

# 基于CRRT的企业二元创新特征选择模型

刘俊迪

上海工程技术大学管理学院, 上海

收稿日期: 2023年4月2日; 录用日期: 2023年5月24日; 发布日期: 2023年5月31日

## 摘要

经济全球化环境中, 企业贯彻新发展理念要坚持以创新为核心的发展战略, 单方面依靠开发式创新或探索式创新均不是最优策略, 结合企业现状均衡进行二元创新更有利于企业发展。企业二元创新能力影响因素众多, 本文结合资源基础理论、二元理论等企业创新理论, 建立企业二元创新能力影响因素模型, 利用CRRT集成特征选择算法对企业二元创新能力影响因素进行重要性评估, 建立重要特征集合子空间。分析得出以下结论: 企业数字化转型程度成为影响探索式创新的最重要因素, 进行数字化转型是新格局下企业进行探索式创新的有效途径。开发式创新影响因素中人力资源相关特征较为重要, 与对探索式创新的影响相比人力资源的投入和数量优势对开发式创新影响更显著, 重要性程度更高。企业规模和资产、得到的政府补助以及高管薪酬则是企业发展二元创新需着重考虑的因素特征。

## 关键词

二元创新, 特征选择, 数字化转型

# Enterprise Ambidextrous Innovation Feature Selection Model Based on CRRT

Jundi Liu

School of Management, Shanghai University of Engineering Science, Shanghai

Received: Apr. 2<sup>nd</sup>, 2023; accepted: May 24<sup>th</sup>, 2023; published: May 31<sup>st</sup>, 2023

## Abstract

In the environment of economic globalization, enterprises should adhere to the development strategy with innovation as the core to implement the new development concept, unilaterally relying on development innovation or exploratory innovation is not the optimal strategy, combined with the current situation of the enterprise to carry out balanced dual innovation is more conducive to the development of enterprises. There are many influencing factors of enterprise ambi-

dextrous innovation ability, this paper combines resource base theory, dual theory and other enterprise innovation theories to establish a model of influencing factors of enterprise ambidextrous innovation ability, and uses CRRT integrated feature selection algorithm to evaluate the importance of influencing factors of enterprise ambidextrous innovation ability, and establishes an important feature set subspace. The analysis concludes that the degree of digital transformation of enterprises has become the most important factor affecting exploratory innovation, and digital transformation is an effective way for enterprises to carry out exploratory innovation in the new pattern. Among the influencing factors of developmental innovation, the characteristics of human resources are more important, and the investment and quantitative advantages of human resources have a more significant impact on developmental innovation and have a higher degree of importance than the impact on exploratory innovation. The size and assets of the enterprise, the government subsidies received, and the executive compensation are the factors that need to be considered in the development of enterprise development.

## Keywords

Ambidextrous Innovation, Feature Selection, Digital Transformation

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

创新是现代企业实现可持续发展的第一动力,关于企业双元创新的研究目前已有许多,相关理论及分析主要围绕影响因素、作用机制及作用效果展开,其中影响因素研究大多从两个角度即内部组织影响和外部环境切入。内部影响因素中,关于企业资源的研究最早提出,Jancennele等[1]将企业资源分为有形和无形两种,基于复杂性理论提出相对无形性即相对资源总和的无形资源份额是企业成功的关键指标。张馨等[2]的研究将人员结构引入创新影响因素分析中,提出科研人员异质性影响。在组织制度方面,股权激励制度的影响受到广泛认可,高层管理者激励[3]和企业员工激励[4]均对创新能力产生作用,员工结构[5]的调节作用也在激励过程中显现。此外,外部影响因素的分析也较为丰富。基本涵盖政策环境[6]、市场环境[7]、技术环境[8]。

综上,各视角下企业创新相关影响因素分析均比较深入透彻。但是,在诸多分析研究中,学者们多数利用双重差分模型或固定效应模型等进行单个变量直接影响分析或调节变量对比分析,较少学者将机器学习模型应用到双元创新影响因素分析中进行大规模变量综合分析,存在研究片面、主观性较强等不足之处。本文利用机器学习特征工程中特征选择算法,建立企业双元创新影响因素模型,对双元创新相关的46个影响因素进行重要性程度综合评估分析,最终得到了含22个重要特征集合空间,在保留高维数据预测准确性优势的同时减少了模型复杂性,提高了企业管理过程中数据价值,为企业双元创新能力提高起到一定的借鉴参考作用。

## 2. 模型建立

### 2.1. 样本选择与数据来源

本文选取2011~2021年中国沪深A股上市公司作为研究样本,初始样本6万余,企业专利及董事会、监事会和高管的个人信息等企业指标数据来自中国研究数据服务平台(CNRDS)、中国经济金融研究数据

平台(CSMAR)以及上市公司企业年报数据,城市层面的数据来自《中国城市统计年鉴》及各省市统计年鉴。对一些研究需要的数据库并未收录的信息,使用 Python 软件通过网页进行数据爬虫抓取,结合手动搜索补充。

## 2.2. 变量选取及计量

### 2.2.1. 双元创新变量

现有关于企业双元创新的研究中,探索式创新和开发式创新的衡量主要集中在专利衡量、投入衡量和文本搜索衡量三方面。专利衡量角度,根据我国专利法专利分为发明专利、实用新型、外观设计三类,学者一般根据专利内容侧重点及审批难度区分,使用侧重新技术新产品、申请难度大的发明专利衡量企业探索式创新,使用侧重技术完善、申请难度较低的实用新型与外观设计衡量开发式创新。投入衡量角度,根据创新时间点不同,企业研发投入可分为研究阶段投资和开发阶段投资,研究阶段创新难度大往往是对新产品的全新创造,符合探索式创新特点;开发阶段则通常是对现有产品及技术的改进完善,符合开发式创新特点。据此使用创新投入衡量双元创新时往往用研发活动的费用化支出衡量探索式创新,资本化支出衡量开发式创新。文本搜索衡量角度,方鑫等[9]对企业年报就分别进行探索式创新、开发式创新相关词汇统计,使用对应词汇频次和度量两类创新水平。

考虑到样本选择、数据获取等因素的影响,本文参考汤莉[10]研究,用专利衡量方式计算上市公司双元创新能力,同时考虑专利申请需要一定时间,使用延后一期的发明专利申请数量代表探索式创新能力,使用延后一期的实用新型与外观设计专利申请数量和代表开发式创新能力。在模型稳健性评估中,使用资本化创新投入和费用化创新投入分别代表探索式创新能力和开发式创新能力替换被解释变量。

### 2.2.2. 解释变量

根据资源基础理论、利益相关者理论、信息不对称理论、企业数字化转型等相关理论,本文按照企业内部影响因素、外部环境影响因素、数字化转型相关因素三大类总结归纳学者们之前的研究,并加入数字经济指数指标,构建企业双元创新能力影响指标体系。

1) 企业内部影响因素。企业内部影响因素细分为基础属性、人员状况、运营状况、治理状况。基于企业创新主体特征及创新行为流程,在基础属性中考虑了企业规模、上市时长、企业年龄、生命周期四指标,在人员分布中考虑了研发人员数量、职工数量、职工学历、岗位分布四指标,运营状况从经济角度出发考虑企业利润、资产周转率、托宾 Q 值、营业收入四指标,在治理状况中考虑了内部控制指数、监管制度、激励制度、决策制度四指标。其中:

企业生命周期。企业生命周期变量参考尹闪等人研究使用企业经营、投资、筹资三维现金流量分类方法计算,计算结果分为增长期、成熟期、衰退期三个阶段,使用哑变量表示。

监管、激励、决策指标。借鉴 Sergey Anokhin [11]、顾乃康等[12]人研究。监管分为董事会监督作用和股权监督作用,董事会监督作用使用独立董事比例与董事会规模表示,股权监督作用使用机构持股比例、股权制衡度(二至五大股东持股比例之和/控股股东持股比例)表示。激励机制选用高管薪酬与高管持股比例来表示。决策主要考虑总经理决策权力,使用董事长职位与总经理职位是否两职合一来表示。

2) 外部环境影响因素。外部环境影响因素分为政策环境、行业环境和外部关注度。根据既有研究可知企业政策环境可使用企业所获得政府补助表示。行业环境根据企业所在省份的全要素生产率以及《中国区域科技创新评价报告 2022》中提到的区域知识创造效用、知识获取效用、企业创新效用、创新环境效用、创新绩效效用表示。外部关注度则使用个人关注度及媒体关注度表示,其中个人关注使用月均超额换手率表示,媒体关注使用总关注度、正面关注度、中性关注度、负面关注度表示,即新闻媒体中出现该企业的新闻总数、正面新闻数、中性新闻数、负面新闻数,该数据来源 CNRDS 数据库。

3) 数字化转型相关影响因素。数字化转型相关影响因素从企业角度和区域角度进行表示。其中企业数字化转型程度参考祁怀锦等[13]的研究,根据企业财务报告所披露的年末无形资产明细项中与数字化转型相关的部分占无形资产总额的比例来度量。即通过查找含有“数字化”“智能网络”“管理系统”“智能平台”等与数字化转型技术相关的关键词的项目或专利,将其定为“数字化技术无形资产”,计算企业当年度所有数字化技术无形资产占本年度无形资产的比例,即为企业数字化转型程度。区域数字经济发展水平主要参考潘为华等[14]的研究,从数字金融使用深度、覆盖广度、数字化程度三方面综合度量,数据来源于《北京大学数字普惠金融指数》及《中国数字经济指数白皮书》。

归纳前人研究得到的企业二元创新能力影响因素变量指标及说明见表 1。

**Table 1.** Variables influencing enterprise ambidextrous innovation

**表 1.** 企业二元创新影响因素变量指标

选取角度	指标	指标说明
基础因素	企业规模	年总资产的自然对数
	上市时长	(当年年份 - 上市年份 + 1)的自然对数
	企业年龄	(当年年份 - 公司成立年份 + 1)的自然对数
	生命周期	见上文
企业内部	研发人员数量	企业当年期末研发人员总数
	职工数量	企业当年期末员工总数
	研究生及以上人数	企业当年期末研究生以上学历人员总数
	本科人数	企业当年期末本科学历人员总数
	专科人数	企业当年期末专科学历人员总数
	生产岗员工数量	企业当年期末生产岗人员总数
	销售岗员工数量	企业当年期末销售岗人员总数
	技术岗员工数量	企业当年期末技术岗人员总数
	资产负债率	年末总负债除以年末总资产
	总资产净利润率	净利润/总资产平均余额
运营状况	净资产收益率	净利润/股东权益平均余额
	总资产周转率	营业收入/平均资产总额
	现金流比率	经营活动产生的现金流量净额除以总资产
治理状况	托宾 Q 值	(流通股市值 + 非流通股股份数 × 每股净资产 + 负债账面值)/总资产
	董事会规模	董事会人数取自然对数
	独立董事比例	独立董事除以董事人数
	机构持股比例	当年机构投资者持有企业股份比例
	两职合一	董事长是否兼任 CEO, 1 代表是, 0 代表否
	Top1	第一大股东持股比例
Top5	前五大股东持股比例	

Continued

		Top10	前十大股东持股比例
		股权制衡度 1	第二大股东持股比例/第一大股东持股比例
		股权制衡度	第二到第十大股东持股比例之和/第一大股东持股比例
		高管持股比例	高级管理人员持股比例之和
		高管薪酬	高级管理人员薪酬平均值
外部 环境	政策环境	政府补助	企业当年财务报告附注中披露的政府补助金额
		企业创新	中国区域科技创新评价报告企业创新效用值
		创新环境	中国区域科技创新评价报告创新环境效用值
	行业环境	创新绩效	中国区域科技创新评价报告创新绩效效用值
		知识创造	中国区域科技创新评价报告知识创造效用值
		知识获取	中国区域科技创新评价报告知识获取效用值
		月均超额换手率	当年股票月均换手率-去年股票月均换手率
		总关注度	新闻媒体中出现该企业的新闻总数
	外部 关注度	正面关注度	新闻媒体中出现该企业的正面新闻数
		中性关注度	新闻媒体中出现该企业的中性新闻数
	负面关注度	新闻媒体中出现该企业的负面新闻数	
数字化	内部	数字化转型程度	企业当年度所有数字化技术无形资产占本年度无形资产的比例
		数字金融使用深度	北大数字普惠金融指数数字金融使用深度
	外部	数字金融数字化程度	北大数字普惠金融指数数字金融数字化程度
		数字金融覆盖广度	北大数字普惠金融指数数字金融覆盖广度

### 2.3. 特征选择模型

在有监督学习问题中，对于每条样本信息都有一个正确的输出，记为  $Y$ ，输入特征的向量  $x = [x_1, x_2, \dots, x_d]$  组成要素空间  $\Omega$ 。在此定义下，特征选择的目标函数可表示为：

$$\min_{\hat{x}, f} \|f(\hat{x}) - y\| \quad (2-1)$$

式(2-1)中， $k < d$ ， $\hat{x} = [x^1, x^2, \dots, x^k]$  组成的特征子空间  $\hat{\Omega} \subseteq \Omega$ ， $f$  表示可以转换输入特征的关于  $x$  的机器学习模型，以最小化预测输出和正确输出之间的距离。距离衡量使用包括欧几里得距离在内的多种计算方式。

特征工程的总体目标是选择最佳特征子空间，为既定的算法提供最佳数据性能支持。因此特征工程实际上依赖于所定处理模型，这意味着不同算法模型的最佳特征子空间可能不同。本文使用多种代表性算法进行特征选择，并通过投票选出多种算法下企业创新模型中共同的重要特征，保证特征选择合理性。

### 2.4. 特征选择方法

特征选择方法主要包括过滤式选择、包裹式选择、嵌入式选择。过滤式选择通常是作为独立于模型的预处理方法，根据通过解释变量与被解释变量的相关性得分选择特征。包裹式选择是根据机器学习的

模型预测性能对特征集进行效用评估，特征子集的错误率越低则结果越好，错误率最低的特征子集被记为最有特征集合，组成特征子空间。嵌入式选择方法是由学习算法引申来的，是学习算法的组成部分，通常在学习过程中引入稀疏诱导正则化或先验到学习算法的目标函数中来起作用，即将选择过程嵌入到学习过程中，在学习不断优化特征集合的权重赋值。

本文选取卡方拟合法(Chi2)、递归消除方法(Random Forest)、随机森林(Random Forest)算法建立 CRRT 特征选择模型进行重要特征评估选择，对其结果进行综合分析。

### 3. 实证分析

#### 3.1. 数据预处理

鉴于样本数据参差不齐，为保证样本可靠性和数据完整性，对于样本值进行初步筛选，筛选标准如下：

- 1) 剔除经营异常的 ST、\*ST 或 PT 以及金融类行业上市公司样本；
- 2) 根据商务数据基本原则，剔除不符合的上市公司样本。即剔除资产负债率不在 0~1 区间的样本；
- 3) 剔除变量数量缺失严重及核心变量缺失的上市公司样本；
- 4) 对连续变量在前后 1%分位处进行双边缩尾处理，避免极端异常值影响模型拟合效果。

最终，获得 809 家上市公司 7065 个样本观测值。

由于样本不同维度数据间存在数量级差别，因此在进行实验前对数据进行预处理。本文使用 Scikit Learn 预处理包的 Standard Scaler 模块针对样本特征进行标准差标准化，标准化公式如式：

$$x^* = \frac{x - \mu}{\sigma} \tag{3-1}$$

式(3-1)中， $\mu$  为所有样本数据均值， $\sigma$  为所有样本数据的标准差。

#### 3.2. 相关性分析

根据皮尔逊相关系数矩阵计算方式，对样本数据进行相关性分析，计算公式为式(3-2)。

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \tag{3-2}$$

其中， $r$  为相关性系数， $x_i$ 、 $y_i$  为变量数据值， $\bar{x}$ 、 $\bar{y}$  为变量数据均值， $n$  为变量个数。如果  $y$  与  $x$  完全正线性相关，则  $r$  等于 1；如果  $y$  与  $x$  完全负线性相关，则  $r$  等于 -1；如果  $y$  与  $x$  完全非线性相关，则  $r$  等于 0。使用 SPSS 软件计算相关性系数，具体结果见表 2。

**Table 2.** Sample data correlation analysis

**表 2.** 样本数据相关性分析

		x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
Invention patents	皮尔森相关	0.068**	0.050**	0.269**	-0.015	0.040**	0.459**	0.063**	0.350**	0.223**	0.419**
	显著性	0	0.001	0	0.346	0.01	0	0	0	0	0
	N	4111	4159	4758	4116	4116	5173	4116	4758	4758	4758
Utility models and designs	皮尔森相关	0.056**	0.041**	0.262**	-0.017	0.025	0.480**	0.055**	0.337**	0.208**	0.404**
	显著性	0	0.009	0	0.281	0.11	0	0	0	0	0
	N	4111	4159	4758	4116	4116	5173	4116	4758	4758	4758

Continued

		x11	x12	x13	x14	x15	x16	x17	x18	x19	x20
Invention patents	皮尔森相关	0.054**	0.328**	0.669**	0.050**	0.380**	0.058**	0.516**	0.027	0.267**	0.245**
	显著性	0.001	0	0	0.001	0	0	0	0.088	0	0
	N	4159	5098	3997	4159	4724	4116	4332	4116	4583	4968
Utility models and designs	皮尔森相关	0.058**	0.372**	0.658**	0.032*	0.410**	0.053**	0.569**	0.014	0.314**	0.262**
	显著性	0	0	0	0.039	0	0.001	0	0.36	0	0
	N	4159	5098	3997	4159	4724	4116	4332	4116	4583	4968
		x21	x22	x23	x24	x25	x26	x27	x28	x29	x30
Invention patents	皮尔森相关	0.379**	0.207**	0.091**	0.012	0.036*	0.041**	0.047**	0.046**	-0.015	0.004
	显著性	0	0	0	0.369	0.01	0.003	0.001	0.001	0.278	0.787
	N	4992	5173	5173	5173	5171	5173	5173	5171	5171	5173
Utility models and designs	皮尔森相关	0.408**	0.217**	0.103**	0.008	0.040**	0.049**	0.043**	0.041**	0.005	0.002
	显著性	0	0	0	0.546	0.004	0	0.002	0.003	0.706	0.911
	N	4992	5173	5173	5173	5171	5173	5173	5171	5171	5173
		x31	x32	x33	x34	x35	x36	x37	x38	x39	x40
Invention patents	皮尔森相关	-0.021	-0.01	-0.008	0.041**	0.034*	-0.01	0.041**	0.021	0.025	0.055**
	显著性	0.13	0.457	0.555	0.003	0.014	0.491	0.003	0.14	0.078	0
	N	5173	5173	5173	5173	5173	5104	5173	5173	5069	5173
Utility models and designs	皮尔森相关	-0.023	-0.013	-0.013	0.046**	0.034*	-0.029*	0.045**	0.011	0.029*	0.058**
	显著性	0.103	0.349	0.36	0.001	0.014	0.037	0.001	0.414	0.039	0
	N	5173	5173	5173	5173	5173	5104	5173	5173	5069	5173
		x41	x42	x43	x44	x45	x46				
Invention patents	皮尔森相关	-0.040**	-0.046**	0.538**	-0.021	-0.091**	0.038**				
	显著性	0.004	0.001	0	0.134	0	0.007				
	N	5079	5173	4474	4961	4012	5055				
Utility models and designs	皮尔森相关	-0.043**	-0.045**	0.552**	-0.032*	-0.088**	0.041**				
	显著性	0.002	0.001	0	0.026	0	0.004				
	N	5079	5173	4474	4961	4012	5055				

\*\*. 相关性在 0.01 层上显著(双尾)。\*. 相关性在 0.05 层上显著(双尾)。

根据表 2 相关性分析可以看出, 绝大部分指标与探索性创新能力及开发式创新能力相关性显著, 表明本章所建立影响因素指标体系科学合理, 可以使用样本数据进行双元创新能力评估分析。

### 3.3. CRRT 特征选择结果分析

使用 Python3.9 对 CRRT 模型进行模拟仿真, 分别得到三种算法下企业双元创新的 20 个重要特征指

标。依据各指标重要度排序赋值得分，三个模型所有指标分数求和后再次排序，得到最终综合指标集。各算法的主要代码见表 3:

**Table 3.** Main codes of feature selection algorithms  
**表 3.** 特征选择算法主要代码

特征选择方法	主要代码
Chi2	<code>sklearn.feature_selection.chi2(X, y), k = 20</code>
RFE	<code>fit = rfe.fit(X, y)</code>
Random Forest	<code>model = ExtraTreesClassifier()</code> <code>model.fit(X, y)</code> <code>print(model.feature_importances_)</code>

### 3.3.1. 探索式创新重要特征

利用发明专利数量作为被解释变量建立探索式创新影响因素选择模型，得到各算法下重要特征集(指标重要性程度至上到下为由高到低)，具体指标见表 4:

**Table 4.** Important features of exploratory innovation sub-model (number of invention patents)  
**表 4.** 探索式创新分模型重要特征(发明专利数量)

模型	Chi2	RFE	Random Forest
指标 (发明专利)	研发人员数量	研发人员数量	研发人员数量
	研究生及以上	总资产净利润率	技术
	本科	现金流比率	数字化转型程度
	政府补助	数字化转型程度	研究生及以上
	数字化转型程度	月均超额换手率	高管薪酬
	技术	政府补助	政府补助
	职工数量	高管薪酬	高管持股比例
	生产	独立董事比例	企业年龄
	销售	高管持股比例	股权制衡度
	负面关注度	第一大股东持股比例	本科
	企业规模	前五大股东持股比例	企业规模
	正面关注度	资产负债率	资产负债率
	总关注度	前十大股东持股比例	托宾 Q 值
	知识创造	机构持股比例	股权制衡度
	独立董事比例	股权制衡度	总资产净利润率
	净资产收益率	总资产周转率	上市时长
	总资产净利润率	董事会规模	净资产收益率
	资产负债率	企业年龄	现金流比率
	高管持股比例	两职合一	机构持股比例
	前十大股东持股比例	上市时长	前十大股东持股比例

从表 4 中可以看出, 以发明专利为探索式创新变量时, 各算法均认为研发人员数量为最重要特征, 研发人员、研究生及以上学历员工通常具有较高的研发创新技能和较强的解决问题的能力, 他们的加入对企业科技水平的提高和推动企业技术创新的发展很有帮助。人力资源素质对探索式创新能力影响重要性程度高也符合探索式创新难度大、重研发的特点。

统计各算法中各指标排名得分, 得到三类算法下重要性程度最高的 20 个指标, 形成发明专利表征下探索式创新能力影响因素重要特征子空间, 具体指标见表 5:

**Table 5.** Comprehensive important features of exploratory innovation (invention patents)

**表 5.** 探索式创新综合重要特征(发明专利)

被解释变量	发明专利
指标	研发人员数量, 政府补助, 数字化转型程度, 总资产净利润率, 高管持股比例, 技术, 净资产收益率, 研究生及以上, 高管薪酬, 资产负债率, 本科, 企业规模, 前十大股东持股比例, 职工数量, 上市时长, 生产, 前五大股东持股比例, 专科, 企业年龄, 第一大股东持股比例

### 3.3.2. 开发式创新重要特征

针对实用新型与外观设计专利和作为被解释变量, 建立开发式创新影响因素选择模型, 使用三类算法分别得到重要特征指标集, 具体指标见表 6:

**Table 6.** Important features of the development innovation sub-model (utility model and design patent)

**表 6.** 开发式创新分模型重要特征(实用新型和外观设计专利)

模型	Chi2	RFE	Random Forest
指标 (实用新型和外观设计)	政府补助	总资产净利润率	政府补助
	职工数量	现金流比率	研究生及以上
	董事会规模	净资产收益率	数字化转型程度
	企业规模	数字化转型程度	本科
	技术	研究生及以上	高管持股比例
	研究生及以上	高管薪酬	知识创造
	股权制衡度	独立董事比例	企业规模
	净资产收益率	研究生及以上	高管薪酬
	总资产净利润率	前五大股东持股比例	技术
	知识创造	高管持股比例	专科
	高管薪酬	第一大股东持股比例	销售
	数字化转型程度	前十大股东持股比例	机构持股比例
	创新环境	机构持股比例	前十大股东持股比例
	知识获取	总资产周转率	股权制衡度
	生产	股权制衡度	股权制衡度
	正面关注度	月均超额换手率	第一大股东持股比例
	股权制衡度	董事会规模	企业年龄
	独立董事比例	两职合一	职工数量
	资产负债率	企业年龄	总资产净利润率
	数字经济指数	上市时长	资产负债率

从表 6 中可以看出，以实用新型和外观设计专利为开发式创新变量时，政府补助以及部分财务指标成为重要性程度较高指标，这符合创新能力对资金要求的特点，与探索式创新对高素质人才要求不同，开发式创新能力影响因素中职工数量及本科人数等重要性程度较高，这些员工通常具有更为专业的技能和更强的操作实践能力，这有利于企业实践应用型的研究及开发新项目，符合开发式创新特点。

统计各算法中各指标排名得分，得到三类算法下重要性程度最高的 20 个指标，形成实用新型和外观设计专利表征下开发式创新影响因素重要特征集合，具体指标见表 7：

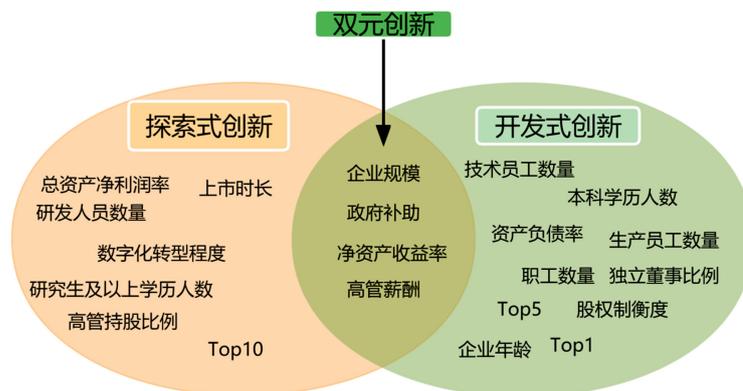
**Table 7.** Comprehensive important features of development-type innovation (utility model and design patents)  
**表 7.** 开发式创新综合重要特征(实用新型和外观设计专利)

被解释变量	实用新型和外观设计
指标	政府补助，本科，技术，资产负债率，研究生及以上，净资产收益率，总资产净利润率，职工数量，高管薪酬，前五大股东持股比例，股权制衡度，企业规模，生产，研发人员数量，企业年龄，第一大股东持股比例，高管持股比例，前十大股东持股比例，数字化转型程度，独立董事比例

### 3.3.3. 双元创新对比

以特征在双元创新影响因素指标集中的重要性排序作为对比依据，若某特征仅在探索式创新指标集中出现或在探索式创新影响因素指标集中的重要性排序高于开发式创新影响因素指标集中的重要性排序，认为该因素指标侧重于影响探索式创新能力；若两集中重要性排序相同，认为该因素指标为标记为双元创新指标；否则认为该因素指标对开发式创新能力影响更大。

由图 1 可以看出，企业规模、净资产收益率、高管薪酬以及政府补助四个因素对探索式创新和开发式创新影响程度相当。从资源基础理论考虑，企业规模、净资产收益率、政府补助均会影响企业的资源丰富度，一般而言，企业规模大、收益高、补助多会直接提高企业资源存量，使得企业在创新发展过程中不至因资源限制而在双元创新选择上厚此薄彼。同理，资产负债率也在一定程度上反映企业资源变化，资产负债高则企业难以进行需要较大投入的探索式创新，开发式创新带来的短期效益则有利于改善企业经济状况。从人力资本讨论，在创新过程中，研发人员因职责所在，更容易进行探索式创新，研究生及以上学历人员则从知识、专业性出发也侧重探索式创新。技术生产员工、本专科学历人员距离实际生产更近，有更多机会进行难度较低的开发式创新。独立董事比例、股权制衡度、股东持股比例等反映董事会监督作用的因素对开发式创新影响更大，这从公司治理角度反映了股东们更倾向于进行容易获得短期大额收益的开发式创新。



**Figure 1.** Comparison of important features of ambidextrous innovation  
**图 1.** 双元创新重要特征对比

对各重要特征排序整理见表 8。

**Table 8.** Combined Important Features of Ambidextrous Innovation (Patents)

**表 8.** 双元创新综合重要特征(专利)

创新维度	影响因素及重要度排序
探索式创新	数字化转型程度(2)、总资产净利润率(3)、研发人员数量(4)、高管持股比例(5)、Top 10 (13)、上市时长(15)、企业年龄(19)
开发式创新	本科学历人数(2)、技术员工人数(3)、资产负债率(4)、研究生及以上(5)、职工数量(7)、Top 5 (10)、股权制衡度(11)、生产(12)、本科学历人数(14)、企业年龄(15)、Top 1 (16)、独立董事比例(20)
双元创新	政府补助(1)、净资产收益率(6)、高管薪酬(9)、企业规模(12)

从表 8 中各指标重要性排序看,数字化转型程度成为影响探索性创新的最重要因素,这反映了数字化时代企业要想进行大幅度、改革化深度创新则需要紧跟时代发展,进行数字化转型是新格局下企业进行探索式创新的有效途径。开发式创新影响因素中,人力相关因素较多影响度排序也更靠前,这与开发式创新的特点相关,开发式创新更多是过程中进行的,在生产、产品改进、业务流程等过程中,人力资源的投入和数量优势可以在一定程度上影响创新能力,对于难度要求较低的开发式创新的影响也更显著。而对于专业性强、难度大的探索式创新,则与研发人员的质量相关性更高。

### 3.3.4. 模型稳健性

使用资本化研发投入与费用化研发投入分别替换发明专利与实用新型和外观设计专利作为双元创新解释变量指标,再次使用所建立特征选择模型进行拟合,检验模型稳健性。替换被解释变量指标后,重复特征选择过程,最终得到的重要特征指标见表 9 和表 10:

**Table 9.** Comprehensive key characteristics of exploratory innovation (capitalized R & D investment)

**表 9.** 探索式创新综合重要特征(资本化研发投入)

被解释变量	资本化投入
指标	数字化转型程度, 高管薪酬, 总资产净利润率, 净资产收益率, 企业规模, 研究生及以上, 技术, 股权制衡度, 知识创造, 股权制衡度, 资产负债率, 第一大股东持股比例, 职工数量, 独立董事比例, 政府补助, 销售, 高管持股比例, 前十大股东持股比例, 本科, 机构持股比例

**Table 10.** Comprehensive key characteristics of development-based innovation (cost-based R & D investment)

**表 10.** 开发式创新综合重要特征(费用化研发投入)

被解释变量	费用化投入
指标	政府补助, 高管薪酬, 数字化转型程度, 企业规模, 研究生及以上技术, 总资产净利润率, 净资产收益率, 股权制衡度, 职工数量, 独立董事比例, 第一大股东持股比例, 股权制衡度, 销售, 高管持股比例, 资产负债率, 知识创造, 本科, 董事会规模, 前五大股东持股比例

分别对比表 5 与表 9、表 7 与表 10 发现,以不同指标表示被解释变量进行特征选择得到的特征选择后指标集合中各指标基本一致,表明所建立特征选择模型稳健性高,重要特征选择有效。

## 4. 结论

使用企业数字化转型程度成为影响探索式创新能力的最重要因素,进行数字化转型是新格局下企业进行探索式创新的有效途径。开发式创新能力影响因素中人力相关特征较为重要,人力资源的投入和数

量优势对开发式创新能力影响较探索式创新能力影响更显著。企业规模和资产、得到的政府补助以及高管薪酬则是企业协同发展二元创新需着重考虑的因素特征。

## 致谢

感谢对本论文完成过程中提供数据的支持者、提供指导和帮助者，同时对本论文应用资料给予转载和引用权的资料、图片、文献、研究思想和设想的所有者表示感谢。

## 参考文献

- [1] Jancenelle, V.E. (2021) Tangible-Intangible Resource Composition and Firm Success. *Technovation*, **108**, Article ID: 102337. <https://doi.org/10.1016/j.technovation.2021.102337>
- [2] 张馨, 彭纪生, 涂海银. 企业创新资源配置结构对创新绩效的影响: 基于不同产权性质视角[J]. 科技进步与对策, 2021, 38(8): 126-135.
- [3] 尹美群, 盛磊, 李文博. 高管激励、创新投入与公司绩效——基于内生性视角的分行业实证研究[J]. 南开管理评论, 2018, 21(1): 109-179.
- [4] 解维敏, 张恒鑫. 自主创新还是技术引进: 业绩型股权激励与企业创新策略[J]. 系统工程理论与实践, 2023, 43(1): 36-57.
- [5] 孟庆斌, 李昕宇, 张鹏. 员工持股计划能够促进企业创新吗?——基于企业员工视角的经验证据[J]. 管理世界, 2019, 35(11): 209-289.
- [6] 史永东, 王彤彤. 政府扶持与企业创新[J]. 系统工程理论与实践, 2022, 42(8): 2002-2016.
- [7] Han, J. and Kang, R. (2021) Market Uncertainty, Innovation of Firms in Alliance and Alliance Partner Characteristics. *European Journal of Innovation Management*, **24**, 1883-1905. <https://doi.org/10.1108/EJIM-05-2020-0195>
- [8] Babkin, A., Mylnikova, E., Chernovalova, G., et al. (2021) Information-Infrastructure Mechanism for Managing Industrial Enterprise Self-Development in the Setting of Digitization. In: Beskopylny, A., Shamtsyan, M., Eds., *XIV International Scientific Conference "INTERAGROMASH 2021". Lecture Notes in Networks and Systems*, **246**, 762-770. [https://doi.org/10.1007/978-3-030-81619-3\\_85](https://doi.org/10.1007/978-3-030-81619-3_85)
- [9] 方鑫, 董静. 管理层能力对创业企业二元创新战略的影响研究[J]. 外国经济与管理, 2022, 44(11): 77-92.
- [10] 汤莉, 余银芳. CEO 学术经历与企业二元创新[J]. 华东经济管理, 2021, 35(10): 59-69.
- [11] Anokhin, S., Peck, S. and Wincent, J. (2016) The Role of Governance Factors. *Business Research*, **69**, 4744-4749.
- [12] 顾乃康, 周艳利. 卖空的事前威慑、公司治理与企业融资行为——基于融资融券制度的准自然实验检验[J]. 管理世界, 2017, 281(2): 120-134.
- [13] 祁怀锦, 曹修琴, 刘艳霞. 数字经济对公司治理的影响——基于信息不对称和管理者非理性行为视角[J]. 改革, 2020, 314(4): 50-64.
- [14] 潘为华, 贺正楚, 潘红玉. 中国数字经济发展的时空演化和分布动态[J]. 中国软科学, 2021, 370(10): 137-147.