

CHAPTER 1

Introduction

Venture Capital (VC) is financial capital provided to early-stage, high-potential, high-growth companies (startups). VC firms have funded many successful companies, such as Google, Apple, Microsoft and Alibaba. Unlike investors in the public markets, VC firms often take a more active role in managing their investments, providing expertise and advice in both managerial and technical areas. VC firms have two primary roles: as scouts, able to identify the potential of new startups, and as coaches, able to help startups realise that potential [1]. In these ways, VC is often critical to the success of startups, and therefore the commercialisation of new technologies more generally. Adoption of the Internet and inexpensive, ubiquitous computing has transformed the VC industry: companies require less funding to launch but more to scale in highly competitive markets [2]. There is now an impetus for VC firms to change the way they operate.

VC firms face the challenge of choosing a few outstanding investments from a sea of hundreds of thousands of potential opportunities. VC firms seek to make investments in companies that can provide a liquidity event that returns many times their investment value within the time frame of their fund. For startups, a liquidity event (also referred to as an ‘exit’) is either an Initial Public Offering (IPO) or an acquisition by a larger competitor. Most VC firms expect their investments to exit within 3-6 years, in accordance with their fund time-frame [3]. When compared to public market investors, this is a long-term investment strategy. However, few companies are capable of maturing from early-stage to exit at this pace. In addition, traditional metrics of performance (e.g. cash-flow, earnings) often do not exist or are unclear [4]. In summary, VC firms must select from a field of many investment candidates, where little information is available on each of them, and only a few will grow at a fast enough rate to be worthwhile – this is why the VC investment screening process is considered difficult.

The VC industry requires better systems and processes to efficiently manage labour-intensive tasks like investment origination and screening. Currently, investment opportunities are either referred or identified through technology scans (e.g. Google searches, patent searches). These manual search processes are time-

consuming for VC firms. Attempts in the literature to solve this problem have three common limitations: small sample size [5, 6, 7, 8, 9, 10, 11, 12], a focus on early-stage investment [13, 5, 14, 15, 12, 16], and a basic feature set [5, 10, 14, 12, 11, 6]. Although individual studies address some of these limitations, none synthesise the findings into software ready for use in industry. The popularity of online databases like AngelList and CrunchBase, which offer information on startups, investments and investors, is evidence of the VC industry’s desire for better, more quantitative methods of assessing startup potential [17]. There is preliminary evidence that mining these data sources may address previous limitations and make VC investment screening more efficient and effective [16, 18].

We believe it is now possible to address previous limitations in this field and produce an improved VC investment screening system. Our system aims to identify startup companies that are likely to raise additional funding, become acquired or have an IPO (or some combination thereof) in a given period of time. This system may assist VC firms to efficiently and effectively screen investment candidates. To be useful in this context, the system must meet the following criteria:

1. **Practicality.** The system must be practical for use in the VC industry and more efficient to use than manual investment screening. The system should be designed to operate with minimal user input and no assumed technical expertise. The system should also be designed to run in reasonable time.
2. **Robustness.** The system must be robust to changes over time. The system should be designed to have minimal variance in performance when training on datasets from different times so investors can trust its ability to make future-looking predictions. The system should also be designed to adapt to the quantum and type of data available from data sources over time.
3. **Versatility.** The system must be versatile in its ability to address a large domain of investment prediction tasks. The system should be designed to make accurate predictions for companies of different developmental stages (e.g. Seed, Series A), for different target outcomes (e.g. Acquisition, IPO) across different forecast windows (e.g. exit in two years, exit in four years).

This thesis is organised as follows:

- **Chapter 2: Literature Review.** We review the theoretical background of startup performance and VC investment and evaluate previous efforts in developing technologies for use in VC investment screening systems.

- Chapter 3: Design. We outline the design of our system architecture, propose a framework that guides our feature and data source selection, and describe the development of our adaptive classification pipeline process.
- Chapter 4: Evaluation. We perform a series of experiments to evaluate our system against three criteria: practicality, robustness and versatility.
- Chapter 5: Discussion. We discuss the merits and limitations of our project and their implications for investors and future research.

Bibliography

- [1] BAUM, J. A., AND SILVERMAN, B. S. 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 (2004), 411–436.
- [2] GRAHAM, P. *Startup Investing Trends*. <http://www.paulgraham.com/invtrend.html/>. Online; accessed 15 May 2017. June 2013.
- [3] GOMPERS, P. A. Optimal investment, monitoring, and the staging of venture capital. *The journal of finance* 50, 5 (1995), 1461–1489.
- [4] SHANE, S., AND CABLE, D. Network ties, reputation, and the financing of new ventures. *Management Science* 48, 3 (2002), 364–381.
- [5] AHLERS, G. K., ET AL. Signaling in equity crowdfunding. *Entrepreneurship Theory and Practice* 39, 4 (2015), 955–980.
- [6] GIMMON, E., AND LEVIE, J. Founder’s human capital, external investment, and the survival of new high-technology ventures. *Research Policy* 39, 9 (2010), 1214–1226.
- [7] DIXON, M., AND CHONG, J. A Bayesian approach to ranking private companies based on predictive indicators. *AI Communications* 27, 2 (2014), 173–188.
- [8] HOENEN, S., ET AL. The diminishing signaling value of patents between early rounds of venture capital financing. *Research Policy* 43, 6 (2014), 956–989.
- [9] YU, Y., AND PEROTTI, V. Startup Tribes: Social Network Ties that Support Success in New Firms. In: *Proceedings of 21st Americas Conference on Information Systems*. 2015.
- [10] AN, J., JUNG, W., AND KIM, H.-W. A Green Flag over Mobile Industry Start-Ups: Human Capital and Past Investors as Investment Signals. In: *PACIS 2015 Proceedings*. AIS Electronic Library, 2015.
- [11] WERTH, J. C., AND BOEERT, P. Co-investment networks of business angels and the performance of their start-up investments. *International Journal of Entrepreneurial Venturing* 5, 3 (2013), 240–256.

- [12] CROCE, A., GUERINI, M., AND UGHETTO, E. Angel Financing and the Performance of High-Tech Start-Ups. *Journal of Small Business Management* (2016).
- [13] BECKWITH, J. Predicting Success in Equity Crowdfunding. Unpublished thesis. Joseph Wharton Research Scholars. Available at http://repository.upenn.edu/joseph_wharton_scholars/25. 2016.
- [14] CHENG, M., ET AL. Collection, exploration and analysis of crowdfunding social networks. In: *Proceedings of the Third International Workshop on Exploratory Search in Databases and the Web*. ACM. 2016, 25–30.
- [15] YUAN, H., LAU, R. Y., AND XU, W. The determinants of crowdfunding success: A semantic text analytics approach. *Decision Support Systems* 91 (2016), 67–76.
- [16] STONE, T. R. Computational analytics for venture finance. Unpublished Ph.D. dissertation. UCL (University College London), 2014.
- [17] PATIL, A. *CrunchBase’s Venture Program Members Are Making Startup Data Better Than Ever*. Ed. by CRUNCHBASE.COM. <https://info.crunchbase.com/2015/01/crunchbases-venture-program-members-are-making-startup-data-better-than-ever/>. Online; accessed 18 05 2015. Jan. 2015.
- [18] BHAT, H., AND ZAELIT, D. Predicting private company exits using qualitative data. In: *Advances in Knowledge Discovery and Data Mining*. Ed. by HUANG, J., CAO, L., AND SRIVASTAVA, J. Vol. 6634. Lecture Notes in Computer Science. Springer, Berlin, 2011, 399–410.