# Annotated Bibliography

## [2008-goodson](C:\\Users\\mark\\OneDrive\\Documents\\University\\Year Four\\CITS4001 Thesis\\honours\\submissions\\honours2-proposal-original\\literature\\2008-goodson.pdf)

**Citation:** Goodson, Robert B., et al. "Entity performance analysis engines." U.S. Patent No. 8,554,709. 8 Oct. 2013.

**Abstract:** A system and method for measuring a performance of an entity and for predicting its future performance is disclosed. Raw information about the entity is collected from internal and external sources. The information is cleaned to exclude false positives. Using performance categories, the information is organized and transformed into meaningful data for the performance analysis engine. The information is normalized by scaling the meaningful data by industry type. A performance score is calculated by the performance analysis engine based on the normalized data. Further, a competitive relationship score is calculated based on the performance score and the normalized data. These scores are reported in a user interface displaying the performance of the entity and are used as inputs, among other factors, to a predictive analysis engine that assesses the future performance of the entities.

**Notes:**

## [2009-block\_sandner](file:///C:\Users\mark\OneDrive\Documents\University\Year%20Four\CITS4001%20Thesis\honours\submissions\honours2-proposal-original\literature\2009-block_sandner.pdf)

**Citation:** Block, Joern, and Philipp Sandner. "What is the effect of the financial crisis on venture capital financing? Empirical evidence from US Internet start-ups." *Venture Capital* 11.4 (2009): 295-309.

**Abstract:** Employing a large dataset of venture capital investments in US Internet firms, we analyse the effect of the current financial crisis on the venture capital market. Using regression analysis, we find that the financial crisis is associated with a 20% decrease in the average amount of funds raised per funding round. This effect, however, can only be found in later funding rounds. We argue that firms in later financing rounds that need capital to survive cannot avoid a deduction induced by the financial crisis, whereas firms that seek initial funding postpone their funding and expansion plans until the capital markets have stabilized. Furthermore, firms in later phases of the venture cycle are more likely to be negatively affected by the weak IPO market than firms seeking initial funding. Our results suggest that the financial crisis can lead to a severe ‘funding gap’ in the financing of technological development and innovation.

**Notes:**

*Validation of CrunchBase dataset:* Block and Sandner (2009) tested the quality of information in CrunchBase by comparing the funding data in the US tech industry with the industry statistics published by the National Venture Capital Association (NVCA, a trade association representing the US VC industry). They found that the number of deals in the CrunchBase figures amounts to about 97% of the internet-related deals as reported by NVCA and estimated a high and statistically significant Pearson correlation coefficient between the time series of new deals related to the two sources (r=0.67; p<0.05).

## [2011-alexy](file:///C:\Users\mark\OneDrive\Documents\University\Year%20Four\CITS4001%20Thesis\honours\submissions\honours2-proposal-original\literature\2011-alexy.pdf)

**Citation:** Alexy, Oliver T., et al. "Social capital of venture capitalists and start-up funding." *Small Business Economics* 39.4 (2012): 835-851.

**Abstract:** How does the social capital of venture capitalists (VCs) affect the funding of start-ups? By building on the rich social capital literature, we hypothesize a positive effect of VCs’ social capital, derived from past syndication, on the amount of money that start-ups receive. Specifically, we argue that both structural and relational aspects of VCs’ social networks provide VCs with superior access to information about current investment objects and opportunities to leverage them in the future, increasing their willingness to invest in these firms. Our empirical results, derived from a novel dataset containing more than 1,500 first funding rounds in the Internet and IT sector, strongly confirm our hypotheses. We discuss the implications of our findings for theories of venture capital and entrepreneurship, showing that the role and effect of VCs’ social capital on start-up firms may be more complex than previously argued in the literature.

**Notes:**

## [2011-martens](file:///C:\Users\mark\OneDrive\Documents\University\Year%20Four\CITS4001%20Thesis\honours\submissions\honours2-proposal-original\literature\2011-martens.pdf)

**Citation:** Martens, David, et al. "Identifying financially successful start-up profiles with data mining." *Expert Systems with Applications* 38.5 (2011): 5794-5800.

**Abstract:** Start-ups are crucial in the modern economy as they provide dynamism and growth. Research on the performance of new ventures increasingly investigates initial resources as determinants of success. Initial resources are said to be important because they imprint the firm at start-up, limit its strategic choices, and continue to impact its performance in the long run. The purpose of this paper is to identify configurations of initial resource bundles, strategy and environment that lead to superior performance in start-ups. To date, interdependencies between resources on the one hand and between resources, strategy and environment on the other hand have been neglected in empirical research. We rely on data mining for the analysis because it accounts for premises of configurational theory, including reversed causality, intra-dimensional interactions, multidimensional dependencies, and equifinality. We apply advanced data mining techniques, in the form of rule extraction from non-linear support vector machines, to induce accurate and comprehensible configurations of resource bundles, strategy and environment. We base our analysis on an extensive survey among 218 Flemish start-ups. Our experiments indicate the good performance of rule extraction technique ALBA. Finally, for comprehensibility, intuitiveness and implementation reasons, the tree is transformed into a decision table.

**Notes:**

*Machine learning techniques:* SVM, C4.5, ALBA (best)



## [2012-liang\_yuan](file:///C:\Users\mark\OneDrive\Documents\University\Year%20Four\CITS4001%20Thesis\honours\submissions\honours2-proposal-original\literature\2012-liang_yuan.pdf)

**Citation:** Eugene, Liang Yuxian, and Soe-Tsyr Daphne Yuan. "Where's the Money? The Social Behavior of Investors in Facebook's Small World." *Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012)*. IEEE Computer Society, 2012.

**Abstract:** Are investing activities dependent on social relationships? In our research, we apply social network analysis to the field of investing behaviors and discover that investors have a tendency to invest in companies that are socially similar to them. While traditional studies on investing behavior tend to focus on factors like psychology, opinions, investing experience etc, they fail to consider social relationship as an important factor. In this paper we provide general rules of thumb that are useful for companies seeking funding from investor. These rules of thumb are generated by analyzing the social relationships between investors and companies found within the small world of Facebook.

**Notes:**

## [2012-song](file:///C:\Users\mark\OneDrive\Documents\University\Year%20Four\CITS4001%20Thesis\honours\submissions\honours2-proposal-original\literature\2012-song.pdf)

**Citation:** Song, Yang. *Network of networks: Uncovering the secrets of entrepreneurs' networks*. 2012.

**Abstract:** The popularity and universal use of online social network sites such as LinkedIn, Facebook and Twitter has become an important phenomenon attracting research in various disciplines. Entrepreneurs also use social networks to access and acquire resources, with online social networks also providing many opportunities for entrepreneurs to share and organize knowledge through their contacts among these networks. The role and importance of social networks for entrepreneurs has been recognized in previous studies. However, the study of online social networking by entrepreneurs is just starting. This study is intended to fill a gap in the literature concerning the structure, characteristics and use of online social networks by entrepreneurs. We developed a novel approach to extract data on the online social networks of entrepreneurs through the use of the Application Programming Interfaces (APIs) of social network sites such as LinkedIn, Facebook and Twitter. The data concerning entrepreneurs’ profiles and network connections was entered into a MySQL database for further analysis to determine the characteristics of entrepreneurs’ online social networks such as size, structure, diversity and the role of these networks in the entrepreneurial process. Based on our findings concerning the structure of these networks, we also developed a simulation model to predict their contribution to entrepreneurial survival.

**Notes:**

## [2012-xiang](file:///C:\Users\mark\OneDrive\Documents\University\Year%20Four\CITS4001%20Thesis\honours\submissions\honours2-proposal-original\literature\2012-xiang.pdf)

**Citation:** Xiang, Guang, et al. "A Supervised Approach to Predict Company Acquisition with Factual and Topic Features Using Profiles and News Articles on TechCrunch." *ICWSM*. 2012.

**Abstract:** Merger and Acquisition (M&A) is a critical corporate strategy for companies to preserve their competitive advantages, and M&A prediction has been an interesting and challenging research topic in the past a few decades. However, past work has only adopted numerical features such as accounting, financial and market variables in building models, and yet the valuable textual information from the great variety of social media sites like news portals and microblogs, which discusses tech trends a lot and is potentially helpful for M&A prediction, has not been touched at all. To fully explore this information, we used the profiles and news articles for companies and people on CrunchBase, the largest public database for the tech world, which anybody can edit. Specifically, we explored topic features via topic modelling techniques, as well as a set of other novel features of our design within a machine learning framework. We conducted experiments of the largest scale in the literature, and our approach achieved a high true positive rate (TP) between 60% to 79:8% with a false positive rate (FP) mostly between 0% and 8:3% over categories with less missing attributes in the CrunchBase profiles.

**Notes:**

## [2013-liang\_yuan](file:///C:\Users\mark\OneDrive\Documents\University\Year%20Four\CITS4001%20Thesis\honours\submissions\honours2-proposal-original\literature\2013-liang_yuan.pdf)

**Citation:** Yuxian, Eugene Liang, and Soe-Tsyr Daphne Yuan. "Investors Are Social Animals: Predicting Investor Behaviour using Social Network Features via Supervised Learning Approach."

**Abstract:** What makes investors tick? In this paper, we explore the possibility that investors invest in companies based on social relationships be it positive or negative, similar or dissimilar. This is largely counter-intuitive compared to past research work. In our research, we find that investors are more likely to invest in a particular company if they have stronger social relationships in terms of closeness, be it direct or indirect. At the same time, if there are too many common neighbours between investors and companies, an investor is less likely to invest in such companies. We use social network features such as those mentioned to build a predictive model based on link prediction in which we attempt to predict investment behaviour.

**Notes:**

## [2013-park](file:///C:\Users\mark\OneDrive\Documents\University\Year%20Four\CITS4001%20Thesis\honours\submissions\honours2-proposal-original\literature\2013-park.pdf)

**Citation:** Park, Hyunseok, Janghyeok Yoon, and Kwangsoo Kim. "Identification and evaluation of corporations for merger and acquisition strategies using patent information and text mining." *Scientometrics* 97.3 (2013): 883-909.

**Abstract:** This paper proposes a framework to identify and evaluate companies from the technological perspective to support merger and acquisition (M&A) target selection decision- making. This employed a text mining-based patent map approach to identify companies which can fulfil a specific strategic purpose of M&A for enhancing technological capabilities. The patent map is the visualized technological landscape of a technology industry by using technological proximities among patents, so companies which closely related to the strategic purpose can be identified. To evaluate the technological aspects of the identified companies, we provide the patent indexes that evaluate both current and future technological capabilities and potential technology synergies between acquiring and acquired companies. Furthermore, because the proposed method evaluates potential targets from the overall corporate perspective and the specific strategic perspectives simultaneously, more robust and meaningful result can be obtained than when only one perspective is considered. Thus, the proposed framework can suggest the appropriate target companies that fulfil the strategic purpose of M&A for enhancing technological capabilities. For the verification of the framework, we provide an empirical study using patent data related to flexible display technology.

**Notes:**

## [2013-spiegel](file:///C:\Users\mark\OneDrive\Documents\University\Year%20Four\CITS4001%20Thesis\honours\submissions\honours2-proposal-original\literature\2013-spiegel.pdf)

**Citation:** Spiegel, Olav, et al. "Going it all alone in web entrepreneurship? A comparison of single founders vs. co-founders." *Proceedings of the 2013 annual conference on Computers and people research*. ACM, 2013.

**Abstract:** There is a considerable amount of entrepreneurial activity in the information technology (IT) industry, especially in the Web. Many claim that the founders are the single most critical factor in startup success. The entrepreneurial self-efficacy theory and the jack-of-all-trades theory suggest that founders with diverse skills and broad experience will be able to successfully create a business alone. Alternatively, social capital theory suggests that founders who have high social capital and who engage in social networking will be able to go alone. Others, however, claim that team building is crucial for startup success as to bringing complementary skills together. In this study, we investigate if having a diverse skill set, higher experience, or more social connections is indeed a determinant of single founders, while co-founders have more specialized (yet complementary) skills. Our results are derived from analysis of 91 Web startups and their 183 (co-) founders. Contradicting existing theory, we could not find determinants for starting a new business alone. However, we find that co-founders do indeed complement each other’s skills. Our results suggest that existing entrepreneurial theory needs to be expanded and revised to accommodate for the contradictions found and that future research is required in this area.

**Notes:**

## 2013-werth\_boeert (requested)

**Citation:** Werth, Jochen Christian, and Patrick Boeert. "Co-investment networks of business angels and the performance of their start-up investments." *International Journal of Entrepreneurial Venturing* 1 5.3 (2013): 240-256.

**Abstract:** The venture capital literature has established the positive impact of co-investment networks on the performance of start-up investments. In early stages, however, often angel financing is the primary source of external equity. Using a novel in-depth dataset of US high technology start-ups, we investigate the effects of business angel networks. Start-ups of better connected angel investors are more likely to receive subsequent funding by venture capitalists and business angels more often exit successfully. Thereby, angel investors seem to rely on their direct contacts, whereas their network position and possibility to act as information brokers plays a far smaller role.

**Notes:**

## [2014-huang\_shi](file:///C:\Users\mark\OneDrive\Documents\University\Year%20Four\CITS4001%20Thesis\honours\submissions\honours2-proposal-original\literature\2014-huang_shi.pdf)

**Citation:** Huang, Jun, and Shi Zhan. "With a Little Help of My (Former) Employer: Past Employment and Entrepreneurs' External Financing." *Academy of Management Proceedings*. Vol. 2015. No. 1. Academy of Management, 2015.

**Abstract:** Securing external financing and attracting value-adding investors are challenging for nascent businesses, largely due to the concern of information asymmetry. We explore how entrepreneurs’ past employment mitigates this concern. Based on the rationale of inherited human and social capital, we hypothesize that: a) for startup founders, their former employers’ prominence signals favourably about the quality of the new ventures, thus increasing the chances of securing funding and attracting prominent investors; b) this signalling is more relevant if the startup’s business is similar to that of the former employers. Empirical tests using a novel dataset of technology companies support our hypotheses. This paper highlights the role of past employment in matching entrepreneurs with prominent investors. Methodologically, we introduce to the management literature a topic-modelling-based measure of business similarity.

**Notes:**

## [2014-shi](file:///C:\Users\mark\OneDrive\Documents\University\Year%20Four\CITS4001%20Thesis\honours\submissions\honours2-proposal-original\literature\2014-shi.pdf)

**Citation:** Shi, Zhan, Gene Moo Lee, and Andrew B. Whinston. "Towards a better measure of business proximity: topic modeling for analyzing M&As." *Proceedings of the fifteenth ACM conference on Economics and computation*. ACM, 2014.

**Abstract:** In this article, we propose a new measure of firms’ dyadic business proximity. Specifically, we analyse the unstructured texts that describe firms’ businesses using the natural language processing technique of topic modelling, and develop a novel business proximity measure based on the output. When compared with the existent methods, our approach provides finer granularity on quantifying firms’ similarity in the spaces of product, market, and technology. We then show our measure’s effectiveness through an empirical analysis using a unique dataset of recent mergers and acquisitions in the U.S. high technology industry. Building upon the literature, our model relates the likelihood of matching of two firms in a merger or acquisition transaction to their business proximity and other characteristics. We particularly employ a class of statistical network analysis methods called exponential random graph models to accommodate the relational nature of the data.

**Notes:**

## [2015-an](file:///C:\Users\mark\OneDrive\Documents\University\Year%20Four\CITS4001%20Thesis\honours\submissions\honours2-proposal-original\literature\2015-an.pdf)

**Citation:** An, Jungkook; Jung, Woojin; and Kim, Hee-Woong, "A Green Flag over Mobile Industry Start-Ups: Human Capital and Past Investors as Investment Signals" (2015). *PACIS 2015 Proceedings*

**Abstract:** Crowdfunding and online start-up platforms are becoming important communication tools for startups and investors. Existing literatures on online start-up platforms usually focus on reward-based crowdfunding platform, which do not offer any equity to backers. In addition, there have not been many empirical researches about equity-based crowdfunding due to the novelty of the regulation. This study analyses the association between funding amount and early stage start-ups’ underlying characteristics, the type of past investors, and influence of investors in the context of equity-based crowdfunding. The distinction of our research is the aspect of approach that we use population data from online start-up platform for the mobile industry. We find that start-up’s funding outcome is positively related to start-up’s human capital and pure investors. Moreover, our study extends theoretical understanding of the importance of human capital and past investors in start-up, and also contributes to the entrepreneurship literature by examining creditable signals for early stage start-up investment.

**Notes:**

## [2015-huang](file:///C:\Users\mark\OneDrive\Documents\University\Year%20Four\CITS4001%20Thesis\honours\submissions\honours2-proposal-original\literature\2015-huang.pdf)

**Citation:** Huang, Lu, et al. "Identifying target for technology mergers and acquisitions using patent information and semantic analysis." 2015 *Portland International Conference on Management of Engineering and Technology (PICMET).* IEEE, 2015.

**Abstract:** Technology plays an increasingly important role in today’s enterprise competition. Technology mergers and acquisitions (Tech M&A), as an effective way to acquire external technology resources rapidly, have attracted attention from researchers because of their potential realization of value through synergy. A big challenge that faces corporate managers and government policy makers is how to identify the appropriate target to support effective technology integration. In this study, we develop a model of target selection of Tech M&A from the perspective of technology relatedness and R&D capability. We present the results relating to M&A in the field of cloud computing in China.

**Notes:**

## [2015-perotti\_yu](file:///C:\Users\mark\OneDrive\Documents\University\Year%20Four\CITS4001%20Thesis\honours\submissions\honours2-proposal-original\literature\2015-perotti.pdf)

**Citation:** Perotti, Victor, and Yang Yu. "Startup Tribes: Social Network Ties that Support Success in New Firms." (2015).

**Abstract:** To answer the question how networking plays a role in entrepreneurial success, we have shown through the lens of social network analysis that online social network structure could be considered as efficient indicators to predict startups overall performance. Social network analysis techniques are applied on a big network, which is generated according the following and followers structure on Twitter.com for 644 ICT operating startup companies. A set of network structure related measures are developed to link with overall company performance. The main findings include the better the startups are connected within the startup community “tribe” and occupying the better “position” within the community, the more successful they are. In addition, no matter within or outside the community, the structure of one’s friends is closely related to the business success also.

**Notes:**

## [2015-tata](file:///C:\Users\mark\OneDrive\Documents\University\Year%20Four\CITS4001%20Thesis\honours\submissions\honours2-proposal-original\literature\2015-tata.pdf)

**Citation:** Tata, Amulya, Daniella Laureiro Martinez, and Stefano Brusoni. "Looking Backward and Forward? Understanding the Role of Temporal Focus on Startup Performance."

**Abstract:** In this study we look at the relationship of perception of time and self with performance in the context of startups. By relying on solid constructs from psychology, namely temporal focus (the degree to which individuals characteristically devote attention to perceptions of the past, present, and future), construals (future events can be construed in a higher-level construal, thus forming abstract representations or low level construals, which are more concrete and specific) and self-categorization which can be (collective or individualistic) we can predict startup performance. Through novel datasets such as CrunchBase and Twitter we analyse perceptions of 2111 startup founders and 620 startup teams. Twitter data provides fine-grained, longitudinal, spontaneous and conversational type of data. From content-analysing this data, we find high past focus, low future focus, high collective self-categorization and lower level construal to be associated with better performance. These results hold true at the team level, demonstrating the importance of perceptional biases in new venture performance.

**Notes:**

## [2015-zhao](file:///C:\Users\mark\OneDrive\Documents\University\Year%20Four\CITS4001%20Thesis\honours\submissions\honours2-proposal-original\literature\2015-zhao.pdf)

**Citation:** Zhao, Xiaoxue, Weinan Zhang, and Jun Wang. "Risk-Hedged Venture Capital Investment Recommendation." *Proceedings of the 9th ACM Conference on Recommender Systems*. ACM, 2015.

**Abstract:** With the increasing accessibility of transactional data in venture finance, venture capital firms (VCs) face great challenges in developing quantitative tools to identify new investment opportunities. Recommendation techniques have the possibility of helping VCs making data-driven investment decisions by providing an automatic screening process of a large number of startups across different domains on the basis of their past investment data. A previous study has shown the potential advantage of using collaborative filtering to catch and predict the VCs' investment behaviours [17]. However, two fundamental challenges in venture finance make conventional recommendation techniques difficult to apply. First, risk factors should be cautiously considered when making investments: for a potential startup, a VC needs to specifically estimate how well this new investment can fit into its holding investment portfolio in such a way that investment risk can be hedged. Second, the investment behaviours are much sparser than conventional recommendation applications and a VC's investments are usually limited to a few industry categories, making it impossible to use a topic-diversification method to hedge the risk. In this paper, we solve the startup recommendation problem from a risk management perspective. We propose 5 risk-aware startup selection and ranking algorithms to catch the VCs' investment behaviours and predict their new investments. Apart from the contribution on the new risk-aware recommendation model, our experiments on the collected CrunchBase dataset show significant performance improvements over strong baselines.

**Notes:**

## [2016-beckwith](file:///C:\Users\mark\OneDrive\Documents\University\Year%20Four\CITS4001%20Thesis\honours\submissions\honours2-proposal-original\literature\2016-beckwith.pdf)

**Citation:** Beckwith, John Jack. "Predicting Success in Equity Crowdfunding." (2016).

**Abstract:** Equity crowdfunding is an increasingly popular means of raising capital for early stage startups. It enables entrepreneurs to finance their companies with smaller contributions from a variety of people. This paper studies the relationship between the characteristics of a given company and its ability to raise funds on an equity crowdfunding platform. A series of statistical and machine learning models are fit to data from a U.S.-based equity crowdfunding website, including a logistic regression, a CART decision tree, a naïve Bayes classifier, and a support vector machine. This study demonstrates that a connection exists between the probability of a company’s crowdfunding success and its previous funding history, Twitter presence, media buzz, size, location, and its founders’ educational backgrounds. As a whole, however, the classification quality of the various models leaves something to be desired. This suggests the need for additional data inputs and more longitudinal research in the field of equity crowdfunding.

**Notes:**

*Missing values:* Multiple imputation with chained equations, as detailed in White, Royston, Wood (2010), proved unable to effectively fill in these missing values.

*Statistic techniques:* A variety of statistical and machine learning models are employed to understand which factors about a company best predict the outcome of its equity crowdfunding campaign. The models considered—logistic regression, a CART decision tree, a naïve Bayes classifier, and a support vector machine—are all suitable to a classification-type problem.

*Feature selection:* “Stability selection” is the chosen method, as outlined in Meinshausen and Bühlmann (2009).

*Feature list:* Previous round of funding (binary) – CrunchBase, Company profile on Twitter (binary) – Twitter, Number of previous press mentions (count) – AngelList, Number of mentions in TechCrunch (count) – AngelList, Number of founders (count) – AngelList, Number of employees between 11-50 (binary) – AngelList, Located in San Francisco (binary) – AngelList, At least one founder has an MBA (binary) – LinkedIn, At least one founder attended a top 20 US university (binary) – LinkedIn, Closing number for the S&P 500 on the day prior to the company’s crowdfunding campaign launch (count) – Yahoo! Finance

*Outcome variable:* Raised funding in an AngelList crowdfunding campaign (binary)

*Potential improvements:* Artificial neural networks, Restricted Bolzmann machines

## [2016-croce](C://Users/mark/OneDrive/Documents/University/Year%20Four/CITS4001%20Thesis/honours/submissions/honours2-proposal-original/literature/2016-croce.pdf)

**Citation:** Croce, Annalisa, Massimiliano Guerini, and Elisa Ughetto. "Angel Financing and the Performance of High Tech Start Ups." *Journal of Small Business Management* (2016).

**Abstract:** In this paper, we investigate what drives the performance of high-tech start-ups receiving angel financing, while taking a closer look at the capabilities (i.e., experience) and investment behaviour of business angels (BAs). We exploit a new data set (extracted from CrunchBase), which consists of 1,933 high-tech start-ups that received at least one financing round from a BA. The results indicate that the experience of BAs in early stage investments is positively associated with additional receipt of follow-on rounds of financing and sequential capital injections from venture capitalists (VCs). Later-stage experience is positively associated with the start-up's success (i.e., probability to be listed or acquired), but reduces the need for new VCs to invest in the start-up. Furthermore, we find consistent evidence that start-ups that combine BA and VC financing experience higher levels of funding amounts, additional VC financing, and an improved likelihood of success. Finally, we find that the co-localization of BA investors and start-ups in the same area facilitates the attraction of VC financing.

**Notes:**

## 2016-liang\_yuan (requested)

**Citation:** Liang, Yuxian Eugene, and Soe-Tsyr Daphne Yuan. "Predicting investor funding behaviour using CrunchBase social network features." *Internet Research* 26.1 (2016): 74-100.

**Abstract:** What makes investors tick? Largely counter-intuitive to past research, this study explores the possibility that investors invest in companies based on social relationships. The purpose of this paper is to build a social network graph using data from CrunchBase, the largest public database with profiles about companies. The authors combine social network analysis with the study of investing behaviour in order to explore how similarity between investors and companies affects investing behaviour through social network analysis. This study crawls and analyses data from CrunchBase and builds a social network graph which includes people, companies, social links and funding investment links. The problem is then formalized as a link (or relationship) prediction task in a social network to model and predict (across various machine learning methods and evaluation metrics) whether an investor will create a link to a company in the social network. Various link prediction techniques such as common neighbours, shortest path, Jaccard Coefficient and others are integrated to provide a holistic view of a social network and provide useful insights as to how a pair of nodes may be related (i.e., whether the investor will invest in the particular company at a time) within the social network. This study finds that funding investors are more likely to invest in a particular company if they have a stronger social relationship in terms of closeness, be it direct or indirect. At the same time, if investors and companies share too many common neighbours, investors are less likely to invest in such companies. The author’s study is among the first to use data from the largest public company profile database of CrunchBase as a social network for research purposes. The authors also identify certain social relationship factors that can help prescribe the investor funding behaviour. The authors’ prediction strategy based on these factors and modelling it as a link prediction problem generally works well across the most prominent learning algorithms and perform well in terms of aggregate performance as well as individual industries. In other words, this study would like to encourage companies to focus on social relationship factors in addition to other factors when seeking external funding investments.

**Notes:**

## [2016-yuan](file:///C:\Users\mark\OneDrive\Documents\University\Year%20Four\CITS4001%20Thesis\honours\submissions\honours2-proposal-original\literature\2016-yuan.pdf)

**Citation:** H. Yuan, et al., The determinants of crowdfunding success: A semantic text analytics approach, *Decision Support Systems* (2016).

**Abstract:** In the era of the Social Web, crowdfunding has become an increasingly more important channel for entrepreneurs to raise funds from the crowd to support their startup projects. Previous studies examined various factors such as project goals, project durations, and categories of projects that might influence the outcomes of the fund raising campaigns. However, textual information of projects has rarely been studied for analysing crowdfunding successes. The main contribution of our research work is the design of a novel text analytics-based framework that can extract latent semantics from the textual descriptions of projects to predict the fund raising outcomes of these projects. More specifically, we develop the Domain-Constraint Latent Dirichlet Allocation (DC-LDA) topic model for effective extraction of topical features from texts. Based on two real-world crowdfunding datasets, our experimental results reveal that the proposed framework outperforms a classical LDA-based method in predicting fund raising success by an average of 11% in terms of F1 score. The managerial implication of our research is that entrepreneurs can apply the proposed methodology to identify the most influential topical features embedded in project descriptions, and hence to better promote their projects and improving the chance of raising sufficient funds for their projects.

**Notes:**