STOCKHOLM SCHOOL OF ECONOMICS
1351 Master's thesis in Business Management
Academic year 2025–2026

Evaluating Founder Signals Under Uncertainty: A Predictive-Modeling Framework Using LLM Policy Induction and Gradient Boosting

Felix Fröndhoff (42804) and Bror Nordström (42793)

This thesis investigates whether founder human- and social-capital signals retain predictive value across time when forecasting early-stage venture funding outcomes. Addressing recent calls for replication and temporal validation in signaling research, the study develops a predictive-modeling framework that combines a novel LLM-based policy-induction model with established machine-learning methods. Using a dataset of more than 50,000 startup founders in Europe and North America, the models are trained on the five yearly cohorts preceding each prediction year (2012–2019) and then evaluated on later, strictly unseen cohorts (2018–2022) to test temporal generalizability under rigorous out-of-sample conditions. The results demonstrate strong predictive stability across cohorts and reveal three consistent patterns. First, prior entrepreneurial experience remains a persistent and influential predictor of high funding attainment. Second, Big Tech experience functions as a rare but highly rewarded signal, consistently outperforming academic credentials such as PhD training or research roles. Third, academic credentials contribute positively only when embedded within strong execution-oriented profiles, and never as standalone drivers of prediction. They function as weak, conditionally relevant signals that the models treat as secondary to operational experience rather than as sources of genuine configurational synergy. Methodologically, the thesis displays how predictive modeling and interpretable LLM-based reasoning can be combined to evaluate theoretical mechanisms with temporal rigor.

Keywords: Venture Capital, Human and Social Capital Signals, Predictive Modeling, Machine Learning Models, LLM based Policy Induction

Supervisor: Sebastian Krakowski
Date examined: 16th December 2025
Discussant(s): Ella Appelholm and Sara Boras
Examiners: Constanze Eib and Roberto Verganti

Acknowledgements

We would like to extend our heartfelt thanks to **Professor Sebastian Krakowski**, whose guidance and thoughtful supervision have been instrumental throughout the journey of this thesis. Our deep appreciation also goes to **Professor Rickard Sandberg**, whose methodological insights greatly strengthened the rigor and clarity of our research. Both Sebastian and Rickard have offered invaluable support, not only in shaping this academic work but also in advising us on our professional paths ahead.

We are equally thankful to **Professor Susanne Sweet** and **Professor Karl Wennberg** for leading the thesis seminars and for creating the academic foundation that enabled us to carry out this project. We would also like to extend our sincere thanks to **Alexander Fred-Ojala** for his continuous support and constructive feedback during the research process. Finally, we wish to express our gratitude to the **Stockholm School of Economics** for its inspiring academic environment, comprehensive training, and seamless logistical support, all of which made both this thesis process and the past two years a deeply enriching and memorable experience.

Abbreviations

Abbreviation	*Meaning*	*Relevance / Usage Context*
AI	Artificial Intelligence	Used in predictive modeling context, particularly in large-language-model applications.
AUC/ AUCPR	Area Under the Curve / Area Under the Precision–Recall Curve	Key model performance metrics for classification and imbalanced datasets.
ROC–AUC	Quantifies how well a classification model separates true positives from false positives across all thresholds.	Used to evaluate overall model discrimination performance, especially for comparing predictive accuracy across models.
ECE	Expected Calibration Error	Metric evaluating how well predicted probabilities align with observed outcomes.
$F_{0.5}$	F0.5 Score	Weighted harmonic mean of precision and recall; emphasizes precision (used due to high false-positive cost in venture contexts).
LLM	Large Language Model	Central to the thesis' methodological innovation - used for structured information extraction and policy induction.
ML	Machine Learning	Core method in predictive modeling and feature extraction.
PDL	People Data Labs	Data source for founder profiles and career information.
RAG	Retrieval-Augmented Generation	Used in extracting institutional affiliations (e.g., universities, accelerators).
RF	Random Forest	One of the predictive ensemble models used.

GB / GBT	Gradient Boosting / Gradient Boosting Tree	Model type for high-dimensional, nonlinear data.
SHAP	Shapley Additive ExPlanations	Model-interpretability technique used to evaluate feature importance and direction.
VC	Venture Capital	Core domain context for funding outcomes and signal interpretation.
T_0	Founding Year Cutoff	The reference point for data truncation - ensures only pre-founding information is included.

Vocabulary (Conceptual & Methodological Core)

1. Theoretical and Conceptual Terms

Information Asymmetry - A market condition where one party (e.g., founders) has more information than another (e.g., investors).

Signaling Theory - Framework explaining how founders communicate private information credibly through observable, costly signals.

Human Capital - Founders accumulated knowledge, education, and experience that signal competence and execution ability.

Social Capital - Founders' networks, affiliations, and reputational ties signaling legitimacy and trustworthiness.

Separating Equilibrium - A state where credible signals allow differentiation between high- and low-quality actors.

Predictive Validation - Testing whether theoretical relationships hold in new, unseen data (evaluates theory's external validity).

Temporal Stability - The consistency of predictive relationships (signal effectiveness) across time.

Configurational Effects - Interactions or combinations of signals that jointly influence outcomes (beyond additive effects).

2. Methodological Terms

Predictive Modeling - Using statistical or ML models to test whether theoretical constructs generalize across contexts.

Out-of-Sample Evaluation - Assessing model performance on data not used for training, ensuring genuine predictive testing.

Feature Extraction - Converting unstructured founder data (e.g., LinkedIn) into structured, theory-aligned variables.

Feature Importance - Quantitative ranking of how much each variable contributes to predictive accuracy.

SHAP Analysis - Decomposition of individual predictions to reveal how each feature influences model output.

Temporal Holdout Design - Method of training on earlier cohorts and testing on later ones to measure generalization over time.

Rolling-Origin Validation - Iterative training/testing method that prevents information leakage and ensures time-consistent analysis.

Calibration - Degree to which predicted probabilities align with real outcomes (tests reliability).

Decision Curve Analysis (DCA) - Evaluates economic usefulness of predictions compared to heuristic strategies

Policy Induction Model - LLM-based framework generating transparent, rule-based decision logic for interpretability.

Fuzzy Matching - A text-comparison technique that identifies similar but not identical strings by measuring approximate rather than exact matches, allowing for variations in spelling or formatting (e.g., “MIT” vs. “Massachusetts Institute of Technology”)

Temperature - A hyperparameter controlling the randomness of model outputs; lower values (e.g., 0.1) produce more deterministic and precise responses, while higher values generate greater diversity and variability in reasoning or predictions.

3. Empirical Constructs

Funding Attainment - Proxy for venture success, measuring cumulative equity raised within five years of founding.

Founder Signal - Observable indicator (education, experience, network) that communicates venture quality.

Big Tech Experience - Employment history at major technology firms; analyzed as a high-cost, high-credibility signal.

Elite University Affiliation - Educational signal implying costly and observable quality differentiation.

Accelerator/Incubator Participation - Institutional signal enhancing legitimacy and network embeddedness.

Cohort-Based Sampling - Grouping founders by founding year to control for macroeconomic and temporal factors.

Table of Contents

Table of Figures

Table of Tables

1. Introduction

1.1 Context

Venture capital is one of the few domains where investors must routinely bet millions of dollars on people long before there is anything solid to judge. Early-stage founders request capital without traction, revenue, or a product. Yet billions of dollars flow to them every year, sustained by the expectation that a small fraction will generate outsized, outlier returns. Unsurprisingly, failure dominates, roughly 90% of startups never return initial capital (Menon & James, 2022). To navigate this environment, investors fall back on signals - credentials, prior experience, elite networks, institutional affiliations of founders to make decisions. In practice, these signals determine how scarce early-stage capital is distributed (Hsu, 2007; Ko & McKelvie, 2018). Signaling theory explains why investors depend on these cues when uncertainty is high (Spence, 1973; Connelly et al., 2011).

Yet empirical findings on which signals matter remain inconsistent. Some studies document strong positive effects of elite education or prior entrepreneurial experience on funding outcomes (Hsu, 2007; Zhang, 2011), while others find weak, null, or context-dependent results (Busenitz et al., 2005; Unger et al., 2011). Recent systematic reviews by Bafera and Kleinert (2023) and Zhang et al. (2023) identify three fundamental gaps. First, most studies examine signals only within their original sample and period, while temporal generalization tests are *"largely absent"* (Bafera & Kleinert, 2023, p. 1847). Second, signals are typically analyzed in isolation and cannot capture whether they combine synergistically or substitute for one another. Consequently, evidence is missing on whether signal to outcome relationships generalize across time and whether configurations of signals jointly influence funding decisions. Third, alternative methods leading to improved replication and generalizability are underexplored.

This thesis addresses these gaps by developing a predictive-modeling framework that evaluates whether founder signals persist in new datasets outside the contexts where they were originally identified. In doing so this study develops and employs a state-of-the-art LLM-based policy-induction model combined with established machine-learning models. These complementary models are trained on founder signals from earlier cohorts (2012–2021) and tested on entirely new founders in later cohorts (2018–2022), enabling a direct assessment of signal generalization. The analysis draws on a dataset of more than 50,000 technology founders across Europe and North America, linking venture-level funding data from PitchBook and Crunchbase with founder-level profiles from LinkedIn to answer the following research question:

To what extent do founder human- and social-capital signals generalize out of sample when predicting venture funding outcomes, both individually and in combination?

To empirically evaluate the research question, the thesis develops three theory-driven predictions that specify how founder signals should generalize across time. The predictions are derived from established work in human- and social-capital signaling. Each prediction specifies how a particular class of founder signals is expected to behave over time, whether its predictive relevance should persist, strengthen, weaken, or arise only in combination with other signals. In doing so, they translate the research question into three focused assessments of temporal generalizability.

1.2 Contribution

This thesis makes two main contributions to signaling theory. First, this study responds directly to recent calls in signaling research for methodologies that evaluate whether founder–signal relationships generalize beyond the samples in which they were originally observed (Bafera & Kleinert, 2023; Zhang et al., 2023). This thesis develops a state-of-the-art LLM-based policy-induction framework alongside established machine-learning models. The hybrid design leverages generative-AI reasoning in a novel way to produce transparent decision rules while simultaneously enabling rigorous out-of-sample prediction. Together, the methodology of this thesis advances how predictive modeling can serve as a tool for theory evaluation in management research.

Second, in response to the research question the thesis tests the generalization of founder signals individually and in combination. Here two main contributions can be outlined. First, this thesis shows that founder signals do not carry universal predictive value across cohorts. Their usefulness depends on whether they address the specific uncertainty investors seek to resolve. In general technology ventures, signals matter only when they reduce the type of uncertainty that dominates the setting, primarily execution risk. This context-dependence generalizes across cohorts. Accordingly, the thesis refines signaling theory by demonstrating that its mechanisms operate not universally, but only when a signal meaningfully reduces the uncertainty relevant to a particular domain. Second, this thesis displays that combining founder signals does not inherently strengthen their predictive power. This refines configurational theorizing by clarifying that signals reinforce one another only when they provide distinct, non-redundant information; when they speak to the same underlying concern, as academic and practical experience do. In this context, they function independently rather than synergistically.

1.3 Delimitation

This study's scope reflects strategic design choices to ensure analytical precision. Five delimitations define what the research addresses and what falls outside its scope.

First, individual founder characteristics are analyzed rather than founding team compositions. While teams represent the norm in venture capital and enable skill complementarity (Kaplan & Strömberg, 2009; Beckman et al., 2007 and beyond) individual-level analysis isolates how personal credentials independently influence funding evaluation. Thus, results generalize most directly to individual credential assessments.

Second, the study tests whether founder signal-outcome relationships maintain out-of-sample predictive validity across yearly cohorts 2018-2022, not whether signals causally determine outcomes. The study documents which characteristics predict funding attainment and whether these generalize temporally, distinguishing stable correlations from context-bound associations. Causal identification meaning whether signals cause success or merely correlate due to confounding falls outside the scope of this study.

Third, the analysis includes only founders who secured at least one equity funding round, excluding unfunded ventures, bootstrapped companies, and those not seeking institutional capital. This reflects a deliberate definition of the prediction task. The findings therefore speak specifically to founder-signal dynamics within the funded population, rather than to the determinants of entering the funding process.

Fourth, the analysis is geographically limited to Europe and North America. These ecosystems share institutional structures, data transparency, and evaluation logics that enable meaningful comparison of founder signals (Bruton et al., 2008).

Lastly, industry filtering restricted the analysis to ventures where early-stage financing follows scalable, market-driven patterns. Sectors characterized by regulatory gating, grant dependence, or project-contract

financing, specifically, biotech, biopharma, pharmaceuticals, healthcare, chemicals, and consulting, are excluded (Baum & Silverman, 2004; Hsu, 2007; Colombo et al., 2019; Ahlers et al., 2015). The empirical setting is limited to general technology ventures whose growth follows standard scaling dynamics.

1.4 Structure

This thesis is organized into six sections starting with this introductory section. Section 2 develops the study's analytical framework. It outlines the core logic of signaling theory, describing how education, experience, and networks indicate founder quality under information asymmetry. It then introduces predictive modeling to test whether these theorized relationships hold beyond their original samples. From this synthesis, the section derives three predictions about how founder signals are predicted to behave across temporal cohorts.

Section 3 outlines the research design, including the cohort-based sampling strategy. The section then explains how founder data is collected, filtered and transformed into theory-consistent features serving as input for the predictive models. It further introduces the machine-learning models and interpretable policy-induction model, followed by the evaluation framework and research quality assessment criteria.

Section 4 presents the empirical results. It first evaluates model performance to confirm the stability of the predictive environment across the yearly cohorts 2018-2022. The analysis then turns to founder signals themselves examining how individual attributes and signal configurations influence predictions over time. Section 5 offers the analysis and discussion. It interprets the empirical results in relation to the three predictions and prior research.

Section 6 concludes the thesis. It synthesizes the theoretical, methodological, and empirical contributions, outlines implications for practitioners and identifies the study's limitations. The section ends by proposing avenues for future research.

2 Conceptual Background

2.1 Human and Social Capital Signaling

Markets with asymmetric information present major challenges for efficient resource allocation. When one party holds private knowledge that others cannot observe, participants face uncertainty about the true value of goods or actors. Akerlof's (1970) "market for lemons" illustrates how this can trigger adverse selection, making high- and low-quality participants indistinguishable and lowering overall market performance. Spence (1973) extended this logic to labor markets, showing that individuals can distinguish themselves through signaling, allowing observers to separate high- and low-quality participants and thereby limit the adverse-selection problem (Riley, 2001).

Entrepreneurship exemplifies an extreme form of this informational problem. New ventures typically lack tangible assets, performance histories, and standardized disclosure, leaving investors unable to assess quality through traditional indicators. Founders, as informed parties, must therefore rely on signals, observable characteristics or actions that credibly convey venture quality to outsiders (Colombo, 2021; Connelly et al., 2011; Dutta & Folta, 2015; Svetek, 2022; Taj, 2016). The earlier the stage of venture funding, the fewer established data points are available to signal performance. Therefore, in early-stage financing, founder signals substitute for performance data, helping entrepreneurs reduce uncertainty and establish legitimacy with investors and partners (Ahlers et al., 2015; Nagy et al., 2012; Plummer et al., 2016; Zimmerman & Zeitz, 2002).

For a signal to function credibly, it must satisfy three conditions:

(1) It must be observable, allowing the receiver to recognize and interpret it,

(2) costly, requiring enough effort, skill, or resources that low-quality actors cannot easily fake it, and

(3) differentiating, creating clear separation between high- and low-quality senders (Spence, 1973)

When these criteria are met, signals create separating equilibria that allow investors to infer underlying quality and allocate resources efficiently (Bergh et al., 2014). Human- and social capital signals are consistently identified as central and well-established signals in early-stage entrepreneurial equity financing (Ko et al., 2018; Bernstein et al., 2017; Colombo, 2021; Spence, 2002; Howell & Nanda, 2019).

Social capital signals mitigate information asymmetry through three distinct mechanisms. First, network ties act as signals. An introduction from a trusted contact implicitly communicates the founder's reputation and competence (Shane & Stuart, 2002). Second, prestigious affiliations enhance credibility. A founder connected to prestigious incubators benefits from institutional validation that substitutes for missing performance records (Shane & Cable, 2002; Stuart, Hoang, & Hybels, 1999). Third, ongoing relationships help investors monitor progress and build trust. Founders who stay closely connected with investors make communication easier and reduce agency risk (Hsu, 2007). Through these mechanisms, social capital sends informational and reputational signals that allow founders to secure financing under uncertainty.

Human capital represents clear signals of a founder's underlying capability and quality. It conveys information about the entrepreneur's skills, knowledge and capacity to execute under uncertainty (Connelly et al., 2011; Colombo et al., 2019). Specifically, education, prior entrepreneurial experience and industry-specific expertise provide such cues of capability (Hsu, 2007; Gimmon & Levie, 2010; Colombo & Grilli, 2005, 2010). Collectively, human capital signals reduce investor uncertainty by demonstrating that the founder possesses the expertise and experience necessary to perform, thereby enhancing the perceived credibility of the venture (Ko et al., 2018; Hsu, 2007; Gimmon & Levie, 2010). Both human and social capital

operationalize signaling theory's core mechanism by translating a founder's unobservable quality into observable, costly and differentiating cues that investors can interpret (Ewens & Townsend, 2020; Colombo et al., 2019; Shane & Stuart, 2002).

As research on the relationship between human- and social capital signals expanded, findings began to diverge across venture stages, industry contexts, and investor types (Colombo & Grilli, 2005, 2010; Hsu, 2007; Busenitz et al., 2005; 2018; Hernández-Carrión et al., 2016; Dudley, 2021; Shao & Sun, 2021). Empirical studies consistently find that founders' education and prior experience influence investors' perceptions and funding decisions, although the magnitude and persistence of these effects vary across contexts (Ko & McKelvie, 2018; Packalen, 2007; Steigenberger & Wilhelm, 2018). Meta-analyses confirm this general pattern but highlight substantial variability across studies (Unger et al., 2011; Colombo & Grilli, 2010).

Bafera and Kleinert (2023) conclude that signaling research has become fragmented, emphasizing statistically significant associations without establishing whether signals reflect enduring mechanisms. They do so by providing the most comprehensive synthesis to date, reviewing 172 signaling studies across management and entrepreneurship. Their central argument is that signaling environments evolve rapidly due to digitalization, regulation, and external shocks, rendering earlier results quickly outdated. They note that replication is almost entirely absent in this literature, identifying only one attempt which failed to reproduce the findings of three earlier signaling studies conducted less than a decade before (Park et al., 2016). To address this instability, the authors call for systematic replication studies and the integration of big-data methodologies capable of validating prior findings at scale. They point to the promise of large-sample analyses, such as Anglin et al.'s (2020) study of 220,646 entrepreneurial loans, as exemplars of how big data can enhance the robustness and generalizability of signaling research by testing the persistence of effects across time, context and measurement conditions. Ultimately, Bafera and Kleinert (2023) argue that few studies test whether signals truly indicate venture quality, because most rely on single-sample explanatory designs.

Zhang et al. (2023) strengthen this critique in their bibliometric review of early-stage entrepreneurial equity financing research over a 36-year period (1987–2022). Drawing on a dataset of 3,713 primary studies and 108,385 secondary documents, they reveal significant conceptual fragmentation across the field. The analysis identifies the rise of emerging technologies, particularly artificial intelligence and big data, as reshaping how venture evaluation is conducted. To address the field's growing complexity, they proposed the use of *"machine-learning models to more accurately forecast investment outcomes"* by integrating traditional and novel signal types. At the same time, they cautioned that algorithmic bias and limited empirical validation remain critical challenges calling for methodological pluralism, specifically machine-learning-assisted inference to uncover which combinations of signals most effectively predict financing success. Bafera and Kleinert (2023) and Zhang et al. (2023) call delineates the methodological gap that this thesis addresses. The field's advancement depends on combining strong theoretical grounding with rigorous and innovative research methods. Research must address configurational complexity, capturing how signals interact in complementary ways beyond additive regression models. Evaluating signal stability under varying conditions therefore becomes essential for testing the robustness of signaling theory. Predictive modeling provides a framework for such evaluation by shifting the focus from explaining past correlations to assessing whether theoretically grounded relationships generalize across new datasets and contexts.

2.2 Predictive Modeling

Predictive modeling enables the evaluation of whether established theoretical relationships maintain predictive power across samples and context, providing a direct test of the empirical generalizability of a theory (Breiman, 2001; Shmueli, 2010). By linking prediction accuracy to theory scope, it transforms theory evaluation from internal model fit to external verification across time and context. Empirical research in management has traditionally relied on explanatory modeling to test theoretical propositions. This approach, as described by Breiman (2001) and Shmueli (2010), assumes that observed data arise from an underlying stochastic process whose parameters reflect the mechanisms posited by theory. Researchers estimate these parameters and test whether the resulting coefficients differ significantly from zero to assess whether the data supports the proposed relationships. Thus, explanatory models link abstract theoretical constructs to observable variables and evaluate how well the structure implied by a theory aligns with the evidence. This process sharpens constructs, strengthens causal reasoning and identifies mechanisms that appear to operate within the observed sample. Yet explanatory modeling primarily evaluates a theory's internal coherence rather than its external reliability. It measures how well a theory fits the data from which it was estimated, not whether the relationships persist across other contexts. Explanatory modeling therefore confirms internal fit but leaves scope untested, providing limited evidence about a theory's ability to generalize beyond the calibration sample (Shmueli, 2010; Donoho, 2017; Shmueli, 2017).

For theory, limited generalization has two key implications. First, it undermines construct reliability. When estimated coefficients change substantially with minor differences in sampling, the theoretical relationships they represent cannot be regarded as empirically reproducible. Second, it obscures boundary conditions within which a theoretical mechanism is expected to operate (Whetten, 1989). Models that fit sample-specific noise capture local rather than stable patterns (Shmueli, 2010; Hawkins, 2004). As a result, researchers cannot easily distinguish context-dependent findings from mechanisms that generalize across settings (Hawkins, 2004; Mullainathan & Spiess, 2017). The outcome is a model and by extension, a theory, that remains context-bound, empirically valid within its estimation environment yet offering limited insight into its applicability elsewhere (Hofman et al., 2017; Yarkoni & Westfall, 2017). Theories evaluated solely by in-sample fit may therefore overstate theoretical support, appearing statistically consistent while relying on relationships that are unstable, incomplete, or contingent on short-lived data patterns (Athey, 2017).

Predictive modeling directly addresses the limitations of explanatory approaches by extending the evaluation of theories from fit within a sample to performance outside of it (Shmueli, 2010; Breiman, 2001). Conceptually, it redefines what constitutes empirical support for a theory. Rather than asking whether a theoretical model explains the observed data, predictive modeling tests whether the same relationships generate accurate forecasts under new conditions (Ioannidis, 2005). In doing so, predictive validation complements explanatory testing without replacing it (Shmueli, 2010). When predictive accuracy declines out-of-sample, the resulting drop in predictive performance offers a measurable indicator of how far explanatory success translates into predictive reliability (Camerer et al., 2018). Identifying this gap converts what was previously an implicit assumption about generalizability into explicit evidence concerning the scope and stability of a theory (Grimmer et al., 2021; Molina & Garip, 2019).

Across business research, scholars employ predictive approaches to identify empirical regularities, refine causal reasoning, and examine the generalizability of theoretical relationships (Hofman et al., 2021; Leavitt et al., 2020; Miric et al., 2023; Aguinis et al., 2024; Hofman et al., 2017; Cerqueira et al., 2020). Many management theories, including signaling theory, argue that outcomes result from combinations of factors rather than single attributes. Recent work in configurational theorizing shows that causal mechanisms often depend on how attributes interact - whether they complement, substitute, or offset one another. Predictive modeling allows to test not only individual signals but also whether these signal combinations generalize across samples (George et al., 2016; Bettis et al., 2014; Furnari et al., 2021; Shrestha et al., 2021; Choudhury et al., 2021; Tidhar & Eisenhardt, 2020; Chou et al., 2023).

Building on this methodological foundation, recent advances in artificial intelligence have opened new possibilities for predictive modeling, most notably through the rapid development of LLMs. Although LLMs demonstrate strong capabilities for capturing semantics and structure in textual data (Brown et al., 2020), their use in business and management research remains limited. Scholars consistently identify this scarcity as a research opportunity, pointing to both technical and methodological challenges in adapting LLMs for structured prediction tasks (Kraus et al., 2020). Key obstacles include high computational cost, training complexity, and the difficulty of integrating numerical and textual data within a single predictive architecture (Chen et al., 2025). As Maarouf et al. (2025) note, "*applications of large language models in predictive modeling are still rare*," while Kashyap and Sinha (2024) emphasize that "*addressing this research gap is essential for unlocking the full potential of LLMs and advancing the state of the art in predictive modeling*". Overall, these observations suggest that LLMs hold considerable promise for predictive analysis but that realizing this potential will require methodological innovation.

2.3 Analytical Framework

Signaling theory explains how founders' observable, costly, and differentiating characteristics communicate private information about venture quality under information asymmetry. In early-stage venture investing, where objective performance data is limited, investors rely on such signals to assess founder capability. Empirical research consistently finds that human-capital signals and social-capital signals correlate with funding outcomes (Colombo & Grilli, 2005, 2010; Hsu, 2007; Shane & Stuart, 2002; Ahlers et al., 2015; Ko & McKelvie, 2018). However, the strength and persistence of these relationships vary across industries, investor types, and time periods.

Recent systematic reviews identify two key methodological gaps. First, signaling studies rarely test whether observed relationships maintain predictive validity beyond their original samples. Replication efforts are "almost entirely absent" (Bafera & Kleinert, 2023). Second, most research examines signals independently, using additive models that deprioritise how signals interact in complementary or substitutive configurations (Bafera & Kleinert, 2023; Zhang et al., 2023). Both reviews call for approaches that combine strong theoretical grounding with innovative methodologies capable of assessing signal stability across time and capturing interaction effects.

Predictive modeling addresses these gaps by treating generalizability as an empirical property rather than an assumption. Instead of evaluating statistical significance for the correlational relationships between human- and social capital and funding attainment, as outlined throughout Section 2, predictive validation tests whether the same signals retain predictive relevance when applied to entirely new datasets. Relationships that generalize across temporal contexts suggest stable signaling mechanisms, whereas those that weaken indicate boundary conditions, contextual sensitivity, or shifts in evaluative norms.

A second conceptual dimension concerns configurational effects (Furnari et al., 2021). Signaling theory often assumes, implicitly or explicitly, that human- and social-capital signals operate in combination rather than isolation. Academic depth may enhance credibility only when paired with practical experience, and strong networks may amplify the value of demonstrated competence. A configurational perspective therefore examines whether combinations of founder attributes produce more credible or informative signals than any single attribute alone, and whether these combinations remain predictive across time.

This leads to the underlying research question stated in the introduction of this thesis:

To what extent do founder human- and social-capital signals generalize out of sample when predicting venture funding outcomes, both individually and in combination?

Building on the research question, three testable predictions guide the analysis. Each prediction is grounded in findings from prior research on human- and social capital as signaling mechanisms towards venture funding attainment as outlined throughout Section 2.1. Together, these predictions address both forms of signal behavior, the persistent relevance of individual signals (Predictions 1 and 2) and the potential configurational value of more complex signals (Prediction 3).

Prediction 1: *Prior entrepreneurial experience is expected to remain a consistently strong predictor of funding success. Accordingly, founders classified as first-time founders should exert a negative influence on the model's predictions, the absence of prior founding experience is expected to reduce, and its presence to increase, the predicted likelihood of securing investment.*

Prior entrepreneurial experience has emerged as a critical determinant of funding outcomes. Founders with previous venture experience demonstrate observable competence, established networks, and familiarity with investor expectations (Hsu, 2004; Delmar & Shane, 2006). Empirical evidence consistently shows that experienced entrepreneurs secure capital more readily and at higher valuations compared to first-time founders (Zhang, 2009; Ucbasaran et al., 2009; Gupta et al., 2024). This pattern persists across contexts because experience serves as a credible signal that is difficult to imitate (Ko & McKelvie, 2010). Given this robust empirical foundation, entrepreneurial experience (e.g., First-time-founder) represents a theoretically grounded and empirically validated factor in investment decision-making and is expected to remain stable or increase over the yearly cohorts 2018-2022.

Prediction 2: *Big-Tech-Experience is expected to be a strong signal and remain a consistently stronger predictor of funding success than other forms of domain expertise, such as PhD training, Academic Researchers, or Corporate Researchers. This means that Big-Tech-Experience exerts a positive influence on the model's predictions, the presence of this signal increases, and its absence decreases the predicted likelihood of securing investment.*

Experience in large technology firms functions as a signal of both competence and legitimacy. Such experience credibly communicates technical expertise, exposure to scalable business models, and access to elite professional networks, traits that are difficult for founders from academia or smaller firms to imitate (Ferrary & Granovetter, 2009; Beckmann et al., 2007; Gimmon & Levie, 2010). Empirical evidence strongly supports this interpretation. Dworak (2022) finds that Big-Tech-Experience increases the likelihood of raising venture capital. Engel (2015) further highlights that Big-Tech tenure provides founders with capital access and process expertise, both recognized as key drivers of startup success.

Although Colombo and Grilli (2005) did not find technology-sector employment to be a significant predictor of funding attainment, more recent research by Balachandran et al. (2024) showcases that corporate experience can generate relational capital and improve navigation of complex organizational landscapes, especially when technology or commercialization focused. Taken together, these findings indicate that Big-Tech-Experience provides a broad and credible signal of founder quality. Thus, Big-Tech-Experience is expected to be one of the strongest predictors of funding success throughout 2018–2022.

By contrast, academic and research-based credentials are expected to be weaker predictors in this context. Their informational value is domain specific. PhD training, academic research roles, and corporate R&D experience signal depth, analytical sophistication, and specialized expertise, but these strengths do not map onto the dominant uncertainty structure of general technology startups. In science-intensive sectors (e.g., biotechnology), academic depth directly addresses the investor's key uncertainty, scientific feasibility (Hsu, 2007; Swift, 2018). However, the ventures examined in this thesis face the opposite configuration. Technology risk is relatively low, while execution risk is high (Hsu, 2007; Swift, 2018). As a result, academic credentials provide limited insight into a founder's ability to hire and scale and thus are expected to have weaker predictive importance.

Consequently, signals predict funding when they address investors' primary concerns. General technology ventures face execution uncertainty. Big-Tech-Experience demonstrates execution capability through verified performance in complex organizations, while academic credentials demonstrate technical depth and analytical sophistication but not commercial execution ability. In contexts where execution dominates evaluation, the more direct signal should persistently outperform the indirect one. Thus, this hierarchy should remain stable across the yearly cohorts 2018-2022.

Prediction 3: *High academic experience, such as holding a PhD or prior work as an academic or corporate researcher is expected to contribute positively to funding attainment only when paired with practical, market-oriented experience such as Big-Tech-Experience, executive roles or board membership. In isolation, these academic signals are expected to have limited or no positive effect. Their value should emerge primarily when combined with market-oriented experience complementary indicators.*

Academic credentials such as PhDs or research experience represent expertise signals that demonstrate analytical capability, discipline-specific knowledge, and scientific credibility (Allison et al., 2017; Ahlers et al., 2015). However, signaling theory stresses that the effectiveness of any signal depends on its interpretability and relevance to the evaluator. Particularly in general technology ventures, purely academic backgrounds are difficult for investors to translate into evidence of practical execution skill or commercial judgement. Academic signals tend to be incomplete on their own (Toole & Czarnitzki, 2008; Zucker et al., 1998). They are valued only when complemented by practical, market-facing, or relational experience. Specifically, founders with academic credentials raise more capital when their expertise is combined with managerial roles, board participation, or industry experience (Hsu, 2007; Gimmon & Levie, 2010; Franco et al., 2021; Keogh & Johnson, 2021; Eesley & Roberts, 2012). Related studies emphasize the advantages of "ambidextrous" founders, who blend research depth with diverse professional experience, showing that such combinations enhance investor confidence (Kaiser et al., 2018; Genin et al., 2023; Noguti et al., 2021; Shetty & Sundaram, 2019). Conversely, academic-only profiles often fail to signal real-world competence or market understanding, limiting their interpretive value in early-stage venture evaluation (Gimmon & Levie, 2010; Allison et al., 2017; Li & Liu, 2025).

This suggests that academic signals create value only when they signal information that practical experience alone cannot. When paired with complementary, market-facing experience, academic credentials can reduce multiple forms of investor uncertainty at once, creating a configuration that is more informative than either signal independently. In general technology ventures, however, academic and operational signals both speak to execution capability. Because they address the same evaluative dimension, the complementarities seen in science-based contexts may not appear. Prediction 3 therefore examines whether such cross-domain reinforcement appears in this context or whether academic and practical signals instead function as independent, non-synergistic indicators of founder quality.

Table 1: Prediction Signals

Signal	**Empirical Foundation in prior Research**	**Rationale (Cost–Differentiation Logic)**	**Expected importance stable / increasing**
Prior entrepreneurial experience	Hsu (2007); Subramanian et al., (2025); Zhang, (2009); Ucbasaran et al. (2009); Gupta et al., (2024); Gottschalk et al., (2014)	Experience is costly, observable, and signals venture-building capability and persistence.	Yes
Big Tech Experience	Jerry Davis et al., (2022); (Ferrary & Granovetter, 2009); Nicas, (2016); Engel, (2015)	Experience in Big Tech is associated with credibility and network access, serving as a strong legitimacy signal.	Yes
Signal configurations (combinations)	Bafera & Kleinert (2023); Zhang et al. (2023); Hsu (2007); Swift (2018); Eesley & Roberts (2012)	Combinations of signals become more credible when academic depth is paired with practical, market-oriented experience, as each addresses a different source of uncertainty.	Conditional.

To test these predictions, Section 3 develops a predictive modeling framework that operationalizes human and social capital signals as structured features, trains models on earlier cohorts, and evaluates whether signal-outcome relationships persist in later cohorts. Feature importance and SHAP analyses quantify the relative contribution of each signal, while interaction analysis tests configurational hypotheses.

3. Method

3.1 Research Design

This chapter outlines the predictive-modeling research design used to evaluate whether founder human- and social-capital signals generalize across time. Rather than relying on a traditional explanatory approach, the study adopts a predictive-modeling framework in which out-of-sample performance serves as the primary test of signal generalization. The design consists of five components that enable theoretical mechanisms to be assessed through temporal generalization rather than in-sample fit.

(1) The prediction task must be clearly defined, including an outcome metric that meaningfully represents the theoretical construct of interest. This outcome must be structured in a way that allows performance to be compared across samples.

(2) Predictive modeling requires temporally valid datasets. Information available at the time predictions would have been made must be cleanly separated from future outcomes. This involves constructing training datasets from earlier periods and prediction datasets from later, non-overlapping periods, inhibiting temporal leakage.

(3) The features used as model inputs must operationalize the theoretical constructs being examined. This requires transforming raw data, structured or unstructured, into measurable values that align with the theory's core concepts.

(4) The research design requires training one or more models on the training data so they can learn relationships between features and the defined outcome. Using complementary model families helps ensure that results do not depend on a particular functional form or modeling approach.

(5) The method requires an evaluation framework capable of assessing both empirical stability and theoretical relevance. Discrimination and calibration metrics determine whether the models produce reliable predictions across samples, while feature-level analysis methods connect the behavior of the models back to the theoretical mechanisms under study.

Together, these components form a research design that operationalizes the analytical framework introduced in Section 2.3.

3.2 The Prediction Task

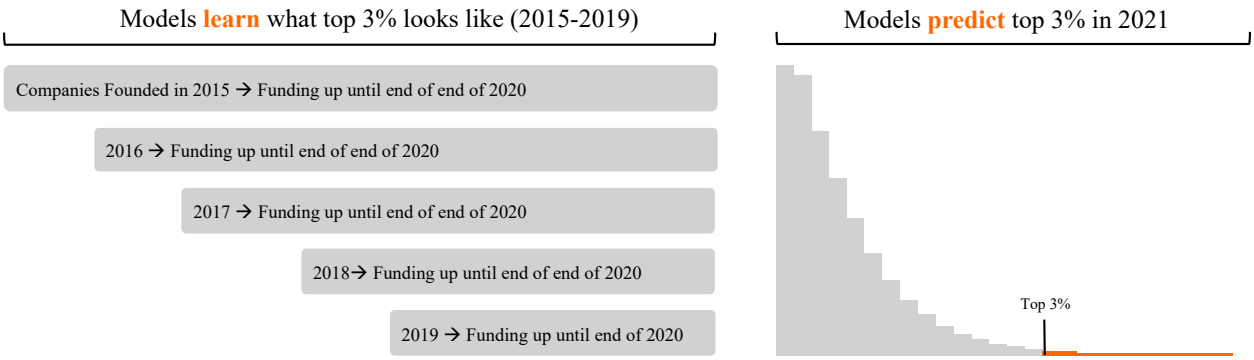
The prediction task focuses on estimating how well founders can secure early-stage funding. For each founder, this is measured by their ranking within their founding-year cohort's funding distribution, based on total equity raised within five years of founding. Predictions are generated for the 2018–2022 cohorts. For each prediction year, only the five-preceding founding-year cohorts are used for training. For example, the 2021 prediction relies on founder-data from the 2015–2019 cohorts. This rolling, cohort-based structure ensures strict temporal separation between training and prediction, prevents information leakage, and allows a direct test of whether signal–outcome relationships that hold in one period remain predictive for new founders under new market conditions.

The outcome metric of this study is cumulative venture funding attainment. Venture funding attainment serves as a robust proxy for early-stage entrepreneurial performance. High-quality ventures credibly differentiate themselves by attracting external capital, while low-quality ventures cannot. Securing venture funding thus functions as both a signal and a test of legitimacy in environments where underlying performance

cannot yet be verified (Plummer, et al., 2016). Empirical research consistently operationalizes funding attainment, whether measured as receiving venture capital or as the total amount of equity raised as a direct reflection of founder quality and market validation (Ko & McKelvie, 2018; Shane & Stuart, 2002). Early-stage financing, particularly within the first five years of a firm's life, captures investors' collective assessment of a venture's potential when internal performance data is limited (Davila et al., 2003; Ahlers et al., 2015). In this sense, cumulative funding attainment within a defined early-stage window functions as a quantifiable market outcome of effective signaling.

The study measures founders' funding-attainment ability by identifying whether their venture reaches the top 3% of cumulative equity funding within its own founding-year cohort over a five-year observation window. This cohort-based approach ensures that each founder is evaluated relative to peers who faced the same market conditions, investment climate, and capital availability in their founding year. Wherever possible, cumulative funding is tracked across the full five years following founding. For the most recent cohorts (2020–2022), funding is measured only up to 2025. Although later cohorts have slightly shorter observation windows, the logic remains consistent.

Figure 1: The Prediction Task



This cohort-based approach normalizes outcomes within founding-year cohorts to account for temporal distortions such as valuation inflation, investment cycles, and changing macroeconomic conditions, while also addressing the extreme right-skewness of venture outcomes (Gornall & Strebulaev, 2019; Hsu, 2007; Guzman & Stern, 2015, 2020; Ewens & Townsend, 2020). Moreover, using percentile ranks instead of raw monetary values captures the inherently selective and power-law nature of venture success, where only a very small share of founders attracts a disproportionate share of capital. Focusing on the uppermost percentiles isolates these outlier cases while holding constant the broader economic conditions faced by their cohort.

3.3 Data Construction and Sampling Framework

This section outlines how the study constructs the training and prediction datasets required for temporal generalization, following a three-step process that defines (1) the venture population, (2) applies contextual sampling criteria, and (3) the specific data set construction and sampling example of the 2021 datasets. For each prediction year from 2018 to 2022, two fully separate datasets are constructed, one training- and one prediction dataset. Models are always trained on the five preceding years and tested on the next one. All training and prediction datasets link venture-level funding histories from Specter, PitchBook and Crunchbase with founder profiles from People Data Labs (PDL) and LinkedIn.

(1) The venture population consists of all companies founded in each calendar year and recorded across the underlying data sources, representing the full universe of ventures before any geographic, industry or funding-based filtering is applied. (2) To define the sample of a given cohort, a set of inclusion and exclusion criteria is applied. These filters define the empirical population from which all training and prediction samples are drawn, establishing the contextual boundaries of the analysis. The geographic scope is limited to ventures headquartered in Europe and North America. These two regions represent mature venture-capital ecosystems with comparable institutional infrastructures, investor practices, and disclosure standards (Colombo & Grilli, 2010; Kerr, Nanda & Rhodes-Kropf, 2014). Including both regions in the population provides variation to evaluate theory generalizability while maintaining data consistency and availability. Each company's location was defined by its headquarters at the time of founding. For each founding-year cohort, outcomes were tracked over a five-year period, starting at the company founding date, following the standard observation window used in prior research.

The focus of this study is on general technology ventures that follow scalable growth patterns. Thus, industry filtering is applied to exclude sectors whose financing dynamics differ fundamentally from this model. Life-science and specific industrial sectors were excluded because their funding structures differ fundamentally from general technology ventures. Financing is concentrated in large, milestone-based rounds tied to regulatory or patent events, followed by multi-year inactivity. Such patterns distort the link between founder signals and funding outcomes in the observation context of this study. Accordingly, biotech, biopharma, pharmaceuticals, healthcare and chemicals are excluded from the datasets. This filtering is consistent with prior venture-finance studies (Baum & Silverman, 2004; Hsu, 2007; Colombo et al., 2019). Only equity-based funding rounds are considered in the analysis, covering the stages Angel, Pre-Seed, Seed, and Series A through Series J. These stages capture the conventional venture-capital financing sequence defining the venture lifecycle (Hellmann & Puri, 2002). Funding data detailing which companies completed which investment rounds were sourced from PitchBook and Crunchbase, standardized in U.S. dollars to ensure comparability across regions and time periods.

(3) The 2021 cohort is used to illustrate the dataset construction process. To generate predictions on the 2021 cohort, there is a prediction and a training dataset. Both follow the same construction logic but differ in their sample composition. All founder data is anonymized.

For the 2021 prediction dataset, the process begins by collecting all ventures founded in 2021 from Specter, PitchBook, and Crunchbase. After applying the exclusion criteria, this yields 244,219 companies founded in that year. From these, only ventures that raised at least one qualifying equity funding round between 2021 and 2025 are retained, resulting in 8,265 funded startups. Founders are then identified by matching company records with individual profiles from People Data Labs and LinkedIn. An individual is classified as a founder if they list the company in their work history and hold a "founder", "co-founder", or similar title starting within twelve months before or after the company's founding date.

Matches are validated through alignment of company domains across LinkedIn and PitchBook. This process yields 5,824 founder–venture pairs and 11,256 total founder profiles, reflecting that many startups are founded by multiple individuals. A stratified random sample of 300 profiles is manually reviewed to confirm extraction accuracy. Each venture's total equity funding raised between 2021 and 2025 is then aggregated and ranked within the 2021 cohort. All founders receive the same rank as their venture. Founders in the top 3 percent of this distribution are labeled as high funding-attainment (TRUE), and all others are labeled as low funding-attainment (FALSE). This results in 538 high-attainment founders and 11,256 low-attainment founders for the 2021 prediction cohort.

Table 2: Prediction Datasets - Descriptive Statistics

Year	Total Companies	Total Founder-Venture Pairs	Total Founder Profiles	top 3% (true)	bottom 97% (false)	Funding Treshold 3% (in $Mio.)
2018	279,994	5,021	9,611	474	9,137	73
2019	269,491	5,328	10,256	542	9,714	70
2020	304,331	6,123	12,057	567	11,490	55.83
2021	244,219	5,824	11,794	538	11,256	49.7
2022	205,847	6,055	12,852	513	12,339	59

The training dataset consists of founders from the five cohorts preceding the prediction year (2015-2019), with each cohort constructed using the same procedure applied to the prediction datasets. This ensures that the models learn founder–signal relationships exclusively from earlier periods before being evaluated on entirely new founders in the 2021 cohort. Within every training cohort, ventures are ranked by their cumulative funding raised within five years of founding, with the observation window ending no later than 2020. The 2020 cohort is excluded because founding dates are recorded at the annual level, making it impossible to distinguish ventures founded early in 2020, whose funding is already observable, from those founded later in the year, which would not yet have had sufficient time to raise capital until the end of 2021. For model training, only the most extreme performers are retained. Founders in the top 3 percent of their cohort's funding distribution were labeled as high funding-attainment (TRUE), and those in the bottom 10 percent were labeled as low funding-attainment (FALSE). The remaining 87 percent were excluded from training to create a sharper decision boundary and reduce noise in model training (Gornall & Strebulaev, 2020). Across the 2015–2019 cohorts, this procedure yielded 2,108 labeled founders: 695 positives and 1,413 negatives.

The labeling scheme intentionally differs between training and testing. During training, the model learns from distinct cases (top 3% vs. bottom 10%), whereas during testing it must identify the top 3 percent within the full founder population (top 3% vs. all others), which reflects the real prediction problem. Each founder appears only once in the dataset and is associated with their most recent venture. Prior ventures are excluded under the assumption that they were discontinued.

Table 3: 2021 Training Dataset - Descriptive Statistics

Year	top 3% (true)	bottom 10% (false)	total
2015	164	310	474
2016	153	287	440
2017	156	279	435
2018	122	283	405
2019	100	254	354
Total	695	1413	2108

In summary, each prediction dataset contains the full founder sample for a given year, labeled by cumulative funding attainment within the following five years. The corresponding training dataset contains only distinct high- and low-performing founders from the five preceding cohorts. This rolling-origin design enables an empirical test of how founder characteristics translated into predictive features at the time of founding predict venture funding outcomes, both individually and in combination across cohorts and time. The descriptive statistics of all training and prediction datasets are provided in Appendices A and B.

3.4 Features

This section explains how raw founder data is transformed into the structured features used as model inputs. Because ML models require numerical and categorical representations of information, the unstructured career histories, educational records, and professional affiliations extracted from LinkedIn and People Data Labs must be converted into measurable indicators. These indicators are called features, each represented as a binary variable that takes the value 1 when its criteria are met and 0 otherwise. The feature set was deliberately designed to cover all founder signals required to evaluate the research question and the three predictions in Section 2.3. Every feature corresponds to a human- or social-capital signal, ensuring comprehensive coverage of the constructs. The feature extraction process converts unstructured founder data into a structured analytical dataset that captures a consistent snapshot of each founder at the moment of venture formation. This dataset forms the analytical foundation on which the predictive models test whether signal–outcome relationships hold across new cohorts over time.

The feature extraction framework combines deterministic Python scripts for structured information with controlled LLM reasoning for unstructured text. In total, 21 features are extracted and grouped into five categories according to their extraction logic: (1) rule-based count features, (2) LLM-based rule features (3) rule-based career features, (4) LLM-based reference-list features, and (5) LLM-based complex reasoning features.

The first group (1) deals with rule-based count features and includes two features derived directly from structured data through deterministic Python scripts. These features rely on simple numeric thresholds and logical conditions. For example, the feature *has_three_or_more_work_experience* returns a value of 1 if the founder held at least three prior professional roles and 0 otherwise.

The second group (2), LLM-based rule features, consists of six features that require context-sensitive interpretation but follow deterministic logical rules. These features were extracted using short, highly specific instruction prompts that guided the model through explicit inclusion and exclusion criteria. For instance, *lived_multiple_countries* identifies founders who have studied or worked in at least two distinct countries and *is_firsttime_founder* checks whether the founder has completed any prior founding roles before the current venture.

The third group (3), rule-based career features, includes five features derived through a Python-based script that calculates employment tenure from start and end dates of past roles. These features quantify each founder's total professional experience and mobility before founding their venture. They categorize founders by cumulative years of work experience and by patterns of job change. For instance, *has_job_hopped* takes the value 1 if a founder has held at least three positions shorter than two years each, and 0 otherwise.

The fourth group (4), LLM-based reference-list features, includes nine features that identify institutional affiliations such as elite universities or venture firms. These were extracted through a retrieval-augmented generation (RAG) process in which the LLM matches text from founder profiles against curated reference lists of known institutions. The model performs fuzzy matching to accommodate name variations, such as "MIT" versus "Massachusetts Institute of Technology", while enforcing strict inclusion boundaries. For example, *top_tier_uni* captures attendance at one of the world's top fifty universities according to QS rankings. All institutional reference lists used for this extraction process are provided in Appendix C.

The final group (5), LLM-based complex reasoning features, consists of nine variables requiring higher-order interpretation of professional and educational histories. Each feature is extracted through a prompt that guides the model to interpret career sequences, role hierarchies, and contextual clues. Examples include *is_board_member*, which detects formal board or director appointments while excluding informal advisory roles. These features capture qualitative aspects of professional signaling that cannot be represented through

simple keyword or pattern matching, enabling the extraction of nuanced information on competence, leadership, and credibility.

Each LLM prompt follows a schema-constrained instruction format consistent with contemporary standards in structured information extraction and prompt engineering (Lu et al., 2022; Qiao et al., 2023; Liu et al., 2026). Prompts explicitly define what qualifies as a match, what must be excluded, and what evidence should be cited. An example-prompt is provided in Appendix H. All LLM calls are executed using GPT-4.1-mini with a temperature setting of 0.1, chosen to maximize determinism while preserving minimal reasoning flexibility. Each founder profile is processed independently across all features, and computations are parallelized both across features and across batches of eight profiles at a time to ensure scalability and efficiency.

Table 4: Features Overview

Feature	Meaning	Extraction Type	Signal Type
top_tier_uni	Attended elite university	LLM – list	Human capital
multiple_degrees	Holds 2+ degrees	LLM - rule-based	Human capital
is_phd	Completed PhD	LLM – reasoning	Human capital
outside_of_us_uni	Studied at university outside the U.S.	LLM – rule-based	Human capital
big_tech_experience	Worked at major tech company (e.g., Google, Apple)	LLM – list	Human capital
is_top_tier_consultant	Worked at leading consultancy (e.g., McKinsey, BCG)	LLM – list	Human capital
is_top_tier_banker	Worked at major investment bank (e.g., Goldman Sachs)	LLM – list	Human capital
researcher_at_big_corp	Research role at major technology firm	LLM – reasoning	Human capital
researcher_at_top_uni	Research or academic position at elite university	LLM – reasoning	Human capital
has_three_or_more_work_experience	Held 3+ prior professional positions	Rule-based	Human capital
has_job_hopped	Held ≥3 short-tenure jobs (<2 years)	Rule-based career	Human capital
one_to_five_years_experience	1–5 years total professional experience	Rule-based career	Human capital
six_to_ten_years_experience	6–10 years total professional experience	Rule-based career	Human capital
more_than_ten_years_experience	>10 years total professional experience	Rule-based career	Human capital
recent_grad	<1 year of professional experience at founding	Rule-based career	Human capital
has_promotions	Received internal promotion	LLM – reasoning	Human capital
exec_experience	Held senior executive (VP/C-level) role outside own venture	LLM – reasoning	Human capital
lived_multiple_countries	Worked/studied in ≥2 countries	LLM – rule-based	Human capital
is_firsttime_founder	No prior completed founding role	LLM – rule-based	Human capital
top_investor	Worked at top-tier VC or investment firm	LLM – list	Social capital
nasdaq_leadership	Leadership role at NASDAQ-listed company	LLM – list + reasoning	Social capital
vc_experience	Worked in venture-capital or private-investment roles	LLM – reasoning	Social capital
is_board_member	Formal corporate board membership	LLM – reasoning	Social capital
comm_experience	Held marketing or communications roles	LLM – reasoning	Social capital

To prevent temporal leakage, all founder profiles are truncated at the company's founding date (T_0). Only pre-T_0 data, such as education, prior employment and affiliations is processed. Post-founding events such as subsequent roles, new affiliations, or later funding rounds are systematically excluded. Because all underlying data sources (LinkedIn, Specter, People Data Labs, PitchBook, and Crunchbase) include time-stamped records, this design guarantees that features reflect only information observable at the time of venture creation. All pre-training and testing data was fully anonymized, and no company IDs, company names, founder IDs, or founder names were included in any dataset.

The resulting feature dataset provides a transparent, theory-aligned representation of founders' human and social capital signals at the moment of venture formation These features constitute the input to the predictive models introduced in the following section.

3.5 Models

The modeling approach integrates two complementary categories of models. First, standard tree-based machine-learning models are deployed to generate predictions. Second, an LLM-based policy-induction model is implemented to generate interpretable, rule-based decision structures. The following section provides detailed descriptions and motivations for both model families. The models are selected to combine well-established standards in predictive modeling with the technological advances enabled by generative AI (Arroyo et al., 2019; Kim et al., 2023; M. W. Powers, 2020; Uddin et al., 2020, Florez-Lopez & Ramon-Jeronimo, 2015).

3.5.1 Machine Learning Models

Three machine learning models were implemented in the study, Gradient Boosting and Random Forest, complemented by a Logistic Regression baseline for benchmarking. First, the Random Forest (RF) model is an established machine learning model that combines simple decision trees to make stronger, more reliable predictions (Breiman, 2001a). Each decision tree in a Random Forest is trained on a slightly different subset of the data, generated through bootstrap resampling of the original dataset (Breiman 2001a, 1996; González et al., 2020). At prediction time, the model aggregates the outputs of all individual trees, using majority voting for classification, to produce a more stable and reliable estimate.

Random Forests can capture the non-linear, complex relationships that characterise founders' human and social capital (Maturo et al., 2025). These attributes combine in complex, non-additive ways that linear models cannot adequately represent. Prior research demonstrates that RF performs particularly well in entrepreneurial and financial prediction settings, where feature heterogeneity and nonlinear structure are common. This makes the model well suited for predicting venture funding outcomes (Schade & Schuhmacher, 2023; Koumbarakis & Volery, 2022; Shi et al., 2023; Lukita et al., 2023).

Second, the Gradient Boosting (GB) model is employed. Instead of training many trees independently, it builds them sequentially. Each new tree is designed to correct the errors made by the previous ones (Natekin & Knoll, 2013). This iterative process, known as boosting, progressively improves the model by assigning greater weight to the data points that were most difficult to predict in earlier rounds (Friedman, 2001; Chen & Guestrin, 2016; Liu et al., 2023; Maturo et al., 2025, Ji & Li, 2025, Acharya et al., 2021).

This thesis employs a GB model because it is particularly effective for complex, high-dimensional data, such as the combination of human- and social-capital indicators that characterize founder profiles (Özkurt, 2024; Davis et al., 2021). These empirical studies show that Gradient Boosting models can achieve very high accuracy in comparable prediction tasks. Because each tree in GB learns directly from the errors of its predecessors, the model is especially effective at capturing subtle interactions between founder characteristics and funding outcomes (Liu et al., 2023; Friedman, 2001)

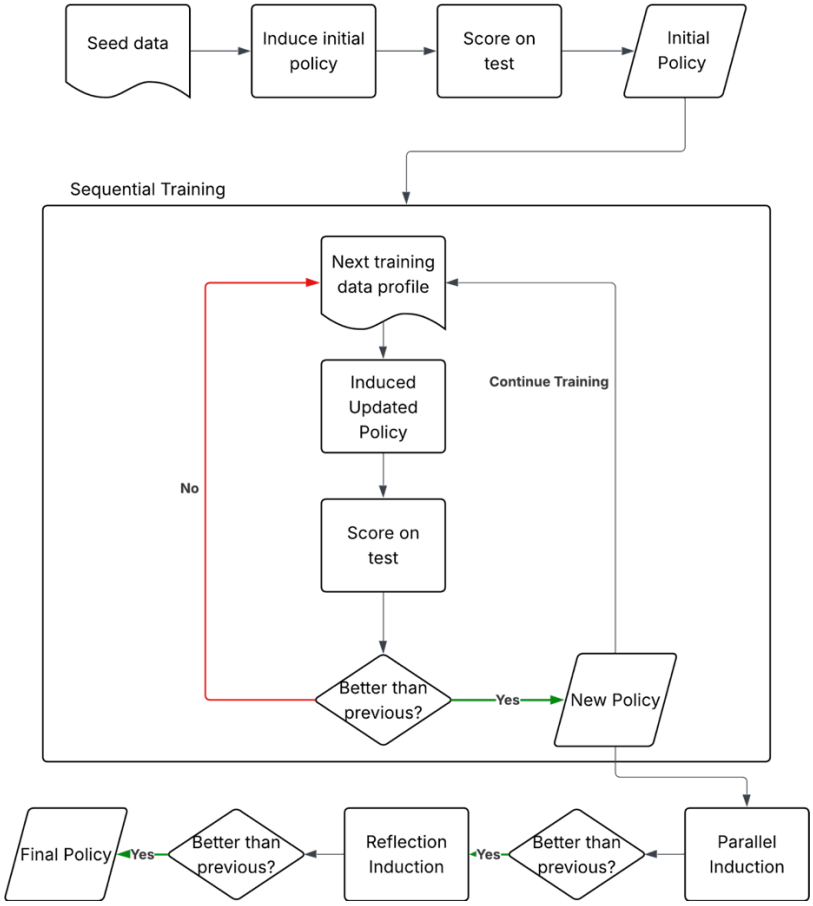
Finally, Logistic Regression is employed as a linear baseline, providing a transparent benchmark for evaluating the added value of the nonlinear models. The model predicts the log-odds of attaining top 3 % venture funding, as a simple linear combination of the input features. This produces coefficients that can be directly interpreted, making it easy to compare how strongly each feature influences the prediction when relationships are assumed to be linear. (Obermann & Waack, 2015; Uddin et al., 2020; Davidsson & Honig, 2003; Ghassemi et al., 2020). Together, these models establish a rigorous baseline for evaluating the generalizability of founder signals and form the analytical foundation for introducing the LLM-based policy-induction model.

3.5.2 Interpretable LLM based Policy-Induction Model

The LLM-based policy-induction model provides a methodological innovation in this thesis by offering an alternative to traditional predictive models. By combining LLM-driven rule induction with traditional predictive modeling, it extends current methodological frontiers in management research, demonstrating how AI-systems can generate transparent, auditable decision rules in natural language that complement ensemble learning models (Mu et al., 2025; Preuveneers et al., 2025). The policy-induction framework is employed due to its strong performance in data-constrained, decision-making contexts. Its efficiency and interpretability make it particularly suitable for testing theoretical mechanisms under uncertainty. (Mu et al., 2025).

The model learns rules through text-based reasoning rather than hidden parameters. It reviews examples, updates its logic through in-context learning, and retains only those revisions that improve measurable performance. All pre-training and testing data was fully anonymized, and no company IDs, company names, founder IDs, or founder names were induced in the model. The training process unfolds in several transparent stages. As outlined in Section 3.2, model training only considers top 3% and bottom 10% cases. Training begins with a small, balanced seed set of founders, 20 high-funding cases - and 20 low-funding cases, presented to the model in text form. From these profiles, the model induces an initial set of "if–then" rules that describe patterns it observes, for example, *"founders with prior successful exits are more likely to raise large funding rounds"*. These preliminary rules are then tested and refined using a larger sample of 100 high-funding cases and 100 low funding cases. After each refinement step, the model keeps only those rule updates that measurably improve predictive performance, evaluated through standard metrics such as precision and the $F_{0.5}$ score**.** Rules that do not increase performance are discarded, ensuring the model improves only through verifiable gains. In a final stage, the model incorporates reflections on recurring patterns across the data, essentially a meta-analysis of its own reasoning, to sharpen and simplify the rule set. This produces a compact, fully interpretable policy consisting of the most informative conditions related to founder experience, education, and career trajectory.

Figure 2: LLM-Policy-Induction-Framework



This results in an interpretable set of decision rules that can be read, audited, and adjusted by human evaluators, addressing one of the central limitations of "black box" models such as neural networks or gradient boosting, which often achieve high accuracy at the cost of transparency (Mu et al., 2025). A full outline of the generated policies and policy analysis can be found in Appendices D, E and F.

3.6 Evaluation Framework

The evaluation framework establishes the empirical foundation for theoretical interpretation. Its purpose is twofold. First, to verify that the predictive environment is statistically reliable and temporally stable. Second, to ensure that any theoretical conclusions drawn from feature dynamics rest on empirically validated model behavior. To achieve this, the framework integrates performance-, feature-level-, and control metrics.

Performance metrics measure empirical stability by capturing how accurately and consistently the models perform across prediction cohorts. Precision and recall assess classification quality under class imbalance, while the $F_{0.5}$ score combines both, weighting precision more heavily to reflect the asymmetric costs and class imbalance of venture investing, where false positives are costlier than false negatives (Hsu, 2007; Breiman, 2001b; Shmueli, 2010).

$$Precision = TP / (TP + FP) \qquad Recall = TP / (TP + FN)$$

where *TP* is true positives and *FP* is false positives.

$$F_{0.5} = (1 + 0.5^2) \times (Precision \times Recall) / (0.5^2 \times Precision + Recall)$$

The Area Under the Precision–Recall Curve (AUCPR) summarizes how well a model balances precision and recall across all classification thresholds, directly measuring how effectively it distinguishes exceptional founders from the large number of less successful cases in regards to the outcome metric funding attainment (Davis & Goadrich, 2006)

$$\text{AUCPR} = \sum_{i=1}^{n-1} (\text{Recall}_{i+1} - \text{Recall}_i) \times \text{Precision}_{i+1}$$

where *n* is the total number of points on the Precision–Recall curve, each corresponding to a different classification threshold and *i* indexes each adjacent pair of PR points, summing the incremental area contributed by each segment.

Empirical stability in this thesis is established when these metrics remain consistent across temporal cohorts, specifically when precision, $F_{0.5}$, and AUCPR values show limited variance (±0.05), and the relative performance of model families remains stable. In this configuration, the model's discrimination ability does not depend on specific sample periods, its ranking performance remains robust despite class imbalance, and any metric drift over time is small and theoretically explainable.

Calibration metrics test whether the model's predicted probabilities correspond to real-world outcomes, ensuring that forecasts can be interpreted as credible likelihoods rather than arbitrary scores. Two complementary indicators capture this. The Brier Score and the Expected Calibration Error (ECE) (Niculescu-Mizil & Caruana, 2005). The Brier Score measures how close predicted probabilities are to actual outcomes. The ECE evaluates systematic bias by checking whether the model is generally overconfident or underconfident.

$$Brier\ Score = (1 / N) \times \sum_{i=1}^{N} (p_i - y_i)^2$$

where p_i is the predicted probability of success for observation *i,* and y_i is the observed binary outcome.

$$ECE = \sum_{k=1}^{K} (n_k / N) \times | acc(B_k) - conf(B_k) |$$

where n_k is the number of samples in bin B_k, $acc(B_k)$ denotes the empirical success rate, and $conf(B_k)$ the mean predicted confidence.

A model is considered *well calibrated* when both metrics remain low and stable across cohorts, Brier ≤ 0.10 and ECE ≤ 0.05 in predictive modeling standards (Brier, 1950; Guo et al., 2017; Shrestha et al., 2021). Under these conditions, predicted probabilities can be interpreted as reliable empirical estimates rather than artifacts of overfitting or imbalance. Calibration metrics such as the Brier score and Expected Calibration Error (ECE) are not applicable to the policy-induction model because it is a rule-based deterministic classifier, not a probabilistic model (Brier, 1950; Niculescu-Mizil & Caruana, 2005; Guo et al., 2017). Instead,

the empirical stability of the policy model is evaluated using discrimination-based metrics, precision, recall, and $F_{0.5}$ (Breiman, 2001b).

Feature-level analysis forms the core of the evaluation framework. To understand how features affect predictions the analysis draws on Shapley Additive ExPlanations (from here on SHAP values) (Lundberg & Lee, 2017). SHAP assigns each feature an additive contribution to every individual prediction, indicating how much that feature pushes the predicted probability upward or downward relative to the model's baseline. By collecting these contributions across the entire dataset, SHAP then quantifies how strongly the model relies on each feature overall. Global feature importance is thus defined as the mean absolute SHAP value, the average magnitude of a feature's influence across all observations. Features with larger mean absolute SHAP values are those the model relies on most strongly, regardless of whether they increase or decrease the predicted likelihood of funding success. Thus, global feature importance captures magnitude, not direction. It identifies the features that the model consistently relies on to differentiate between higher and lower predicted outcomes. This ranking establishes a first layer of interpretability, showing which signals, such carry the greatest explanatory weight.

$$SHAP\colon f(x) = \varphi_0 + \sum_{i=1}^{M} \varphi_i$$

where φ_i denotes the SHAP value for feature i, representing its marginal contribution to the prediction, and φ_0 is the model baseline.

Examining the distribution of SHAP values across observations outlines how and when each feature influences the model's predictions. The mean SHAP value captures the overall direction of the effect, whether the feature generally increases or decreases the predicted likelihood of success, while the median SHAP value reflects the typical size of that effect, minimizing the impact of outliers. SHAP interaction values measure non-additive relationships, instances where the combined effect of two features on a prediction differs from the sum of their independent additive effects. Finally, all feature-level results are analyzed across temporal cohorts to assess how each signal's predictive relevance changes over time.

Because the policy-induction model operates on natural-language decision heuristics rather than numerical features, traditional SHAP or feature-importance analysis is not applicable. To identify which founder signals most strongly influences its predictions, the model is evaluated through policy ablation analysis. In this approach, each policy, an individual "if–then" rule learned by the model, is temporarily removed, and the model's $F_{0.5}$ performance are remeasured. The decrease in performance caused by the removal of a given rule quantifies its marginal contribution to predictive accuracy. Policies producing the largest performance drops are interpreted as the most influential signals within the model's decision logic. This produces a ranked list from 1 to 20 outlining which policies contribute most to the model's predictions. The approach offers a direct and interpretable alternative to feature-importance measures, enabling systematic evaluation of the human-readable founder attributes and signal combinations that drive predictive outcomes.

3.7 Quality Discussion

Because this study employs predictive modeling to test temporal generalization, research quality is evaluated according to criteria appropriate for predictive rather than explanatory designs. Quality is assessed across three dimensions:

(1) Reliability, meaning consistent and transparent procedures across analyses;

(2) Validity, including predictive validity (whether relationships generalize to future cohorts) and construct validity (whether features accurately represent theoretical founder characteristics); and

(3) Replicability, meaning that results can be independently reproduced with equivalent data and methods.

Together, these dimensions ensure that the study identifies temporally stable patterns, uses interpretable constructs, and applies transparent, verifiable procedures.

Reliability concerns the internal consistency of the analysis, whether the models produce stable and consistent results over time. All preprocessing, feature extraction, and validation procedures are standardized and deterministic, minimizing random variation and subjective bias. The evaluation framework confirms that core performance metrics remain stable across the 2018–2022 cohorts, while consistent feature-importance rankings and SHAP interaction patterns show that founder-signal effects persist over time. Models are always trained on four consecutive years and evaluated on a strictly later cohort, ensuring separation between in-sample learning and out-of-sample prediction. Outcomes are evaluated using internal percentile rankings rather than absolute funding amounts, reducing distortion from shifting market conditions and enabling comparable assessments across years.

Predictive validity assesses whether the models forecast future outcomes rather than reproduce past data. It is established through strict out-of-sample evaluation. All features are restricted to information observable at or before the founding year (T_0), preventing temporal leakage and aligning predictions with what investors could realistically know at that moment. Rolling-origin validation shows that predictive performance generalizes consistently across time, indicating that the models capture stable informational patterns rather than artifacts of a particular historical period.

Construct validity ensures that model inputs map cleanly onto signaling-theoretic constructs. Human-capital features represent competence signals (education, experience), while social-capital features represent legitimacy signals (networks, affiliations, investor ties). These conceptual categories are translated into measurable, time-stamped founder attributes. A manual review of 300 randomized profiles across the 2018–2022 cohorts demonstrated 99% agreement between automated classifications and human interpretation, indicating that the features accurately capture the signals they are designed to measure.

Replicability concerns whether independent researchers could reproduce the results using equivalent data and procedures. This study integrates information from four proprietary databases, Crunchbase, PitchBook, Specter and People Data Labs which requires licensed access for full dataset reconstruction. However, the data-construction pipeline itself is fully deterministic and documented. Scripts use fixed parameters, predefined random seeds, and a constant LLM temperature (0.1), enabling near-identical outputs across runs. Model outlines, variable dictionaries, and preprocessing steps are provided in the appendices to support procedural replication even without direct access to the purchased data.

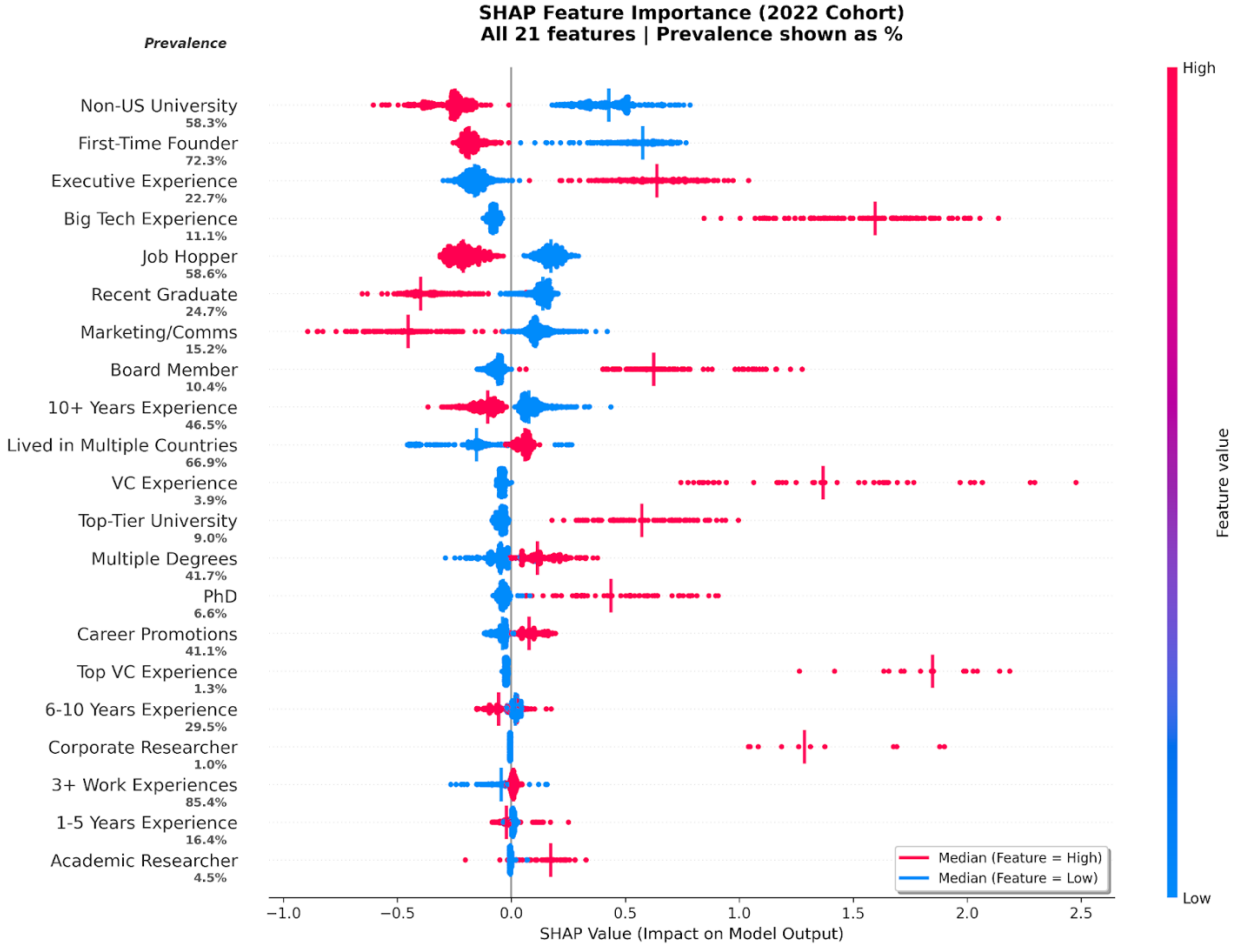
A limitation arises from the dynamic nature of platforms such as LinkedIn, where founder profiles may change or disappear over time, potentially preventing perfect reconstruction of the original raw data. Additionally, changes in the structure or provisioning of the proprietary databases may affect exact replication of the dataset used here. The large-language-models used for feature extraction introduce a version-specific dependency, but deterministic prompting and temperature control mitigate variation. While different LLMs may yield small differences in classification, the methodology itself remains reproducible as long as the underlying model version is accessible.

4. Results

This section presents the empirical results of the study. It begins by illustrating how SHAP values impact model prediction, followed by the assessment of model reliability and calibration. This establishes the empirical stability required for interpreting founder signals. The section then continues by examining the signals themselves for each respective prediction1-3 introduced in Section 2.3.

To illustrate how individual founder signals influence model predictions, Figure 3 presents a SHAP beeswarm plot for the 2022 cohort. Each point represents a single founder, and its horizontal position shows whether a given feature pushed the model's predicted likelihood of high funding attainment upward (positive SHAP value) or downward (negative SHAP value). The color of each point reflects the founder's feature value. Blue indicates a low feature value, while red indicates a high feature value. When red points appear on the right, high values of that feature increase predicted success. When red points appear on the left, high values decrease it. Features are ordered vertically by global importance, with the most influential features at the top. Below each feature label, the prevalence of that feature is shown in percentage terms for the given cohort, providing immediate context for how common or rare each signal is. All SHAP beeswarm plots for the 2018–2022 cohorts are included in Appendix G.

Figure 3: 2022 SHAP beeswarm plot



The calibration and discrimination metrics confirm that the predictive environment is empirically stable and statistically reliable across all cohorts. Performance measures, precision, recall, $F_{0.5}$, and AUCPR, remain consistent across prediction cohorts. Gradient Boosting, Random Forest and the policy models exhibit minimal variance (±0.05) in $F_{0.5}$ and AUCPR values, with Logistic Regression showing stable but lower scores, confirming temporal robustness (Hsu, 2007; Breiman, 2001b; Shmueli, 2010).

Table 5: Discrimination and Calibration Metrics

2018					2019				
Model	Logistic Regression	Random Forest	Gradient Boosting	Policy Model	**Model**	Logistic Regression	Random Forest	Gradient Boosting	Policy Model
$F_{0.5}$	10.93%	17.89%	19.31%	23.70%	**$F_{0.5}$**	13.10%	18.08%	20.42%	23.75%
Precision	14.57%	18.98%	19.80%	25.33%	**Precision**	15.89%	18.44%	21.34%	25.42%
Recall	5.47%	14.55%	17.55%	18.86%	**Recall**	7.70%	16.75%	17.43%	18.80%
Accuracy	91.20%	92.40%	94.40%	93.57%	**Accuracy**	87.44%	91.20%	95.30%	93.23%
AUCPR	0.407	0.492	0.494	0.511	**AUCPR**	0.417	0.513	0.599	0.499
Brier	0.049	0.053	0.05	0.089	**Brier**	0.046	0.057	0.046	0.094
ECE	0.033	0.046	0.04	0.064	**ECE**	0.038	0.046	0.042	0.053

2020					2021				
Model	Logistic Regression	Random Forest	Gradient Boosting	Policy Model	**Model**	Logistic Regression	Random Forest	Gradient Boosting	Policy Model
$F_{0.5}$	12.21%	19.08%	20.54%	23.60%	**$F_{0.5}$**	12.68%	20.00%	21.30%	24.24%
Precision	13.22%	19.55%	20.87%	24.55%	**Precision**	14.32%	20.01%	21.78%	25.66%
Recall	9.34%	17.42%	19.33%	20.44%	**Recall**	8.69%	19.98%	19.56%	19.86%
Accuracy	88.89%	90.24%	96.30%	95.36%	**Accuracy**	89.45%	86.55%	94.30%	93.30%
AUCPR	0.411	0.587	0.613	0.532	**AUCPR**	0.411	0.547	0.637	0.494
Brier	0.047	0.049	0.048	0.097	**Brier**	0.046	0.049	0.0531	0.089
ECE	0.041	0.052	0.044	0.048	**ECE**	0.038	0.052	0.05	0.053

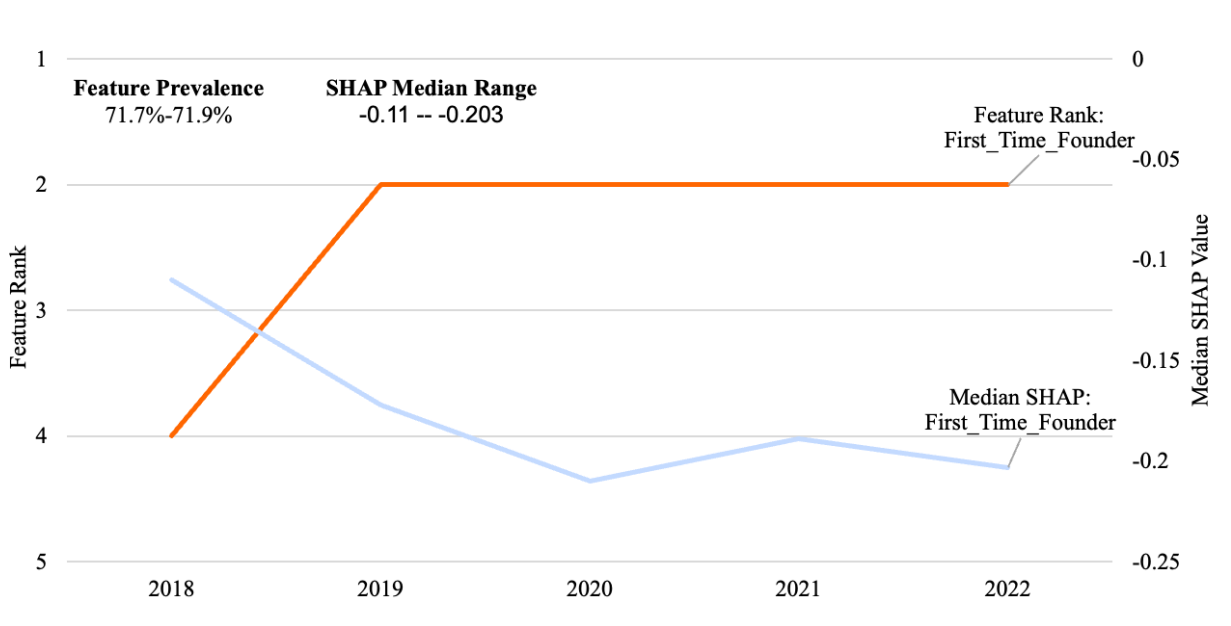
2022				
Model	Logistic Regression	Random Forest	Gradient Boosting	Policy Model
$F_{0.5}$	12.52%	18.34%	22.10%	25.11%
Precision	14.67%	18.55%	22.24%	26.30%
Recall	7.90%	17.54%	21.55%	21.25%
Accuracy	88.98%	83.54%	93.67%	94.20%
AUCPR	0.411	0.587	0.621	0.563
Brier	0.047	0.048	0.047	0.087
ECE	0.038	0.051	0.049	0.051

Calibration results confirm that the models' predictions are trustworthy. The Brier scores (0.046–0.053) and ECE values (below 0.052) fall within the thresholds outlined in Section 3.6. Across all years, the ranking of model performance remains the same, Gradient Boosting performs best, followed by Random Forest, then Logistic Regression, with the policy model outperforming the ML models. Thus, the models are not reacting to noise in individual cohorts but reflecting stable underlying patterns in the data. Together, the strong calibration and consistent performance rankings demonstrate that the predictive environment is empirically stable. Due to their superior performance, the Gradient Boosting model - and the interpretable policy-induction model are used to evaluate the research question and three predictions outlined in Section 2.3.

Prediction 1 proposes that prior entrepreneurial experience remains a consistently strong predictor of funding success. Specifically, founders identified as *First_time_founder* are expected to display stable or increasing importance in the model from 2018 to 2022, while their lack of prior founding experience should lower the predicted likelihood of receiving investment.

Feature importance analysis displays that entrepreneurial experience is one of the strongest predictors of funding success. Figure 4 displays that, across all cohorts from 2018–2022, *First_time_founder* consistently ranks between second and fourth place in importance. This indicates that the model consistently treats prior entrepreneurial experience, or the absence of it, as one of the most influential features for making predictions.

Figure 4: *First_Time_Founder* Feature Analysis



Directional effects reinforce this pattern**.** In every cohort, the median SHAP value for *First_time_founder* is negative (–0.20 to –0.11), meaning that lacking prior entrepreneurial experience consistently lowers the model's predicted likelihood of securing funding. Since roughly 72% of founders in each cohort are first-time founders, this negative effect applies to most of the population and remains virtually unchanged over time. In contrast, founders with prior ventures, the remaining ~28%, show equally stable positive effects on predicted funding success. The persistence of both the penalty for first timers and the advantage for not being a first-time founder indicates that this is not a temporary pattern, but a structural feature of how the model evaluates founder experience.

Policy ablation analysis outlines that entrepreneurial experience remains one of the most consistently emphasized signals in the policy model across all cohorts from 2018 to 2022. In 2018 and 2019, eight to nine rules explicitly reference founder history, establishing prior entrepreneurial activity as a core structural feature of the model's reasoning.

Table 6: Example - Policy Ablation Results 2021 - Top 3 out of 20

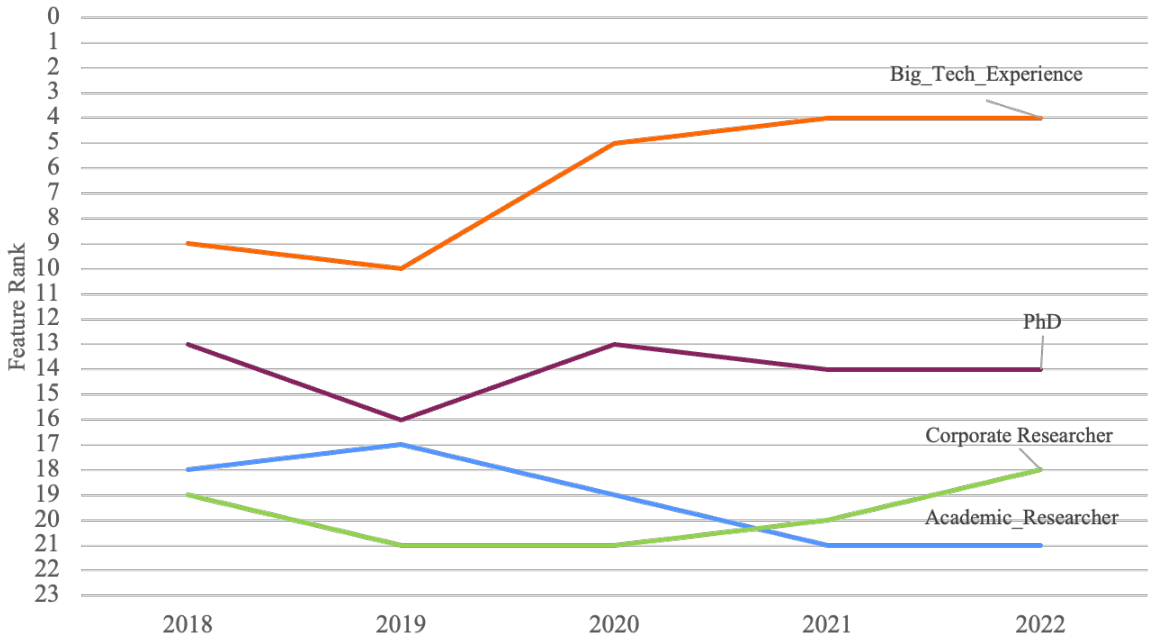
Importance Ranking	Policy Rule	Importance Score
1	Long tenure in junior roles, or lack of advancement, is a negative signal unless offset by clear entrepreneurial or leadership activity.	0.0301148
2	Multiple distinct leadership, founder, or principal roles (co-founder, chief, partner, director, advisor, board member) across different organizations, sectors, or regions before or at company inception are strong positive signals.	0.0158825
3	Non-traditional education or lack of degree can be positive if accompanied by significant, varied work or leadership experience, especially across sectors or regions.	0.0158825

In 2020, the number dips slightly, but the underlying penalty for first-time founders persists, now embedded within a more operational and cross-functional framing. By 2021, the emphasis returns to 2019 levels, reaffirming the stability of this signal. The trend culminates in 2022, where twelve of twenty rules reference entrepreneurial experience - the strongest representation across all years - indicating that founder track record becomes increasingly central to how the model evaluates funding potential.

Prediction 2 examines whether *Big_Tech_Experience* functions as a consistently stronger predictor of funding attainment than other forms of domain expertise. *Big_Tech_Experience* ranks among the top predictors in every cohort (positions 9 to 4), while academic and research-oriented features - *PhD*, *Academic_Researcher*, *Corporate_Researcher* - remain at the bottom of the importance ranking. This indicates that Big Tech backgrounds carry substantially greater evaluative weight than academic credentials.

Feature prevalence helps clarify this pattern. Although still relatively rare, *Big_Tech_Experience* becomes more common over time, increasing from roughly 5% of founders in 2018 to more than 11% in 2022. By contrast, academic and research credentials remain both infrequent and stable across cohorts. The SHAP beeswarm plots reflect this distribution. For most founders who lack Big Tech backgrounds, its absence produces a small negative effect on predicted funding likelihood. For the minority where it is present, however, the effect is strongly positive.

Figure 5: Prediction 2 Feature Analysis



This pattern becomes clear when comparing median and mean SHAP values. The median SHAP values are slightly negative (–0.136 to –0.079) simply because nearly all founders lack Big Tech experience. But the mean SHAP value is strongly positive (≈ +0.20), revealing the model's impact for the small minority of cases where *Big_Tech_Experience* is present, giving a large boost to the predicted funding likelihood. Thus, *Big_Tech_Experience* is a rare but powerful signal, too uncommon to lift most founders' predictions, but strong enough that for those who do have it, it increases the predicted likelihood of high funding attainment. The academic features consistently show minimal predictive influence across all cohorts. Their SHAP values cluster near zero, indicating that they do not meaningfully affect the model's predicted likelihood of receiving funding in general technology ventures.

It is important to note that the policy-induction model operates on anonymized founder profiles. Company names and employer identities are removed to avoid leakage and to ensure generalizable rule induction. As a result, the model cannot explicitly recognize *Big_Tech_Experience*. Nonetheless five to eight mid- and high-ranking rules reward technical depth, cross-functional mobility, and product-building experience, while academic credentials appear only twice per year and only as weak or conditional signals in 2018 and 2019. This aligns with the internal logic of Prediction 2 although *Big_Tech_Experience* is not explicitly mentioned. In 2020, the pattern persists, with operational signals remaining prominent while academic ones continue to appear in low-ranked and negative positions. By 2021, nine rules reflect this, further strengthening the model's emphasis on execution over academic specialization. The trend peaks in 2022, with twelve rules rewarding operational and leadership profiles while academic features remain marginal. Collectively, these results show a convergence on execution, operational scale, and cross-functional versatility as central evaluative signals, with academic backgrounds consistently deprioritized across time. This divergence and the growing relevance of *Big_Tech_Experience* over time outlined by the ML model - suggests that investors systematically prioritize signals of scalable execution capability over those signaling analytical or technical depth in general technology ventures.

Prediction 3 focuses on the conditional value of highly academic experience as a signal for funding attainment. It expects academic features *PhD*, *Academic_Researcher* and *Corporate_Researcher*, to contribute positively only when combined with practical career signals such as *Board_Member*, *Multiple_Work_Experiences*, *Big_Tech_Experience* or *Executive_Experience*. Academic experience on its own is not expected to have strong predictive weight.

SHAP interaction analysis across all cohorts (2018–2022) shows that founder signals are evaluated independently rather than in combination. Most feature pairs have interaction values extremely close to zero, meaning the model treats their effects as additive, not synergistic (Appendices I and J).

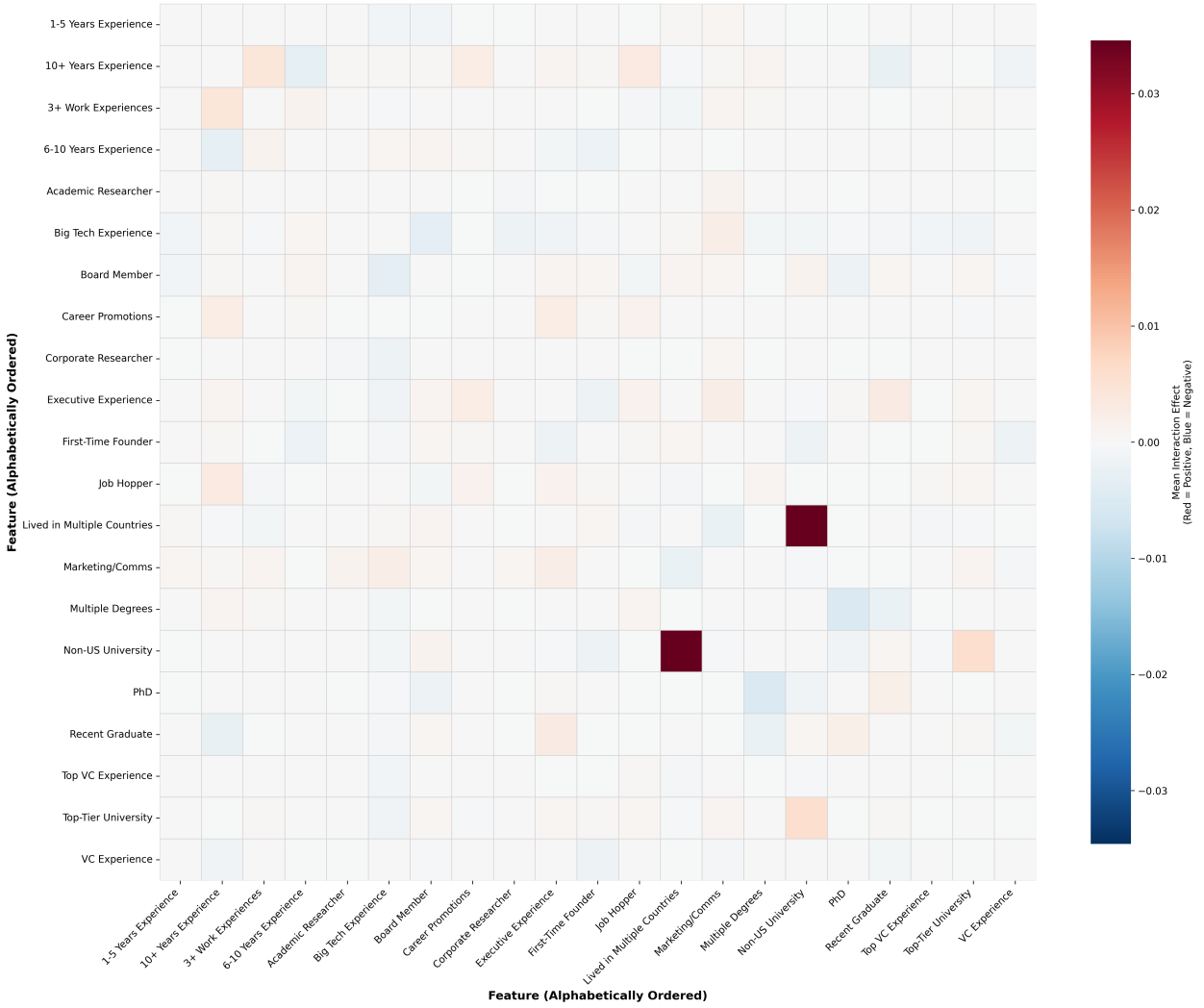
The few meaningful synergies that do appear, occur within the same signal category. For example, international background indicators reinforce each other (*Lived_in_Multiple_Countries* × *Non_US_University* averages +0.026). These patterns are small but consistent, showing stable, category-internal interactions over time.

By contrast, Prediction 3's expected complementarity between academic (*PhD*, *Academic_Researcher*, *Corporate_Researcher*) and practical signals (*Board_Member*, *Big_Tech_Experience*, *Executive_Experience*) does not appear in the SHAP interaction matrices over the years. Almost all academic–practical feature pairings show interaction values below 0.001, effectively zero, and several trend slightly negative. Only about one-third show any positive sign and even those are too small to indicate meaningful synergy. *PhD* × *Board_Member* becomes increasingly negative in later years, and *Corporate_Researcher* × *Big_Tech_Experience* declines to –0.002 by 2022. The only consistent positive interaction is *PhD* × *Recent_Graduate*.

Across all cohorts from 2018 to 2022, the interaction between academic credentials and execution signals is consistently reflected in both the content and ranking positions of the policy rules. In 2018 and 2019, academic-related rules appear only twice per year and always in mid-to-low ranking slots, showing weak standalone influence and contributing positively only when paired with strong execution markers such as technical depth, leadership progression, or founder roles. In 2020, this becomes even clearer. The single conditional-positive academic rule sits in a lower-importance tier, while two negative rules appear near the bottom of the set, reinforcing the minimal predictive weight of academic-only backgrounds. By 2021, academic rules remain confined to mid- and lower-tier positions, indicating that degrees or research experience matter only when accompanied by leadership or entrepreneurial activity. The 2022 rules provide the strongest evidence of this pattern, with all academic-related policies clustered in mid-to-low positions even when

framed as conditionally positive, confirming that academic experience remains a secondary, complementary signal rather than a primary driver of prediction.

Figure 6: SHAP Interaction Matrix - 2022 cohort



Overall, SHAP interaction values are near zero or negative, indicating that academic credentials function as low-impact, standalone signals rather than becoming more predictive when paired with practical experience, and thus show no systematic academic–practical complementarity. The policy results, however, offer more nuanced insights. They reveal that academic backgrounds can contribute positively when combined with operational or leadership experience.

5. Analysis & Discussion

This thesis tests whether founder signals maintain predictive validity across 2018-2022 cohorts, both individually and in combination. Three patterns emerged. First, operational experience signals (*First_Time_Founder* and *Big_Tech_Experience*) demonstrate persistent importance with temporal generalization. Second, signal hierarchies favor operational experience over academic credentials, with *Big_Tech_Experience* consistently outranking academic experience. Third, signals operate independently rather than synergistically, academic-practical combinations show no complementarity.

Taken together, these results display that founder signals are stable, hierarchically ordered toward operational experience, and additive rather than synergistic, with no evidence that academic and practical credentials amplify each other's effects. Additionally, the study also demonstrates how predictive-modeling techniques, including both traditional machine-learning models and an interpretable LLM-based policy framework, can be used to evaluate the temporal generalizability and structure of founder signals.

Prediction 1 proposes that founders without prior entrepreneurial experience would face systematically lower predicted funding likelihood across the yearly cohorts 2018-2022. *First_Time_Founder* remains one of the top predictors in every cohort from 2018–2022, and its directional effect is consistently negative, indicating a stable and generalizable penalty for founders lacking prior venture-building experience. The consistency across five different cohorts displays that the pattern is stable across samples.

The policy model supports this across all cohorts. From 2018 to 2022, between eight and twelve rules per year explicitly reference prior founding or leadership experience, making entrepreneurial history one of the most consistently emphasized signals in the rule set. Although the salience of these rules fluctuates slightly across cohorts, the underlying penalty for first-time founders persists and becomes even more pronounced in 2022, where founder-experience rules reach their highest representation. This mirrors the feature-importance trends and aligns with the policy ablation results, which show that removing entrepreneurial-experience rules produces some of the largest declines in model performance.

The persistence of this penalty aligns closely with prior research suggesting that entrepreneurial experience functions as a stage-dependent signal. Ko and McKelvie (2018) demonstrate that experience is most influential in the earliest funding rounds, when uncertainty is highest and firm-level indicators are not yet available. Social-capital mechanisms also matter, as experienced founders more effectively mobilize prior professional ties and assemble stronger early teams (Gupta et al., 2024). This interpretation is further supported by the research of Delmar & Shane, (2006), highlighting that the benefits of entrepreneurial experience are threshold-based. Moving from no experience to some experience provides significant advantages, whereas additional ventures yield diminishing marginal effects. As ventures mature and firm-level performance information accumulates, later-stage investors increasingly differentiate among types of experience. Its relevance, domain fit, and prior success become more important (Hsu, 2007; Zhang, 2009). Thus, the value of entrepreneurial experience becomes more nuanced the more a venture matures.

Since previous research outlines that the value of entrepreneurial experience becomes more nuanced as ventures mature, the temporal stability observed in this study may seem paradoxical. This discrepancy may stem from path dependence, where each funding stage depends on successfully clearing the preceding one. Early rounds supply the capital, momentum, and credibility needed to access larger later rounds, meaning that founders who struggle at Seed or Series A due to inexperience rarely accumulate the multi-round funding required to enter the top 3% of their cohort. As a result, the *First_Time_Founder* penalty persists across five-year cumulative funding outcomes measured in this study because early-stage gatekeeping constrains downstream funding attainment.

Consequently, the results show that what stays stable over time is the way investors judge founders in the earliest stages, not that experience matters equally at every stage of fundraising. The findings complement rather than contradict the referenced stage-specific research. This study establishes temporal generalizability of this gatekeeping function across the yearly cohorts 2018-2022.

Prediction 2 proposes that *Big_Tech_Experience* would function as a consistently stronger predictor of funding attainment than *PhD*, *Academic_Researcher* and *Corporate_Researcher* across cohorts. *Big_Tech_Experience* consistently ranks among the model's most influential predictors of funding success across all cohorts from 2018–2022. Academic and research-oriented features occupy the lowest positions in the feature rankings and exert minimal influence on the model's predictions. This temporal consistency indicates that the model systematically assigns greater predictive weight to operational and organizational experience than to academic or research-based expertise when forecasting funding outcomes.

Across all cohorts, the policy model offers convergent evidence for this pattern. Although it cannot identify *Big_Tech_Experienc*e explicitly, the policy rules consistently reward the same underlying competencies, such as technical depth, cross-functional mobility, rapid advancement, and operational leadership, traits strongly associated with *Big_Tech_Experience*. These rules grow more prominent across cohorts, while academic-related rules remain few, low-ranked, and largely conditional on execution experience. The alignment between policy induction and the feature-importance results reinforces Prediction 2.

One core reason for this persistent hierarchy lies in the information gaps that early-stage investors must overcome in general technology investing. Early-stage investors must assess whether a founder can execute under uncertainty, build a product, assemble a team, and scale an organization. The model's consistently high weighting of *Big_Tech_Experience* reflects how directly this signal speaks to the core execution-focused criteria investors use in general technology ventures. Such experience signals familiarity with complex, high-velocity environments, exposure to scalable product architectures, and experience working within structured organizational processes. This signal is both costly and difficult to imitate, making it highly diagnostic towards execution capability.

Prior research reinforces this interpretation. Dworak (2022) finds that Big-Tech alumni are significantly more likely to secure early-stage funding because their backgrounds reduce perceived uncertainty among investors. Similarly, work by Beckman et al. (2007) and Gimmon & Levie (2010) shows that managerial and commercially oriented experience meaningfully improves resource acquisition and venture survival. In contrast, academic credentials signal a very different type of capability. PhD training, academic research roles, and corporate R&D experience convey depth, analytical skill, and specialized technical competence. Yet, in general tech contexts, these competencies are not the primary source of uncertainty for investors (Engel, 2015).

Importantly, this contrast does not imply academic credentials are universally weak. Rather, their importance may be domain specific. In science-intensive sectors, such as biotechnology, pharmaceuticals or life science for example, academic depth is central because the core uncertainty investors face is scientific feasibility (Hsu, 2007; Swift, 2018). The signal hierarchy observed in this study suggests that investors in general technology ventures treat execution uncertainty as the dominant risk. Thus, *Big_Tech_Experience* carries far more weight to the predictions, while academic credentials offer little additional information and therefore remain weak signals in comparison.

Overall, we demonstrate that *Big_Tech_Experienc*e functions as a rare but highly rewarded signal that generalizes across time and cohorts, while academic credentials remain consistently weak in this setting. The temporal generalizability of this hierarchy suggests that investors in general tech consistently value execution ability over academic depth and look to operational track records to reduce early-stage uncertainty.

This finding refines signaling theory by displaying how sector-specific evaluation priorities shape the informational value of founder signals and by illustrating that the strength of a signal depends not only on its costliness or observability, but also on its alignment with the dominant uncertainties of the domain.

Prediction 3 proposes that academic credentials would contribute positively to funding attainment only when paired with practical, market-oriented experience. This expectation reflects a longstanding theoretical assumption that deep scientific or analytical expertise becomes valuable only when complemented by signals of execution capability. Prediction 2 already demonstrated that practical, operational signals carry substantially more evaluative weight than academic ones when considered individually. Prediction 3 therefore focuses on whether academic credentials gain relevance not on their own, but when combined with the operational signals that investors already prioritize.

Across the empirical results, however, the feature-level predictive models show no evidence of such complementarity. SHAP interaction values across all cohorts remain near zero or slightly negative, indicating that academic and practical signals operate independently. When academic credentials are paired with *Executive_Experience* or *Big_Tech_Experience*, the combined effect is no stronger than the effect of practical experience alone. Conversely, academic signals do not become more informative when accompanied by operational backgrounds. Academic credentials therefore act as low-impact, standalone indicators, and practical experience retains its influence regardless of whether academic depth is present.

The policy model provides complementary - and methodologically insightful - evidence for this pattern. Across all cohorts from 2018 to 2022, academic-related rules appear only sparsely and consistently occupy mid-to-low ranking positions. When they contribute positively, it is always conditional on strong execution markers such as leadership progression, founder roles, or technical depth. Importantly, these conditional rules do not imply genuine interaction effects in the statistical sense; rather, they reflect the policy model's tendency to encode execution capability as the primary evaluative criterion and to treat academic credentials as relevant only insofar as they indirectly proxy for that capability. Academic-only profiles are repeatedly encoded as weak or negative signals, and this pattern becomes more explicit over time. The policy ablation results reinforce this interpretation. Removing execution-oriented rules meaningfully harms performance, whereas removing academic-only rules produces little measurable change.

Together, the ML interaction analysis and the policy ablation results generate an important methodological insight. While the ML model evaluates feature relationships categorically and therefore captures no academically meaningful interaction effects, the policy-induction model reveals small but systematic conditional patterns that a purely tabular representation cannot express. These rule-based structures offer a more nuanced and interpretable account of why predictions are reached, showing how academic credentials may add limited but context-dependent value when embedded within strong operational profiles. This suggests that policy induction can complement conventional predictive modelling by exposing subtle dependencies that might be functionally relevant even when too weak or sparse to appear as formal statistical interactions. In this sense, the integration of LLM-based interpretability represents a methodological contribution, expanding the analytical vocabulary available to predictive entrepreneurship research.

Several theoretical perspectives help explain the empirical pattern. Research on scientific human capital by Zucker et al. (1998) shows that academic expertise is embodied, highly specialized, and domain specific. Its value often lies in tacit frontier knowledge that does not automatically scale when combined with managerial or corporate experience. In this view, academic and practical experience solve different problems only in certain contexts - most notably in science-intensive sectors, where investors face both scientific uncertainty and commercialization uncertainty. In such settings, academic credentials address the technical feasibility question, while practical experience addresses execution. When these uncertainties refer to distinct evaluative challenges, the signals can complement each other.

This logic does not apply in the general-technology context studied here. Software ventures typically face execution uncertainty, not scientific uncertainty. The technology risk is comparatively low, while the commercial, organizational, and scaling risks are high. As a result, both academic and practical signals end up addressing the same investor question: *Can this founder execute?* Academic credentials provide only indirect evidence of execution capability, while operational backgrounds provide direct, verifiable evidence. Because the signals map onto the same evaluative concern, the more direct and interpretable practical signal dominates, rendering academic depth redundant rather than complementary.

Research on hybrid academic–industry careers reinforces this interpretation. Eesley and Roberts (2012) demonstrate that academic–practical combinations create value only under conditions of knowledge adjacency - when the founder's industry experience closely matches their scientific domain. The founders in this study do not operate under such adjacency conditions, reducing the likelihood of meaningful complementarity. Likewise, Toole and Czarnitzki (2008) document a persistent "translation gap" in which academic expertise enhances discovery but does not automatically enhance commercialization unless paired with experience specifically geared toward bridging science and markets. Generic operational experience does not reliably provide this bridge, which helps explain the independence observed here.

Taken together, the findings for Prediction 3 suggest that academic–practical complementarity is not a universal mechanism but one that appears only when signals speak to different kinds of uncertainty. In general technology ventures, academic depth and operational experience both function as answers to the same evaluative question - founder execution capability - and therefore combining them adds little new information. Investors appear to view academic signals as domain-specific, useful mainly when scientific feasibility is the primary uncertainty. Beyond the prediction-guided analyses, the combined ML and policy results highlight subtle dependencies that warrant further investigation and illustrate how LLM-based policy induction can enhance predictive modelling by revealing interpretable patterns that traditional tabular ML methods alone cannot surface.

6. Conclusions

This thesis set out to address the methodological and replication gaps in founder-signaling research by treating generalizability as an empirical property rather than an assumption. To evaluate whether founder human- and social-capital signals persist across new cohorts, the study employed a novel LLM-based policy-induction model alongside established machine-learning approaches addressing the research question:

To what extent do founder human- and social-capital signals generalize out of sample when predicting venture funding outcomes, both individually and in combination?

The results show clear evidence of temporal generalization. Models trained on earlier cohorts (2012–2019) reliably predicted funding attainment in later cohorts (2018–2022), demonstrating that several founder signals retain stable predictive value across time. Taken together, the findings demonstrate that the predictive power of founder human and social capital signals is not sample-bound but reproducible across temporal contexts, clarifying which mechanisms of signaling theory persist in real-world venture evaluation and which remain context-dependent or non-generalizable. The sections that follow synthesize these insights into (1) theoretical contributions, (2) managerial implications, (3) limitations, and (4) directions for future research.

6.1 Theoretical Contributions

This thesis contributes to signaling theory and predictive modeling through three theoretical and two methodological advances.

This thesis makes three main contributions to signaling theory. First, it directly answers recent calls in the signaling literature for methods that test whether founder–signal relationships generalize beyond the samples and periods in which they were originally estimated (Bafera & Kleinert, 2023; Zhang et al., 2023). To address this gap, the study develops a hybrid predictive-validation framework that combines a state-of-the-art LLM-based policy induction with established machine-learning models. The LLM-based policy model contributes a novel methodological capability to management research by producing transparent, human-readable decision rules through iterative natural-language reasoning, allowing explicit inspection of the logic underlying predictions. The results outline that the policies deliver additional insights for theory validation. In parallel, the machine-learning models enable rigorous out-of-sample temporal prediction, providing empirical tests of whether theoretically grounded signals retain predictive value when applied to new founders across the yearly cohorts 2018-2022.

Together, these complementary model families demonstrate how predictive modeling, augmented with interpretable LLM architectures, can be used as a tool for theory evaluation, not merely prediction. They demonstrate how temporal generalizability, signal persistence, and configurational independence can be assessed empirically, expanding the available methods for testing the scope, boundaries, and robustness of signaling theory in venture contexts. This contribution advances the field by offering a replicable framework for assessing whether theoretical mechanisms travel across time, contexts, and populations.

Second, the study demonstrates that founder human- and social-capital signals generalize out of sample across time. The analysis reveals that signal–outcome relationships are not sample-specific but retain predictive relevance when applied to entirely new founders in new periods. This establishes temporal generalization as an empirical property of key founder signals. Moreover, temporal generalization is asymmetric. Signals rooted in operational experience, especially Big-Tech backgrounds, remain consistently strong predictors across all temporal cohorts, indicating that investors repeatedly rely on these cues even as market conditions evolve. In contrast, academically oriented credentials, PhD training, academic research roles, corporate research experience, remain consistently weak and show no increase in predictive relevance. The

stability of this hierarchy across five independent test cohorts suggests that these patterns reflect durable mechanisms of investor evaluation, not correlations tied to a particular period or sample. Thus, founder signals do not possess universal predictive value. Their effectiveness depends on whether they address the specific uncertainty investors seek to resolve. In general technology ventures, where the dominant evaluative question is whether a founder can execute and scale, operational experience signals consistently retain predictive power, while academic credentials do not. This contrast demonstrates that signals matter only when they meaningfully reduce the type of uncertainty that characterizes the domain. In refining signaling theory, the thesis shows that its mechanisms generalize not universally but only when the informational content of a signal aligns with the sector's core evaluation priorities.

Third, the thesis specifies configurational scope boundaries. Academic-practical signal combinations, which demonstrate synergies in science-intensive contexts (Hsu, 2007; Swift, 2018), exhibit independence in general technology ventures. Academic credentials seem to not gain substantive predictive value when paired with practical experience. This finding challenges the generalizability of complementarity documented in more science-heavy contexts and suggests configurational effects are sector-specific. One explanation is dimensional overlap. In general tech, both academic and practical signals may address the same investor concern, execution capability, creating redundancy rather than synergy. In more science-heavy contexts, by contrast, academic credentials address technical feasibility while practical experience addresses commercialization, different investor uncertainties producing genuine complementarity. This refines configurational theory (Furnari et al., 2021) by specifying when signal combinations amplify versus operate independently. Complementarity requires that signals provide distinct information and address different types of uncertainty. When signals target the same evaluation dimension, as academic and practical credentials appear to in general tech, they remain additive rather than multiplicative. This distinction specifies structural limits of configurational effects across entrepreneurial contexts.

6.2 Practical Implications

The findings carry practical relevance for both founders and investors. Founders should prioritize building operational experience when shaping their career trajectories. Big-Tech experience consistently emerges as one of the strongest predictors of funding success, exerting a substantial positive influence whenever present. Academic credentials alone, by contrast, display near-zero predictive value in general technology ventures. Thus, a founder choosing between a PhD program and a role at a major tech company should recognize that these paths carry markedly different signaling weight for investors. This does not imply academic routes are inherently inferior, but founders pursuing them should understand they do not accumulate funding-relevant credibility to the same degree as Big-Tech experience does.

Moreover, the predictive models consistently assign lower funding likelihoods to first-time founders, indicating that raising institutional capital is systematically more difficult for founders without prior entrepreneurial experience. This disadvantage arises at the earliest stages of evaluation, where investors rely most heavily on experience-based signals to reduce uncertainty, and where early outcomes shape the entire cumulative funding trajectory. Consequently, first-time founders must deliberately compensate for the absence of founding experience, by recruiting experienced advisors, joining reputable accelerators, or partnering with serial entrepreneurs.

Lastly, investors can refine screening decisions by prioritizing signals that consistently predict funding outcomes. Across all cohorts (2018–2022), operational experience, such as Big Tech roles, prior entrepreneurial ventures, and executive positions, shows strong and stable predictive value, whereas academic credentials contribute little information in general technology contexts. This persistence indicates durable evaluation priorities rather than transient market shifts. Accordingly, investors should avoid overweighting prestigious academic credentials which may indicate intelligence but do not reliably signal execution capability.

6.3 Limitations

Several important limitations of this study warrant consideration. The research deliberately focuses on individual founder characteristics rather than founding team dynamics. Although venture capital typically evaluates teams, this individual-level approach enables precise measurement of how personal credentials independently influence investor assessments. Consequently, the findings speak most directly to individual signal effects and may not fully capture the complementarity and synergies that emerge when diverse founder backgrounds combine within teams.

Geographically, the analysis concentrates on ventures in Europe and North America - regions characterized by mature venture capital ecosystems and relatively homogeneous institutional norms. These shared characteristics facilitate meaningful temporal comparison but limit generalizability to emerging markets where entrepreneurial ecosystems often feature distinct credential hierarchies, network structures, and investor decision-making logics shaped by different cultural and institutional contexts. The sectoral scope further constrains generalizability. This thesis examines general technology ventures while excluding science-intensive industries. In such sectors, academic credentials often carry greater weight because they signal necessary technical capability. The finding that academic signals are not top ranked may therefore not extend to contexts where scientific expertise is a core competitive advantage.

The dataset includes only founders who secured at least one equity funding round, excluding unfunded ventures, bootstrapped companies, and those not seeking institutional capital. This introduces selection bias by omitting founders who never obtained initial investment. The analysis therefore addresses which signals distinguish high performers within the funded population rather than which characteristics determine entry into that population. Additionally, although the study employs rigorous predictive validation through rolling-origin forecasting, it establishes temporal associations rather than causal relationships. Operational experience may predict funding not because it directly enhances venture quality but because it correlates with unmeasured attributes such as resilience, network access, or problem-solving ability. Disentangling correlation from causation would require experimental or quasi-experimental designs.

The reliance on proprietary data sources introduces replicability constraints, as external researchers cannot independently replicate the analysis without equivalent institutional access. Additionally, the dynamic nature of online data sources such as LinkedIn poses challenges, as founder profiles can be updated, deleted, or made private over time. The analysis spans 2018 through 2022, a relatively short window in the broader history of venture capital. Signal hierarchies that appear stable across these five years may shift over longer horizons as technological paradigms evolve, labor markets transform, and investor preferences adapt. Finally, the study's focus on predictive accuracy prioritizes external validity over mechanistic explanation. Understanding why operational experience matters more than academic credentials, or why signal interactions remain weak, requires complementary qualitative research exploring how investors interpret credentials and integrate multiple information sources into holistic judgments.

Moreover, the study’s temporal window (2018–2022) necessarily includes the COVID-19 years, during which venture-capital markets experienced unusual volatility, shifts in funding pace, and rapid changes in investor risk preferences. These macro-level disruptions may have temporarily altered the distribution of funding outcomes and the relative weight investors assign to founder signals. Because these dynamics fall outside the study’s control and are not explicitly modeled, the observed patterns may partially reflect period-specific conditions rather than purely structural signal effects. Lastly, the analysis evaluates generalizability only for founders who started their ventures between 2018 and 2022 and measures funding outcomes over a maximum five-year window, extended only until 2025 due to data availability constraints. This design ensures comparability across cohorts but limits the ability to assess whether signal hierarchies persist in more recent founding periods or over longer funding horizons. Ventures founded after 2022, and ventures requiring more than five years to reach meaningful funding milestones, fall outside the scope of

observation. As a result, the findings may not fully capture post-2022 shifts in investor behavior or long-term signal value trajectories.

6.4 Future Research

This thesis outlines several promising directions for future research. (1) extending predictive-validation designs across sectors, geographies, and outcomes, (2) integrating policy-induction insights into feature engineering, (3) establishing causal mechanisms behind predictive signals, (4) assessing algorithmic fairness in founder evaluation and (5) developing the additional signal patterns surfaced beyond the three focal predictions.

(1) Applying the predictive-modeling framework across new contexts represents an important next step. Academic credentials may gain predictive relevance in science-intensive sectors. Founder evaluation norms may differ in emerging markets. Shifting from individual founders to full founding teams could reveal diversity and complementarity effects not observable here. Extending the outcome variable to time-to-exit, survival, or later-stage growth would also clarify whether the signal hierarchy observed in early funding persists or evolves over the entire venture lifecycle.

(2) A promising avenue is to integrate the policy-induction model and the machine-learning models into a unified hybrid architecture. Embedding high-performing policy rules directly into the feature space of traditional ML models would allow researchers to test whether rule-derived configurations improve temporal generalization and capture higher-order decision logic that standalone models miss.

(3) Introducing more granular founder attributes would allow sharper tests of configurational effects. Examples include PhD discipline, executive function (CEO, CMO), Big Tech role type (engineering, commercial), and prior-venture outcomes. This refinement would clarify whether academic–practical combinations truly lack complementarity, or whether synergies arise only in domain-specific pairings that binary indicators cannot detect.

(4) A key next step is identifying the causal mechanisms behind the predictive patterns. Quasi-experimental settings, such as exogenous Big Tech layoffs or regulatory shocks offer opportunities to test whether operational experience truly increases funding prospects, whether academic credentials remain irrelevant once endogeneity is addressed, and whether the first-time-founder penalty reflects skill deficits, network disadvantages, or investor bias. Clarifying these mechanisms would reveal whether investors respond to genuine capability differences or to correlated but non-causal cues.

(5) In addition to the three predictions examined in this thesis, the models also surfaced several additional patterns that fall outside the core analytical scope but point to promising avenues for future research. Several signals, such as career promotions, multiple work experiences, board membership, top-tier university attendance, and VC experience consistently rank among the most influential predictors, even though they were not part of the focal predictions. Their prominence suggests that investors may use broader heuristics related to status, career trajectory, or institutional validation, not just operational capability, when evaluating founders. Early-cohort interaction patterns (2018–2019) likewise reveal dynamics that look different from the later, more stable years. For example, board membership often shows negative interactions unless paired with strong operational experience, and job-hopping penalties appear particularly strong for founders with mid-career tenure. At the same time, some early interactions, such as combinations reflecting global exposure or academic depth paired with broader work experience show positive effects. These early, interaction-sensitive patterns contrast with the mostly additive evaluation logic that dominates from 2020 onward. Together, these findings suggest that investor decision heuristics may have shifted over time, potentially due to changing market conditions, evolving investor preferences, or broader macroeconomic shocks.

Future research should explore whether these patterns represent structural changes in how founders are evaluated or simply reflect temporary fluctuations in the entrepreneurial environment.

Bibliography

Acharya, S., Pustokhina, I. V., Pustokhin, D. A., Geetha, B. T., Joshi, G. P., Nebhen, J., Yang, E., & Seo, C. (2021). An improved gradient boosting tree algorithm for financial risk management. *Knowledge Management Research & Practice*, 1–12. https://doi.org/10.1080/14778238.2021.1954489

Aguinis, H., Beltran, J. R., & Marshall, J. D. (2024). Performance: Confirming, refining, and refuting theories. *Journal of Management Scientific Reports*, *2*(2). https://doi.org/10.1177/27550311241247487

Ahlers, G. K. C., Cumming, D., Günther, C., & Schweizer, D. (2015). Signaling in Equity Crowdfunding. *Entrepreneurship Theory and Practice*, *39*(4), 955–980. https://doi.org/10.1111/etap.12157

Akerlof, G. A. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, *84*(3), 488–500. https://doi.org/10.2307/1879431

Anglin, A. H., Short, J. C., Ketchen, D. J., Allison, T. H., & McKenny, A. F. (2019). Third-Party Signals in Crowdfunded Microfinance: The Role of Microfinance Institutions. *Entrepreneurship Theory and Practice*, *44*(4), 623–644. https://doi.org/10.1177/1042258719839709

Arroyo, J., Corea, F., Jimenez-Diaz, G., & Recio-Garcia, J. A. (2019). Assessment of Machine Learning Performance for Decision Support in Venture Capital Investments. *IEEE Access*, *7*, 124233–124243. https://doi.org/10.1109/ACCESS.2019.2938659

Athey, S. (2017). Beyond prediction: Using big data for policy problems. *Science*, *355*(6324), 483–485. https://doi.org/10.1126/science.aal4321

Athey, S., & Imbens, G. W. (2019). Machine Learning Methods That Economists Should Know About. *Annual Review of Economics*, *11*(1), 685–725. https://doi.org/10.1146/annurev-economics-080217-053433

Bafera, J., & Kleinert, S. (2022). Signaling Theory in Entrepreneurship Research: A Systematic Review and Research Agenda. *Entrepreneurship Theory and Practice*, *47*(6), 104225872211384. https://doi.org/10.1177/10422587221138489

Balachandran, S. (2024). The inside track: Entrepreneurs' corporate experience and startups' access to incumbent partners' resources. *Strategic Management Journal*, *45*(6). https://doi.org/10.1002/smj.3576

Baum, J. A. C., & Silverman, B. S. (2004). Picking winners or building them? Alliance, intellectual, and human capital as selection criteria in venture financing and performance of biotechnology startups. *Journal of Business Venturing*, *19*(3), 411–436. https://doi.org/10.1016/s0883-9026(03)00038-7

Becker-Foss, E. (2024). Performance and reporting predictability of hedge funds. *Journal of Forecasting*, *43*(6). https://doi.org/10.1002/for.3122

Beckman, C. M., Burton, M. D., & O'Reilly, C. (2007). Early teams: The impact of team demography on VC financing and going public. *Journal of Business Venturing*, *22*(2), 147–173. https://doi.org/10.1016/j.jbusvent.2006.02.001

Bergh, D. D., Connelly, B. L., Ketchen, D. J., & Shannon, L. M. (2014). Signalling Theory and Equilibrium in Strategic Management Research: An Assessment and a Research Agenda. *Journal of Management Studies*, *51*(8), 1334–1360. https://doi.org/10.1111/joms.12097

Bergmeir, C., & Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences*, *191*(3), 192–213. https://doi.org/10.1016/j.ins.2011.12.028

Bernstein, S., Korteweg, A., & Laws, K. (2017). Attracting Early-Stage Investors: Evidence from a Randomized Field Experiment. *The Journal of Finance*, *72*(2), 509–538. https://doi.org/10.1111/jofi.12470

Bettis, R., Gambardella, A., Helfat, C., & Mitchell, W. (2014). Quantitative empirical analysis in strategic management. *Strategic Management Journal*, *35*(7), 949–953. https://doi.org/10.1002/smj.2278

Breiman, L. (2001a). Random Forests. *Machine Learning*, *45*(1), 5–32.

Breiman, L. (2001b). Statistical Modeling: The Two Cultures. *Statistical Science*, *16*(3), 199–215. JSTOR. https://doi.org/10.2307/2676681

Brown, T., Mann, B. F., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R.,

Ramesh, A., Ziegler, D. M., Wu, J. C. S., Winter, C., & Hesse, C. (2020). Language Models Are Few-Shot Learners. *34th Conference on Neural Information Processing Systems (NeurIPS 2020)*, *34*. https://doi.org/10.48550/arxiv.2005.14165

Bruton, G. D., Ahlstrom, D., & Obloj, K. (2008). Entrepreneurship in Emerging Economies: Where Are We Today and Where Should the Research Go in the Future. *Entrepreneurship Theory and Practice*, *32*(1), 1–14. https://doi.org/10.1111/j.1540-6520.2007.00213.x

Busenitz, L. W., Fiet, J. O., & Moesel, D. D. (2005). Signaling in Venture Capitalist-New Venture Team Funding Decisions: Does It Indicate Long-Term Venture Outcomes? *Entrepreneurship Theory and Practice*, *29*(1), 1–12. https://doi.org/10.1111/j.1540-6520.2005.00066.x

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B. A., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., & Isaksson, S. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour*, *2*(9), 637–644. https://doi.org/10.1038/s41562-018-0399-z

Cerqueira, V., Luís Torgo, Jasmina Smailović, & Igor Mozetič. (2017, October 1). A Comparative Study of Performance Estimation Methods for Time Series Forecasting. *2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. https://doi.org/10.1109/dsaa.2017.7

Chen, Q., Wang, L., Zheng, B., & Song, G. (2025). DAGPrompT: Pushing the Limits of Graph Prompting with a Distribution-aware Graph Prompt Tuning Approach. *WWW '25: Proceedings of the ACM on Web Conference 2025*, 4346–4358. https://doi.org/10.48550/arxiv.2501.15142

Chou, Y., Chuang, H. H., Chou, P., & Oliva, R. (2023). Supervised machine learning for theory building and testing: Opportunities in operations management. *Journal of Operations Management*, *69*(4). https://doi.org/10.1002/joom.1228

Choudhury, P., Allen, R. T., & Endres, M. G. (2020). Machine learning for pattern discovery in management research. *Strategic Management Journal*, *42*(1), 30–57. https://doi.org/10.1002/smj.3215

Chuang, H. H., Chou, Y., & Oliva, R. (2021). Cross-item Learning for Volatile Demand forecasting: an Intervention with Predictive Analytics. *Journal of Operations Management*, *67*(7), 828–852. https://doi.org/10.1002/joom.1152

Colombo, M. G., & Grilli, L. (2005). Founders' human capital and the growth of new technology-based firms: A competence-based view. *Research Policy*, *34*(6), 795–816. https://doi.org/10.1016/j.respol.2005.03.010

Colombo, M. G., & Grilli, L. (2010). On growth drivers of high-tech start-ups: Exploring the role of founders' human capital and venture capital. *Journal of Business Venturing*, *25*(6), 610–626. https://doi.org/10.1016/j.jbusvent.2009.01.005

Colombo, O. (2020). The Use of Signals in New-Venture Financing: A Review and Research Agenda. *Journal of Management*, *47*(1), 014920632091109. https://doi.org/10.1177/0149206320911090

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling Theory: a Review and Assessment. *Journal of Management*, *37*(1), 39–67.

Davis, J., Devos, L., Reyners, S., & Schoutens, W. (2021). Gradient boosting for quantitative finance. *The Journal of Computational Finance*. https://doi.org/10.21314/jcf.2020.403

De Bock, K. W., Bogaert, M., & du Jardin, P. (2025). Ensemble learning for operations research and business analytics. *Annals of Operations Research*, *353*(2), 419–448. https://doi.org/10.1007/s10479-025-06852-w

De Caigny, A., Coussement, K., & De Bock, K. W. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. *European Journal of Operational Research*, *269*(2), 760–772. https://doi.org/10.1016/j.ejor.2018.02.009

Delmar, F., & Shane, S. (2006). Does experience matter? The effect of founding team experience on the survival and sales of newly founded ventures. *Strategic Organization*, *4*(3), 215–247. https://doi.org/10.1177/1476127006066596

Donoho, D. (2017). 50 Years of Data Science. *Journal of Computational and Graphical Statistics*, *26*(4), 745–766. https://doi.org/10.1080/10618600.2017.1384734

Dudley, E. (2021). Social capital and entrepreneurial financing choice. *Journal of Corporate Finance*, *70*(102068), 102068. https://doi.org/10.1016/j.jcorpfin.2021.102068

Dutta, S., & Folta, T. B. (2015). Information Asymmetry and Entrepreneurship. In *Wiley Encyclopedia of Management* (pp. 1–5). Wiley. https://doi.org/10.1002/9781118785317.weom030055

Dworak, D. (2022, April 1). *Analysis of Founder Background as a Predictor for Start-up Success in Achieving Successive Fundraising Rounds*. https://backend.production.deepblue-documents.lib.umich.edu/server/api/core/bitstreams/5849f0a1-1de3-4e67-999c-a96a8f9e480b/content

Eesley, C. E., & Roberts, E. B. (2012). Are You Experienced or Are You Talented?: When Does Innate Talent versus Experience Explain Entrepreneurial Performance? *Strategic Entrepreneurship Journal*, *6*(3), 207–219. https://doi.org/10.1002/sej.1141

Eggers, A. C., Tuñón, G., & Dafoe, A. (2023). Placebo Tests for Causal Inference. *American Journal of Political Science*, *68*(3). https://doi.org/10.1111/ajps.12818

Engel, J. S. (2015). Global Clusters of Innovation: Lessons from Silicon Valley. *California Management Review*, *57*(2), 36–65. https://doi.org/10.1525/cmr.2015.57.2.36

Ewens, M., & Townsend, R. R. (2019). Are early stage investors biased against women? *Journal of Financial Economics*, *135*(3), 653–677. https://doi.org/10.1016/j.jfineco.2019.07.002

Ferrary, M., & Granovetter, M. (2009). The role of venture capital firms in Silicon Valley's complex innovation network. *Economy and Society*, *38*(2), 326–359. https://doi.org/10.1080/03085140902786827

Florez-Lopez, R., & Ramon-Jeronimo, J. M. (2015). Enhancing accuracy and interpretability of ensemble strategies in credit risk assessment. A correlated-adjusted decision forest proposal. *Expert Systems with Applications*, *42*(13), 5737–5753. https://doi.org/10.1016/j.eswa.2015.02.042

Franco, S., Cappa, F., & Pinelli, M. (2021). Founder Education and Start-Up Funds Raised. *IEEE Engineering Management Review*, *1*(1), 1–1. https://doi.org/10.1109/emr.2021.3077966

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, *29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Furnari, S., Crilly, D., Misangyi, V. F., Greckhamer, T., Fiss, P. C., & Aguilera, R. (2020). Capturing Causal Complexity: Heuristics for Configurational Theorizing. *Academy of Management Review*, *46*(4). https://doi.org/10.5465/amr.2019.0298

George, G., Osinga, E. C., Lavie, D., & Scott, B. A. (2016). Big Data and Data Science Methods for Management Research. *Academy of Management Journal*, *59*(5), 1493–1507. https://doi.org/10.5465/amj.2016.4005

Gimmon, E., & Levie, J. (2010). Founder's human capital, external investment, and the survival of new high-technology ventures. *Research Policy*, *39*(9), 1214–1226. https://doi.org/10.1016/j.respol.2010.05.017

Gornall, W., & Strebulaev, I. A. (2020). Squaring venture capital valuations with reality. *Journal of Financial Economics*, *135*(1). https://doi.org/10.1016/j.jfineco.2018.04.015

Gottschalk, S., Greene, F. J., Hoewer, D., & Mueller, B. (2014). If You Don't Succeed, Should You Try Again? The Role of Entrepreneurial Experience in Venture Survival. *ZEW - Centre for European Economic Research Discussion Paper No. 14-009*. https://doi.org/10.2139/ssrn.2387508

Griffin, B., Vidaurre, D., Koyluoglu, U., Ternasky, J., Alican , F., & Ihlamur , Y. (2025, May 30). *Random Rule Forest (RRF): Interpretable Ensembles of LLM-Generated Questions for Predicting Startup Success*.

Grimmer, J., Roberts, M. E., & Stewart, B. M. (2021). Machine Learning for Social Science: An Agnostic Approach. *Annual Review of Political Science*, *24*(1). https://doi.org/10.1146/annurev-polisci-053119-015921

Gupta, A., Qian, F., & Sun, Y. (2024). Entrepreneur Experience and Success: Causal Evidence from Immigration Wait Lines. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4813330

Guzman, J. A., & Stern, S. (2015). Nowcasting and Placecasting Entrepreneurial Quality and Performance. *NBER Working Paper Series*, *20954*. https://doi.org/10.3386/w20954

Guzman, J., & Stern, S. (2020). The State of American Entrepreneurship: New Estimates of the Quantity and Quality of Entrepreneurship for 32 US States, 1988–2014. *American Economic Journal: Economic Policy*, *12*(4), 212–243. https://doi.org/10.1257/pol.20170498

Hannigan, T. R., Haans, R. F. J., Vakili, K., Tchalian, H., Glaser, V. L., Wang, M. S., Kaplan, S., & Jennings, P. D. (2019). Topic Modeling in Management Research: Rendering New Theory from Textual Data. *Academy of Management Annals*, *13*(2), 586–632. https://doi.org/10.5465/annals.2017.0099

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning : Data Mining, Inference, and Prediction, Second Edition*. Springer New York.

Hawkins, D. M. (2004). The Problem of Overfitting. *Journal of Chemical Information and Computer Sciences*, *44*(1), 1–12. https://doi.org/10.1021/ci0342472

Hellmann, T., & Puri, M. (2002). Venture Capital and the Professionalization of Start-Up Firms: Empirical Evidence. *The Journal of Finance*, *57*(1), 169–197.

Hernández-Carrión, C., Camarero-Izquierdo, C., & Gutiérrez-Cillán, J. (2016). Entrepreneurs' Social Capital and the Economic Performance of Small Businesses: The Moderating Role of Competitive Intensity and Entrepreneurs' Experience. *Strategic Entrepreneurship Journal*, *11*(1), 61–89. https://doi.org/10.1002/sej.1228

Hofman, J. M., Sharma, A., & Watts, D. J. (2017). Prediction and explanation in social systems. *Science*, *355*(6324), 486–488. https://doi.org/10.1126/science.aal3856

Hofman, J. M., Watts, D. J., Athey, S., Garip, F., Griffiths, T. L., Kleinberg, J., Margetts, H., Mullainathan, S., Salganik, M. J., Vazire, S., Vespignani, A., & Yarkoni, T. (2021). Integrating explanation and prediction in computational social science. *Nature*, *595*(7866), 181–188. https://doi.org/10.1038/s41586-021-03659-0

Howell, S. T., & Nanda, R. (2019, November 1). *Networking Frictions in Venture Capital, and the Gender Gap in Entrepreneurship*. National Bureau of Economic Research. https://doi.org/10.3386/w26449

Hsu, D. H. (2004). What Do Entrepreneurs Pay for Venture Capital Affiliation? *The Journal of Finance*, *59*(4), 1805–1844. https://doi.org/10.1111/j.1540-6261.2004.00680.x

Hsu, D. H. (2007). Experienced entrepreneurial founders, organizational capital, and venture capital funding. *Research Policy*, *36*(5), 722–741. https://doi.org/10.1016/j.respol.2007.02.022

Ioannidis, J. P. A. (2005). Contradicted and Initially Stronger Effects in Highly Cited Clinical Research. *JAMA*, *294*(2), 218. https://doi.org/10.1001/jama.294.2.218

Ji, L., & Li, S. (2025). A dynamic financial risk prediction system for enterprises based on gradient boosting decision tree algorithm. *Systems and Soft Computing*, *7*(3), 200189. https://doi.org/10.1016/j.sasc.2025.200189

Kaiser, U., Kongsted, H. C., Laursen, K., & Ejsing, A.-K. (2018). Experience matters: The role of academic scientist mobility for industrial innovation. *Strategic Management Journal*, *39*(7), 1935–1958. https://doi.org/10.1002/smj.2907

Kaplan, S. N., & Strömberg, P. (2001). Venture Capitalists as Principals: Contracting, Screening, and Monitoring. *The American Economic Review*, *91*(2), 426–430. JSTOR. https://doi.org/10.2307/2677802

Kashyap, Y., & Sinha, A. (2024). *LLM is All You Need: How Do LLMs Perform on Prediction and Classification Using Historical Data*. https://www.ijfmr.com/papers/2024/3/23438.pdf

Keogh, D., & K.N. Johnson, D. (2021). Survival of the funded: Econometric analysis of startup longevity and success. *Journal of Entrepreneurship, Management and Innovation*, *17*(4), 29–49. https://doi.org/10.7341/20211742

Kim, J., Kim, H., & Geum, Y. (2023). How to succeed in the market? Predicting startup success using a machine learning approach. *Technological Forecasting and Social Change*, *193*(10), 122614. https://doi.org/10.1016/j.techfore.2023.122614

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2017). Human Decisions and Machine Predictions. *The Quarterly Journal of Economics*, *133*(1). https://doi.org/10.1093/qje/qjx032

Kleinberg, J., Ludwig, J., Mullainathan, S., & Obermeyer, Z. (2015). Prediction Policy Problems. *American Economic Review*, *105*(5), 491–495. https://doi.org/10.1257/aer.p20151023

Ko, E.-J., & McKelvie, A. (2018). Signaling for more money: The roles of founders' human capital and investor prominence in resource acquisition across different stages of firm development. *Journal of Business Venturing*, *33*(4), 438–454. https://doi.org/10.1016/j.jbusvent.2018.03.001

Koumbarakis, P., & Volery, T. (2022). Predicting New Venture Gestation Outcomes With Machine Learning Methods. *Journal of Small Business Management*, *61*(2), 1–34. https://doi.org/10.1080/00472778.2022.2082453

Kraus, M. (2020). Deep learning in business analytics and operations research: Models, applications and managerial implications. *European Journal of Operational Research*, *281*(3), 628–641. https://doi.org/10.1016/j.ejor.2019.09.018

Leavitt, K., Schabram, K., Hariharan, P., & Barnes, C. M. (2020). Ghost in the machine: On organizational theory in the age of machine learning. *Academy of Management Review*, *46*(4). https://doi.org/10.5465/amr.2019.0247

Li, Z., & Liu, X. (2025). Top Executives with Academic Work Experience, Stakeholder-friendly Engagement, and Firm Value. *Abacus*. https://doi.org/10.1111/abac.12359

Lindner, T., Puck, J., & Verbeke, A. (2022). Beyond Addressing multicollinearity: Robust Quantitative Analysis and Machine Learning in International Business Research. *Journal of International Business Studies*, *53*(7), 1307–1314. https://doi.org/10.1057/s41267-022-00549-z

Liu, J., Li, C., Ouyang, P., Liu, J., & Wu, C. (2023). Interpreting the prediction results of the tree-based gradient boosting models for financial distress prediction with an explainable machine learning approach. *Journal of Forecasting*, *42*(5). https://doi.org/10.1002/for.2931

Liu, Y.-Y., Zheng, Z., Zhang, F., Feng, J.-C., Fu, Y.-Y., Zhai, J.-D., He, B.-S., Zhang, X., & Du, X.-Y. (2026). A comprehensive taxonomy of prompt engineering techniques for large language models. *Frontiers of Computer Science*, *20*(3). https://doi.org/10.1007/s11704-025-50058-z

Lukita, C., Ninda Lutfiani, Saulina, R., None Bhima, Untung Rahardja, & Muhamad Lutfi Huzaifah. (2023). Harnessing The Power of Random Forest in Predicting Startup Partnership Success. *2023 Eighth International Conference on Informatics and Computing (ICIC)*, 1–6. https://doi.org/10.1109/icic60109.2023.10381988

Lundberg, S., Allen, P., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *31st Conference on Neural Information Processing Systems (NIPS 2017)*. https://papers.nips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

M. W. Powers, D. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *International Journal of Machine Learning Technology*, *2*(1). https://doi.org/10.48550/arxiv.2010.16061

Maarouf, A., Feuerriegel, S., & Pröllochs, N. (2024). A fused large language model for predicting startup success. *European Journal of Operational Research*, *322*(1). https://doi.org/10.1016/j.ejor.2024.09.011

Maturo, F., Riccio, D., Mazzitelli, A., Bifulco, G. M., & Paolone, F. (2025). Explainable gradient boosting for corporate crisis forecasting in Italian businesses. *Annals of Operations Research*, *353*(2). https://doi.org/10.1007/s10479-025-06570-3

Menon, R., & James, L. (2022). Understanding Startup Valuation and its Impact on Startup Ecosystem. *Journal of Business Valuation and Economic Loss Analysis*, *0*(0). https://doi.org/10.1515/jbvela-2022-0020

Miric, M., Jia, N., & Huang, K. G. (2022). Using Supervised Machine Learning to Create Categorical Variables for Use in Management Research: The Case for Identifying Artificial Intelligence Patents. *Strategic Management Journal*, *44*(2). https://doi.org/10.1002/smj.3441

Molina, M., & Garip, F. (2019). Machine Learning for Sociology. *Annual Review of Sociology*, *45*(1), 27–45. https://doi.org/10.1146/annurev-soc-073117-041106

Molnar, C. (2025). *Interpretable Machine Learning*. Christoph Molnar.

Mu, X., Ternasky, J., Alican, F., & Ihlamur, Y. (2025). Policy Induction: Predicting Startup Success via Explainable Memory-Augmented In-Context Learning. *IEEE 2025*. https://doi.org/10.48550/arxiv.2505.21427

Mullainathan, S., & Spiess, J. (2017). Machine Learning: an Applied Econometric Approach. *Journal of Economic Perspectives*, *31*(2), 87–106.

Nagy, B. G., Pollack, J. M., Rutherford, M. W., & Lohrke, F. T. (2012). The Influence of Entrepreneurs' Credentials and Impression Management Behaviors on Perceptions of New Venture Legitimacy. *Entrepreneurship Theory and Practice*, *36*(5), 941–965. https://doi.org/10.1111/j.1540-6520.2012.00539.x

Obermann, L., & Waack, S. (2015). Demonstrating non-inferiority of easy interpretable methods for insolvency prediction. *Expert Systems with Applications*, *42*(23), 9117–9128. https://doi.org/10.1016/j.eswa.2015.08.009

Özkurt, C. (2024). Enhancing Financial Decision-Making: Predictive Modeling for Personal Loan Eligibility with Gradient Boosting, XGBoost, and AdaBoost. *Information Technology in Economics and Business*, *1*(1). https://doi.org/10.69882/adba.iteb.2024072

Packalen, K. A. (2007). Complementing Capital: The Role of Status, Demographic Features, and Social Capital in Founding Teams' Abilities to Obtain Resources. *Entrepreneurship Theory and Practice*, *31*(6), 873–891. https://doi.org/10.1111/j.1540-6520.2007.00210.x

Park, U. D., Borah, A., & Kotha, S. (2016). Signaling revisited: The use of signals in the market for IPOs. *Strategic Management Journal*, *37*(11), 2362–2377. https://doi.org/10.1002/smj.2571

Plummer, L. A., Allison, T. H., & Connelly, B. L. (2016). Better Together? Signaling Interactions in New Venture Pursuit of Initial External Capital. *Academy of Management Journal*, *59*(5), 1585–1604. https://doi.org/10.5465/amj.2013.0100

Qiao, S., Ou, Y., Zhang, N., Chen, X., Yao, Y., Deng, S., Tan, C., Huang, F., & Chen, H. (2023, January 1). Reasoning with Language Model Prompting: A Survey. *Proceedings of the 61st Annual Meeting*

of the Association for Computational Linguistics (Volume 1: Long Papers). https://doi.org/10.18653/v1/2023.acl-long.294

Schade, P., & Schuhmacher, M. C. (2023). Predicting entrepreneurial activity using machine learning. *Journal of Business Venturing Insights*, *19*(C), e00357. https://doi.org/10.1016/j.jbvi.2022.e00357

Shane, S., & Cable, D. (2002). Network Ties, Reputation, and the Financing of New Ventures. *Management Science*, *48*(3), 364–381. https://www.jstor.org/stable/822571

Shane, S., & Stuart, T. (2002). Organizational Endowments and the Performance of University Start-ups. *Management Science*, *48*(1), 154–170. https://doi.org/10.1287/mnsc.48.1.154.14280

Shao, Y., & Sun, L. (2021). Entrepreneurs' social capital and venture capital financing. *Journal of Business Research*, *136*(C), 499–512. https://doi.org/10.1016/j.jbusres.2021.08.005

Shi, Y., Ekaterina Eremina, & Long, W. (2023). Machine learning models for early-stage investment decision making in startups. *MDE. Managerial and Decision Economics/Managerial and Decision Economics*, *45*(3). https://doi.org/10.1002/mde.4072

Shmueli, G. (2010). To Explain or to Predict? *Statistical Science*, *25*(3), 289–310. https://doi.org/10.1214/10-sts330

Shmueli, G. (2016). Analyzing Behavioral Big Data: Methodological, Practical, Ethical, and Moral Issues. *Quality Engineering*, *29*(1). https://doi.org/10.1080/08982112.2016.1210979

Shrestha, Y. R., He, V. F., Puranam, P., & von Krogh, G. (2021). Algorithm Supported Induction for Building Theory: How Can We Use Prediction Models to Theorize? *Organization Science*, *32*(3). https://doi.org/10.1287/orsc.2020.1382

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374. https://doi.org/10.2307/1882010

Steigenberger, N., & Wilhelm, H. (2018). Extending Signaling Theory to Rhetorical Signals: Evidence from Crowdfunding. *Organization Science*, *29*(3), 529–546. https://doi.org/10.1287/orsc.2017.1195

Steyerberg, E. W., Harrell, F. E., Borsboom, G. J. J. M., Eijkemans, M. J. C., Vergouwe, Y., & Habbema, J. Dik. F. (2001). Internal validation of predictive models. *Journal of Clinical Epidemiology*, *54*(8), 774–781. https://doi.org/10.1016/s0895-4356(01)00341-9

Stuart, T. E., Hoang, H., & Hybels, R. C. (1999). Interorganizational Endorsements and the Performance of Entrepreneurial Ventures. *Administrative Science Quarterly*, *44*(2), 315. https://doi.org/10.2307/2666998

Svetek, M. (2022). Signaling in the context of early-stage equity financing: review and directions. *Venture Capital*, *24*(1), 1–34. https://doi.org/10.1080/13691066.2022.2063092

Swift, T. (2018). PhD scientists in the boardroom: the innovation impact. *Journal of Strategy and Management*, *11*(2), 184–202. https://doi.org/10.1108/jsma-06-2017-0040

Taj, S. A. (2016). Application of signaling theory in management research: Addressing major gaps in theory. *European Management Journal*, *34*(4), 338–348. https://doi.org/10.1016/j.emj.2016.02.001

Tidhar, R., & Eisenhardt, K. M. (2020). Get rich or die trying… finding revenue model fit using machine learning and multiple cases. *Strategic Management Journal*, *41*(7), 1245–1273. https://doi.org/10.1002/smj.3142

Toole, A. A., & Czarnitzki, D. (2008). Exploring the Relationship Between Scientist Human Capital and Firm Performance: The Case of Biomedical Academic Entrepreneurs in the SBIR Program. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.966125

Ucbasaran, D., Westhead, P., & Wright, M. (2009). The extent and nature of opportunity identification by experienced entrepreneurs. *Journal of Business Venturing*, *24*(2), 99–115. https://doi.org/10.1016/j.jbusvent.2008.01.008

Uddin, M. S., Chi, G., Al Janabi, M. A. M., & Habib, T. (2020). Leveraging random forest in micro-enterprises credit risk modelling for accuracy and interpretability. *International Journal of Finance & Economics*, *27*(3). https://doi.org/10.1002/ijfe.2346

Unger, J. M., Rauch, A., Frese, M., & Rosenbusch, N. (2011). Human capital and entrepreneurial success: A meta-analytical review. *Journal of Business Venturing*, *26*(3), 341–358.

Varian, H. R. (2014). Big Data: New Tricks for Econometrics. *Journal of Economic Perspectives*, *28*(2), 3–28. https://doi.org/10.1257/jep.28.2.3

Whetten, D. A. (1989). What Constitutes a Theoretical Contribution? *The Academy of Management Review*, *14*(4), 490–495. https://doi.org/10.2307/258554

Yarkoni, T., & Westfall, J. (2017). Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. *Perspectives on Psychological Science*, *12*(6), 1100–1122. https://doi.org/10.1177/1745691617693393

Zhang, J. (2009). The advantage of experienced start-up founders in venture capital acquisition: evidence from serial entrepreneurs. *Small Business Economics*, *36*(2), 187–208. https://doi.org/10.1007/s11187-009-9216-4

Zhang, R., Tian, Z., McCarthy, K. J., Wang, X., & Zhang, K. (2022). Application of machine learning techniques to predict entrepreneurial firm valuation. *Journal of Forecasting*, *42*(2), 402–417. https://doi.org/10.1002/for.2912

Zhang, Y., Wang, D., Yang, M. M., Li, J., & Liu, S. (2023). A bibliometric review of early-stage entrepreneurial equity financing: an overview and future research agenda. *Venture Capital*, *27*(3), 1–46. https://doi.org/10.1080/13691066.2023.2295252

Zimmerman, M. A., & Zeitz, G. J. (2002). Beyond Survival: Achieving New Venture Growth by Building Legitimacy. *The Academy of Management Review*, *27*(3), 414. https://doi.org/10.2307/4134387

Zucker, L. G., Darby, M. R., & Brewer, M. B. (1998). Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises. *The American Economic Review*, *88*(1), 290–306. https://www.jstor.org/stable/116831

Appendix

Appendix A: Cohort Prediction Datasets – Descriptive Statistics

Year	2018	2019	2020	2021	2022
Total Companies	279,994	269,491	304,331	244,219	205,847
Total Founder-Venture Pairs	5,021	5,328	6,123	5,824	6,055
Total Founder Profiles	9,611	10,256	12,057	11,794	12,852
Profiles Europe	4,458	4,850	5,803	5,577	5,355
Profiles North America	5,153	5,406	6,254	6,217	7,497
top 3% (true)	474	542	567	538	513
bottom 97% (false)	9,137	9,714	11,490	11,256	12,339
Funding Treshold 3% (Mio USD)	73	70	55.83	49.7	59

Appendix B: Training Datasets – Descriptive Statistics

2018 Cohort Prediction: Training Dataset

Year	top 3% (true)	bottom 10% (false)	total
2012	142	210	352
2013	148	233	381
2014	156	261	417
2015	129	253	382
2016	87	199	286
Total	662	1156	1818

2019 Cohort Prediction: Training Dataset

Year	top 3% (true)	bottom 10% (false)	Total
2013	85	139	224
2014	93	157	250
2015	90	166	256
2016	68	133	201
2017	56	112	168
Total	392	707	1099

2020 Cohort Prediction: Training Dataset

Year	top 3% (true)	bottom 10% (false)	total
2014	163	261	424
2015	151	285	436
2016	126	252	378
2017	137	246	383
2018	92	223	315
Total	669	1267	1936

2021 Cohort Prediction: Training Dataset

Year	top 3% (true)	bottom 10% (false)	total
2015	164	310	474
2016	153	287	440
2017	156	279	435
2018	122	283	405
2019	100	254	354
Total	695	1413	2108

2022 Cohort Prediction: Training Dataset

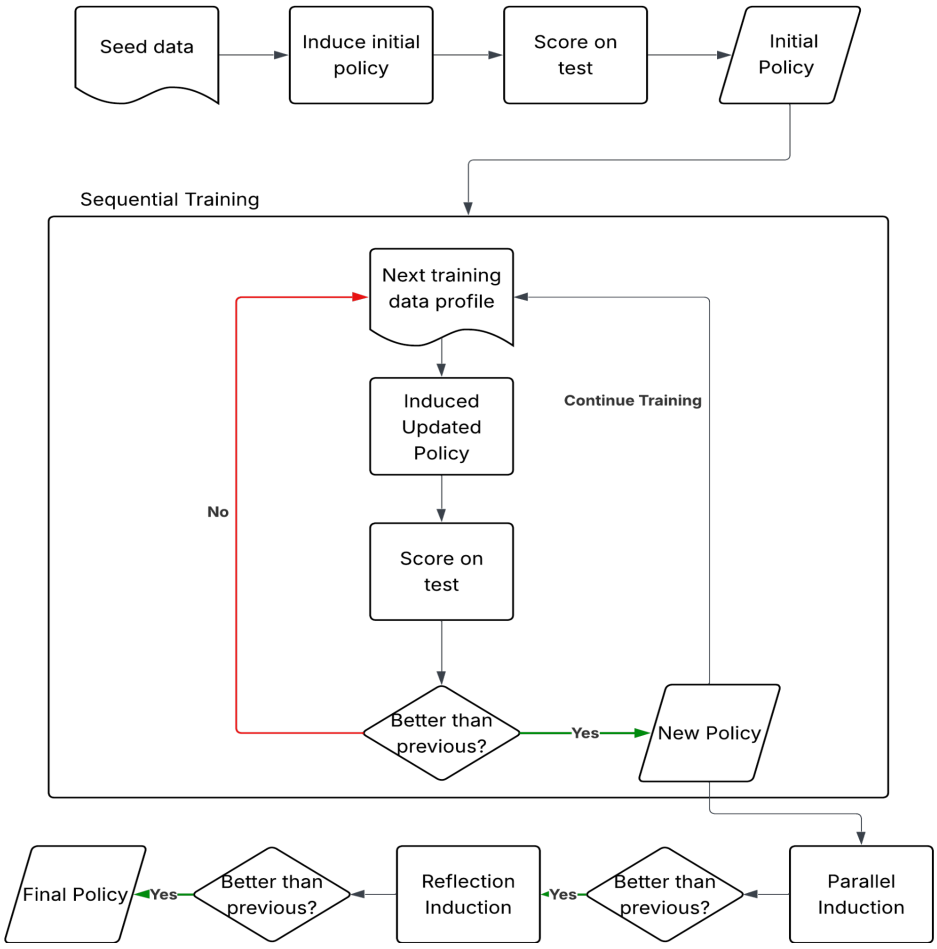
Year	top 3% (true)	bottom 10% (false)	total
2017	68	143	211
2018	79	140	219
2019	73	168	241
2020	75	122	197
2021	69	135	204
Total	364	708	1072

Appendix C: Feature Extraction Lists

Nasdaq Companies	
https://www.cnbc.com/nasdaq-100/	
Ticker	**Name**
ADBE	Adobe Inc
AMD	Advanced Micro Devices
ABNB	Airbnb Inc
GOOGL	Alphabet Inc (Class A)
GOOG	Alphabet Inc (Class C)
AMZN	Amazon.com Inc
AEP	American Electric Power Co Inc
AMGN	Amgen Inc
ADI	Analog Devices Inc
AAPL	Apple Inc
AMAT	Applied Materials Inc
APP	AppLovin Corporation
ARM	Arm Holdings plc
ASML	ASML Holding N.V.
AVGO	Broadcom Inc
AXON	Axon Enterprise Inc
AZN	AstraZeneca plc
BKNG	Booking Holdings Inc
CCEP	Coca-Cola Europacific Partners
CSCO	Cisco Systems Inc
CTSH	Cognizant Technology Solutions Corp
CMCSA	Comcast Corporation
etc.	

Banks	**Consultancies**	**Big Tech Companies**
https://markets.ft.com/data/league-tables/tables-and-trends	https://www.forbes.com/sites/haniya-rae/2025/08/27/meet-the-worlds-best-management-consulting-firms-2025/	https://companies-marketcap.com/tech/largest-tech-companies-by-market-cap/
Name	**Name**	**Name**
Goldman Sachs	Roland Berger	NVIDIA
JPMorgan Chase	Boston Consulting Group	Microsoft
Morgan Stanley	McKinsey & Company	Apple
Bank of America	Bain & Company	Alphabet (Google)
Centerview Partners	Kearney	Amazon
Evercore	Oliver Wyman	Meta Platforms (Facebook)
Lazard	Booz Allen Hamilton	Broadcom
PJT Partners	Accenture	Tesla
Moelis & Company	L.E.K. Consulting	TSMC
Qatalyst Partners	IBM Global Business Services	Oracle
Parella Weinberg	Alvarez & Marsal	Tencent
Guggenheim Securities	Gartner	Netflix
Jane Streed	AlixPartners LLP	Palantir
Citadel	Analysis Group	Alibaba
Citadel Securities	Cornerstone Research	Samsung
Blackstone	Mercer LLC	ASML
etc.	etc.	etc.

Appendix D: Policy Induction Framework



Appendix E: Policy Induction Model Decision Policies

```
# Policy Rules: 2018 Cohort

## Overview

This document outlines the policy rules used for evaluating founder potential in the
2018 cohort. These rules are based on career trajectory, experience patterns, and
leadership indicators.

---

## Positive Indicators

1. **Early evidence of founding or co-founding roles**, even if brief or concurrent
with other roles, is a strong positive signal.

2. **Rapid progression to leadership or founder titles** (within 1-3 years of starting
career) increases likelihood of success.

3. **Experience spanning multiple industries or functions**, especially with at least
one technical or product-focused role, is a positive indicator.

4. **Short stints (≤1 year) in high-learning roles** (internships, consulting, venture
development) followed by founder/leadership roles is a strong positive pattern.

5. **Technical or product-building experience** (engineering, development, pro-
gram/product management) prior to founding is highly predictive.

6. **Multiple instances of concurrent or overlapping leadership/founder roles** indi-
cate high initiative and are positive.

7. **Gaps or lack of formal education** do not negatively impact prediction if compen-
sated by early founder/leadership or technical roles.

8. **Advanced degrees** (masters, PhD, MBA) are only positive if paired with
founder/leadership or technical/product roles.

9. **International or cross-regional work or education experience** before or at
founding is a positive signal.

10. **Early cross-functional experience** (e.g., technical + business, or product +
sales/marketing) is a strong positive signal.

11. **Current role as founder/co-founder/owner**, even with sparse prior experience,
is a positive indicator.
```

```
12. **Multiple, sequential founder/owner roles** (serial entrepreneurship) are strong positive signals.

13. **Founder/leadership roles in technical, high-growth, or innovation-driven industries** are more predictive than in traditional or service industries.

14. **Board/advisory roles** are only positive if accompanied by founder/leadership or technical/product experience.

15. **Lack of any work or degree information** is not a negative if a founder/owner title is present.

---

## Negative Indicators

16. **Long tenure (>5 years) in non-leadership, single-function roles** without evidence of progression or initiative is a negative indicator.

17. **Multiple short, unrelated roles** without progression to leadership or founder titles is a negative signal.

18. **Repeated roles in the same function/industry** without upward mobility or initiative are negative indicators.

19. **Absence of any founder/co-founder/owner/entrepreneur title** by early/mid-career is a negative indicator.

20. **Multiple degrees without corresponding leadership or founder roles** do not increase likelihood of success.

---

*Policy rules for 2018 cohort evaluation*


# Policy Rules: 2019 Cohort

## Overview

This document outlines the policy rules used for evaluating founder potential in the 2019 cohort. These rules emphasize rapid advancement, cross-functional experience, and leadership progression.

---

## Positive Indicators
```

```
1. **Early founder, co-founder, or C-level titles** (especially before age 30 or within 7 years of career start) are strong positive signals.

2. **Multiple founder/co-founder or C-level roles**, even across unrelated industries, indicate high founder potential.

3. **Rapid advancement to leadership/executive roles** (director, VP, CTO, CEO) within 5-7 years of career start is a strong positive.

4. **Experience spanning both technical and non-technical roles** (e.g., engineering + business, research + operations) is a strong indicator.

5. **Demonstrated progression from individual contributor → manager → executive** is a positive signal, especially if achieved quickly.

6. **Breadth of experience across multiple industries or geographies**, especially international work or study, is a positive.

7. **Advanced degrees** (master's, doctorate, or professional) are only positive when paired with significant non-academic or operational work.

8. **Multiple degrees in different disciplines or regions** (e.g., science + business, engineering + law) increase likelihood of success.

9. **Board membership, advisory, or president roles** concurrent with operational roles are strong positive signals.

10. **Repeated leadership in project-based or cross-functional environments** (e.g., product manager, principal engineer) is positive.

11. **Extended tenure (2+ years) in at least one role**, combined with shorter stints elsewhere, signals both depth and breadth.

12. **Early entrepreneurial activity** (founder, owner, president) before significant corporate experience is a strong positive.

13. **International mobility** (work or education in 2+ continents) is a positive signal.

14. **Nonlinear or unconventional career paths** (e.g., field switching, diverse roles, mid-career education) are a positive signal, but only when paired with leadership or founder experience.

---

## Negative Indicators
```

```
15. **Absence of upward trajectory**, or exclusively academic/research roles without operational or leadership experience, is negative.

16. **Long tenure (5+ years) in a single non-leadership role** without progression is a negative signal.

17. **Profiles with only consulting, analyst, support, or creative/teaching/research roles** and no leadership or founder titles are negative.

18. **Multiple incomplete or ongoing degrees** without substantial work experience is a negative indicator.

19. **Lack of any founder, co-founder, or C-level title** at or before company inception is a strong negative signal.

20. **Profiles with only sales or account management roles** and no leadership or founder experience are less likely to succeed.

---

*Policy rules for 2019 cohort evaluation*

# Policy Rules: 2020 Cohort

## Overview

This document outlines the policy rules used for evaluating founder potential in the 2020 cohort. These rules emphasize technical depth, operational experience, and clear progression patterns.

---

## Positive Indicators

1. **Multiple founder, co-founder, or C-level roles** across distinct organizations, industries, or regions are strong positive signals, especially with evidence of building and scaling new initiatives.

2. **Clear, rapid progression from technical or specialist roles to executive leadership** in technology or product-driven fields is a strong positive indicator.

3. **Sustained tenure (4+ years) in technical, engineering, or product roles** at high-impact or growth-oriented companies, followed by transition to founder/executive roles, is a strong positive signal.
```

```
4. **Experience spanning both technical and business/operational roles** (e.g., engi-
neering lead and product manager, or CTO and business development) is a strong posi-
tive indicator, especially when paired with leadership.

5. **Prior investor, advisor, or entrepreneur-in-residence experience** is a strong
positive signal only when combined with technical or operational leadership roles.

6. **International work or education experience** across multiple continents/regions
is a positive indicator if paired with founder or executive roles.

7. **Evidence of rapid advancement** (multiple promotions or increased responsibility
in short timeframes) is a strong positive signal, especially when paired with tech-
nical or operational depth.

8. **Non-linear or unconventional career paths** (e.g., military, academia, creative
industries) are positive only if followed by operational or technical leadership roles
with clear execution responsibility.

9. **Multiple degrees or ongoing education in diverse fields** are positive only if
followed by founder or executive roles with demonstrated operational impact.

10. **Multiple concurrent roles** (advisor, investor, executive) are positive only if
at least one is a founder/executive position with operational responsibility.

11. **Repeated transitions between industries** are positive only if accompanied by
upward progression into leadership roles in high-growth or technology sectors.

12. **Evidence of both technical and customer-facing experience** (sales, marketing,
business development) is a strong positive signal only when combined with leadership
roles.

---

## Negative Indicators

13. **Absence of any founder, executive, or leadership roles**, or exclusively aca-
demic/consulting/non-profit experience without operational or technical roles, is a
strong negative indicator.

14. **Multiple short tenures (<1 year) in roles** without clear progression, leader-
ship, or deepening expertise are strong negative signals.

15. **Profiles limited to student status, internships, or entry-level roles**, without
evidence of leadership or entrepreneurial activity, are negative indicators.

16. **Profiles with only academic or research roles**, even at senior levels, without
operational or leadership experience, are strong negative indicators.
```

```
17. **Multiple degrees or long academic tenure** without transition to leadership or operational roles is a negative indicator.

18. **Profiles with only non-profit, government, or civic roles**, without technical or operational leadership, are negative indicators.

19. **Incomplete education (no degree)** is a negative indicator unless paired with strong founder or executive experience and clear operational progression.

20. **Mid-level management or technical roles** (e.g., manager, team lead) without founder, executive, or clear upward progression are negative signals unless followed by founder/executive roles with operational impact.

---

*Policy rules for 2020 cohort evaluation*

# Policy Rules: 2021 Cohort

## Overview

This document outlines the policy rules used for evaluating founder potential in the 2021 cohort. These rules emphasize cross-functional versatility, entrepreneurial activity, and diverse experience patterns.

---

## Positive Indicators

1. **Multiple distinct leadership, founder, or principal roles** (co-founder, chief, partner, director, advisor, board member) across different organizations, sectors, or regions before or at company inception are strong positive signals.

2. **Rapid progression from technical, product, or analytical roles to leadership or founder positions**, especially with cross-functional and cross-sector exposure, is highly predictive of founder success.

3. **Demonstrated cross-functional experience** (technical, business, marketing, finance, operations) and ability to operate in multiple domains is a strong positive indicator.

4. **Short-to-medium tenures (1–5 years) in several roles**, showing mobility, breadth, and upward trajectory, are more predictive than long, static tenures or repeated lateral moves.

5. **Experience spanning multiple industries, sectors, or geographic regions** is a positive signal, especially when paired with leadership or founder roles.
```

```
6. **Evidence of entrepreneurial activity prior to company inception** (side projects, self-employment, advisory, board, or investor roles) is a strong positive, even if ventures were small or unconventional.

7. **Advanced degrees are only positive** when paired with significant, diverse, non-academic work experience and clear transitions into leadership or entrepreneurial roles; academic-only or primarily academic profiles are negative.

8. **Non-traditional education or lack of degree** can be positive if accompanied by significant, varied work or leadership experience, especially across sectors or regions.

9. **Multiple concurrent roles** (advisor, board, partner, investor, or operator) are strong positive signals, especially when combined with founder or leadership titles.

10. **Demonstrated ability to transition between industries or functions** (e.g., engineering to product to business, or sector switches) is a strong positive indicator.

11. **Early involvement in high-impact, strategic, or global projects** (product launches, acquisitions, international expansion) or responsibility for P&L is a positive signal.

12. **Multiple roles with "founder," "co-founder," or "chief" in the title**, even in small, early, or non-traditional ventures, are strong positive indicators.

13. **Experience as an advisor, mentor, or investor** in addition to operating roles is a positive signal, especially if concurrent with other leadership roles.

14. **Profiles showing both technical and customer-facing experience** (sales, marketing, business development, operations) are more likely to succeed.

15. **Repeated, clear evidence of initiative** (starting new projects, switching industries, building new teams, launching ventures, or taking on new responsibilities) is a strong positive indicator.

---

## Negative Indicators

16. **Repeated internships, entry-level, or research roles** without rapid progression to higher responsibility, cross-functional exposure, or entrepreneurial activity is a negative signal.

17. **Profiles limited to a single sector, function, or primarily academic/research/creative/freelance work** without leadership or entrepreneurial initiative are less likely to become successful founders.
```

```
18. **Long tenure in junior roles, or lack of advancement**, is a negative signal unless offset by clear entrepreneurial or leadership activity.

19. **Absence of any work experience, or only current academic enrollment**, is a negative signal unless offset by clear entrepreneurial activity or leadership outside academia.

20. **The most predictive profiles combine leadership, cross-functional, and entrepreneurial roles** across regions, sectors, and functions, with clear upward trajectory, repeated initiative, and non-academic work experience; static or single-sector profiles, even with "CEO" or "founder" titles, are less predictive—especially when lacking cross-sector or cross-functional exposure.

---

*Policy rules for 2021 cohort evaluation*


# Policy Rules: 2022 Cohort

## Overview

This document outlines the policy rules used for evaluating founder potential in the 2022 cohort. These rules emphasize repeated initiative, measurable impact, and sustained leadership responsibility.

---

## Positive Indicators

1. **Repeated founding or co-founding roles**—especially across multiple domains, geographies, or over time—are highly predictive, particularly when paired with measurable leadership outcomes.

2. **Sustained, increasing leadership responsibility** (e.g., CEO, founder, P&L owner, team lead) across organizations or industries is strongly positive, especially if roles show growing scope or complexity.

3. **Rapid advancement or significant increases in responsibility**—especially spanning diverse roles, organizations, or sectors—signal strong founder potential; slow, lateral, or stagnant moves are negative.

4. **Demonstrated ability to bridge technical and business/strategic functions**, or clear cross-functional versatility, is a strong positive, especially when paired with ownership or initiative.
```

```
5. **Repeated creation or launch of new initiatives, products, or teams** (not just maintenance), in multiple contexts, is highly predictive; single or student-only projects are insufficient.

6. **Non-linear or unconventional career paths** (e.g., industry switches, overlapping roles, entrepreneurship during education), when paired with upward mobility or impact, are strong signals.

7. **International or cross-cultural experience** is positive only when paired with clear initiative, leadership, or impact.

8. **Sustained balancing of multiple demanding commitments** (work, study, entrepreneurship, volunteer leadership) over time, with measurable outcomes, is a strong positive.

9. **Consistent engagement in entrepreneurial, self-initiated, or high-impact projects** outside traditional employment, with clear initiative and follow-through, is predictive.

10. **Early exposure to high-responsibility, high-velocity, or high-stakes environments** (startups, military, rapid promotions), with visible impact, is a strong positive.

11. **Demonstrated ability to lead and inspire teams**, shown by repeated formal/informal leadership and team outcomes, is a strong indicator.

12. **Multi-year, multi-role engagement in a single organization** is positive only if paired with clear evidence of initiative, new product/team creation, or rapid advancement.

13. **Volunteer or extracurricular leadership** is positive when sustained and accompanied by measurable outcomes or organizational impact, but insufficient alone.

14. **Serial entrepreneurship**, especially with demonstrated impact, scaling, or exits across multiple organizations or markets, is among the strongest positive indicators.

---

## Negative Indicators

15. **Purely linear, corporate, or specialist trajectories** without entrepreneurial signals, ownership, or initiative—even at senior levels—are negative indicators.

16. **Long tenures in technical, creative, or operational roles** without evidence of initiative, leadership, or cross-functional activity are negative signals.
```

```
17. **Founding experience limited to a single, recent, or student project**—without
other signals of ownership, initiative, or leadership—is insufficient on its own.

18. **Advisory, investor, or board roles** are neutral unless paired with prior found-
ing/leadership signals or direct operational experience.

19. **Absence of any signals of initiative, ownership, or leadership**—regardless of
education or employer prestige—is a strong negative indicator.

20. **Academic achievement or elite education** is only weakly positive unless paired
with entrepreneurial or leadership signals; advanced degrees alone are not predictive.

---

*Policy rules for 2022 cohort evaluation*
```

Appendix F: Policy Ablation Analysis

Note: The following tables correspond to the policy ablation results for each prediction year. The line number in each table corresponds to the line number of the policy in each prediction year as outlined in Appendix 4. So, for instance, in 2018 line_number 1 corresponds to "*Early evidence of founding or co-founding roles, even if brief or concurrent with other roles, is a strong positive signal."*.

2018 Prediction Cohort - Policy Ablation Analysis

line number	im-portance score	baseline f05	ablated f05	ablated precision	ablated recall	ablated accuracy	delta pre-cision	delta recall
1	0.0168	0.1773	0.2736	0.2127	0.3333	0.9424	-0.0190	0.0000
2	0.0253	0.1887	0.2566	0.2320	0.2222	0.8656	0.0299	0.0000
3	0.0213	0.1887	0.2683	0.1457	0.1111	0.8453	0.0198	0.1111
4	0.0063	0.1773	0.1600	0.1812	0.4444	0.9352	-0.0070	0.0000
5	0.0253	0.2898	0.2075	0.1697	0.4444	0.8500	0.0151	0.0000
6	0.0239	0.2898	0.1521	0.1217	0.4444	0.8654	0.0086	0.0000
7	0.0160	0.1887	0.2013	0.1628	0.2222	0.8864	0.0165	0.0000
8	0.0055	0.2898	0.1849	0.2478	0.2222	0.9029	-0.0200	0.1111
9	0.0066	0.1773	0.2655	0.2401	0.4444	0.8670	0.0141	0.0000
10	0.0194	0.1887	0.1982	0.2182	0.4444	0.9115	0.0198	0.0000
11	0.0168	0.2898	0.2184	0.1270	0.2222	0.9341	-0.0222	0.1111
12	0.0208	0.2898	0.2495	0.1634	0.4444	0.8759	0.0068	0.0000
13	0.0213	0.2898	0.1903	0.1796	0.4444	0.9232	-0.0093	0.0000
14	0.0159	0.2898	0.2245	0.1934	0.3333	0.8450	0.0041	0.1111
15	0.0075	0.1773	0.1619	0.1804	0.4444	0.9306	0.0052	0.1111
16	0.0063	0.1887	0.2310	0.1711	0.1111	0.8780	0.0173	0.0000
17	0.0154	0.1773	0.1972	0.1270	0.1111	0.9182	-0.0193	0.0000
18	0.0176	0.2898	0.2116	0.1947	0.4444	0.8454	-0.0121	0.0000
19	0.0079	0.1887	0.2405	0.1698	0.4444	0.8372	-0.0025	0.1111
20	0.0201	0.2898	0.1703	0.2287	0.2222	0.8538	0.0106	0.1111

2019 Prediction Cohort - Policy Ablation Analysis

line number	im-portance score	baseline f05	ablated f05	ablated preci-sion	ablated recall	ablated accuracy	delta pre-cision	delta recall
10	0.0269	0.1773	0.1504	0.1290	0.4444	0.8392	0.0225	0.1111
11	0.0139	0.1773	0.1634	0.1389	0.5556	0.8241	0.0126	0.0000
1	0.0095	0.1773	0.1678	0.1429	0.5556	0.8291	0.0087	0.0000
4	0.0000	0.1773	0.1773	0.1515	0.5556	0.8392	0.0000	0.0000
2	0.0000	0.1773	0.1773	0.1515	0.5556	0.8392	0.0000	0.0000
16	0.0000	0.1773	0.1773	0.1515	0.5556	0.8392	0.0000	0.0000
13	0.0000	0.1773	0.1773	0.1515	0.5556	0.8392	0.0000	0.0000
5	0.0000	0.1773	0.1773	0.1515	0.5556	0.8392	0.0000	0.0000
20	0.0000	0.1773	0.1773	0.1515	0.5556	0.8392	0.0000	0.0000
9	0.0000	0.1773	0.1773	0.1515	0.5556	0.8392	0.0000	0.0000
8	0.0000	0.1773	0.1773	0.1515	0.5556	0.8392	0.0000	0.0000
3	-0.0052	0.1773	0.1825	0.1563	0.5556	0.8442	-0.0047	0.0000
7	-0.0052	0.1773	0.1825	0.1563	0.5556	0.8442	-0.0047	0.0000
15	-0.0052	0.1773	0.1825	0.1563	0.5556	0.8442	-0.0047	0.0000
19	-0.0052	0.1773	0.1825	0.1563	0.5556	0.8442	-0.0047	0.0000
6	-0.0107	0.1773	0.1880	0.1613	0.5556	0.8492	-0.0098	0.0000
17	-0.0107	0.1773	0.1880	0.1613	0.5556	0.8492	-0.0098	0.0000
18	-0.0107	0.1773	0.1880	0.1613	0.5556	0.8492	-0.0098	0.0000
12	-0.0107	0.1773	0.1880	0.1613	0.5556	0.8492	-0.0098	0.0000
14	-0.0165	0.1773	0.1938	0.1667	0.5556	0.8543	-0.0152	0.0000

2020 Prediction Cohort - Policy Ablation Analysis

line number	importance score	baseline f05	ablated f05	ablated precision	ablated recall	ablated accuracy	delta precision	delta recall
6	0.0943	0.1887	0.0943	0.0909	0.1111	0.9095	0.0909	0.1111
9	0.0943	0.1887	0.0943	0.0909	0.1111	0.9095	0.0909	0.1111
5	0.0866	0.1887	0.1020	0.1000	0.1111	0.9146	0.0818	0.1111
18	0.0776	0.1887	0.1111	0.1111	0.1111	0.9196	0.0707	0.1111
14	0.0247	0.1887	0.1639	0.1538	0.2222	0.9095	0.0280	0.0000
20	0.0132	0.1887	0.1754	0.1667	0.2222	0.9146	0.0152	0.0000
16	0.0132	0.1887	0.1754	0.1667	0.2222	0.9146	0.0152	0.0000
10	0.0132	0.1887	0.1754	0.1667	0.2222	0.9146	0.0152	0.0000
2	0.0132	0.1887	0.1754	0.1667	0.2222	0.9146	0.0152	0.0000
19	0.0000	0.1887	0.1887	0.1818	0.2222	0.9196	0.0000	0.0000
15	0.0000	0.1887	0.1887	0.1818	0.2222	0.9196	0.0000	0.0000
11	0.0000	0.1887	0.1887	0.1818	0.2222	0.9196	0.0000	0.0000
12	0.0000	0.1887	0.1887	0.1818	0.2222	0.9196	0.0000	0.0000
8	0.0000	0.1887	0.1887	0.1818	0.2222	0.9196	0.0000	0.0000
4	0.0000	0.1887	0.1887	0.1818	0.2222	0.9196	0.0000	0.0000
13	-0.0154	0.1887	0.2041	0.2000	0.2222	0.9246	-0.0182	0.0000
7	-0.0154	0.1887	0.2041	0.2000	0.2222	0.9246	-0.0182	0.0000
17	-0.0154	0.1887	0.2041	0.2000	0.2222	0.9246	-0.0182	0.0000
3	-0.0335	0.1887	0.2222	0.2222	0.2222	0.9296	-0.0404	0.0000
1	-0.0572	0.1887	0.2459	0.2308	0.3333	0.9196	-0.0490	-0.1111

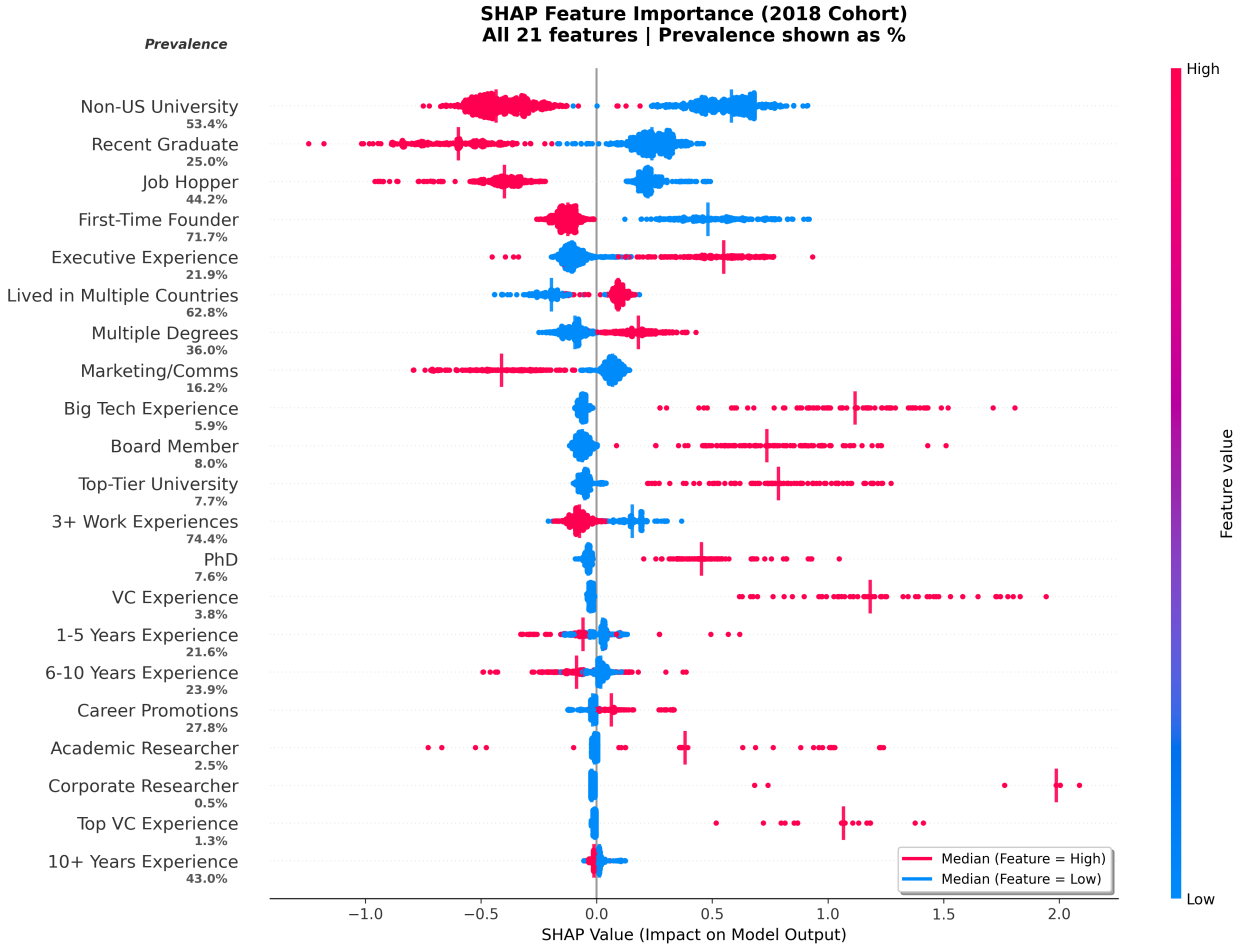
2021 Prediction Cohort - Policy Ablation Analysis

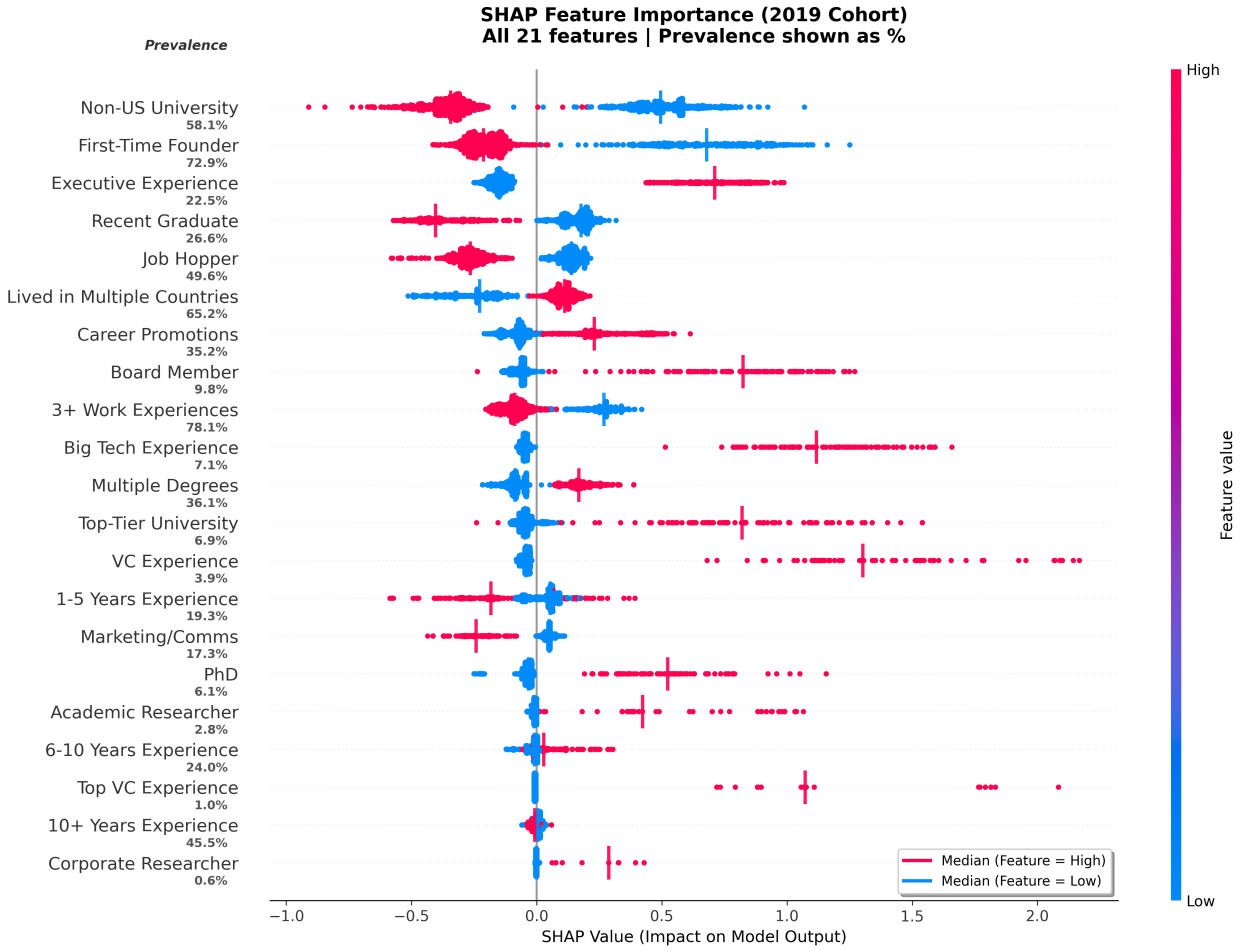
line number	importance score	baseline f05	ablated f05	ablated precision	ablated recall	ablated accuracy	delta precision	delta recall
14	0.0301	0.2899	0.2597	0.2353	0.4444	0.9095	0.0314	0.0000
1	0.0159	0.2899	0.2740	0.2500	0.4444	0.9146	0.0167	0.0000
8	0.0159	0.2899	0.2740	0.2500	0.4444	0.9146	0.0167	0.0000
19	0.0159	0.2899	0.2740	0.2500	0.4444	0.9146	0.0167	0.0000
16	0.0159	0.2899	0.2740	0.2500	0.4444	0.9146	0.0167	0.0000
12	0.0159	0.2899	0.2740	0.2500	0.4444	0.9146	0.0167	0.0000
10	0.0159	0.2899	0.2740	0.2500	0.4444	0.9146	0.0167	0.0000
11	0.0159	0.2899	0.2740	0.2500	0.4444	0.9146	0.0167	0.0000
7	0.0159	0.2899	0.2740	0.2500	0.4444	0.9146	0.0167	0.0000
4	0.0159	0.2899	0.2740	0.2500	0.4444	0.9146	0.0167	0.0000
3	0.0159	0.2899	0.2740	0.2500	0.4444	0.9146	0.0167	0.0000
9	0.0000	0.2899	0.2899	0.2667	0.4444	0.9196	0.0000	0.0000
2	0.0000	0.2899	0.2899	0.2667	0.4444	0.9196	0.0000	0.0000
6	0.0000	0.2899	0.2899	0.2667	0.4444	0.9196	0.0000	0.0000
13	0.0000	0.2899	0.2899	0.2667	0.4444	0.9196	0.0000	0.0000
5	0.0000	0.2899	0.2899	0.2667	0.4444	0.9196	0.0000	0.0000
15	0.0000	0.2899	0.2899	0.2667	0.4444	0.9196	0.0000	0.0000
17	0.0000	0.2899	0.2899	0.2667	0.4444	0.9196	0.0000	0.0000
18	0.0000	0.2899	0.2899	0.2667	0.4444	0.9196	0.0000	0.0000
20	-0.0178	0.2899	0.3077	0.2857	0.4444	0.9246	-0.0190	0.0000

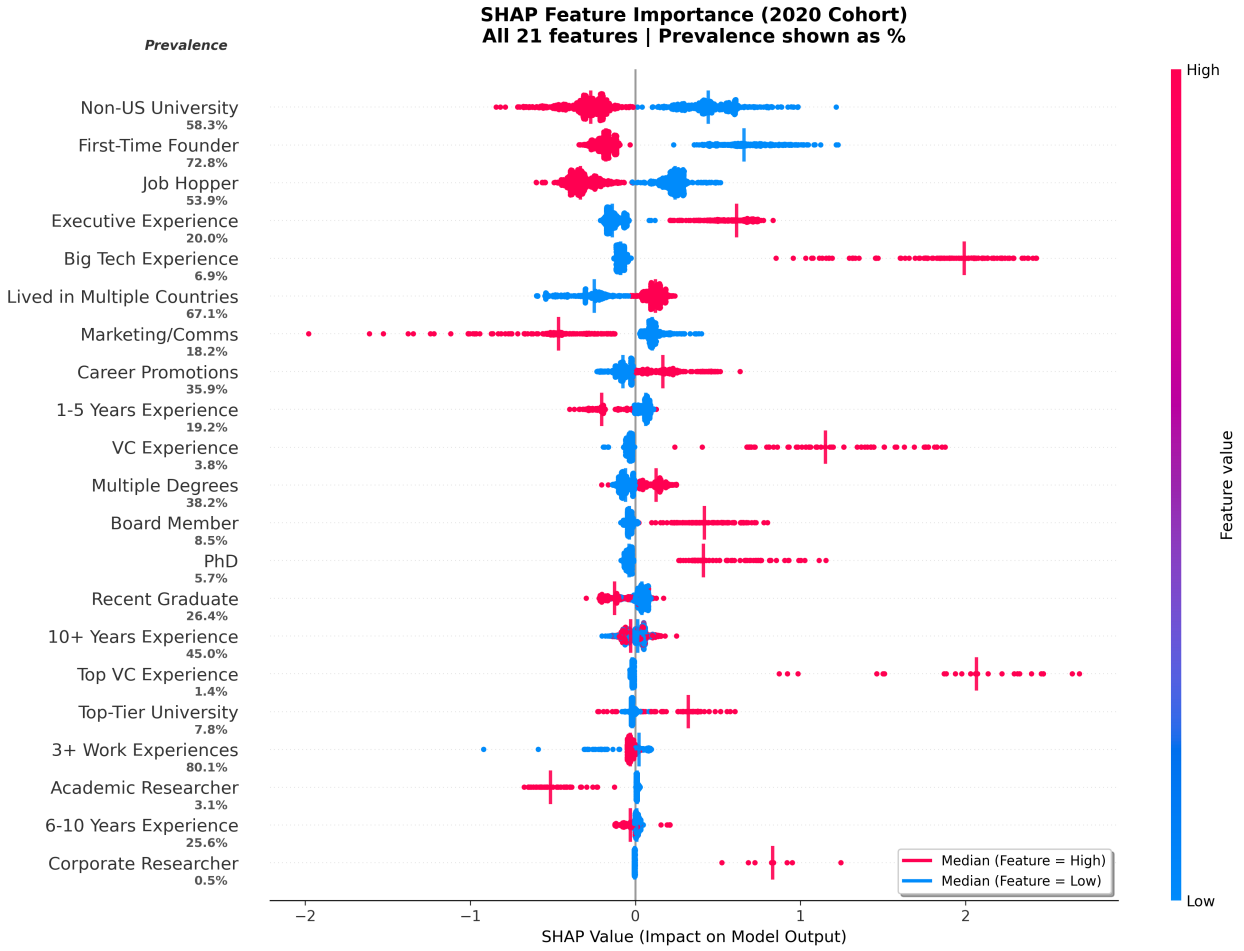
2022 Prediction Cohort - Policy Ablation Analysis

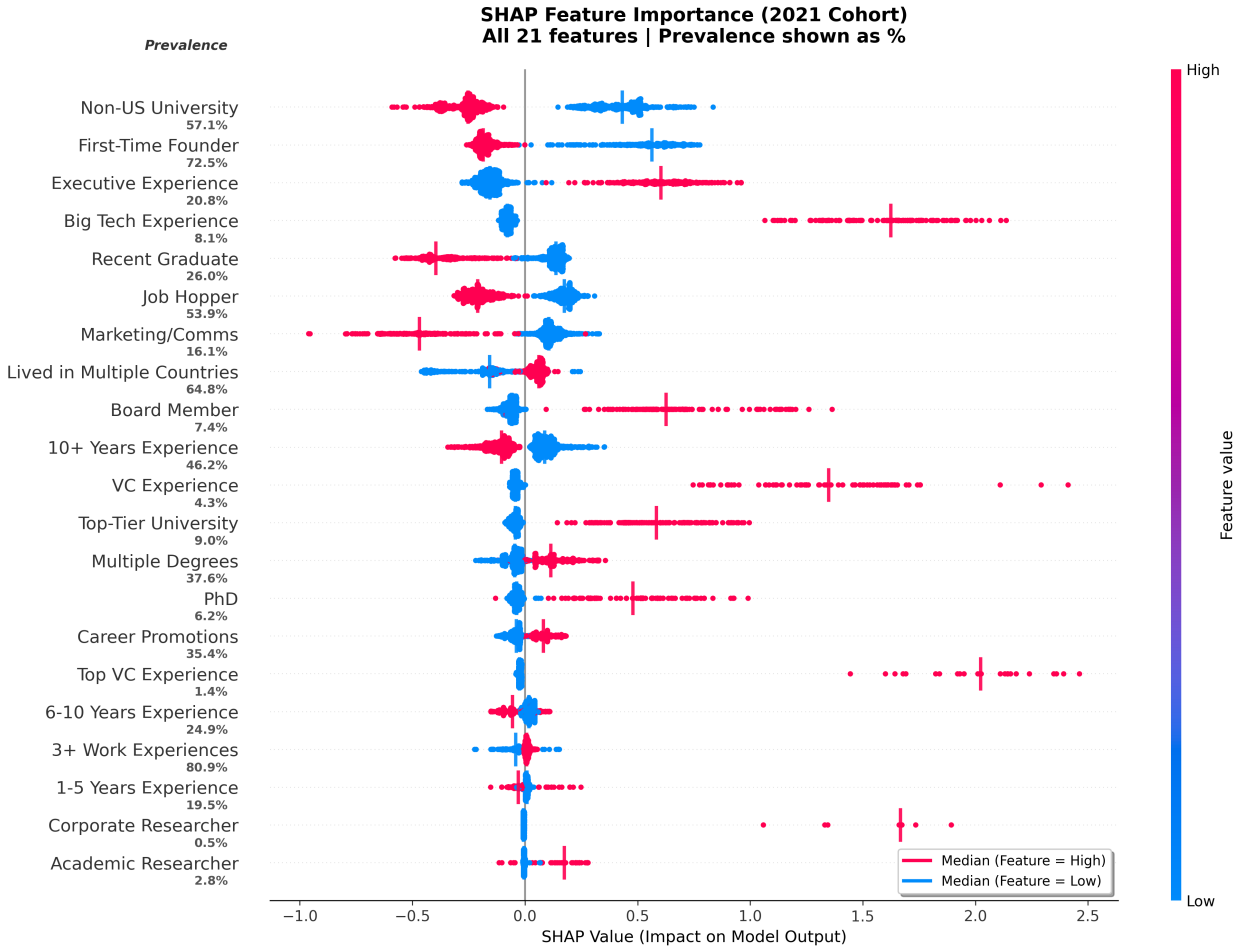
line number	importance score	baseline f05	ablated f05	ablated precision	ablated recall	ablated accuracy	delta precision	delta recall
1	0.0135	0.1773	0.2764	0.1367	0.3333	0.9186	0.0027	0.0000
2	0.0297	0.1773	0.1512	0.1611	0.3333	0.8987	-0.0206	0.0000
3	0.0221	0.2898	0.2793	0.1744	0.3333	0.9163	-0.0258	0.1111
4	0.0162	0.2898	0.2291	0.2199	0.2222	0.9340	-0.0114	0.0000
5	0.0102	0.2898	0.1585	0.2011	0.4444	0.8920	0.0121	0.1111
6	0.0247	0.2898	0.1543	0.2480	0.2222	0.8683	-0.0026	0.1111
7	0.0249	0.2898	0.1997	0.2410	0.2222	0.8470	-0.0287	0.1111
8	0.0070	0.1887	0.2741	0.1609	0.1111	0.9113	-0.0288	0.0000
9	0.0230	0.2898	0.1997	0.1827	0.3333	0.8877	0.0015	0.1111
10	0.0157	0.1773	0.2656	0.1973	0.4444	0.9214	-0.0287	0.1111
11	0.0058	0.1887	0.1620	0.1743	0.3333	0.8779	-0.0140	0.0000
12	0.0269	0.1887	0.2722	0.1297	0.4444	0.8510	0.0056	0.0000
13	0.0145	0.1887	0.2537	0.2018	0.4444	0.9321	0.0157	0.0000
14	0.0208	0.2898	0.1586	0.2189	0.2222	0.9068	-0.0113	0.1111
15	0.0113	0.1887	0.1974	0.2043	0.4444	0.8326	-0.0151	0.1111
16	0.0145	0.2898	0.2784	0.1369	0.1111	0.8454	-0.0048	0.1111
17	0.0224	0.1773	0.1630	0.2471	0.2222	0.8856	-0.0029	0.0000
18	0.0272	0.1887	0.2517	0.1304	0.3333	0.8909	-0.0133	0.0000
19	0.0268	0.2898	0.2245	0.1654	0.3333	0.8778	-0.0249	0.0000
20	0.0199	0.2898	0.1698	0.2300	0.3333	0.9055	-0.0173	0.1111

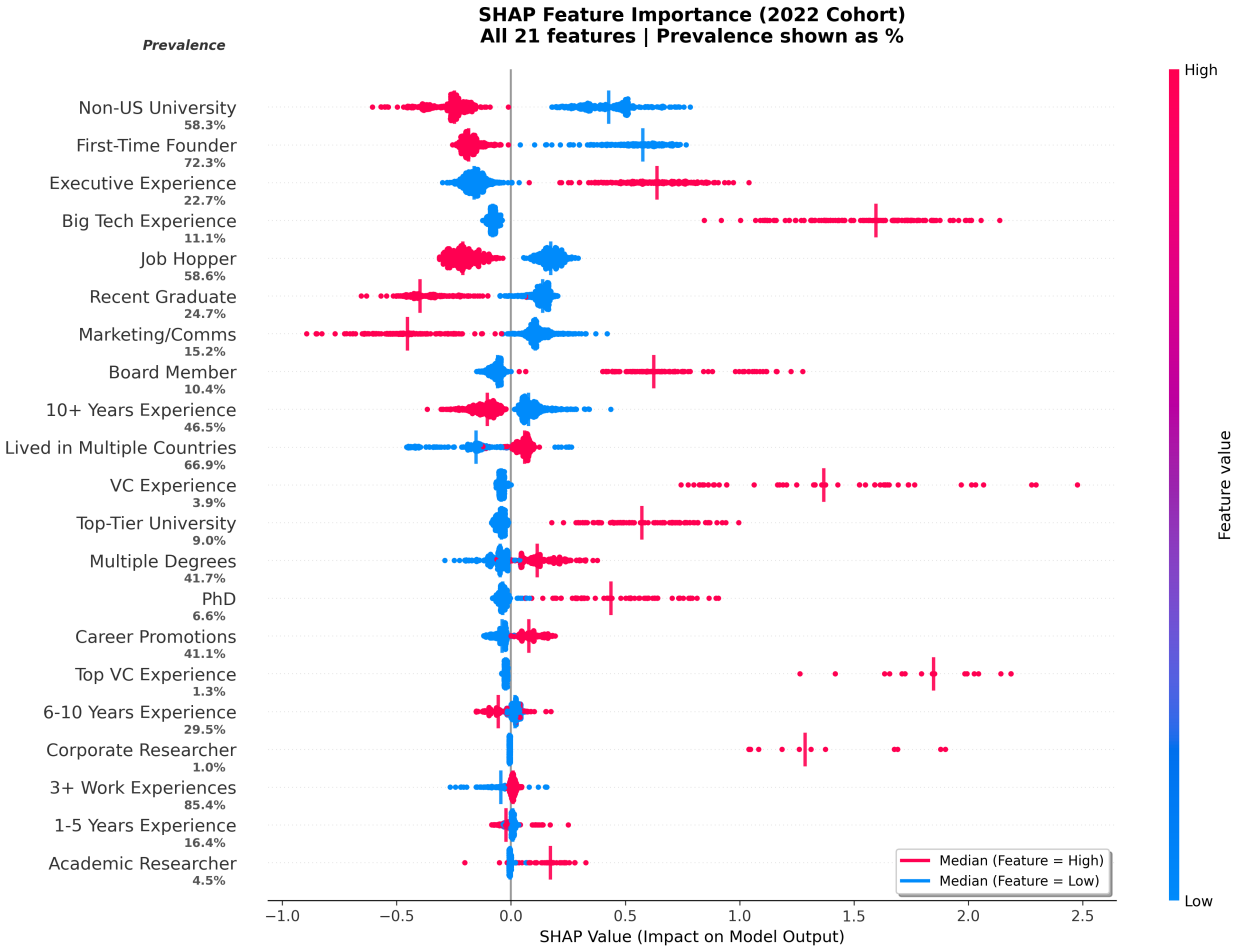
Appendix G: SHAP Beeswarm Plots (2018-2022)




SHAP Feature Importance (2019 Cohort)
All 21 features | Prevalence shown as %
Prevalence
Non-US University
58.1%
First-Time Founder
72.9%
Executive Experience
22.5%
Recent Graduate
26.6%
Job Hopper
49.6%
Lived in Multiple Countries
65.2%
Career Promotions
35.2%
Board Member
9.8%
3+ Work Experiences
78.1%
Big Tech Experience
7.1%
Multiple Degrees
36.1%
Top-Tier University
6.9%
VC Experience
3.9%
1-5 Years Experience
19.3%
Marketing/Comms
17.3%
PhD
6.1%
Academic Researcher
2.8%
6-10 Years Experience
24.0%
Top VC Experience
1.0%
10+ Years Experience
45.5%
Corporate Researcher
0.6%
Median (Feature = High)
Median (Feature = Low)
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
SHAP Value (Impact on Model Output)
High
Feature value
Low


SHAP Feature Importance (2020 Cohort)
All 21 features | Prevalence shown as %
Prevalence
Non-US University
58.3%
First-Time Founder
72.8%
Job Hopper
53.9%
Executive Experience
20.0%
Big Tech Experience
6.9%
Lived in Multiple Countries
67.1%
Marketing/Comms
18.2%
Career Promotions
35.9%
1-5 Years Experience
19.2%
VC Experience
3.8%
Multiple Degrees
38.2%
Board Member
8.5%
PhD
5.7%
Recent Graduate
26.4%
10+ Years Experience
45.0%
Top VC Experience
1.4%
Top-Tier University
7.8%
3+ Work Experiences
80.1%
Academic Researcher
3.1%
6-10 Years Experience
25.6%
Corporate Researcher
0.5%
Median (Feature = High)
Median (Feature = Low)
−2
−1
0
1
2
SHAP Value (Impact on Model Output)
High
Feature value
Low


SHAP Feature Importance (2021 Cohort)
All 21 features | Prevalence shown as %
Prevalence
Non-US University
57.1%
First-Time Founder
72.5%
Executive Experience
20.8%
Big Tech Experience
8.1%
Recent Graduate
26.0%
Job Hopper
53.9%
Marketing/Comms
16.1%
Lived in Multiple Countries
64.8%
Board Member
7.4%
10+ Years Experience
46.2%
VC Experience
4.3%
Top-Tier University
9.0%
Multiple Degrees
37.6%
PhD
6.2%
Career Promotions
35.4%
Top VC Experience
1.4%
6-10 Years Experience
24.9%
3+ Work Experiences
80.9%
1-5 Years Experience
19.5%
Corporate Researcher
0.5%
Academic Researcher
2.8%
Median (Feature = High)
Median (Feature = Low)
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
SHAP Value (Impact on Model Output)
High
Feature value
Low


SHAP Feature Importance (2022 Cohort)
All 21 features | Prevalence shown as %
Prevalence
Non-US University
58.3%
First-Time Founder
72.3%
Executive Experience
22.7%
Big Tech Experience
11.1%
Job Hopper
58.6%
Recent Graduate
24.7%
Marketing/Comms
15.2%
Board Member
10.4%
10+ Years Experience
46.5%
Lived in Multiple Countries
66.9%
VC Experience
3.9%
Top-Tier University
9.0%
Multiple Degrees
41.7%
PhD
6.6%
Career Promotions
41.1%
Top VC Experience
1.3%
6-10 Years Experience
29.5%
Corporate Researcher
1.0%
3+ Work Experiences
85.4%
1-5 Years Experience
16.4%
Academic Researcher
4.5%
Median (Feature = High)
Median (Feature = Low)
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
SHAP Value (Impact on Model Output)
High
Feature value
Low

Appendix H: LLM-based Feature Extraction Prompt

```
# Feature Extraction: LLM Prompts & RAG Context Documentation

**Model Used:** `gpt-4.1-mini-2025-04-14` (OpenAI)
**Temperature:** 0.1 (for deterministic, rule-based extraction)
**Output Format:** JSON with structured reasoning

---

## Executive Summary

This document details all 16 LLM-powered feature extraction prompts used in our
founder evaluation system. The system extracts 21 binary features total.

---

## System Architecture

### Global System Prompt (Applied to All LLM Calls)

**For Complex Reasoning Features:**
```
You are a data analyst extracting {feature_name}. STRICT RULES:
1) Base answer ONLY on explicit information provided.
2) When reference lists are provided, ONLY match items from those EXACT lists - NO ex-
ceptions.
3) If something seems similar but is NOT in the reference list, return 0.
4) Do NOT make assumptions, guesses, or judgments about what 'should' be included.
5) If information is unclear or missing, return 0.
```

**For LLM-Based Rule Features:**
```
You are a data analyst extracting {feature_name}. STRICT RULES:
1) Base answer ONLY on explicit information provided.
2) When reference lists are provided, ONLY match items from those EXACT lists - NO ex-
ceptions.
3) If something seems similar but is NOT in the reference list, return 0.
4) Do NOT make assumptions, guesses, or judgments about what 'should' be included.
5) If information is unclear or missing, return 0.
```

### Standard RAG Context Injected for All Features

**Founder Data Structure:**
```python
```
```

```
{
    "name": str,                  # Founder's full name
    "position": str,              # Current role/title
    "company": str,               # Company name
    "founded_year": int,          # Year company was founded
    "current_country": str,       # Current country of residence
    "education": list[dict],      # Education history (JSON)
    "work_history": list[dict],   # Work experience (JSON)
    "skills": str,                # Comma-separated skills
    "company_description": str    # Optional company description
}
```

---

1. LLM-Based Rule Features (6 Features)

These features use LLM for complex logical reasoning with messy data.

1.1 `lived_multiple_countries`

Description: Has the founder lived, worked, or studied in 2+ different countries?

LLM Prompt:
```
EXTRACT: Has this person lived/worked/studied in exactly 2 or more different coun-
tries?

STEP-BY-STEP LOGIC:
1. Find education_country from all education entries
2. Find country from all work_history entries
3. Find current_country from the provided current_country field
4. Normalize country names ("United States" = "USA" = "US" = "united states")
5. Count distinct countries across all sources
6. DECISION LOGIC:
   - If distinct countries = 0 or 1 → Return 0
   - If distinct countries ≥ 2 → Return 1
7. DOUBLE CHECK: Count your distinct countries and verify the return value matches

EVIDENCE REQUIRED: "Education: [countries], Work: [countries], Current: [current_coun-
try] → Distinct: X countries → Decision: Return Y"
```

RAG Context:
- `education` (full list with `education_country`)
- `work_history` (full list with `country`)
- `current_country`

Appendix I: SHAP Interaction Values

2022 Top 20 SHAP Interaction Values

```
================================================================================================
TOP 20 STRONGEST POSITIVE INTERACTIONS
================================================================================================

Rank   Feature 1                        Feature 2                        Interaction
------------------------------------------------------------------------------------------------
1      Lived in Multiple Countries      Non-US University                  +0.034489
2      Non-US University                Top-Tier University                +0.005753
3      10+ Years Experience             3+ Work Experiences                +0.004090
4      10+ Years Experience             Job Hopper                         +0.003090
5      Executive Experience             Recent Graduate                    +0.002718
6      10+ Years Experience             Career Promotions                  +0.002662
7      Career Promotions                Executive Experience               +0.002545
8      Executive Experience             Marketing/Comms                    +0.002256
9      Big Tech Experience              Marketing/Comms                    +0.002204
10     PhD                              Recent Graduate                    +0.002095
11     Career Promotions                Job Hopper                         +0.001494
12     Executive Experience             Job Hopper                         +0.001451
13     Academic Researcher              Marketing/Comms                    +0.001308
14     3+ Work Experiences              6-10 Years Experience              +0.001174
15     Board Member                     Non-US University                  +0.001124
16     Board Member                     Executive Experience               +0.001053
17     Board Member                     Lived in Multiple Countries        +0.001016
18     10+ Years Experience             Multiple Degrees                   +0.001006
19     10+ Years Experience             Executive Experience               +0.000997
20     Job Hopper                       Multiple Degrees                   +0.000970
================================================================================================
TOP 20 STRONGEST NEGATIVE INTERACTIONS
================================================================================================

Rank   Feature 1                        Feature 2                        Interaction
------------------------------------------------------------------------------------------------
1      Multiple Degrees                 PhD                                -0.005201
2      Big Tech Experience              Board Member                       -0.003321
3      10+ Years Experience             6-10 Years Experience              -0.003045
4      10+ Years Experience             Recent Graduate                    -0.002530
5      Multiple Degrees                 Recent Graduate                    -0.002233
6      Lived in Multiple Countries      Marketing/Comms                    -0.002228
7      6-10 Years Experience            First-Time Founder                 -0.002011
8      Executive Experience             First-Time Founder                 -0.001863
9      First-Time Founder               Non-US University                  -0.001815
10     First-Time Founder               VC Experience                      -0.001717
11     Big Tech Experience              Corporate Researcher               -0.001713
12     Board Member                     PhD                                -0.001652
13     Non-US University                PhD                                -0.001608
14     Big Tech Experience              Top-Tier University                -0.001608
15     Big Tech Experience              Executive Experience               -0.001450
16     10+ Years Experience             VC Experience                      -0.001395
17     1-5 Years Experience             Board Member                       -0.001322
18     1-5 Years Experience             Big Tech Experience                -0.001151
19     Big Tech Experience              Top VC Experience                  -0.001105
20     Recent Graduate                  VC Experience                      -0.000994

================================================================================================
END OF INTERACTION MATRIX REPORT
================================================================================================
```

2021 Top 20 SHAP Interaction Values

```
================================================================================================
TOP 20 STRONGEST POSITIVE INTERACTIONS
================================================================================================

Rank   Feature 1                        Feature 2                          Interaction
------------------------------------------------------------------------------------------------
1      Lived in Multiple Countries      Non-US University                    +0.034588
2      PhD                              Recent Graduate                      +0.007124
3      Big Tech Experience              Board Member                         +0.006603
4      Non-US University                Top-Tier University                  +0.006087
5      10+ Years Experience             Job Hopper                           +0.004396
6      10+ Years Experience             3+ Work Experiences                  +0.003620
7      Executive Experience             Marketing/Comms                      +0.003077
8      Executive Experience             Recent Graduate                      +0.002737
9      Big Tech Experience              Multiple Degrees                     +0.002562
10     Recent Graduate                  Top-Tier University                  +0.001690
11     Career Promotions                Job Hopper                           +0.001676
12     Board Member                     Executive Experience                 +0.001569
13     10+ Years Experience             Multiple Degrees                     +0.001558
14     10+ Years Experience             Career Promotions                    +0.001391
15     Job Hopper                       Top-Tier University                  +0.001385
16     3+ Work Experiences              Marketing/Comms                      +0.001264
17     3+ Work Experiences              Top-Tier University                  +0.001241
18     Job Hopper                       Multiple Degrees                     +0.001123
19     Executive Experience             Job Hopper                           +0.000965
20     10+ Years Experience             Marketing/Comms                      +0.000945
================================================================================================
TOP 20 STRONGEST NEGATIVE INTERACTIONS
================================================================================================

Rank   Feature 1                        Feature 2                          Interaction
------------------------------------------------------------------------------------------------
1      Multiple Degrees                 PhD                                  -0.006436
2      First-Time Founder               VC Experience                        -0.005352
3      Multiple Degrees                 Recent Graduate                      -0.003843
4      Big Tech Experience              Executive Experience                 -0.003197
5      Non-US University                PhD                                  -0.002802
6      10+ Years Experience             6-10 Years Experience                -0.002649
7      Recent Graduate                  VC Experience                        -0.002584
8      Executive Experience             VC Experience                        -0.002134
9      10+ Years Experience             Recent Graduate                      -0.002133
10     Board Member                     Top-Tier University                  -0.002114
11     Board Member                     PhD                                  -0.002103
12     First-Time Founder               Non-US University                    -0.002080
13     10+ Years Experience             VC Experience                        -0.001772
14     First-Time Founder               Marketing/Comms                      -0.001590
15     3+ Work Experiences              Big Tech Experience                  -0.001576
16     Executive Experience             First-Time Founder                   -0.001407
17     10+ Years Experience             Big Tech Experience                  -0.001363
18     Executive Experience             Top-Tier University                  -0.001316
19     Board Member                     Lived in Multiple Countries          -0.001293
20     Lived in Multiple Countries      Marketing/Comms                      -0.001268

================================================================================================
END OF INTERACTION MATRIX REPORT
================================================================================================
```

2020 Top 20 SHAP Interaction Values

```
================================================================================================
TOP 20 STRONGEST POSITIVE INTERACTIONS
================================================================================================

Rank   Feature 1                        Feature 2                          Interaction
------------------------------------------------------------------------------------------------
1      Lived in Multiple Countries      Non-US University                    +0.031852
2      3+ Work Experiences              Executive Experience                 +0.007399
3      Career Promotions                Job Hopper                           +0.003717
4      Board Member                     First-Time Founder                   +0.003114
5      First-Time Founder               Marketing/Comms                      +0.002725
6      Big Tech Experience              Job Hopper                           +0.002454
7      Career Promotions                Non-US University                    +0.002402
8      Job Hopper                       Non-US University                    +0.002330
9      10+ Years Experience             Career Promotions                    +0.002123
10     Executive Experience             Job Hopper                           +0.002091
11     10+ Years Experience             Multiple Degrees                     +0.002040
12     10+ Years Experience             Executive Experience                 +0.002034
13     Non-US University                VC Experience                        +0.001932
14     Job Hopper                       Recent Graduate                      +0.001894
15     1-5 Years Experience             Non-US University                    +0.001885
16     1-5 Years Experience             Career Promotions                    +0.001711
17     3+ Work Experiences              Big Tech Experience                  +0.001542
18     First-Time Founder               Non-US University                    +0.001339
19     Career Promotions                Recent Graduate                      +0.001314
20     Job Hopper                       VC Experience                        +0.001133
================================================================================================
TOP 20 STRONGEST NEGATIVE INTERACTIONS
================================================================================================

Rank   Feature 1                        Feature 2                          Interaction
------------------------------------------------------------------------------------------------
1      Non-US University                PhD                                  -0.006904
2      10+ Years Experience             Recent Graduate                      -0.005373
3      Top VC Experience                VC Experience                        -0.003627
4      Lived in Multiple Countries      Multiple Degrees                     -0.003042
5      Marketing/Comms                  Top VC Experience                    -0.003003
6      3+ Work Experiences              Lived in Multiple Countries          -0.002853
7      Multiple Degrees                 Non-US University                    -0.001848
8      Multiple Degrees                 PhD                                  -0.001841
9      Board Member                     Career Promotions                    -0.001474
10     Marketing/Comms                  Top-Tier University                  -0.001440
11     Corporate Researcher             Multiple Degrees                     -0.001359
12     10+ Years Experience             Marketing/Comms                      -0.001234
13     Academic Researcher              Multiple Degrees                     -0.001194
14     Big Tech Experience              Marketing/Comms                      -0.001148
15     Big Tech Experience              Non-US University                    -0.000973
16     Job Hopper                       Multiple Degrees                     -0.000949
17     1-5 Years Experience             6-10 Years Experience                -0.000820
18     3+ Work Experiences              Job Hopper                           -0.000798
19     Board Member                     Lived in Multiple Countries          -0.000759
20     10+ Years Experience             Academic Researcher                  -0.000723

================================================================================================
END OF INTERACTION MATRIX REPORT
================================================================================================
```

2019 SHAP Top 20 Interaction Values

```
================================================================================================
TOP 20 STRONGEST POSITIVE INTERACTIONS
================================================================================================

Rank   Feature 1                        Feature 2                        Interaction
------------------------------------------------------------------------------------------------
1      3+ Work Experiences              Job Hopper                         +0.030275
2      Lived in Multiple Countries      Non-US University                  +0.019053
3      3+ Work Experiences              Multiple Degrees                   +0.004940
4      Board Member                     First-Time Founder                 +0.004663
5      1-5 Years Experience             Career Promotions                  +0.004497
6      Academic Researcher              PhD                                +0.003306
7      Board Member                     Executive Experience               +0.003228
8      3+ Work Experiences              Big Tech Experience                +0.003000
9      Board Member                     Job Hopper                         +0.002681
10     Career Promotions                Job Hopper                         +0.002562
11     PhD                              Recent Graduate                    +0.002485
12     Executive Experience             Job Hopper                         +0.002472
13     Big Tech Experience              Non-US University                  +0.002379
14     Non-US University                Top-Tier University                +0.002194
15     3+ Work Experiences              First-Time Founder                 +0.002141
16     Multiple Degrees                 PhD                                +0.002118
17     1-5 Years Experience             Non-US University                  +0.002042
18     Executive Experience             Non-US University                  +0.002001
19     First-Time Founder               VC Experience                      +0.001898
20     Job Hopper                       Top-Tier University                +0.001784
================================================================================================
TOP 20 STRONGEST NEGATIVE INTERACTIONS
================================================================================================

Rank   Feature 1                        Feature 2                        Interaction
------------------------------------------------------------------------------------------------
1      Executive Experience             Recent Graduate                    -0.007072
2      3+ Work Experiences              Lived in Multiple Countries        -0.004064
3      1-5 Years Experience             First-Time Founder                 -0.003738
4      Non-US University                Recent Graduate                    -0.003731
5      First-Time Founder               Non-US University                  -0.003071
6      Career Promotions                First-Time Founder                 -0.002769
7      Big Tech Experience              First-Time Founder                 -0.002505
8      Board Member                     VC Experience                      -0.001972
9      First-Time Founder               Top VC Experience                  -0.001622
10     1-5 Years Experience             Recent Graduate                    -0.001520
11     Academic Researcher              Non-US University                  -0.001463
12     10+ Years Experience             Recent Graduate                    -0.001441
13     3+ Work Experiences              Career Promotions                  -0.001360
14     Career Promotions                Top-Tier University                -0.001342
15     Multiple Degrees                 Recent Graduate                    -0.001333
16     1-5 Years Experience             Executive Experience               -0.001211
17     Academic Researcher              Multiple Degrees                   -0.001200
18     1-5 Years Experience             Top-Tier University                -0.001191
19     1-5 Years Experience             VC Experience                      -0.001178
20     Lived in Multiple Countries      PhD                                -0.001145

================================================================================================
END OF INTERACTION MATRIX REPORT
================================================================================================
```

2018 Top 20 SHAP Interaction Values

```
================================================================================================
TOP 20 STRONGEST POSITIVE INTERACTIONS
================================================================================================

Rank   Feature 1                        Feature 2                        Interaction
------------------------------------------------------------------------------------------------
1      Lived in Multiple Countries      Non-US University                  +0.012110
2      Non-US University                Recent Graduate                    +0.009546
3      Executive Experience             Recent Graduate                    +0.008889
4      Academic Researcher              Executive Experience               +0.006559
5      Academic Researcher              Lived in Multiple Countries        +0.004741
6      Multiple Degrees                 PhD                                +0.003851
7      Multiple Degrees                 Recent Graduate                    +0.003081
8      First-Time Founder               Marketing/Comms                    +0.003080
9      Career Promotions                Non-US University                  +0.002981
10     3+ Work Experiences              6-10 Years Experience              +0.002956
11     3+ Work Experiences              Multiple Degrees                   +0.002726
12     3+ Work Experiences              VC Experience                      +0.002579
13     Executive Experience             Non-US University                  +0.002435
14     1-5 Years Experience             Top-Tier University                +0.002401
15     3+ Work Experiences              Big Tech Experience                +0.002336
16     3+ Work Experiences              Executive Experience               +0.002220
17     6-10 Years Experience            Board Member                       +0.001816
18     Executive Experience             Multiple Degrees                   +0.001647
19     Multiple Degrees                 Non-US University                  +0.001456
20     Executive Experience             Lived in Multiple Countries        +0.001402

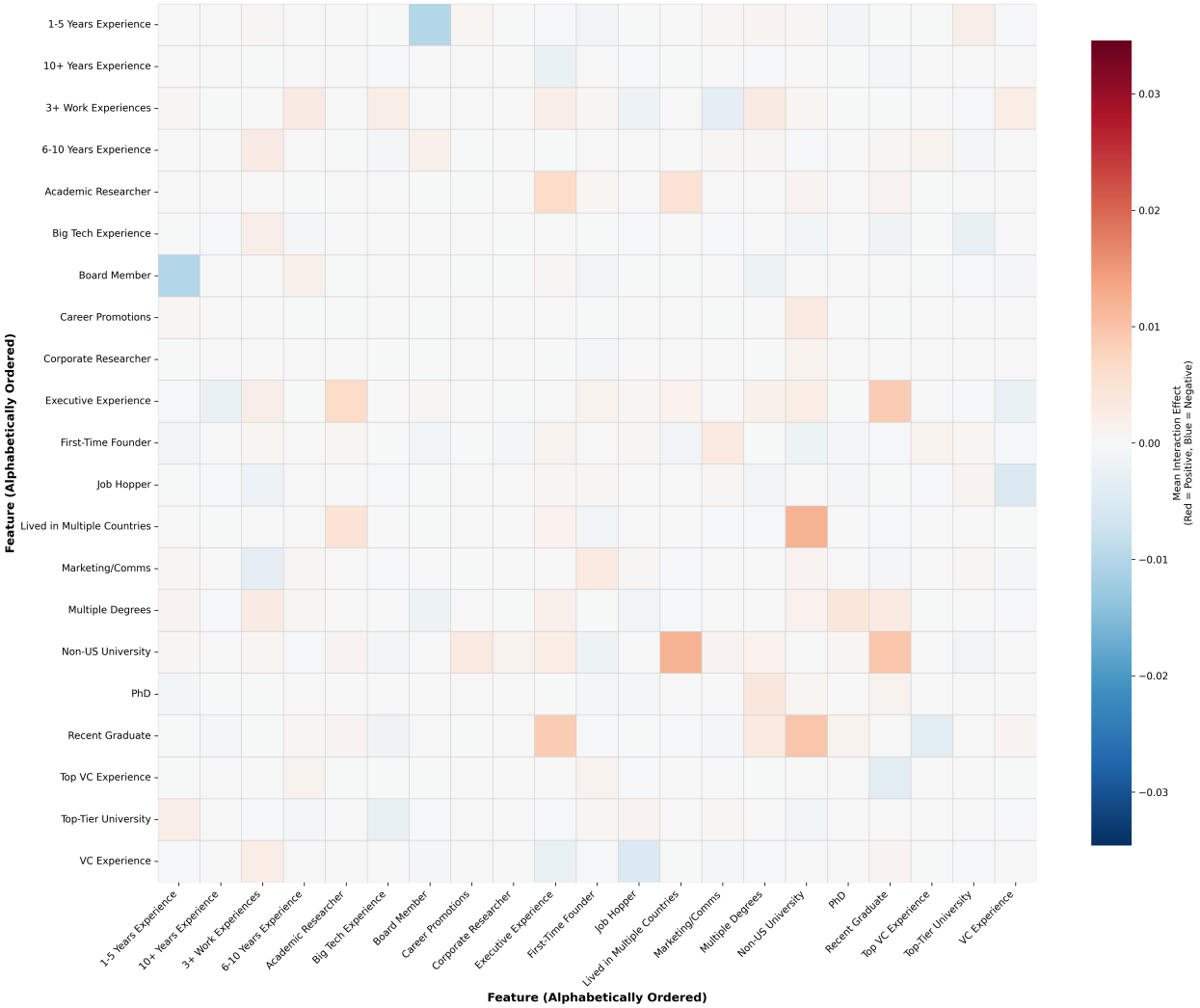
================================================================================================
TOP 20 STRONGEST NEGATIVE INTERACTIONS
================================================================================================

Rank   Feature 1                        Feature 2                        Interaction
------------------------------------------------------------------------------------------------
1      1-5 Years Experience             Board Member                       -0.010000
2      Job Hopper                       VC Experience                      -0.005215
3      Recent Graduate                  Top VC Experience                  -0.003933
4      3+ Work Experiences              Marketing/Comms                    -0.003289
5      Big Tech Experience              Top-Tier University                -0.002687
6      Executive Experience             VC Experience                      -0.002424
7      10+ Years Experience             Executive Experience               -0.002188
8      3+ Work Experiences              Job Hopper                         -0.001896
9      First-Time Founder               Non-US University                  -0.001881
10     Board Member                     Multiple Degrees                   -0.001856
11     Big Tech Experience              Recent Graduate                    -0.001593
12     Board Member                     First-Time Founder                 -0.001313
13     First-Time Founder               Lived in Multiple Countries        -0.001165
14     Non-US University                Top-Tier University                -0.001149
15     Job Hopper                       Multiple Degrees                   -0.001125
16     1-5 Years Experience             First-Time Founder                 -0.001120
17     Big Tech Experience              Non-US University                  -0.001040
18     1-5 Years Experience             PhD                                -0.000881
19     Job Hopper                       PhD                                -0.000810
20     6-10 Years Experience            Big Tech Experience                -0.000803

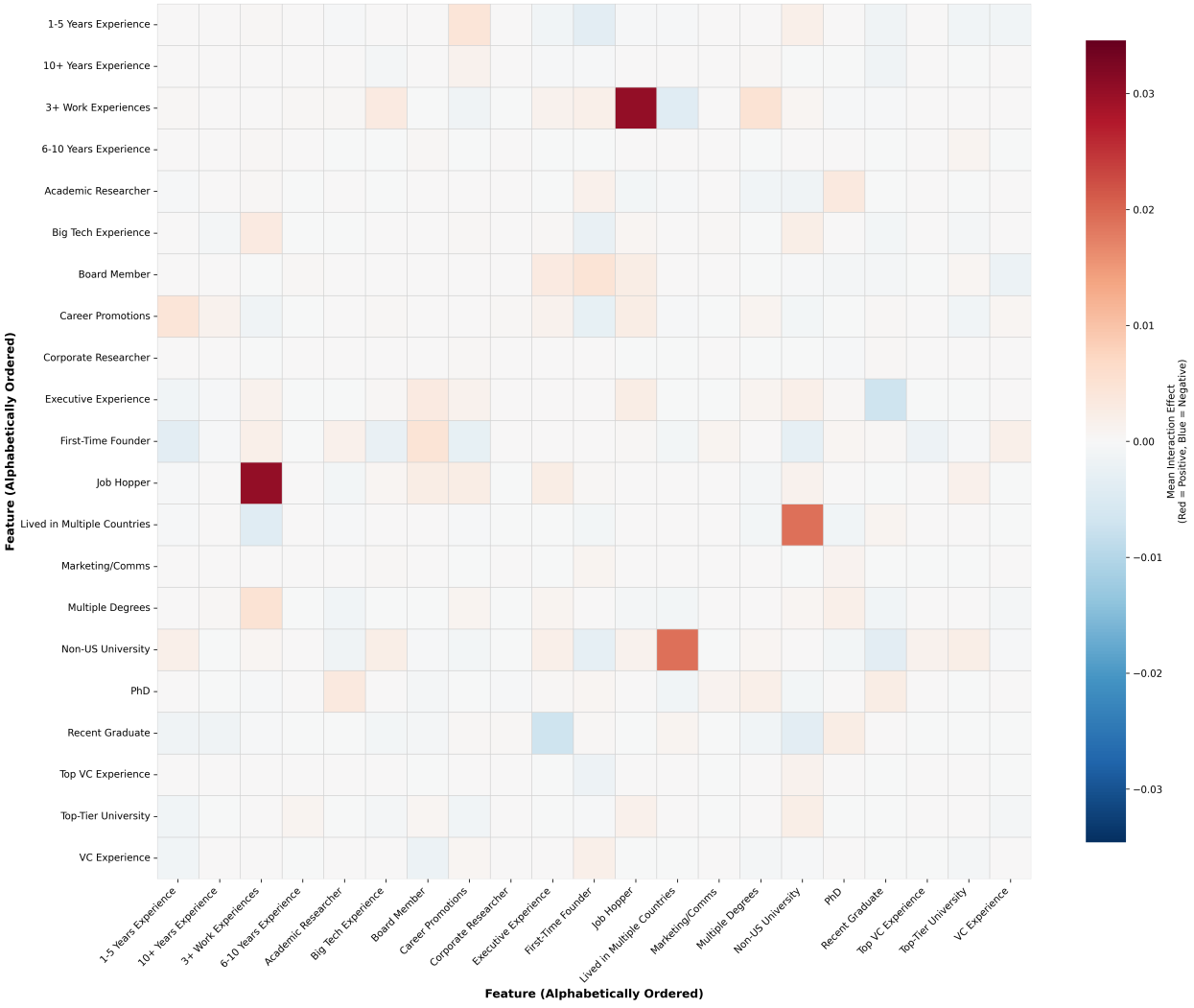
================================================================================================
END OF INTERACTION MATRIX REPORT
================================================================================================
```

Appendix J: SHAP Interaction Matrices

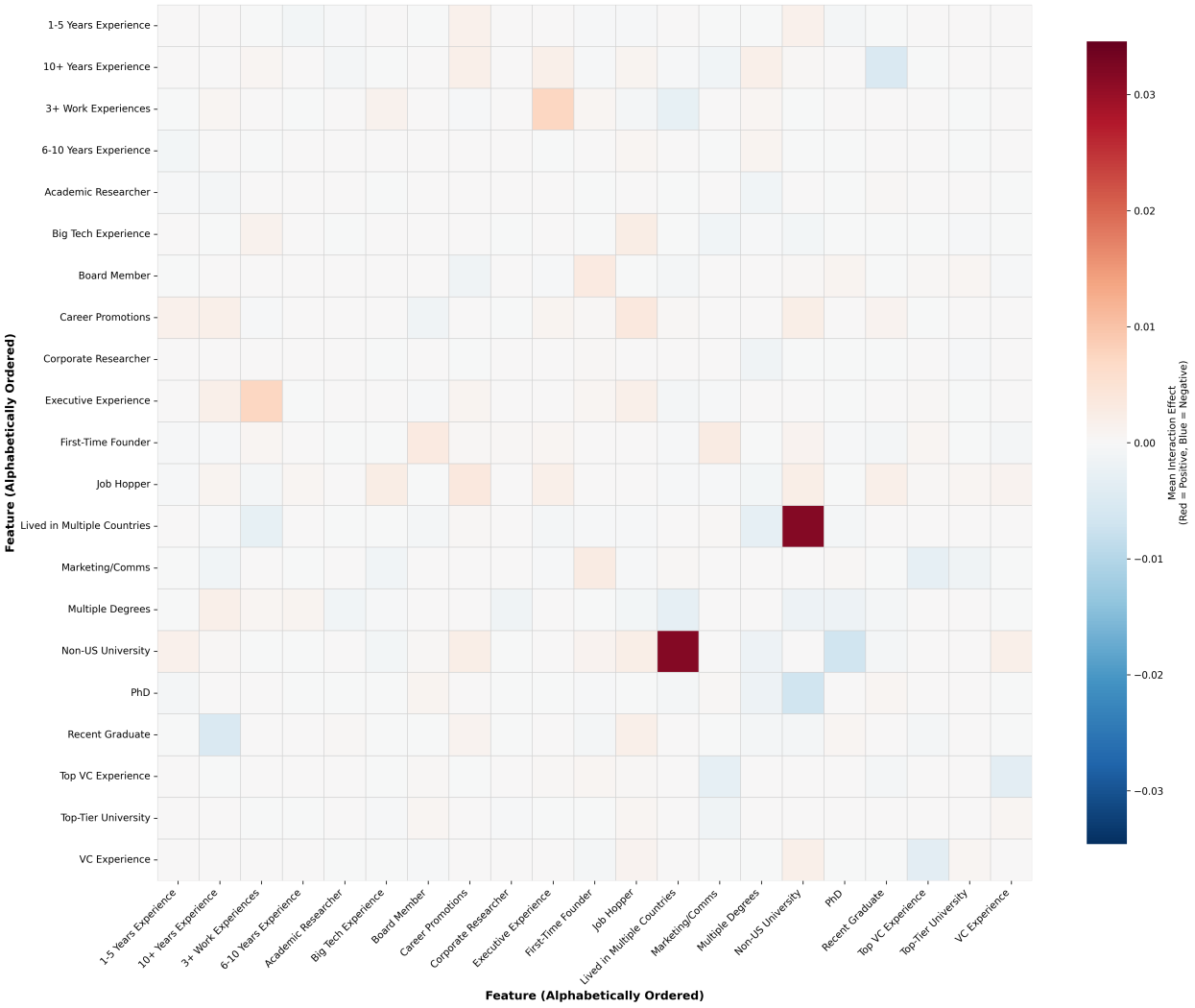
2018 SHAP Interaction Matrix



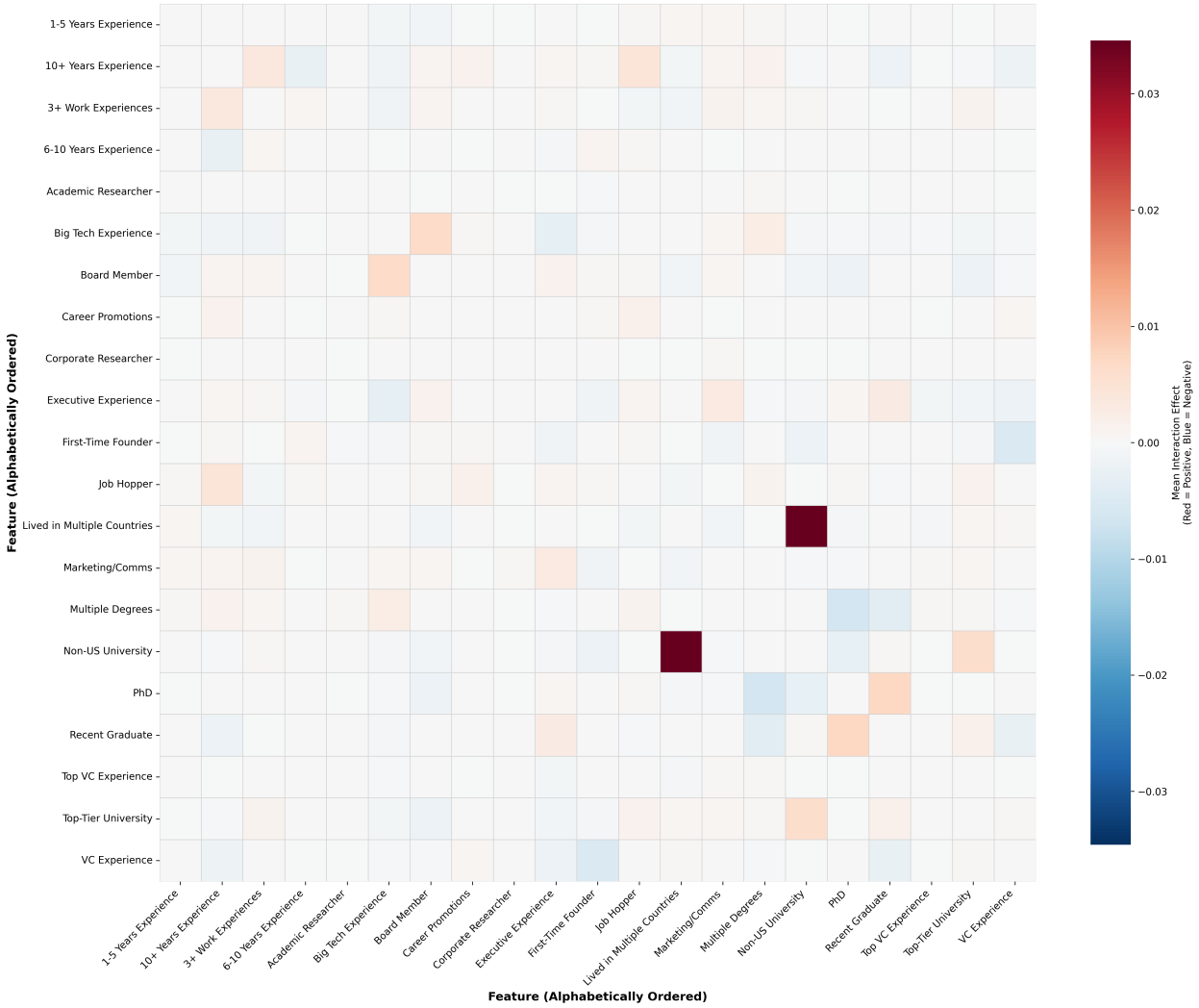
2019 SHAP Interaction Matrix



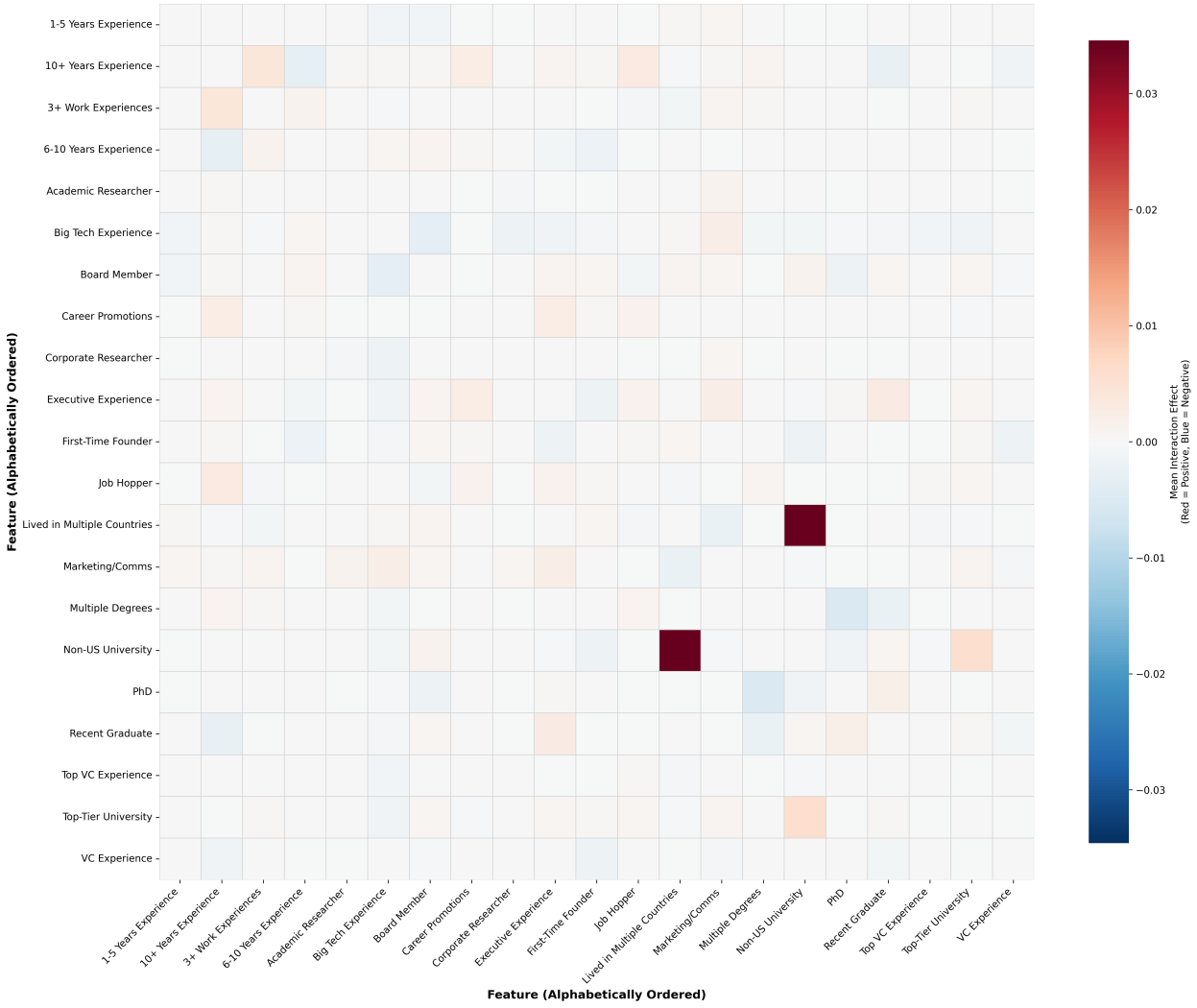
2020 SHAP Interaction Matrix



2021 SHAP Interaction Matrix



2022 SHAP Interaction Matrix



Appendix K: General SHAP Tables

Absolute Mean SHAP Value across cohorts 2018-2022

Feature	2018	2019	2020	2021	2022
1-5_years_experience	0.051	0.0863	0.0807	0.0157	0.015
10+_years_experience	0.0163	0.0125	0.0486	0.1052	0.1022
3+_Work_Experiences	0.0953	0.1289	0.0328	0.0189	0.0187
6-10_Years_Experience	0.0472	0.0245	0.0212	0.0306	0.0318
Academic_Researcher	0.0256	0.0258	0.0254	0.0082	0.011
Big_Tech_Experience	0.1196	0.1257	0.2102	0.1983	0.2414
Board_Member	0.1152	0.1377	0.0705	0.108	0.1274
Career_Promotions	0.0324	0.1423	0.1046	0.058	0.061
Corporate_Researcher	0.0235	0.003	0.0073	0.0132	0.0194
Exec_Experience	0.1949	0.2769	0.2225	0.2506	0.266
First_Time_Founder	0.2346	0.3372	0.3192	0.2785	0.2826
Job_Hopper	0.3076	0.2014	0.2818	0.1904	0.1924
Lived_in_multiple_countries	0.1378	0.1591	0.1731	0.109	0.1004
Marketing & Comms	0.1283	0.0802	0.1716	0.1707	0.1682
Multiple Degrees	0.1295	0.1178	0.0772	0.0769	0.0849
Non_US_Uni	0.4834	0.421	0.3571	0.3326	0.3278
PhD	0.0702	0.0659	0.0678	0.063	0.063
Recent_Grad	0.3417	0.2231	0.0672	0.1935	0.1914
Top_Tier_Uni	0.109	0.1024	0.0388	0.0947	0.0924
Top_VC_Experience	0.023	0.0199	0.0449	0.0514	0.0472
VC_Experience	0.0692	0.0967	0.0785	0.0989	0.0948

Feature Prevalence (%) across cohorts 2018-2022

Feature	2018	2019	2020	2021	2022
1-5_years_experience	21.6	19.3	19.2	19.5	16.4
10+_years_experience	43	45.5	45	46.2	46.5
3+_Work_Experiences	74.4	78.1	80.1	80.9	85.4
6-10_Years_Experience	23.9	24	25.6	24.9	29.5
Academic_Researcher	2.5	2.8	3.1	2.8	4.5
Big_Tech_Experience	5.9	7.1	6.9	8.1	11.1
Board_Member	8	9.8	8.5	7.4	10.4
Career_Promotions	27.8	35.2	35.9	35.4	41.1
Corporate_Researcher	0.5	0.6	0.5	0.5	1
Exec_Experience	21.9	22.5	20	20.8	22.7

First_Time_Founder	71.7	72.9	72.8	72.5	72.3
Job_Hopper	44.2	49.6	53.9	53.9	58.6
Lived_in_multiple_countries	62.8	65.2	67.1	64.8	66.9
Marketing & Comms	16.2	17.3	18.2	16.1	15.2
Multiple Degrees	36	36.1	38.2	37.6	41.7
Non_US_Uni	53.4	58.1	58.3	57.1	58.3
PhD	7.6	6.1	5.7	6.2	6.6
Recent_Grad	25	26.6	26.4	26	24.7
Top_Tier_Uni	7.7	6.9	7.8	9	9
Top_VC_Experience	1.3	1	1.4	1.4	1.3
VC_Experience	3.8	3.9	3.8	4.3	3.9

Median SHAP Value across cohorts 2018-2022

Feature	**2018**	**2019**	**2020**	**2021**	**2022**
Recent_Grad	0.256	0.157	0.051	0.159	0.167
Job_Hopper	0.209	0.02	-0.0255	-0.178	-0.186
Lived_in_multiple_countries	0.07	0.129	0.135	0.075	0.055
Marketing & Comms	0.057	0.089	0.136	0.152	0.146
1-5_years_experience	0.032	0.091	0.067	-0.0009	0.0004
6-10_Years_Experience	0.011	-0.002	0.0008	0.0011	0.0007
10+_years_experience	0.002	0.0001	0.0001	0.082	0.064
Top_VC_Experience	-0.003	-0.001	0.003	-0.058	-0.036
Career_Promotions	-0.004	-0.098	-0.042	-0.059	-0.037
Academic_Researcher	-0.005	-0.005	0.001	0.0001	-0.0002
Corporate_Researcher	-0.006	0.0008	-0.00001	-0.0004	-0.0004
VC_Experience	-0.035	-0.078	-0.048	-0.078	-0.061
PhD	-0.043	-0.067	-0.05	-0.069	-0.045
Top_Tier_Uni	-0.075	-0.08	-0.0028	-0.075	-0.059
Board_Member	-0.077	-0.099	-0.049	-0.081	-0.09
Big_Tech_Experience	-0.079	-0.081	-0.136	-0.089	-0.11
Multiple Degrees	-0.081	-0.083	-0.028	-0.064	-0.029
3+_Work_Experiences	-0.082	-0.101	-0.003	0.0007	0.0005
Exec_Experience	-0.09	-0.156	-0.161	-0.167	-0.189
First_Time_Founder	-0.11	-0.172	-0.21	-0.189	-0.203
Non_US_Uni	-0.241	-0.272	-0.24	-0.21	-0.22

Global Feature Importance based on Absolute Mean SHAP value across cohorts

Feature	2018	2019	2020	2021	2022
1-5_years_experience	15	14	9	19	20
10+_years_experience	21	20	15	10	9
3+_Work_Experiences	12	9	18	18	19
6-10_Years_Experience	16	18	20	17	17
Academic_Researcher	18	17	19	21	21
Big_Tech_Experience	9	10	5	4	4
Board_Member	10	8	12	9	8
Career_Promotions	17	7	8	15	15
Corporate_Researcher	19	21	21	20	18
Exec_Experience	5	3	4	3	3
First_Time_Founder	4	2	2	2	2
Job_Hopper	3	5	3	6	5
Lived_in_multiple_countries	6	6	6	8	10
Marketing & Comms	8	15	7	7	7
Multiple Degrees	7	11	11	13	13
Non_US_Uni	1	1	1	1	1
PhD	13	16	13	14	14
Recent_Grad	2	4	14	5	6
Top_Tier_Uni	11	12	17	12	12
Top_VC_Experience	20	19	16	16	16
VC_Experience	14	13	10	11	11

Appendix L: AI Usage Report

Throughout the development of this thesis, several AI tools were used in a controlled, transparent, and narrowly defined manner. Generative AI models—including ChatGPT (GPT-5.1), Claude, and Gemini 3—were employed exclusively for linguistic refinement, while Grammarly and Google Docs spell-check were used continuously to identify grammatical errors, spelling issues, and minor stylistic inconsistencies. None of these tools were used to generate theoretical ideas, empirical insights, methodological logic, or interpretations of predictive-model results. Their role was confined to improving clarity, conciseness, and terminological consistency. All AI-generated suggestions were critically reviewed and incorporated only when fully aligned with the authors' intended meaning and analytical reasoning.

GPT-5.1, Claude, and Gemini were applied in three strictly limited ways. First, they assisted in rephrasing individual sentences or short passages to improve readability without altering substantive content. For example, during revision of the methodological section describing the construction of training and prediction datasets, GPT-5.1 was used to refine dense passages developed through multiple restructuring rounds. One such passage began: "For the 2021 prediction dataset, the process begins by collecting all ventures founded in 2021 from Specter, PitchBook, and Crunchbase…." GPT-5.1 produced stylistic alternatives that tightened sequencing, clarified exclusion criteria, and reduced redundancy, while preserving technical accuracy. Second, generative models were used to provide formal synonyms or alternative formulations of specific methodological terms to ensure precision and consistency across chapters. When standardizing terminology in the Theory section, a prompt was used to request alternatives to the term "predictive stability," accompanied by the sentence "Predictive stability is essential for determining whether founder signals retain explanatory value across previously unseen cohorts." AI-generated suggestions such as "temporal robustness" and "out-of-sample consistency" were incorporated selectively when they fit the conceptual intent. Third, generative AI supported refinement of terminology where definitional accuracy mattered for core constructs such as "human-capital signals," "configurational effects," and "temporal generalization." When refining phrasing in the Analytical Framework section, a prompt requested alternatives to "temporal holdout design," using the sentence "Predictive modeling requires temporally valid datasets…" as context. Suggestions emphasizing "forward-consistent sampling" or "founding-year-anchored generalization tests" were integrated only when conceptually appropriate. In all cases, outputs were treated strictly as optional stylistic proposals. No conceptual mechanisms, no theoretical insights, and no interpretations of empirical findings were generated by AI; all substantive reasoning—including the formulation of predictions, feature-engineering logic, temporal holdout design, SHAP-based interpretation, and theoretical contributions—remains entirely the authors' work.

AI tools also improved academic prose by removing redundancies, smoothing transitions, and aiding consistency of terminology. For example, when refining phrasing around theoretical constructs, prompts such as "Propose alternative formulations for the term 'predictive stability' suitable for an academic methodology section" were used. Suggestions were adopted only when they enhanced clarity without altering theoretical content. Likewise, GPT-5.1 and Claude were occasionally used to generate synonyms for formal expressions (e.g., "hereinafter") to ensure stylistic coherence, again under strict manual oversight. Gemini 3 was additionally used in a verificatory role to double-check computational steps such as descriptive summary statistics and Python logic in the data-processing pipeline, ensuring that code execution aligned with the methodological descriptions. Importantly, Gemini did not participate in statistical interpretation or model evaluation; its role was limited to verifying implementation accuracy.

In isolated cases, GPT-5.1 and Claude were used to refine wording for example-based explanations—for instance, phrasing short illustrative descriptions in the SHAP-interpretation section. Even there, the exam-

ples themselves were entirely derived from empirical results; AI assisted only with smoothing the surrounding language. Across the entire thesis, AI tools never generated content related to predictive-model design, theoretical argumentation, feature logic, policy-induction interpretation, or empirical conclusions.

Using AI tools surfaced several risks and led to purposeful mitigation strategies. The primary risk was distortion of analytical meaning: generative models occasionally proposed stylistic revisions that subtly shifted theoretical or methodological intent. This was mitigated by manually reviewing every suggestion and rejecting those that altered substance. A second risk was fabricated or ungrounded content. To avoid this, AI was never used to generate theoretical ideas, empirical claims, or interpretations; all such content was author-produced. A third risk concerned over-reliance on AI phrasing, potentially diluting the authors' academic voice. To address this, AI use was restricted to sentence-level refinement, with no paragraph- or section-level generation allowed. Finally, because LLMs may introduce stylistic homogeneity, all outputs were critically evaluated for tone consistency and adjusted where necessary.

Several insights emerged from using AI tools in the thesis-writing process. First, generative AI is highly effective at micro-editing—clarifying complex methodological descriptions, reducing redundancy, and enhancing conceptual readability—but it is unsuitable for macro-level reasoning or theory development. Second, the tools were particularly helpful in maintaining terminological precision across a thesis that spans predictive modeling, signaling theory, and configurational analysis. Third, the process underscored that AI is most valuable when used as a linguistic assistant under close human supervision, not as a source of conceptual content. Ultimately, the experience demonstrated that, when applied responsibly, AI can enhance the communicative quality of academic writing without compromising originality, theoretical rigor, or analytical integrity.