# 大学生毕业生就业问题的预测与评价模型

## 摘要

近年来,大学生毕业生就业难问题愈演愈烈。及时把握大学生就业趋势和特征,能提升积极就业政策对大学毕业生的针对性。本文提供了有效的数学模型,通过收集大学生就业统计数据,建立模型对 2024 年大学生毕业生的就业环境进行预测。并通过建立竞争力评价模型评估在校学生的竞争力,为增加就业提供了一定的参考价值。

针对问题一,对于常规因子,引入偏最小二乘法来解决多个因变量的回归问题。在 2015-2023 年的时间跨度内,首先收集大学生毕业生就业率作为衡量大学生毕业生就业情况的指标,即将就业率作为偏最小二乘法中的 Y 变量;随后将收集到的国民生产总值,各类价格指数和平均工资等全部作为 X 变量,通过因子载荷系数表得到变量的重要性程度用以衡量是否作为主要影响因素。对于波动性大的因子,仅使用近三年的数据,并结合《中国本科生就业报告》确定其重要程度。最终得到所有的主要影响因素是国民总收入,国内生产总值,第三产业增加值,工厂价格指数,企业平均工资和院校层次,地区和专业。

**针对问题**二,将 2024 年的就业环境预测分为就业率预测和其他宏观因素预测。对于**就业率预测**,采用**灰色预测模型 GM(1,1)**得到 2024 年的就业率为 **93.071%**。对于其他宏观因素,首先采用**主成分分析法对分大类 GDP,价格指数,行业增加值进行降维处理**,随后同样采用**灰色预测模型 GM(1,1),**得到标志性的指标,即 2024 年的预期月平均工资为 **6825.4** 元。

**针对问题三**,在参考大量机构通过问卷调查法等方式得出的高校学生竞争力评价模型的基础上,我们基于猎聘大数据,智联招聘等专业机构近三年的就业数据统计分析,首先将所有数据处理为统一的**描述性统计结果**。在此基础上,我们进一步将在校学生就业竞争力划分为实践与学识、专业职业技能、人际关系、职业素养**四个维度**,并借助CRITIC 权重法进行建模。

**针对问题四**, 收集 2024 年的毕业生情况、实际行业热点和招聘热门等数据, 放入到问题三建立的模型中, 并根据定量结果对 2024 当年增加就业和提高大学毕业生个人就业竞争力提供解决方案。

关键词:偏最小二乘法 灰色预测模型 主成分分析法 CRITIC 权重法

## 一、问题重述

### 1.1 问题背景

近年来,大学毕业生的就业率呈现缓慢下降的趋势。随着应届大学毕业生不断增多,解决以大学毕业生为主体的青年就业问题成为当务之急。

大学生就业难的问题是由多种因素导致的,通过研究就业趋势和特征确定各类影响 因子,能有针对性地提高大学毕业生就业技能,获得良好的职业发展结果,是积极促进 大学生就业的重要方法。

## 1.2 问题要求

**问题一**:在本问题中,需要分类收集影响大学生毕业生就业的主要因素,并通过不同因素与大学生毕业生就业率的相关度分析,来确定主要影响因子。

**问题二:**本问题在问题一的基础上,对就业率采用灰色预测模型进行多阶预测,得到 2024 年的就业率预测结果。同时对其他因子在原有的偏最小二乘模型上进行

问题三:在本问题中,需要通过处理收集到的大量数据并处理为为统一的描述性统计结果。在此基础上,将在校学生就业竞争力划分为四个维度,并借助 CRITIC 权重法进行建模并衡量。

**问题四:** 在本问题中,需要综合问题一、二、三和 2024 年的部分已知就业偏好数据对增加就业率提供建议,并结合实际情况为增强个人就业竞争力提供解决方案。

## 二、问题分析

本题的研究对象是大学生毕业生群体,研究的内容为建立评估大学生就业环境和就业竞争力的数学模型。它属于自主收集数据,具有多影响因子的预测类问题。为了解答这个问题,我们需要在多渠道收集相关数据的基础上,将各类影响因子放到统一的量化标准下进行分析,并提供合理化建议。

## 2.1 问题一的分析

问题一属于在多个因变量回归模型中确定因子载荷系数的问题,基于偏最小二乘算法,确定常规因子的载荷系数,以此为依据确定其是否是主要影响因素。即将就业率作为单一评价指标,也就是 Y 变量,将其他因素作为 X 变量,用十年的周期来确定累计投影重要性,并得到因子载荷系数。将常规因子中的重要影响因素与搜集到的近三年的热点因子结合,得到题目要求的影响大学生毕业生就业的主要因素。

#### 2.2 问题二的分析

问题二属于多个因素的预测问题。将 2024 年的就业环境预测分为就业率预测和其他宏观因素预测。对于其他宏观因素,要先采用主成分分析法降维处理。通过建立灰色预测模型 GM(1,1)得到 2024 年的预期就业率和月平均工资水平,以此作为 2024 年的就业环境预测指标。

#### 2.3 问题三的分析

问题三属于评价类模型的建模问题。需要对收集到的大量数据进行数据预处理,简化数据维度,并进行建模。随后要通过模型得到定量的结果。

#### 2.4 问题四的分析

问题四属于实际做法类问题。根据前面得到的定量计算结果和 2024 年的实际环境数据对相关部门和个人的就业竞争力发展提供有效的建议。

## 三、模型假设

- 1. 假设文中的大学生毕业生就业率等同于大学生本科生就业率。
- 2. 假设从当前到 2024 年结束这一段时期内没有重大社会事件发生。

## 四、模型的建立与求解

#### 4.1 问题一的数据收集与求解

影响大学生毕业生就业的因素是多方面的,在收集数据收集之前,我们首先把可能的影响大学生毕业生就业的因素分为几个方面分别讨论。

- 1. 宏观经济结构(GDP, 国民总收入, 各类价格指数, 平均工资, 行业增加值)
- 2. 个人能力(院校层次,留学经历,学历,实习经历)
- 3. 市场需求/(薪资,热点行业/行业发展趋势,地区)
- 4. 政策(创业补贴,产业结构调整)

在实际收集和处理过程中,由于大部分宏观经济指标,即国民生产总值,价格指数,平均工资等能直接和整体就业率进行拟合比较,且整体变化趋势较为稳定,缺失值较少,我们收集近十年的数据,对其采用偏最小二乘法来建模分析,将因子载荷大的作为影响就业情况的主要因素。

而对于市场发展需求,学历要求等波动较大且缺失值较多的指标,我们采用近三年的数据,并结合《中国本科生就业报告》,各大分析平台等综合得到其中影响就业的主要因素。

#### 4.1.1 相关影响因素收集和算法配置

就业率 95.5% 95.4% 95.7%

收集得到 2015-2022 年的大学生毕业生就业率,由于就业率可较好地简单反映就业情况,我们将其简化为衡量大学生毕业生就业情况的唯一指标。在偏最小二乘回归算法中,将大学生毕业生就业率作为 Y 变量。

· · · · · · · · · · · · · · · · · · ·								
年份	2015	2016	2017	2018	2019	2020	2021	2022

95.8%

95.6%

表 1 大学生毕业生就业率

\*数据来源:《2023年中国本科生就业报告》

94.8%

收集得到经济社会层面的各类宏观指标,包括 2015-2022 年的国民总收入,国内生产总值,第一、二、三产业增加值,不同单位就业人员平均工资,各类价格指数等,作为偏最小二乘法的 X 变量。数据详见附件 1 Q1. x1sx。

\*数据来源:国家统计局

94.3%

93.1%

## 算法配置:

变量 Y: { 就业率(%) };

变量 X: { 国民总收入(万亿元), 国内生产总值(万亿元), 第一产业增加值(千亿元), 第二产业增加值(万亿元), 第三产业增加值(万亿元), 人均国内生产总值(千元), 城镇单位就业人员平均工资(千元), 国有单位就业人员平均工资(千元), 城镇集体单位就业人员平均工资(千元), 股份合作单位就业人员平均工资(千元), 联营单位就业人员平均工资(千元), 有限责任公司就业人员平均工资(千元), 股份有限公司就业人员平均工资(千元), 其他单位就业人员平均工资(千元), 港、澳、台商投资单位就业人员平均工资(千元), 外商投资单位就业人员平均工资(千元), 居民消费价格指数(上年=100), 城市居民消费价格指数(上年=100), 农村居民消费价格指数(上年=100), 商品零售价格指数

(上年=100), 工业生产者出厂价格指数(上年=100), 工业生产者购进价格指数(上年 =100) }

### 4.1.2 通过方差解释情况和 VIP (累积投影重要性) 参考最大主成分数量

表 2 展示潜在因子的信息综合解释能力。其中,累计的 X 方差代表对自变量信息的 提取,累计的Y<sup>2</sup>(R<sup>2</sup>)代表对因变量信息的提取,可以以此为依据确定参数最大主成分数 量。因子对方差解释情况表的结果显示,前2个潜在因子就可解释自变量80%的信息。

潜在因子 X 方差 累计的 X 方差 Y 方差 累计的 Y 方差(R2) 调整后的 R2 0.786 0.786 0.64 0.64 0.58 2 0.166 0.951 0.757 0.66 0.117 3 0.041 0.993 0.069 0.826 0.695 4 0.998 0.958 0.005 0.133 0.903 0.002 0.041 0.999 1 1

表 2 因子方差解释情况表

表3展示VIP(累积投影重要性)的情况,它表示成分个数不同时,X对于Y的解 释重要性力度,也可以用于参考最大主成分数量。其中对于 VIP 很大(大于1)的自变 量,它在解释潜在因子(从而在解释因变量)时作用相对更大一些。

耒	3	白亦昰 VID	(累积投影重要性)	汇兑表
72	· O			7 W XX

ス O 口文里 VII へ赤小川	スポノ王	女 1工 /	/二/心1	<u>~</u>	
· 变量	因子	因子	因子	因子	因子
· · · · · · · · · · · · · · · · · · ·	1	2	3	4	5
国民总收入(万亿元)	1.11	1.024	0.981	0.916	0.903
国内生产总值(万亿元)	1.117	1.029	0.985	0.919	0.905
第一产业增加值(千亿元)	1.255	1.232	1.208	1.163	1.207
第二产业增加值(万亿元)	1.142	1.05	1.006	0.94	0.954
第三产业增加值(万亿元)	1.076	1.005	0.963	0.931	0.913
人均国内生产总值(千元)	1.123	1.034	0.99	0.92	0.907
城镇单位就业人员平均工资(千元)	1.154	1.062	1.02	0.948	0.932
国有单位就业人员平均工资(千元)	1.125	1.034	0.992	0.929	0.929
城镇集体单位就业人员平均工资(千	1 1/11	1 0/19	1 005	0.935	0 924
元)	1.171	1.043	1.005	0.555	0.524
股份合作单位就业人员平均工资(千	1 037	0 994	0 958	0.954	0 934
元)	1.007	0.55	0.550	0.55 .	0.55
联营单位就业人员平均工资(千元)	1.131	1.04	0.997	0.967	0.95
有限责任公司就业人员平均工资(千	1 135	1.043	1	0.954	0 935
元)	1.100	2.0.0	_	0.55 .	0.555
股份有限公司就业人员平均工资(千	1.192	1.109	1.073	1.023	1.004
元)					
其他单位就业人员平均工资(千元)	0.993	0.973	0.938	1.235	1.231
港、澳、台商投资单位就业人员平均	1.223	1.16	1.127	1.088	1.072
工资(千元)	0		<b></b> ,		<b>_</b>

## 图 1 自变量 VIP (累积投影重要性) 图

## 4.1.3 通过成分矩阵表得到主成分的组成情况

由表 4 可得到主成分分析降维后的成分矩阵表。

表 4 成分矩阵表

变量	因子1	因子 2	因子3	因子4	因子 5
国民总收入(万亿元)	0.237	0.015	0.027	0.069	0.138
国内生产总值(万亿元)	0.238	0.003	0.009	0.053	0.123
第一产业增加值(千亿元)	0.267	-0.269	-0.41	-0.319	-0.542
第二产业增加值(万亿元)	0.244	-0.044	-0.034	-0.075	0.232
第三产业增加值(万亿元)	0.229	0.064	0.084	0.178	0.123
人均国内生产总值(千元)	0.24	-0.007	-0.006	0.027	0.118
城镇单位就业人员平均工资(千元)	0.246	-0.064	-0.112	-0.021	-0.105
国有单位就业人员平均工资(千元)	0.24	-0.023	-0.064	0.046	-0.183
城镇集体单位就业人员平均工资 (千元)	0.243	-0.04	-0.05	0.022	-0.119
股份合作单位就业人员平均工资 (千元)	0.221	0.124	0.182	0.26	0.118
联营单位就业人员平均工资(千元)	0.241	-0.046	-0.072	0.135	-0.038
有限责任公司就业人员平均工资 (千元)	0.242	-0.034	-0.056	0.108	0.079
股份有限公司就业人员平均工资 (千元)	0.254	-0.128	-0.217	-0.208	-0.149
其他单位就业人员平均工资(千元)	0.212	0.154	0.206	0.573	0.467
港、澳、台商投资单位就业人员平均工资(千元)	0.261	-0.187	-0.295	-0.273	-0.232
外商投资单位就业人员平均工资 (千元)	0.254	-0.127	-0.219	-0.278	-0.15
居民消费价格指数(上年=100)	-0.054	0.448	0.022	-0.018	-0.18
城市居民消费价格指数(上年=100)	-0.068	0.428	-0.037	-0.262	-0.25
农村居民消费价格指数(上年=100)	-0.035	0.424	0.048	0.348	-0.001
商品零售价格指数(上年=100)	0.173	0.235	0.075	-0.472	0.434
工业生产者出厂价格指数(上年 =100)	0.098	0.36	0.83	0.118	-0.181
工业生产者购进价格指数(上年	0.115	0.272	0.711	0.001	-0.241

## 4.1.4 通过因子载荷系数表得到变量的重要性

表 5 为因子载荷系数表,分析了到每个因子中显变量的重要性。

表 5 因子载荷系数表

变量	因子1	因子 2	因子3	因子 4	因子5
国民总收入(万亿元)	0.243	0.034	-0.006	0.016	0.117
国内生产总值(万亿元)	0.243	0.029	-0.011	0.01	0.106
第一产业增加值(千亿元)	0.237	-0.079	-0.13	-0.015	-0.412
第二产业增加值(万亿元)	0.242	-0.013	0.023	-0.165	0.268
第三产业增加值(万亿元)	0.242	0.068	-0.02	0.128	0.054
人均国内生产总值(千元)	0.243	0.024	-0.01	-0.013	0.112
城镇单位就业人员平均工资(千元)	0.242	0.018	-0.067	0.056	-0.096
国有单位就业人员平均工资(千元)	0.241	0.045	-0.069	0.146	-0.202
城镇集体单位就业人员平均工资	0 2/12	0 006	-0.031	U U80	-0 1/17
(千元)	0.242	0.000	0.031	0.003	0.147
股份合作单位就业人员平均工资	0 2/1	0 086	0.014	Λ 1Q1	0.004
(千元)	0.241	0.000	0.014	0.131	0.004
联营单位就业人员平均工资(千元)	0.239	0.013	-0.089	0.195	-0.112
有限责任公司就业人员平均工资	0.242	0.02	-0.072	0 112	U U38
(千元)	0.242	0.02	0.072	0.112	0.030
股份有限公司就业人员平均工资	0 242	-0 003	-0.068	-0 107	-0 059
(千元)	0.242	0.003	0.000	0.107	0.055
其他单位就业人员平均工资(千元)	0.235	0.115	-0.092	0.407	0.262
港、澳、台商投资单位就业人员平	0 241	-0 037	-0.085	-0 122	-0 122
均工资(千元)	0.2-1	0.037	0.005	0.122	0.122
外商投资单位就业人员平均工资	0.242	0	-0.047	-0 186	-0 041
(千元)	0.212	J	0.017	0.100	0.011
居民消费价格指数(上年=100)	0.003	0.716	-0.332	0.042	-0.169
城市居民消费价格指数(上年=100)	-0.014	0.729	-0.297	-0.192	-0.14
农村居民消费价格指数(上年=100)	0.019	0.658	-0.41	0.385	-0.144
商品零售价格指数(上年=100)	0.206	0.35	0.057	-0.739	0.607
工业生产者出厂价格指数(上年	0 147	-0.062	0.596	-0 079	-0 23
=100)	0.147	0.002	0.550	0.075	0.23
工业生产者购进价格指数(上年	0.152	-0 11	0.575	-0.15	-0 235
=100)					
	-0.194	0.274	0.197	1.108	0.886

表 6 展示了本次模型结果,主要包括模型的系数,用于分析自变量 X 对于因变量 Y 的影响关系情况。

表 6 模型系数结果表

就业率(%)

常数	95.025
国民总收入(万亿元)	0.151
国内生产总值(万亿元)	0.116
第一产业增加值(千亿元)	-0.966
第二产业增加值(万亿元)	0.053
第三产业增加值(万亿元)	0.274
人均国内生产总值(千元)	0.079
城镇单位就业人员平均工资(千元)	-0.189
国有单位就业人员平均工资(千元)	-0.165
城镇集体单位就业人员平均工资(千元)	-0.139
股份合作单位就业人员平均工资(千元)	0.39
联营单位就业人员平均工资(千元)	0.039
有限责任公司就业人员平均工资(千元)	0.113
股份有限公司就业人员平均工资(千元)	-0.455
其他单位就业人员平均工资(千元)	1.012
港、澳、台商投资单位就业人员平均工资(千元)	-0.62
外商投资单位就业人员平均工资(千元)	-0.528
居民消费价格指数(上年=100)	-0.039
城市居民消费价格指数(上年=100)	-0.362
农村居民消费价格指数(上年=100)	0.48
商品零售价格指数(上年=100)	-0.086
工业生产者出厂价格指数(上年=100)	0.198
工业生产者购进价格指数(上年=100)	-0.018

## 4.1.5 求解偏最小二乘回归的标准化公式

基于上述分析,得到了模型的标准化公式为:

就业率(%) = 95.025+0.151 \* 国民总收入(万亿元)+0.116 \* 国内生产总值(万亿元)-0.966 \* 第一产业增加值(千亿元)+0.053 \* 第二产业增加值(万亿元)+0.274 \* 第三产业增加值(万亿元)+0.079 \* 人均国内生产总值(千元)-0.189 \* 城镇单位就业人员平均工资(千元)-0.165 \* 国有单位就业人员平均工资(千元)-0.139 \* 城镇集体单位就业人员平均工资(千元)+0.39 \* 股份合作单位就业人员平均工资(千元)+0.039 \* 联营单位就业人员平均工资(千元)+0.113 \* 有限责任公司就业人员平均工资(千元)-0.455 \* 股份有限公司就业人员平均工资(千元)+1.012 \* 其他单位就业人员平均工资(千元)-0.455 \* 股份有限公司就业人员平均工资(千元)+1.012 \* 其他单位就业人员平均工资(千元)-0.62 \* 港、澳、台商投资单位就业人员平均工资(千元)-0.528 \* 外商投资单位就业人员平均工资(千元)-0.039 \* 居民消费价格指数(上年=100)-0.362 \* 城市居民消费价格指数(上年=100)+0.48 \* 农村居民消费价格指数(上年=100)-0.018 \* 工业生产者购进价格指数(上年=100)

## 4.1.6 宏观因素中的主要影响因素

综合模型系数结果表看,剔除掉部分偏差过大的负相关因素,相关性较大的即为主要影响因素。由此得到主要影响因素是国民总收入,国内生产总值,第三产业增加值,工厂价格指数,企业平均工资。

#### 4.1.7 其他因子中的主要影响因素

综合《中国本科生就业报告》,我们采用毕业落实率来进行衡量。毕业去向落实率:本科生的毕业去向落实率=已就业本科毕业生数/本科毕业生总数。其中已就业人群包括"受雇工作"、国内外读研等五类。

以2023年的报告为例。

数据显示,2022届本科毕业生毕业半年后毕业去向落实率为86.0%,其中"双一流"院校毕业去向落实率为91.1%,高于地方本科院校(85.0%)。

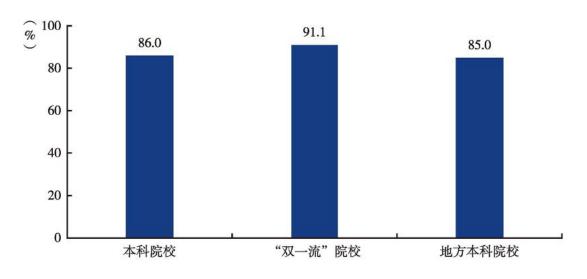


图 2

从不同地区来看,2022 届东部地区本科院校毕业生毕业半年后的毕业去向落实率 (88.0%) 最高;其次是西部地区 (85.9%) (见图 2-3),其中西南地区的毕业去向落实率 (86.3%) 略高于西北地区 (85.5%)。

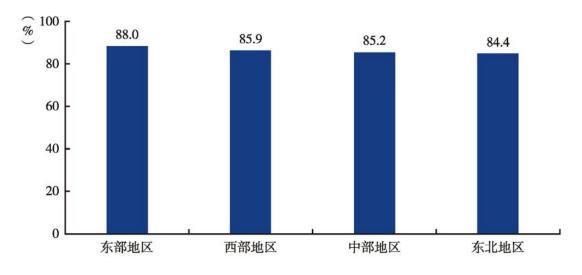


图 3

进一步从专业来看,也存在很大差异。

			单位: %
本科学科门类名称	毕业去向落实率	本科学科门类名称	毕业去向落实率
工学	89.5	历史学	84.2
教育学	87.0	经济学	83.3
农学	86.6	艺术学	83.1
理学	85.6	文学	82.7
管理学	85.3	法学	78.2
医学	85.2		
全国本科	86.0	全国本科	86.0

注: 个别学科门类因为样本较少,没有包括在内。

#### 图 4

因此,我们可以简单认为,院校层次,地区和专业是其他因素中的主要影响因素。 由此我们得到所有的主要影响因素是国民总收入,国内生产总值,第三产业增加值, 工厂价格指数,企业平均工资和院校层次,地区和专业。

## 4.2 问题二的模型建立与求解

首先尝试引入灰色预测模型 GM(1,1),对 2024 年的就业率进行预测。

#### 4.2.1 时间序列的级比检验

在建立灰色预测模型 GM(1,1)前,需要对时间序列进行级比检验。若通过级比检验,即所有的级比值都位于区间( $e^{-(-2/(n+1))}$ , $e^{-(2/n+1)}$ )内,则说明该序列适合构建灰色模型,若不通过级比检验,则需要对序列进行"平移转换",从而使得新序列满足级比值检验。

将 2015-2022 年的就业率统计结果按时间序列进行级比检验,得到级比检验结果表,如表 7 所示。

表 7 级比检验结果表

索引项	原始值	级比值
2015年	95.5	-
2016年	95.4	1.001
2017年	95.7	0.997
2018年	95.8	0.999
2019年	95.6	1.002
2020年	94.8	1.008
2021年	94.3	1.005
2022年	93.1	1.013

从上表分析可以得到,原序列的所有级比值都位于区间(0.801, 1.249)内,说明原序列适合构建灰色预测模型。

#### 4.2.2 构建灰色模型

构建灰色模型 GM(1,1)对 2024年的就业率进行预测,代码见附录 2。

	_
==	O
7	~

发展系数 a	灰色作用量 b	后验差比c值
0.004	96.678	0.283

上表格展示了发展系数、灰色作用量、后验差比值。由发展系数和灰色作用量可以构建灰色预测模型。发展系数表示数列的发展规律和趋势,灰色作用量反映数列的变化关系。后验差比值可以验证灰色预测的精度,后验差比值越小,则说明灰色预测精度越高。一般后验差比值 C 值小于 0.35 则模型精度高,C 值小于 0.5 说明模型精度合格,C 值小于 0.65 说明模型精度基本合格,如果 C 值大于 0.65,则说明模型精度不合格。

从上表分析可以得到,后验差比值为0.283,模型精度高。

## 4.2.3 模型拟合分析

将建立的灰色预测模型与原有数据进行拟合,得到表9。

表 9 模型拟合结果表

索引项	原始值	预测值	残差	相对误差(%)
2015年	95.5	95.5	0	0
2016年	95.4	96.102	-0.702	0.736
2017年	95.7	95.718	-0.018	0.019
2018年	95.8	95.335	0.465	0.485
2019年	95.6	94.954	0.646	0.676
2020年	94.8	94.574	0.226	0.238
2021年	94.3	94.196	0.104	0.11
2022年	93.1	93.819	-0.719	0.773

计算得到模型平均相对误差仅为 0.38%, 意味着模型拟合效果良好。



图 5 模型拟合预测图

## 4.2.3 模型预测结果

表 10 模型预测结果表

预测阶数	预测值
1	93.444
2	93.071
3	92.699
4	92.328
5	91.959
6	91.591

7 91.225

故得到二阶预测结果,即 2024 年就业率(%)预测数据点:93.071得到 2024 年大学生就业环境预测的一个关键值。

#### 4.2.4 KMO 和 Bartlett 的检验

要预测 2024 年就业环境的其他关键指标,首先我们在问题一的基础上,针对问题一中收集到的数据类别过多,分类过细不易处理的问题,运用主成分分析法对原有数据进行降维处理。在实际过程中,我们分别对 GDP,价格指数,行业增加值进行了降重聚合处理。下述过程以 GDP,即国内生产总值(万亿元)、第一产业增加值(千亿元)、第二产业增加值(万亿元)、第三产业增加值(万亿元)、第三产业增加值(万亿元)。

下表展示了 KMO 检验和 Bartlett 球形检验的结果,用来分析是否可以进行主成分分析。

表 11		
KMO 检验和 Bartlett 的检	验	
KMO 值		0.49
	近似卡方	183.235
Bartlett 球形度检验	df	6
	P	0.000***

注: \*\*\*、\*\*、\*分别代表 1%、5%、10%的显著性水平

KMO 检验的结果显示, KMO 的值为 0.49, 同时, Bartlett 球形检验的结果显示, 显著性 P 值为 0.000\*\*\*, 水平上呈现显著性, 拒绝原假设, 各变量间具有相关性, 主成分分析有效, 程度为极不适合。

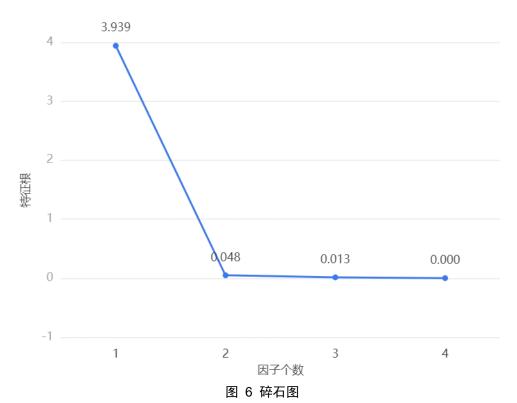
#### 4.2.5 确认主成分数量

通过分析方差解释表格和碎石图,确定主成分的数量方差解释表格。表 12 为总方差解释表格,主要是看主成分对于变量解释的贡献率(可以理解为究竟需要多少主成分才能把变量表达为 100%),一般都要表达到 90%以上才可以,否则就要调整因子数据。一般情况下,方差解释率越高,说明该主成分越重要,权重占比也应该越高。

在下表中,在主成分2时,总方差解释的特征根低于1,变量解释的贡献率达到99.678

表 12 方差解释表格						
成分	特征根					
)JX;/J	特征根	方差解释率(%)	累积方差解释率(%)			
1	3.939	98.484	98.484			
2	0.048	1.194	99.678			
3	0.013	0.322	100			
4			100			

根据各主成分对数据变异的解释程度绘制碎石图如图 3,根据特征值下降的坡度,结合方差解释表来确认需要选择的主成分个数。



# 4.2.6 求解因子载荷

	表 13	
	因子载荷系数表	
	因子载荷系数	共同度(公因子方差)
	主成分1	共四及(公凶丁万左)
国内生产总值(万亿元)	0.999	0.997
第一产业增加值(千亿元)	0.982	0.964
第二产业增加值(万亿元)	0.994	0.988
第三产业增加值(万亿元)	0.995	0.99

图 4 为载荷矩阵热力图,可以分析到每个主成分中隐变量的重要性。并进行各因子的 隐变量分析。

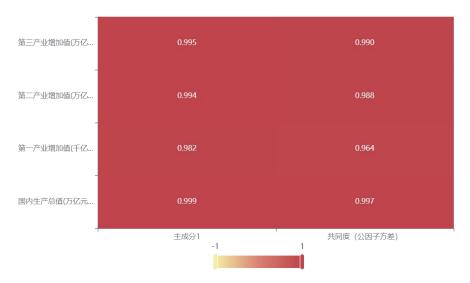


图 7 因子载荷矩阵热力图

## 4.2.7 求解成分矩阵

表 14 成分矩阵表

~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~	
成分矩阵表	
名称	成分 成分 1
国内生产总值(万亿元)	0.253
第一产业增加值(千亿元)	0.249
第二产业增加值(万亿元)	0.252
第三产业增加值(万亿元)	0.253

表 14 为成分矩阵表, 意在说明各个成分的所包含的因子得分系数(主成分载荷), 用于计算出成分得分, 得出因子公式, 其计算公式为: 线性组合系数\*(方差解释率/累积方差解释率), 最后将其归一化即为因子权重得分。

故得到模型的公式:

F1=0.253×国内生产总值(万亿元)+0.249×第一产业增加值(千亿元)+0.252×第二产业增加值(万亿元)+0.253×第三产业增加值(万亿元)

且由上可以得到:

F=F1

#### 4.2.8 主成分分析法降维合并后的结果

合并后的 GDP (代表)=0.253×国内生产总值(万亿元)+0.249×第一产业增加值(千亿元)+0.252×第二产业增加值(万亿元)+0.253×第三产业增加值(万亿元)

合并后的行业增加值指数=0.123×农林牧渔业增加值指数(上年=100)+0.2×工业增加值指数(上年=100)+0.028×建筑业增加值指数(上年=100)+0.203×批发和零售业增加值指数(上年=100)+0.196×交通运输、仓储和邮政业增加值指数(上年=100)+0.185×住宿和餐饮业增加值指数(上年=100)+0.014×金融业增加值指数(上年=100)+0.118×房地产业增加值指数(上年=100)+0.166×其他行业增加值指数(上年=100)

合并后的价格指数=0.339×居民消费价格指数(上年=100)+0.335×城市居民消费价格指数(上年=100)+0.334×农村居民消费价格指数(上年=100)

在此基础上,我们选择最佳指标即月平均工资,同样将其放入到灰色预测模型 GM (1,1)中进行处理,最终得到 2024 年月平均工资预测值是 6825.4 元。

于是我们得到了对 2024 年就业环境预测的两个关键值, 即就业率 93.071%和月平均工资水平 6825.4 元。

## 4.3 问题三的模型建立与求解

在参考大量机构通过问卷调查法等方式得出的高校学生竞争力评价模型的基础上,我们基于猎聘大数据,智联招聘等专业机构近三年的就业数据统计分析,首先将所有数据处理为统一的描述性统计结果。在此基础上,我们进一步将在校学生就业竞争力划分为实践与学识、专业职业技能、人际关系、职业素养四个维度,并借助 CRITIC 权重法进行衡量。

## 4.3.1 描述性结果统计

表 15 描述性统计结果

变量名	最大值	最小值	均值	标准差	中位数	方差
对专业知识的掌握程度	5	3	4.665	0.538	5	0.289
大学期间的实践经验支撑表现	5	3	4.665	0.538	5	0.289
如何	3	3	4.005	0.550	3	0.203
学术背景对未来职业发展的积	5	3	4.632	0.522	5	0.273
极影响						
在专业领域发展规划的清晰度	5	3	4.639	0.495	5	0.245
解决复杂问题的能力及创新能	5	3	4.652	0.554	5	0.306
力					-	
大学期间的实习经验对进入职	5	3	4.665	0.500	5	0.250
场的支撑度						
是否具备与他人合作的团队能	5	3	4.619	0.538	5	0.289
<b>力</b>						
是否掌握了必要的职业技能以	5	3	4.645	0.531	5	0.282
胜任自身所选择的行业						
能否熟练使用与自身专业相关	5	3	4.600	0.565	5	0.319
的工具与技术						
沟通表达能力能否在职场中脱	5	3	4.671	0.498	5	0.248
颖而出						
是否具备建立和维护良好人际 关系的能力	5	3	4.626	0.524	5	0.275
面对工作的挑战和冲突 ,能否有						
效解决	5	3	4.639	0.533	5	0.284
双 <i>件(</i> 能否在团队中发挥协调与领导						
作用	5	3	4.626	0.536	5	0.288
对多元文化环境中适应能力的						
评价	5	3	4.574	0.546	5	0.298
人际关系技巧在职场中的竞争	5	3	4.581	0.557	5	0.310
	3	3	1.501	0.557	<b>J</b>	0.510

是否具备良好的职业操守和道	5	3	4.600	0.542	5	0.294
德标准	3	3	4.000	0.542	3	0.254
能否有效地进行职业规划及目	5	3	4.632	0.535	5	0.286
标设定	5	3	4.032	0.555	5	0.280
能否在工作中有效地管理时间	5	3	4.613	0.539	5	0.291
和任务	5	3	4.013	0.559	5	0.291
是否对行业动态和职业发展趋	5	3	4.658	0.515	5	0.265
势保持关注	5	3	4.036	0.515	5	0.205
是否具备在职场中持续学习及	5	2	4 710	0.492	-	0.222
适应环境变化的能力	Э	3	4.710	0.483	5	0.233

#### 4.3.2 相关性分析

通过 Spearman 相关性分析探究了实践与学识、专业职业技能、人际关系、职业素 养这四个关键维度之间的关系。如表 16 所示,四个维度在 0.01%的水平上呈显著正 相关关系,毕业生在实践经验及学术方面的能力与其在专业职业技能、人际关系、职业 素养方面的表现密切相关。

实践与学识\*\* 专业职业技能 人际关系\*\* 职业素养\*\*\* \* \* \* 实践与学识 1.000 0.842 0.829 0.805 专业职业技能 0.842 1.000 0.884 0.847 人际关系 0.829 0.884 1.000 0.856 职业素养 0.805 0.847 0.856 1.000

表 16 相关性分析

注: \* \* \* 代表 0.01 的显著性水平

#### 4.3.2 权重法衡量

为更好地区分各维度在毕业生就业竞争力中的重要程度,借助 CRITIC 权重法进行 衡量。如表 17 所示,实践与学识维度在所有维度中权重最大(28.429%),说明实践 经验与学术知识对毕业生的就业竞争力具有至关重要的影响,毕业生通过积累实践经验 及深刻理解专业知识能够更好地适应职场环境。专业职业技能维度权重占总权重的 24. 331%,说明在职业环境中掌握一门专业技术能够有效提高毕业生的就业竞争力。 职业素养维度权重占总权重的 24. 292%, 说明培养毕业生就业竞争力时注重提高其职

业素养同样重要。尽管人际关系维度的权重最小,但依然占有22.948%的比重,说明良 好的人际关系能够帮助高校毕业生更好地融入职场,促进团队协作与职业发展。

衣 17 仗里纪未						
项	指标变异性	指标冲突性	信息量	权重/%		
实践与学识	2. 229	0.602	1.341	28. 429		
专业职业技能	2. 205	0. 521	1.148	24. 331		
人际关系	2. 266	0.478	1.083	22.948		
职业素养	2. 147	0.534	1.146	24. 292		

丰 17 切香姓里

#### 4.4 问题四建议及解决方案

#### 4.4.1增加就业的建议

综合问题一、二、三,由问题三的定量分析可知,大学生毕业生的个人素质方面在不断提高。而分析问题一、二,可知导致大学生就业难的根本原因在于供需的不平衡,主要体现为行业需求与现有应聘人才的不匹配,行业缺口与大学生毕业生期望就业的不匹配。

- 1. 优化高等教育结构:政府应引导高校根据行业需求调整专业设置和课程内容,确保教育与市场需求紧密对接。鼓励高校与行业合作,开展实习实训项目,提升学生实践能力和职业素养。
- 2. 加强职业教育与培训:加大对职业教育的投入,提供更多与市场需求相匹配的职业技能培训。鼓励企业参与职业教育,共同培养符合行业需求的高素质人才。
- 3. 完善就业政策与服务:制定更加灵活的就业政策,鼓励大学生到二三线及以下城市、农村地区就业。提供针对性的就业指导服务,帮助大学生了解行业动态、明确职业定位。
- 4. 促进产学研深度融合:支持高校、科研机构与企业之间的产学研合作,推动科技创新成果转化。鼓励企业在高校设立研发中心或实验室,为学生提供更多实践机会。
- 5. 引导大学生转变就业观念:通过宣传、教育等方式,引导大学生树立正确的就业观,鼓励其到基层、到一线锻炼成长。提倡多元化就业,鼓励大学生自主创业、灵活就业。
- 6. 加强行业与人才需求预测:建立行业与人才需求预测机制,定期发布行业发展趋势和人才需求报告。根据预测结果,及时调整高等教育和职业教育的人才培养方向。
- 7. 搭建校企合作平台: 政府应搭建校企合作平台, 促进企业与高校之间的信息交流、 人才共享。举办定期的校企对接活动, 让大学生有更多机会了解企业需求、展示自己。

### 4.4.2 提高就业竞争力的解决方案

首先,针对2024年的行业需求情况,可以得到以下结论:

1. 互联网行业对毕业生需求遥遥领先

IT/互联网/游戏行业毕业生需求旺盛,近年来一直处于各行业的首位,释放 20%-25%的毕业生职位需求。2024年 Q1 新发校招职位占 比前五的行业分别为 IT/互联网/游戏、电子/通信/半导体、专业服务、金融、机械/制造,共释放 56.8%的需求。

表 18			
近三年 IT/互联网/游戏行业新发校招职			
位占比			
2021	24.70%		
2022	21. 20%		
2023	21.90%		

表 19 2024 年新发校招职位行业分布

IT/互联网/游戏	23.90%
电子/通信/半导体	9.30%
专业服务	8.60%
金融	7.60%
机械/制造	7.40%
医疗健康	6.90%
消费品	6.00%

能源/化工/环保	6.00%
房地产/建筑	5.30%
汽车	5.00%
交通/物流/贸易/零售	4.30%
政府/非营利组织/其他	2.90%
广告/传媒/文化/体育	2.50%
教育培训/科研	2.30%
生活服务	2.00%

## 3. 制造业对毕业生需求稳中有升



图 8

## 4. 新质生产力需要更多硕博人才

不同行业对于不同学历毕业生需求有一定的差异性,对于硕博人才需求较多的行业 主要为新赛道,与发展新质生产力高度相关的行业,如电子/半导体/集成电路、互联网、 新能源、人工智能等。

## 5. 发达城市释放大量需求

经济发展是创造就业机会的根本动力,经济强市 释放更多毕业生就业机会。从新发校招职位来看,2024年新发校招职位城市 TOP10 同样是 GDP 十强城市。

表 20 2024GDP 十强城市				
排名		城市	2024年 Q1GDP (亿 元)	同比增速(%)
	1	上海	11098.46	5. 00%
	2	北京	10581.40	6.00%
	3	深圳	8314.98	6.40%
	4	重庆	7232.03	6. 20%
	5	广州	7161.14	3.60%

17

6	苏州	5549.00	7.90%
7	成都	5518.20	5.00%
8	杭州	5112.00	5. 10%
9	武汉	4532.16	5.60%
10	南京	4359.56	3.80%

### 6. 重点城市群机会丰富

城市群发展创造就业新机遇。从新发校招职位来看,相较于京津冀、粤港澳、成渝等重点城市群,长三角释放毕业生需求最多,超过三分之一。京津冀和粤港澳新发校招职位呈现增长趋势,释放越来越多高校毕业生需求。

#### 7. 毕业生就业机会有下沉趋势

2024年新发校招职位集中在一线和新一线城市,占比分别为 36.96%、33.20%。对比来看,一线和新一线新发校招职位占比有所下降,其他城市新发校招职位占比有所上升。从新发校招职位数量同比变化来看,一线和新一线城市新发校招职位数处于下降状态,其他城市新发校招职位数有所上升。

#### 具体建议为:

- 1. 学历方面的建议:明确目标定位:根据自身的兴趣、专业背景和职业规划,明确学历提升的目标和方向。是选择继续攻读本专业的研究生,还是跨专业深造?是追求学术型硕士还是专业型硕士?明确的目标定位有助于更加高效地制定学习计划和职业规划。
- 2. 制定合理计划:根据目标定位,制定切实可行的学习计划。合理安排时间,确保既能完成学业任务,又能兼顾个人生活和兴趣爱好。同时,注重培养自主学习能力和时间管理能力,为未来的学习和工作打下坚实的基础。
- 3. 关注行业动态与市场需求: 在学历提升的过程中,密切关注行业动态和市场需求的变化。了解未来就业市场的趋势和热点领域,有助于更加精准地选择学习方向和职业路径。同时,也可以通过实习、兼职等方式积累实践经验,提升个人综合素质和就业竞争力。
- 4. 利用资源拓展人脉:在学历提升的过程中,积极利用学校、社会等资源拓展人脉。参加各类学术活动、社团活动和社会实践等,结识志同道合的朋友和行业专家。这些人脉资源不仅有助于个人成长和发展,还能为未来的求职和职业发展提供有力的支持。
- 5. 加强专业知识的学习和实践,确保自己在专业领域具有扎实的技能和经验。
- 培养跨领域的技能和兴趣爱好,增加自己的综合素质和适应能力。
- 积极参与校内外的实习、志愿服务、社团活动等,积累实际经验和展示自己的能力。
- 6. 灵活调整职业规划:根据就业市场的变化,灵活调整自己的职业规划。不要仅局限于某一特定行业或职位,而是要保持开放的思维,探索更多的可能性。
- 考虑创业或自由职业等非主流职业路径,利用互联网和新技术创造自己的就业机会。
- 7. 利用网络和社交媒体: 充分利用互联网和社交媒体平台,关注行业动态、企业招聘信息和职业发展资源。建立和维护个人职业网络,与校友、行业专家、招聘人员等保持联系,获取更多的职业信息和机会。
- 8. 准备充分的求职材料:精心准备简历和求职信,突出自己的教育背景、技能、经验和成就。根据不同职位和公司的要求,定制化求职材料,展示自己与职位的匹配度和价值。9. 保持积极心态和持续学习:面对就业市场的挑战和变化,保持积极的心态和乐观的态度。相信自己的能力和价值,不断寻找和创造机会。持续学习和提升自己的能力,关注新技术、新知识和新技能的发展,保持自己的竞争力和适应力。

# 五、参考文献

- [1] 梁北辰, 戴景民. 偏最小二乘法在系统故障诊断中的应用[J]. 哈尔滨工业大学学报, 2020, 52(3):156-164. DOI:10.11918/201805149.
- [2] 邓聚龙. 灰色预测与灰决策[M]. 武汉: 华中科技大学出版社, 2002.
- [3] 何晓群. 多元统计分析. 北京: 中国人民大学出版社, 2012.
- [4] 刘高见,刘伦华,王有斌. 影响高校毕业生就业的因素和对策研 究 [J]. 中外企业家,2019(07):175 176.
- [5] 唐诗潮. 当前我国大学生就业问题及对策研究[J]. 高教学刊, 2018(16): 194 196.

# 附录 附录一 支撑文件列表 问题一 Q1.xlsx 问题二 Q2. x1sx 增加值指数.xls 近十年 GDP. x1sx 近十年大学生毕业人数.xlsx GDP+价格指数+人数.x1sx 问题三 2024 高校毕业生就业数据报告一精准对接制胜未来-猎聘-202406-报告汇. pdf 2024年大学生就业力调研报告-智联招聘-202405-报告汇.pdf 2024年新发校招职位 TOP10 城市. x1sx 2024年新发校招职位行业分布.x1sx 附录二 程序代码 问题二(数据处理.py) import numpy as np import pandas as pd import matplotlib.pyplot as plt #灰色预测模型 GM(1,1) class gml1: def init (self): self.a = 0 # 灰色作准指数 self.b = 0 # 灰色作准常数 self.data0 = None # 初始数据序列 self.data cum = None # 累加数据序列 def fit(self, data): self. data0 = datan = 1en(data)self.data cum = np.cumsum(data) # 累加生成数列 Z = 0.5 \* (self.data\_cum[:-1] + self.data\_cum[1:]) # 紧邻均值生成数 列

B = np. vstack((-Z, np. ones(n - 1))). T

```
Y = data[1:]
        U = np. dot (B. T, B)
        self.a, self.b = np.dot(np.linalg.inv(U), np.dot(B.T, Y)) # 最小二
乘估计参数
        return self.a, self.b
    def predict(self, n):
        x \text{ predict} = np. zeros(n)
        x predict[0] = self.data0[0]
        for i in range (1, n):
             x \text{ predict}[i] = (\text{self. data0}[0] - \text{self. b} / \text{self. a}) * (\text{np. exp}(-\text{self. a})
* i)) + self.b / self.a
        x_pred_ress = np. zeros(n)
        x pred ress[0] = x predict[0]
        for i in range (1, n):
             x pred ress[i] = x predict[i] - x predict[i - 1]
        return x_pred_ress
df = pd. read_excel("就业率. xlsx")
sim = []
plt.rc("font", family='Microsoft YaHei')
plt. figure (figsize=(10, 28))
for i in range(1, len(df.columns)):
    n=df.columns[i]
    data = np. array(df[df. columns[i]]. values)
    model = gm11()
    a, b = model.fit(data)
    sim.append((df.columns[i],a, b))
    m = 3 # 预测未来3个数据点
    X pred future = model.predict(len(data) + m)[len(data):]
    if n=="就业率(%)":
        for i in range (0, m):
            print(f"{n}预测第{i + 1}个数据点: {X pred future[i]:.2f}")
    k=["2023年","2024年","2025年"]
    plt.plot(df[df.columns[0]].values, data, 'o-', label=f' {n}Original')
                    X pred future, 'x--', label=f' {df.columns[i]}Future
    plt. plot (k,
Prediction')
    plt. ylim(20, 200)
plt.legend()
plt. show()
```

```
a_std=sim[0][1]
b_std=sim[0][2]
for i in range(len(sim)):
    for j in range(len(sim)-1,i,-1):
        if sim[j-1][1]>sim[j][1]:
            sim[j-1], sim[j]=sim[j], sim[j-1]

print("就业率影响因素相关度", end=":")
for i in range(len(sim)-2,0,-1):
    print(sim[i][0], end=">")
print(sim[0][0])
```