smartCVC: Data Science Meets Corporate Venture Capital

TREO Talk Paper

Ruiyun Xu
City University of Hong Kong
ruiyunxu2-c@my.cityu.edu.hk

Hailiang Chen
City University of Hong Kong
hailchen@cityu.edu.hk

J. Leon Zhao
City University of Hong Kong
jlzhao@cityu.edu.hk

Abstract

Corporate venture capital (CVC) is a practice where large non-financial firms (i.e., corporate investors) invest directly in private entrepreneurial companies (i.e., portfolio companies). CVC has experienced a strong global growth over the past decades and has drawn increasing attention in academic literature. According to the Global CVC report by CBInsights, the global investments in corporate venture programs reached a historic high in 2017, up to $31.2 billion across 1,791 deals. As a critical first step for effective CVC investment, the selection of appropriate portfolio companies is very challenging and difficult due to the high level of uncertainty about potential financial and strategic benefits. In current CVC practices, the selection of target portfolio companies is usually done by manual screening. The selection quality heavily depends on the experience and selection capability of the CVC program. To the best of our knowledge, no prior research has addressed the problem of target selection in CVC investment through a data science approach.

Although many Mergers and Acquisitions (M&A) prediction models have been developed to facilitate the target selection in M&A activities, the differences between CVC investments and M&A activities make these models not directly applicable in this context. Unlike M&A, which results in a legal consolidation of two entities into one entity and is therefore an exclusive transaction between two firms, CVC investment allows multiple enterprises to invest in the same portfolio company. CVC is also different from typical VC investment, because CVCs pursue both strategic goals and financial returns for their parent companies, while the sole investment goal of VC firms is to achieve high financial returns.

The primary goal of this study is to apply data science to develop a prediction model, smartCVC, to facilitate CVC investments by identifying a list of highly promising investees from a large pool of portfolio companies for a CVC investor. The business implications of solving this problem are enormous and can lead to substantial financial gains and improved operational efficiency for CVCs. Portfolio companies can also benefit from the suggestions provided by smartCVC as they can proactively reach out to potential CVCs to secure additional funding and maximize the strategic fit with their CVCs.

Several key features of smartCVC are described as follows. First, we apply topic models on the textual descriptions of investors/companies and construct a fine-grained business proximity measure to assess the closeness between two companies in the space of products, technologies, and markets. Second, we construct a novel feature—similar firms' investment choices—to tap into the wisdom of the CVC community by leveraging the aggregate opinion of all CVC investors in the market regarding whether to invest in a portfolio company or not. This novel feature relies on the history of prior CVC investments and gives higher weights to the decisions made by more similar CVCs. Third, we develop a strategic alignment feature to quantify the fit between corporate objectives and the business of the portfolio company based on the theory of strategic consistency. In addition, we include standard firm characteristics (e.g., number of employees, number of prior funding rounds etc.) into the model.

Experimental results on the CrunchBase dataset show that the proposed model significantly outperforms the benchmark method. The evaluation measures are calculated by comparing the suggested list of potential investees with the real investment choices of corporate investors on a hold-out sample.