisis de Estados Financieros; Gestión 2000*: Barcelona, Spain, 1990.
22. Erdoğan, E.; Ferrer-Sapena, A.; Jiménez-Fernández, E.; Sánchez-Pérez, E.A. Index spaces and standard indices in metric modelling. *Nonlinear Anal. Model. Control* **2022**, *27*, 803–822. [[CrossRef](#)]
23. Ferrer-Sapena, A.; Erdogan, E.; Jiménez-Fernández, E.; Sánchez-Pérez, E.A.; Peset, F. Self-defined information indices: Application to the case of university rankings. *Scientometrics* **2020**, *124*, 2443–2456. [[CrossRef](#)]
24. Di Franco, G.; Santurro, M. Machine learning, artificial neural networks and social research. *Qual. Quant.* **2021**, *55*, 1007–1025. [[CrossRef](#)]
25. Dobson, A.J.; Barnett, A.G. *An Introduction to Generalized Linear Models*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2018.
26. Gangwani, D.; Zhu, X. Modeling and prediction of business success: A survey. *Artif. Intell. Rev.* **2024**, *57*, 44. [[CrossRef](#)]
27. Blank, S.; Dorf, D. *The Startup Owner's Manual: The Step-by-Step Guide for Building a Great Company*; John Wiley & Sons: Hoboken, NJ, USA, 2020.
28. Kerr, W.R.; Nanda, R.; Rhodes-Kropf, M. Entrepreneurship as experimentation. *J. Econ. Perspect.* **2014**, *28*, 25–48. [[CrossRef](#)]
29. Koumbarakis, P.; Volery, T. Predicting new venture gestation outcomes with machine learning methods. *J. Small Bus. Manag.* **2023**, *61*, 2227–2260. [[CrossRef](#)]
30. Gujarathi, A.; Shukla, T.; Nirban, V. Probabilistic Prediction for a Start-Up Success Through Bayesian Networks-Based Machine Learning Approach. In Proceedings of the 2024 International Conference on Emerging Innovations and Advanced Computing (INNOCOMP), Sonipat, India, 25–26 May 2024; IEEE: Piscataway, NJ, USA, 2024; pp. 513–518.
31. Huang, K.; Zhou, Y.; Yu, X.; Su, X. Innovative entrepreneurial market trend prediction model based on deep learning: Case study and performance evaluation. *Sci. Prog.* **2024**, *107*, 00368504241272722. [[CrossRef](#)] [[PubMed](#)]
32. Köseoğlu, S.D.; Patterson, A. Introduction to startup valuation: From idea to, I.P.O. In *A Practical Guide for Startup Valuation: An Analytic Approach*; Springer Nature: Cham, Switzerland, 2023; pp. 7–42.
33. McCarthy, P.X.; Gong, X.; Stephany, F.; Braesemann, F.; Rizoio, M.A.; Kern, M.L. The science of startups: The impact of founder personalities on company success. *arXiv* **2023**, arXiv:2302.07968. [[CrossRef](#)]
34. Thirupathi, A.N.; Alhanai, T.; Ghassemi, M.M. A machine learning approach to detect early signs of startup success. In Proceedings of the Second ACM International Conference on AI in Finance, Virtual Event, 3–5 November 2021; pp. 1–8.
35. Dworak, D. Analysis of Founder Background as a Predictor for Start-Up Success in Achieving Successive Fundraising Rounds. Doctoral Thesis, University of Michigan, Ann Arbor, MI, USA, 2022.
36. Risku, J. Improving the performance of early-stage software startups: Design and creativity viewpoints. *arXiv* **2021**, arXiv:2108.00521.
37. Polese, F.; Sirianni, C.A.; Guazzo, G.M. How Startups Attained Resilience During COVID-19 Pandemic Through Pivoting: A Case Study. In *The International Research & Innovation Forum*; Springer International Publishing: Cham, Switzerland, 2022; pp. 519–527.
38. Zhang, J.; Yu, B.; Lu, C. Exploring the effects of innovation ecosystem models on innovative performances of startups: The contingent role of open innovation. *Entrep. Res. J.* **2021**, *13*, 1139–1168.

39. Park, J.; Choi, S.; Feng, Y. Predicting startup success using two bias-free machine learning: Resolving data imbalance using generative adversarial networks. *J. Big Data* **2024**, *11*, 122. [[CrossRef](#)]
40. Shi, Y.; Eremina, E.; Long, W. Machine learning models for early-stage investment decision making in startups. *Manag. Decis. Econ.* **2024**, *45*, 1259–1279. [[CrossRef](#)]
41. Sevilla-Bernardo, J.; Sanchez-Robles, B.; Herrador-Alcaide, T.C. Success factors of startups in research literature within the entrepreneurial ecosystem. *Adm. Sci.* **2022**, *12*, 102. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.