

# Discussion on Application of Big Data in Venture Capital

Jin Kang

Shandong Shuchuang Community Science and Technology Service Co., Ltd., Jinan, Shandong, 250000, China

## Abstract

This paper starts from the problems faced by the venture capital industry, combines the latest policies of the big data industry and the venture capital industry, and analyzes the causes of the problems through the aspects of business cost, risk control and project exit. By elaborating the role of big data on venture capital, the application of big data in venture capital is introduced, and the application direction of big data is described in detail in four stages including fund raising stage, project investment stage, post-investment management stage and project exit stage, focusing on the whole chain of venture capital "solicitation, investment, management and withdrawal". Cloud computing, OCR recognition technology, Spiders acquisition technology, machine learning, neural network and natural language processing technology are pointed out, and relevant application examples are given. Finally, the industry trend outlook is carried out.

## Keywords

big data; venture capital; investment decision

# 大数据在创业投资中的应用探讨

康金

山东数创共同体科技服务有限公司, 中国·山东·济南 250000

## 摘要

论文从创业投资行业面临的问题出发, 结合大数据行业和创业投资行业最新政策, 通过对业务成本、风险控制和项目退出等方面分析产生问题的原因。通过阐述大数据对创业投资的作用, 引出大数据在创业投资中的应用, 从而围绕创业投资“募投管退”全链条, 从资金募集阶段、项目投资阶段、投后管理阶段和项目退出阶段等四个阶段详细描述了大数据应用的方向, 指出采用的云计算、OCR识别技术、Spiders采集技术、机器学习、神经网络和自然语言处理技术等相关技术, 并进行相关应用举例说明。最后进行行业趋势展望。

## 关键词

大数据; 创业投资; 投资决策

## 1 引言

创业投资, 俗称风险投资, 是指向创业企业进行股权投资, 以期所投资创业企业发展到一定阶段后通过转让股权等方式获得资本增值收益的投资方式。创业投资, 风险与收益并存, 其中, 创业企业多数处在早期, 甚至个别处在种子期, 发展过程中存在着巨大的不确定性, 风险较大; 同时, 投资这类企业一旦成功, 收益颇丰。

近年来大数据行业兴起, 越来越多地应用到各个行业中, 大数据在创业投资中的应用也越来越受到重视。大数据、云计算、人工智能、区块链等技术正在颠覆金融业, 越来越多的新兴技术被应用于投资过程, 创业投资也受到影响<sup>[1]</sup>。2024年6月, 国务院办公厅印发了《促进创业投资高质量发展

发展的若干政策措施》, 这将促进未来中国创业投资市场的进一步活跃。利用大数据及其相关技术, 提升效率、降低风险和追求收益, 是创业投资行业的重要关注点。

## 2 创业投资行业面临的问题

市场上专门从事创业投资业务的公司, 一般称为创业投资机构。当前, 创业投资行业普遍存在业务成本较高、项目风险较大、项目退出较难等问题。

### 2.1 业务成本较高

多年来, 创业投资机构人员成本普遍较高。除此之外, 业务流程较长, 带来业务成本居高不下。业务流程通常包括以下步骤:

①寻找投资标的: 创业投资机构通过行业研究确定投资领域, 通过同行推荐、项目路演等多种渠道寻找有潜力的项目标的。

②项目筛选: 创业投资机构会对收集到的项目进行初步筛选, 判断项目是否符合投资要求, 以及是否具有市场

【作者简介】康金(1985-), 男, 中国山东新泰人, 硕士, 中级经济师, 从事大数据与金融、金融科技、金融风险模型研究。

前景。

③走访调研：通过初审筛选的项目，创业投资机构会与标的企业主要负责人进行面谈，并对企业进行初步走访调研，收集企业相关材料，了解企业的团队、业务模式、市场情况等。

④内部立项：经过面谈和走访调研，创业投资机构召开项目立项会，对项目进行初步评审，决定是否将该项目立项跟进。

⑤项目估值：创业投资机构会根据标的企业的财务状况、市场前景、竞争格局、团队能力等多个方面，以确定标的企业的估值和投资金额。

⑥项目谈判：确定投资金额和估值后，创业投资机构会与标的企业主要负责人（一般为实际控制人）进行初步谈判，双方商定投资条款内容。

⑦尽职调查：谈判完成后，达成一致后，创业投资机构会进行尽职调查，对企业的团队、财务、业务、法律等方面进行全面的审查、分析与核实，以期全面了解企业情况。

⑧投资决策：尽职调查完成后，创业投资机构会形成详尽的尽职调查报告，连同其他相关材料，报创业投资机构投资决策部门进行评审。

⑨签署投资协议：通过投资决策的项目，创业投资机构与标的企业签署投资协议，打款，工商变更，完成投资交易。

⑩投后管理：投资完成后，创业投资机构会积极协助企业的管理和运营，为企业提供更多的资源，多方面赋能企业发展，帮助企业快速成长。

⑪项目退出。创业投资机构一般会在两到三年后会通过股权转让、IPO等方式择机退出，实现投资回报。

不同创业投资机构的投资流程可能略有不同，但基本如此。每个环节都需要付出较大的精力和时间，一个项目往往仅是前期跟踪就有数月甚至一到两年，一个完整的项目往往需要好几年，这也是业务成本较高的原因。

## 2.2 项目风险较大

创业投资项目存在较大不确定性，成功率较低，往往十个项目难成其一。创投项目风险是指在创业投资过程中，由于政策、市场以及企业内部管理等各种不确定性因素而导致投资失败、收益较低甚至损失的风险。常见的创业投资风险有：

①政策风险：政策风险指的是因国家宏观政策（如行业政策、财政政策、地区发展政策等）发生变化带来的风险。

②市场风险：由于市场需求波动、竞争环境等变化，导致商品销售波动，影响收益和回报。

③技术风险：由于技术发展与创新，面临技术缺陷、技术过时、技术被替代等风险。

④管理风险：企业管理者的管理能力和经验不足，难以应对各种挑战和问题。

⑤财务风险：企业面临资金不足、资金链断裂等财务风险，导致企业无法维持正常运转。

⑥法律风险：企业在生产经营过程中可能面临知识产权、合同纠纷、法律诉讼等法律风险。

## 2.3 项目退出较难

创业投资机构最理想的退出方式就是标的企业上市并度过限售期，择期一次性或者分批有序退出。然而，上述情况属实较少。通常是标的公司未能如期上市或者发展未达预期，此时创投基金就面临退出难的问题，大多通过股权转让、股权溢价回购以及清算等方式退出。创业机构投资的底层资产一般为非上市企业股权，较股票二级市场，流动性较差，退出较为复杂，周期较长。

## 3 大数据对创业投资的作用

①大数据能够帮助创业投资机构提高效率、降低成本。

投资机构需要收集项目所有的相关数据，还需要对这些数据进行深度分析。大数据技术在数据采集、整理、加工和分析方面具有无可比拟的优势，通过对来自各种渠道的数据进行整合和分析，能够高效快捷得出更准确的结论和决策。通过大数据，可以大大提高效率、降低人员成本。

②大数据能够帮助创业投资机构辅助决策。

通过收集和分析大量的市场数据、行业数据和企业数据，投资机构可以更方便地了解市场趋势、行业发展情况以及潜在投资项目的价值，从而提高决策水平和效率。

③大数据能够帮助创业投资者提高风险管理水平。

大数据技术可以通过收集和分析大量的政策信息、市场数据、行业数据、经济指标、财务数据等，通过分析各数据之间的逻辑勾稽关系，提示各种异常情况，帮助创业投资机构提升风险管理水平。

## 4 大数据在创业投资中的应用

目前，国内创业投资机构多数是私募基金管理公司，围绕创业投资“募投管退”全链条开展创业投资活动。

### 4.1 资金募集阶段

①筛选客户群体。在守法合规，不侵犯客户隐私的情况下，创业投资机构可以通过网络调查问卷、线上直播等多种方式和渠道采集客户在投资、消费、理财等多方面的数据，运用大数据手段，形成客户精准画像，为实施精准营销提供支持。

②客户精准营销。基于不同客户群体和个人，通过大数据分析，制定个性化营销方案，提高客户转化率。目前，人工智能技术在该领域的主要应用场景是挖掘与客户数据强相关的向量，从而实现个性化推荐。

### 4.2 项目投资阶段

#### 4.2.1 行业研究

通过大量信息检索、采集市场数据、行业数据和企业数据，通过大数据结合机器学习、自然语言处理和知识图谱

等技术,可以帮助投资机构了解市场趋势、行业发展情况以及潜在投资机会。自然语言处理技术可以运用于加工提取相关行业新闻、政策法规、媒体热点等中的非结构化数据,多维度分析数据。通过挖掘分析数据间的潜在关系,加以线性和非线性分析,全方位立体化构建建立知识图谱,助力创业投资机构对相关信息进行研究,大大提高研究效率,同时也避免了过去传统分析中过度依赖分析人员水平和受制于数据样本容量有限带来的不足。例如,美国风险投资公司 Connetic,通过自动化预筛选平台 Wendal,不仅降低了私募股权投资的筛选成本,提高了筛选效率,也有效降低了私募股权投资中的偏见<sup>[2]</sup>。在图像数据采集、整理过程中,过去只能采用人工识别后录入、整理,极大浪费人力,通达大数据结合 OCR 识别技术,可以快速高精度识别图像数据,有效解决了数据分析中的图标相关数据的分析。利用已经设定的模型公式、参数和重要提示性数值,可以实现普通数据和图片数据等各类数据的自动采集、整理、加工和分析,甚至可以实时动态展现数据情况,生成动态图表,大大提高投资行业研究的分析效率、准确度,同时也降低了成本。

#### 4.2.2 尽职调查

在尽职调查阶段,通过人工收集的材料,可以使用 OCR 技术,直接输入数据系统,快速完成数据的录入,导入系统分析工具;通过 Spiders 采集标的企业的公开信息和数据,进行加工整理后一并导入系统;通过 API 接口的形式,接入第三方大数据,查询标的企业的基本信息、主要人员、知识产权、法律纠纷、财务数据、纳税数据、社会信用等一系列信息。通过系统模型处理,可以快速输出尽职调查中所需的信息,大大提高尽职调查的效率,提升尽职调查报告的质量,能够更加全面展现标的企业的情况,为创业投资机构决策提供必要依据。

#### 4.2.3 反欺诈

此外,通过上述数据结合标的企业提供的数据进行对比分析,可以快速发现数据之间存在的差异和不合理性,为反欺诈调查提供精准的数据支持。

### 4.3 投后管理阶段

#### 4.3.1 运营情况监测

利用大数据及其相关技术,可以实现跨部门数据及信息的共用共享,通过明确数据管理职责、完善数据治理机制、推广数据管理标准等措施,在技术驱动下实现了多维数据的互联互通,打破以往存在的数据壁垒和信息孤岛,构建以大

数据为中心的业务协同系统<sup>[3]</sup>,运用到项目和业务管理中,解决了过去流程复杂、数据掌握不全面等问题。通过设定数据标签,引入机器学习、神经网络和自然语言处理技术,构建各个数据链接的关系图谱,自动匹配相关产能数据之间的联系,可以有效提高系统的产能分析能力,便于找出生产经营过程中存在的短板与不足,为改进措施提供参考依据。

#### 4.3.2 风险监测预警

风险控制是创业投资机构的工作重心,风险控制能力的高低直接决定了项目的成败,甚至关系到整个创业投资机构的生死存亡。当前,创业投资风控体系逐渐趋于成熟,但风控人员的主观分析评价占居主导地位,因而风控人员的水平直接影响着创业投资机构的风险控制水平。客观上,宏观经济环境、行业波动、财务情况等多维度的信息往往掌控不足或不全,因而风险控制决策水平受到制约。通过大数据、人工智能等金融科技手段,引入云计算和大模型技术,将风控数据信息进行采集全部或大部分由机器完成,“采集—整理—加工—分析”一条龙,运用 AI 技术进行机器自主学习,划定风险管理的上下限阈值和警戒线,通过不间断收集互联网、第三方资讯数据源、社交媒体等渠道的数据,实现进行分析监测,及时、准确、智能化监控,大大提高风险监测能力。

### 4.4 项目退出阶段

投资的退出是创业投资业务流程的核心环节。怎么退,何时退直接关系到项目的收益大小。通过大数据及机器学习等相关技术,结合行业研究情况和标的企业实际情况,制定退出方案。利用数据的及时性和有效性,不断优化调整退出方案,筛选出有利退出时机,从而提升投资收益。

## 5 行业趋势展望

大数据技术在创业投资行业的应用会越来越多,甚至改变原有的业务运营模式和投资决策方式。大数据及其相关技术的发展,必将会给整个行业带来更多机遇与挑战,精通大数据技术与金融的复合型人才将更受欢迎<sup>[4]</sup>。

### 参考文献

- [1] 舒蕾.大数据技术在私募股权投资行业的应用及影响[J].科技金融,2022(4).
- [2] Eagle Alpha. “2021 Alternative Data Report: Year in Review”,2021.
- [3] 黄琦,赵大伟,代会琳,等.人工智能“火爆”背景下的金融安全“冷”思考[J].金融发展研究,2023(11):77-81.
- [4] 张乐.大数据技术在私募股权投资中的应用与革新[J].中国管理信息化,2024(2).