Hacking the Venture Industry: An Early-Stage Startups Investment Framework for Data-Driven Investors

Short Abstract

Investing in early-stage companies is incredibly hard, especially when no data are available to support the decision process. Venture capitalists often rely on gut feeling or heuristics to reach a decision, which is biased and potentially harmful. This work proposes a new data-driven framework to help investors being more effective in selecting companies with a higher probability of success. We built upon existing research and augment it with some further analysis to eventually compile a smart checklist of 21 relevant features that should be considered when evaluating a new deal. The framework represents, therefore, a new additional instrument in the investors' toolbox.

Longer Abstract

This study wants to be an attempt to unify the existing knowledge on how to effectively perform due diligence on early-stage companies and uses analytics to augment some well-known results.

The generality of the framework makes it a powerful tool for VC investors to assess and evaluate companies in the initial stages of their lives. Its generality also makes it partially incomplete, because companies belonging to specific industries, adopting specific business models or using particular technologies may need further adjustments and variables to be taken in considerations. In fact, we believe the framework achieves a good balance between being general enough to allow investors to set a baseline for companies that deserve investments and at the same time gives them room to improve it and personalize it at wish.

The ESI Framework does not want to be, in fact, the only tool an investor uses in her decision-making process, but rather the first test for a potential investment. Moreover, we believe that every VC can select the risk-return profile that better fits her needs and strategy (and thus the corresponding score), but also that it might be possible to find an optimal threshold above which a company can be funded.

However, in addition to be useful to single investors, the framework introduces opportunities for the entire industry: new superfunds are making the work of growth funds extremely hard, competitive, and less remunerative. This may push funds to change their strategies and look for companies at earlier stages before institution like the Vision Fund come in. The framework can, in fact, facilitate this transition toward earlier-stage investments.

Machine learning was not also set up here to fully replace the job of any VC investor, but rather to augment her skills. Humans still have a relevant impact in what makes a business successful and in how to spot one that might become the next big hit, and we imagine that further studies may show the importance of creating hybrid human-machine models (Dellermann et al., 2017).

The rationale for this work was then demystifying funding clichés and helping investors seeing past extreme risks, relying less on their gut feel and more on hard data. Gut-feel is, in fact, an *intuiting process* (Huang, 2019) that can make you or break you (Huang and Pearce, 2015), and this is why having data supporting your initial thoughts is essential when it comes to investing in early-stage companies. In this paper, we decided to optimize the probability of success and identifying the determinants of success for a startup. As mentioned in the introductory paragraphs, this is only one of the uses of machine learning in the venture space and of course does not guarantee that an investor gets access to the best deals. In fact, even having a perfect way to predict the success of a company,

investors need to convince the best entrepreneurs to partner and work together, because money is most of the time only part of the additional value brought to the table.

This also creates a second interesting future avenue of research, which we previously identified to be a problem for the industry. We know that VCs influence the growth of a company (Croce et al., 2013) much more than banks (Cole et al, 2016), and the same is partially true for corporate venture capitals (Colombo and Murtinu; 2017), but we do not to what extent and in which way. It is of course incredibly hard to disentangle the different effects (Bertoni et al., 2011; Chemmanur et al., 2011; Croce et al., 2013), but investigating further the difference and the magnitude of the "*sorting* (or selection) effect" (cherry-picking skills) and "*treatment* effect" (providing better value-adding services) represents a future possible step (Sorensen, 2007).

Another interesting stream of research may interest the calculation of the optimal time-to-exit for a company. Intuitively, a VC would prefer the shortest holding period for any investment, but this is not true at all and actually holding periods shorter than three years might be harmful (Capizzi, 2015). This could be intrinsically connected to the probability of success of a company, so it may deserve further attention.

We are also aware that betting on the horse could be more rewarding than betting on the jockey (Kaplan et al., 2009). We believe our framework shows that people and team characteristics are extremely relevant (even more than pure business features) at the initial stage of development of a company, but we also intuitively think there comes a point where this may change. The study of this discontinuity, which corresponds also to the point where the valuation methods jump from "real option" to "NPV" (see Figure 1 and 3), has not been investigated yet and could provide investors with new insights.

The importance of the ecosystem is also paramount and usually not investigated in depth. Hong et al. (2018) show theoretically and proves empirically that a more competitive VC market increases the likelihood of successful exits for low-quality projects and lowers the one for good companies. This implies that an ecosystem might indeed affect the probability of a company to be successful much more than traditional variables, and it might deserve additional attention. Moreover, it is known that public perception of a company (i.e., the words used to describe it in the news) can be used to augment the predictive power of the current methods (Xiang et al., 2012), although a causation study has not yet been performed on the topic.

Creating a structured model that emulates the investment decision-making process of a VC is also of interest. Many studies have tried different approaches, such as fuzzy theory (Lin, 2009; Afful-Dadzie and Afful-Dadzie, 2016; Zhang, 2012; Aouni et al., 2014; Afful-Dadzie et al., 2015; Tian et al., 2018), goal programming (Aouni et al., 2013; Colapinto and La Torre, 2015), or Probabilistic Latent Factor model (Zhong et al., 2016), but the connection between those theoretical models and the probability of identifying a successful company remains still unexplored.

Finally, a strong component which is still weak is the psychological one. We believe founders' (and VCs) personality influence the likelihood of starting a company and then make it a success (Kessler et al., 2012), and we are aware that some of the variables included in the framework are proxy of some personality trait. However, a more detailed analysis would bring more complete and useful information that could be used to understand whether the founder is a good fit for that specific project at that specific time. This concept could even be stretched up to the point of trying to predict the right set of skills, personality traits, and inclinations that could make a person a better founder (Ng and Stuart, 2016; Levine and Rubinstein, 2016).