Predicting startup success with Artificial Intelligence

KatchLabs is a Toronto-based research lab operating out of York University. We are researching and creating AI-powered tools that analyze and score ventures so investors, accelerators, and entrepreneurs can make better investment decisions. Our team combines extensive knowledge in venture funding, investor decision making, and machine learning to bring forth innovative solutions that aim to disrupt the investment and entrepreneur landscape.

We intend to use venture scoring models and heuristics, such as the CFA, developed by Dr. Andrew Maxwell, which boasts a successful track record of ranking over 20,000 ventures, and being used worldwide in the private sector, academic insitatutes, and governmental organizations.

This purpose of this whitepaper is to summarize and establish the background and foundation of our research, while giving the reader context and insight into our research goal and methodologies.

Predicting venture success

“Startup success” is a loosely defined term that does not have an agreed upon definition among entrepreneurs and investors. To begin, we must perhaps firstly define what “success” means generally. Oxford Dictionary breaks down success into two types: countable and uncountable[1]. While uncountable success is subjective and related to the initial aim or goal of the subject aiming at success, countable success is defined as “a person or a thing that has achieved a good results and been successful”. One group of researchers[2] defined “startup success”, or rather, *good results*, as a startup that has raised venture funding, or managed to generate 100,000 EUROS per anum. That figure has been realized through the surveying of dozens of venture capital groups and hundreds of entrepreneurs. Although I personally believe that this number is quite low, we can accept it as a “minimum viable proof of success” for the basis of my work.

In the case of Dr. Andrew Maxwell’s PhD thesis, and award winning paper, *Business Angel Early Stage Decision Making*[3], success meant receiving a funding offer from a group of business angels on the popular TV show called Dragons’ Den (comparable to Shark Tank in the US). This research identified that *“BAs focus on eight critical venture criteria that inform the assessment of*

---

[1] [success noun | Oxford Advanced American Dictionary]()
[2] [Econometric Estimation of the Factors That Influence Startup Success]()
[3] [Business angel early stage decision making - ScienceDirect]()

the investment return and resident risk component of the investment risk". Dr. Maxwell's model was accurately able to predict the business angels' decision roughly 90% of the time. More on the eight critical factors can be found at the Canadian Innovation Center's Study Proposal, The Critical Factor Assessment: *Planning for Venture Success*[4]

The Decision Making Process of Venture Funding

In Malcom Gladwell's second book Blink: The Power of Thinking Without Thinking, he highlights various academic studies that reveal that lack of understanding within human decision making, as well as the power of unconscious thought. The paper *Gender Differences in Mate Selection: Evidence From a Speed Dating Experiment*[5] brings forth an incredible finding - individuals' decision making process, and their explanation of their own decision making process often did not match. Gladwell then argues that humans often do not truly understand what drives their decision making.

Almost all venture funding decision making is still done by humans, though some work has been done in identifying and bridging the gap between decision making that can be done by humans, versus decisions that can be made by AI.

Research such as Finding the Unicorn: Predicting Early Stage Startup Success through a Hybrid Intelligence Method[6], How Do Venture Capitalists Make Decisions?[7], and others provides good framework for my work to identify potential areas of improvement, where decision making is still best handled by humans (Soft Signal). See Table 1, taken from the aforementioned paper. A potential area of improvement is to perhaps create a model where initial judgement is done by a human, and is codified for AI analysis, thus making it both a Soft and Hard signal.

[4] Critical Factor Assessment: Planning for Venture Success
[5] GENDER DIFFERENCES IN MATE SELECTION: EVIDENCE FROM A SPEED DATING EXPERIMENT* The choice of a marriage partner is one of the mo
[6] https://arxiv.org/ftp/arxiv/papers/2105/2105.03360.pdf
[7] How Do Venture Capitalists Make Decisions?

Input: Taxonomy of Signals

Category	Signals	Reference Examples	Hard Signal	Soft Signal	Machine Input Metric	Crowd-readable Format
Meta	B2B vs B2C	Böhm et al. (2017)	■		• Categorical (Type)	Textual
	Industry	Hoenig and Henkel (2015)	■		• Categorical (Type of industry)	Textual
	Firm Age	Zahra et al. (2003)	■		• Numeric (# of years)	Numeric
	Business Model DNA	Böhm et al. (2017)	■		• Categorical (Cluster=Freemium Platforms/Experience Crowd Users/Long Tail Subscribers/Affiliate Markets/Mass Customizing Orchestrators/Innovative Platforms/E-Commercer/Add-On Layers/Crowdsourcing Platforms/Customized Layers/Hidden Revenue Markets)	Textual
Value Proposition	Product Innovativeness	Li (2001); Nambisan (2016)		■		Graphic and Textual
	Technological Hype	Maxwell et al. (2011)	■		• Categorical (Phase in Gartner Hypecycle)	Graphic and Textual
	Proof of Concept	Maxwell et al. (2011);	■		• Dummy (Existence of prototype=yes/no)	Textual
	Scalability	Huang et al. (2017); Nambisan (2016)		■		Textual
Market	Competition	Landström (1998); Shi et al. (2016)	■		• Numeric (# of competitors) • Numeric (MEAN proximity to competitor based on Shi et al. (2016))	Graphic and Textual
	Revenue Model	Kirsch et al. (2009); Afuah and Tucci (2001); Böhm et al. (2017)	■		• Categorical (Type of revenue model=commission-based/fee-for-service/advertising/subscription/referral/production/mark-up based/other)	Textual
Resources	Capital Raised	Baum and Silverman (2004)	■		• Numeric (Amount of capital raised in USD)	Numeric
	Team Size	Mason and Stark (2002); Kirsch et al. (2009)	■		• Numeric (# of team members)	Numeric
	Team Constellation	Kirsch et al. (2009)	■	■	• Numeric (# of field background)	Textual
	Entrepreneurial Experience	Franke et al. (2006); Maxwell et al. (2011)	■		• Dummy (Previous founded ventures=yes/no)	Textual
	Entrepreneurial Vision	Sudek (2006); Huang and Pearce (2015)		■		Textual
	Entrepreneurial Education	Baum and Silverman (2004); Franke et al. (2006); Maxwell et al. (2011)	■		• Categorical (Level of education=high-school, bachelor, master, PhD)	Graphic and Textual
Commitment of 3. Party Support	Knowledge Support	Maxwell et al. (2011); Song et al. (2008); Zahra and Bogner (2000)	■		• Categorical (Type of support=incubator/accelerator/business angel/university)	Textual
	Financial Support	Baum and Silverman (2004)	■		• Categorical (Type of funding=state support/equity backed/bank financing)	Textual
	Proof of Value	Shepherd and Zacharakis (1999); Mason and Stark (2002)	■	■	• Numeric (# of pilot customers)	Numeric and Textual
	Web Analytics	Silver (2012)	■	■	• Website visits/average duration/backlinks/bounce rate	Numeric and Textual
	Social Media Analytics	Silver (2012)	■	■	• Twitter follower/# of tweets/sentiment of tweets	

Table 1. Taxonomy of Signals for Prediction Input

Our new research and development inteds to use existing research, such as *Econometric Estimation of the Factors That Influence Startup Success*[8], as well as much of Dr. Maxwell's successful research and commercialized work (see Canadian Center for Innovation, used by Canada's IRAP over 20,000 times) when designing new decision making tools, particularly when codifying investment factors into ML-ready data.

Additional areas of concern and improvement that we hope to advance with our research include diversity, equity, and inclusion. Papers such as:

What you are is what you like—similarity biases in venture capitalists' evaluations of start-up teams[9]

How gender biases drive venture capital decision-making: exploring the gender funding gap[10]

Highlight important findings of similarity bias that we may unconsciously have, such as men, who make up 93% of the VC space, preferring to invest in men-led startups. We hope that robust AI models look at the hard facts of venture success, and that those facts overlook identity, personality, or any other characteristics that are not important to the success of a venture.

Lastly, it is important to note that with globalization and the increasing access to resources and information, the number of variables and complexities when it comes to venture funding go up. It is becoming increasingly difficult for venture capitalists and business angels to make decisions, and it is of utmost importance to develop heuristics and AI models that are able to filter information faster and more effectively.

---

[8] Econometric Estimation of the Factors That Influence Startup Success

[9] What you are is what you like—similarity biases in venture capitalists' evaluations of start-up teams - ScienceDirect

[10] How gender biases drive venture capital decision-making: exploring the gender funding gap | Emerald Insight

Artificial Intelligence Research

Existing research of various degrees of success has paved the way to early findings of what methodology might work and others that definitely will not. Some points that researches must consider when developing an ML model for scoring startups include:

- Data diversity - ensuring data is not biased by industry, decision making process, geography, or other influences. This can be mitigated through the use of various data sources to train a model
- Age - to be able to study the full scope of a startup lifecycle, ideal data sources must include studied and recorded ventures from inception, venture funding, and an exit event, such as a merger, acquisition, or IPO
- Decision making model - there is no one standard method for evaluating ventures. Methodologies, culture, intrinsic beliefs, background knowledge, and subjective decision making result in a varied landscape of data depending on the business angel or venture capital group. As a result, most existing research is limited, and often cannot be commercially adapted beyond a niche use for the environment under which the data was studied

In the paper *Startup Investment Decision Support: Application of Venture Capital Scorecards Using Machine Learning Approaches*[11], a team of researchers were able to predict which startups would receive venture funding from a VC group with a 78% accuracy. While using a scorecard evaluation method, the researched model's top two influential and predicting features matched traditional, non-machine learning research[12]. In this case, "Team Management" and "Planning Strategy" were the two most important decision making features that influenced the model's predictions.

These promising results inspired us to continue building on top of existing insights and findings to develop our own ML models, especially combining them with traditional decision making research, namely, the Critical Factor Assessment.

---

[11] Startup Investment Decision Support: Application of Venture Capital Scorecards Using Machine Learning Approaches
[12] How Do Venture Capitalists Make Decisions?

Our Research in Action

Critical Factors Assessment

"The CFA Snapshot is a simplified version of our flagship Critical Factor Assessment (CFA) service. The CFA Full Assessment is an assessment of your venture by experts against the 42 criteria that our research shows are important factors affecting venture success."[13]

An important pillar of our research involves the collection of CFA data for model training. We have began collecting data from startups seeking venture funding through a partner channel. The simplified version of the CFA used for this data collection can be seen below:

1. Feature & Benefits - Does your proposed product or service offer performance advantages compared to currently deployed solutions?

2. Readiness - How far away are you from being able to deliver completed products or services to your first revenue customer?

3. Barrier To Entry - What is unique or patentable about your product that represents a barrier to entry for potential competitors?

4. Adoption - Can you demonstrate that customers in your target market will purchase your product or service when it is available?

5. Supply Chain - Can you provide confirmation that there are no success barriers either about your supply chain, or distribution channel?

6. Market Size - Is the overall size of the market and your likely market share, sufficient to generate the envisaged revenues? Further, is the overall market forecast to be large enough to be interesting?

7. Entrepreneur Experience - Do you or members of your team have any direct or relevant (entrepreneurial, industry, business) experience that can be directly applied to the challenges facing this business?

8. Financial Expectations - Do your financial projections present a persuasive argument that your company can achieve cash-flow neutrality, based on your own investment, money you can borrow, and money you can raise from external investors?

---

[13] [CFA Snapshot | Canadian Innovation Centre]

Reaching for New Horizons

While much of ML research relies on codified data entered by humans, our lab aims to analyze ventures' presentations and pitch decks, that often influence the investment decision making process. The ability to storytell and sell the vision of a company often goes far beyond numbers and facts that can be presented on a balance sheet or found in a startup's data room.

We believe that with the power of Optical Character Recognition, Natural Language Process, and Sentiment Analysis, we can begin to pave the way for a more robust venture scoring that more closely emulates that of the "Golden Standard", A.K.A human analysis.

Our Team

Yan Katcharovski
Research Lead

Yan Katcharovski is a serial entrepreneur, tech geek and marketing visionary. He is the co-founder and CTO of Schoolio, a Google Accelerator alumnus. He is also a partner at VirtusVC, a Toronto-based venture-capital firm. He writes for Entrepreneur Magazine about startups, technology, psychology and entrepreneurship. He is currently pursuing a Masters of Engineering from York University, and has a Bachelor of Computer Science.

Dr. Andrew Maxwell
Research Supervisor

Andrew is the Bergeron Chair In Technology Entrepreneurship, and Director Bergeron Entrepreneurs in Science and Technology. He received his Ph.D. in Technology Entrepreneurship from the University of Waterloo in 2011, winning the Academy of Management's Heizer Award for the top PhD in his field. He is also a journal editor for the Journal of Business Venturing. While at Waterloo, Andrew taught the capstone technology entrepreneurship class, helping numerous technology entrepreneurs get their start. Prior to this Andrew worked for three years in the technology transfer office of the University of Toronto (also teaching at Rotman and UTM). Andrew's work experience includes founding four technology companies and working in two technology multinationals. He has an MBA from London Business School, and a B.Sc. (Eng.) in Electrical Engineering, from Imperial College London.

Dr. James Elder
Supervising Committee

James Elder is a member of the Centre for Vision Research and a professor in the Departments of Electrical Engineering and Computer Science and Psychology at York University, and is appointed as a joint Lassonde School of Engineering/Faculty of Health Tier I Chair. Dr. James Elder received his PhD in Electrical Engineering from McGill University in 1996. His research interests include the development of novel and useful computer vision algorithms and machine vision systems through a better understanding of visual processing in biological systems.Professor Elder's research has won a number of awards and honours, including the Premier's Research Excellence Award and the Young Investigator Award from the Canadian Image Processing and Pattern Recognition Society.

Ron Tal
Research Advisor

Ron Tal is machine learning expert, boasting senior positions in WalmartLabs, Uber, Coinbase, and CloudTrucks. Ron holds a Masters of Engineering from York University, where he conducted extensive research on computer vision and 3D modelling.

Next Steps

Join our journey of shaping the future of machine learning, venture success prediction, and investment decision making. We are seeking:

1. Business partners seeking to get involved in research & development of AI

2. Commercialization partners seeking to integrate our tools in their business

3. Academic and research partners, including individuals, labs, or institutions

Contact us:

yank@yorku.ca
https://www.yankatch.com/machine-learning