

doi:10.16055/j.issn.1672-058X.2015.0009.019

基于 AHP 的大学生就业信心指数组合预测*

周润娟, 蔡金平, 胡长新

(安徽工程大学 电气工程学院, 安徽 芜湖 241000)

摘要:大学生就业信心指数预测结果的准确性关系到就业政策的制定与实施的效果.提出综合利用各单模型预测信息的组合预测思路,采用层次分析法确定单模型的权重值,构建组合预测模型.结果显示,组合预测模型在拟合期的表现与神经网络模型接近,优于其他两种模型;在预测期远超过其他模型的预测效果.组合预测模型的拟合性能和泛化性能优越,预测信息可作为高校制定相关政策时的重要参考依据.

关键词:大学生就业;信心指数;组合预测;层次分析法

中图分类号:TP391

文献标志码:A

文章编号:1672-058X(2015)09-0076-05

准确掌握大学生就业信心情况,有助于及时把握大学生思想动态,为高校制定人才培养方案和应对就业问题,提供科学依据,具有较强的现实意义^[1,2].据人力资源和社会保障部最新统计资料显示,2014 年高校毕业生的就业规模将达到 727 万,比 2013 年增加 28 万人,再创历史新高.运用科学方法及时预测出大学生的就业信心,对高校学生工作的开展尤为重要^[3].

严春红等借鉴经济领域中的消费信心指数,编制了浙江省高职院校的调查问卷,测度出大学生的就业信心指数,结果显示浙江省高职生的整体就业信心指数偏低^[4].杨光军等采用问卷调查方式研究了山东省德州市 3 所高校的就业信心指数,并用灰色模型法构建了基于时间序列的就业信心指数预测模型,可为高校制定长远计划提供参考依据^[5].然而在文献^[5]中,所构建的灰色模型仅对已有数据进行了拟合,缺少对模型的检验环节,模型的预测性能尚未得出.根据金菊良等的研究成果显示^[6],灰色模型法对具有趋势性的时间序列预测性能较好,但对周期性数据泛化能力较差,不如自回归法和神经网络法.

此处提出综合灰色模型法的趋势性预测性能和自回归法、神经网络法的周期性预测性能,利用层次分析法(AHP)求解 3 种预测模型的权重,构建出基于 AHP 的大学生就业信息指数组合预测模型(Combination Forecasting based on AHP method,CF-AHP).

1 CF-AHP 模型构建

CF-AHP 模型主要思路是组合多个模型的预测性能,提高单模型的适用性.选用具有代表的自回归法(TAR)、神经网络法(BP)和灰色模型法(GM),分别运行预测模型,根据层次分析法(AHP)计算出单模型的权重值,组合得到组合预测模型结果.因此,CF-AHP 模型的构建共包括如下 6 个步骤.

步骤 1: 数据预处理.首先,需要对原始数据(X_1, X_2, \dots, X_n)进行无量纲化,消除量纲效应.

$$X_i = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}}, i = 1, 2, \dots, n \quad (1)$$

收稿日期:2015-03-08;修回日期:2015-04-20.

* 基金项目:教育部人文社会科学研究专项任务项目(12JDSZ3041);安徽高校省级科学研究项目(2011SK637);安徽工程大学青年科研基金项目(2013YQ38)

作者简介:周润娟(1984-),女,安徽太湖人,硕士,助教,从事大学生就业与环境意识教育研究.

其次,根据待预测时间序列的数据分布特征,确定模型输入值、输出值的个数,构建样本对 $(X_{n-k}, \dots, X_{n-1}, X_n)$,其中 X_n 为待预测年份的输出值, X_{n-k}, \dots, X_{n-1} ,表示用前 k 年数据预测第 n 年的数据.模型构建过程,一般可应用数理统计中的自相关法,确定 k 值,考虑到大学生就业信心指数数据的小样本特点,实际使用中, k 值可直接取 2~4.

步骤 2: 自回归模型预测.自回归模型(TAR)能有效地描述具有周期性、跳跃性、相依性等复杂现象的非线性动态系统,非常适合处理具有周期性的系统预测问题.其基本思想:在观测时序 $\{X(i)\}$ 的取值范围内引入 $L-1$ 个门限值 $(r(j), j=1, 2, \dots, L-1)$,将该范围分成 L 个区间,并根据延迟步数 k 将 $\{X(i)\}$ 按 $\{X(i-k)\}$ 值的大小分配到不同的门限区间内,再对不同区间内的 $X(i)$ 采用不同的 AR 模型来描述,这些 AR 模型的总和完成了对时序 $\{X(i)\}$ 整个非线性动态系统的描述.模型应用时需确定延迟步数 k 、门限个数 L ,及各个 AR 模型的系数,可用最小二乘法或遗传算法等优化算法求解.限于篇幅,此处不再展开,可参考文献[6].

步骤 3: 神经网络模型预测.BP 神经网络是用 BP 算法训练的一种多层前馈型非线性映射网络,网络中各神经元接受前一级的输入,并输出到下一级,网络中没有反馈联接^[7].BP 神经网络通常可以分为不同的层(级),第 j 层的输入仅与第 $j-1$ 层的输出联接.BP 算法是目前应用最为广泛且较成功的一种算法,在各行各业都有着广泛的应用,适合预测样本中含有最大、最小值的插值式预测,对趋势性数据的外延性能较差.BP 神经网络在应用时,需确定隐层节点个数,各层之间的连接系数可通过 BP 网络算法求解.

步骤 4: 灰色系统模型预测.灰色系统预测模型来源于邓聚龙等提出的灰色系统,即用微分方程描述事物发展的连续过程,实现预测未来发展的目的,因此适合于预测具有趋势性特征的时间序列数据.杨光军等基于大学生就业信心指数数据,建立了 GM 预测模型,模型运行结果显示,灰色系统模型可准确描述就业信心指数序列.与步骤 2、3 类似,对灰色系统模型预测更具体的内容可参考文献[5].

步骤 5: AHP 法计算单模型权重.AHP 算法是一种定性与定量相结合的决策分析方法,通过将复杂问题分解为若干层次和若干因素,可采用专家打分法,构建出层次结构^[8],如图 1 所示.CF-AHP 模型构建中,首先对 TAR、BP、GM 模型进行两两比较,在 1-9 标度中根据打分结果得到两两比较的判断矩阵,进而按照方根法求解各模型权重和一致性比例 CR,若 $CR < 1.0$,表示据此判断矩阵计算出的权重值较为可靠,否则需重新打分确定判断矩阵,直到获得可靠的 CR 为止.

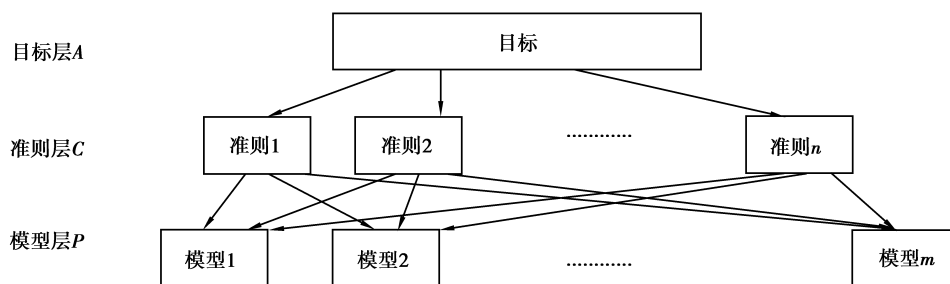


图 1 层次分析法的结构

步骤 6: 组合模型预测.假设对同一就业信心指数有 m 个预测模型分别进行预测,记实际观测值为 $y_d(t)$,第 i 个预测模型的预测值为 $y(i, t)$,第 i 个预测模型的权重为 $w(i)$.其中,时刻 $t=1 \sim n$, n 为样本容量;模型序号 $i=1 \sim m$; $w(i) \geq 0, \sum_{i=1}^m w(i) = 1$.则组合预测模型的预测值为

$$c(t) = \sum_{i=1}^m w(i)y(i, t) \tag{2}$$

组合预测模型的预测误差绝对值为

$$e(t) = |c(t) - y_d(t)| \tag{3}$$

根据组合模型预测结果,可与拟合期的 TAR、BP、GM 模型预测结果进行比较,分析组合 3 种单模型的 CF-AHP 模型的拟合性能.然后,用没有参加拟合计算的数据样本,分别检验 4 种模型的预测性能,得出更为

可靠的大学生就业信心指数预测结果,进而科学指导高校政策制定实践.

2 实例应用

2.1 数据来源

为便于比较,以文献[5]中采集的山东省德州市 3 所高校的就业信心指数数据为例,验证 CF-AHP 模型的性能,预测大学生就业信心的变化趋势.文献[5]的就业信心指数位于 0~200 之间,数值越小表示越没信心,越大表示越有信心,数值 100 为“消极”与“积极”的临界值.图 2 显示 2000-2012 年大学生就业信心指数与经济景气指数的关系.

由图 2 可知,就业信息指数与经济景气指数的变化方向不一致,经济景气指数高的时候,就业信心指数相反处于低位,呈负相关关系,相关系数 $R=-0.254$,且就业信息指数的波动性明显高于经济景气指数的,说明仅用经济景气指数预测就业信心指数远远不够,可从就业信心指数的周期性出发加以预测.

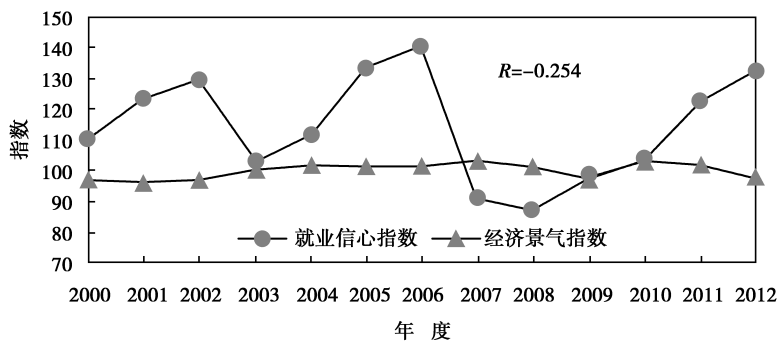


图 2 大学生就业信心指数与经济景气指数的关系

2.2 单模型预测

2.2.1 TAR 模型预测

由于样本数据较少,模型的延迟步数 $k=2$,TAR 模型的门限个数 $L=1$.根据最小二乘法计算结果,门限阈值为 106.60,整理得到预测模型如下:

$$X(t) = \begin{cases} 0.54 + 1.25X(t-1) - 0.18X(t-2), & X(t-1) \leq 106.60 \\ 7.05 + 3.10X(t-1) - 2.46X(t-2), & X(t-1) > 106.60 \end{cases} \quad (4)$$

模型拟合期预测结果如表 1 中第 Y_{TAR} 列所示,拟合期误差绝对值和为 92.75.分别用 2009、2010 年数据预测 2011,2010、2011 年数据预测 2012 年,对就业信心指数进行检验.预测结果显示,2011 年实测数据为 122,预测结果为 111;2012 年实测数据为 132,预测结果为 131.4,说明所建立的 TAR 模型在拟合期和检验期都有很好的预测性能,可作为组合预测模型的输入模型.

2.2.2 BP 模型预测

BP 模型网络结构中,输入层节点个数为 2,隐层节点个数取 3,输出层节点个数为 1,将 2000-2010 年数据输入 BP 网络中,2011,2012 年数据用作检验.预测结果如表 1 中第 Y_{BP} 列所示,拟合期误差绝对值和为 34.93.2011,2012 年的检验结果显示,2011 年实测数据为 122,预测结果为 129;2012 年实测数据为 132,预测结果为 139.3,说明所建立的 BP 模型在拟合期和检验期也有较好的预测性能,可作为组合预测模型的输入模型.相对于 TAR 模型,BP 模型在拟合期的表现更加优秀,但检验期的表现不如 TAR 模型,这与 BP 模型的拟合性能较强,泛化能力较弱有关.

表 1 CF-AHP 组合模型法预测结果误差分析

时间/年份	X_{n-2}	X_{n-1}	实测	Y_TAR		Y_BP		Y_GM		Y_AHP		
				输出	误差	输出	误差	输出	误差	输出	误差	
2002	110	123	129	117.2	11.77	129.7	0.67	134.4	5.42	127	2.05	
2003	123	129	103	103.8	0.79	100.3	2.68	125.9	22.86	103.9	0.89	
2004	129	103	111	105.3	5.7	96.3	14.71	93.8	17.22	98.4	12.64	
2005	103	111	133	97.3	35.71	132.9	0.13	124.9	8.13	122.8	10.16	
拟合期	2006	111	133	140	145.8	5.76	138.4	1.64	136.2	3.78	140	0.05
	2007	133	140	91	113.2	22.25	96.9	5.86	120.5	29.53	103.6	12.57
	2008	140	91	87	88.3	1.31	95.6	8.57	86.4	0.62	92.7	5.73
	2009	91	87	98	92.3	5.66	97.6	0.41	99	1.04	96.4	1.62
	2010	87	98	103	106.8	3.8	102.7	0.26	115.6	12.62	105.1	2.13
拟合误差绝对值和					92.75	34.93	101.2	47.84				
检验期	2011	98	103	122	111	10.99	129	7.03	108.9	13.11	122.3	0.27
	2012	103	122	132	131.4	0.62	139.3	7.34	123.9	8.1	135.7	3.67
	预测误差绝对值和					11.61	14.37	21.20	3.94			

注: Y_TAR 表示门限自回归法的模型输出值, Y_BP 对应神经网络法, Y_GM 对应灰色模型法, Y_AHP 对应此处构建的组合模型法, 其中 Y_GM 模型输出值源自文献[5].

2.2.3 灰色系统模型预测

表 1 中第 Y_GM 列出了文献[5]中灰色系统模型的预测结果, GM 模型在拟合期、检验期的误差绝对值和均大于相应的 TAR, BP 模型, 说明 GM 模型在应对具有周期波动性的大学生就业信心指数预测中, 表现不佳, 与灰色系统模型适用于描述趋势性特征有关.

2.3 模型权重计算

采用专家打分法, 对 TAR, BP, GM 模型的性能进行两两比较, 得到 1-9 标度的判断矩阵:

$$\begin{matrix}
 & \text{TAR} & \text{BP} & \text{GM} \\
 \text{TAR} & 1 & \frac{1}{3} & 3 \\
 \text{BP} & 3 & 1 & 5 \\
 \text{GM} & \frac{1}{3} & \frac{1}{5} & 1
 \end{matrix}$$

应用方根法计算出 3 种模型的权重值 $W = \{0.258, 0.637, 0.105\}$, 为了检验权重值的可靠性, 进一步计算判断矩阵的一致性. 根据权重值 W , 计算得最大特征 $\lambda_{\max} = 3.04$, 判断矩阵的一致性系数 $C.I. = (\lambda_{\max} - n) / (n - 1) = (3.04 - 3) / (3 - 1) = 0.0193$, 其中 $n = 3$ 为判断矩阵的阶数. 为消除矩阵阶数对一致性系数的影响, 还需考虑判断矩阵阶数对应的平均随机一致性系数 $R.I. (n)$, 取 $n = 3$, 查 AHP 法计算表^[8]可得 $R.I. (3) = 0.58$, 因此求出一致性比例 $C.R. = C.I. / R.I. = 0.0193 / 0.58 = 0.033 < 0.1$, 说明判断矩阵具有较好的一致性, 据此求解出的权重值是可靠的.

2.4 CF-AHP 组合预测

将 3 种模型 ($i = 1, 2, 3$) 的拟合期输出值 $y(i, t)$ 和权重值 $w(i)$ 代入式(2)中, 得到组合预测模型的拟合期输出值, 列入表 1 中第 Y_AHP 列. CF-AHP 模型的拟合期误差绝对值和 47.84 劣于 BP 模型的 34.93, 但优于 TAR 和 GM 模型的 92.75 和 101.2. 说明组合 3 种模型后的 CF-AHP 具有与 BP 模型接近的优秀拟合性能.

将 3 种模型的检验期 ($t = 2011, 2012$) 输出值和权重值代入式(2)中, 求出 CF-AHP 模型的检验期预测

值,其误差绝对值和仅为 3.94,远小于 GM 模型的 21.2 和 BP 模型的 14.37,也优于 TAR 模型的 11.61,显示出较强的泛化预测能力.

组合模型与单模型预测结果比较如图 3 所示(其中 2000-2010 年为拟合期,2011、2012 为检验期).进一步说明,综合多种单模型的预测信息后,CF-AHP 模型具有更加强大的适用性,尤其适合于预测大学生信心指数,或类似的时间序列预测问题.

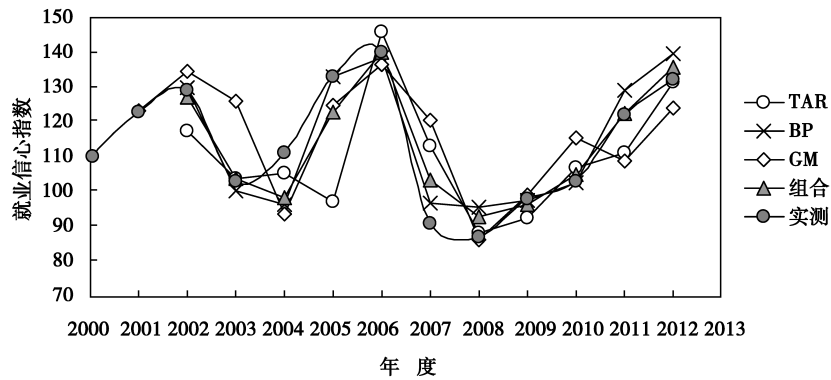


图 3 组合模型与单个模型预测结果比较

2.5 预测结果讨论

从图 3 的 CF-AHP 模型预测结果可知,虽然近年就业压力较大,但学生对就业的信心仍较高,且呈逐年上升趋势.这与国家层面对大学生就业问题的重视和制定的各种鼓励自主创业政策的实施,具有很大的关系.高校在面对当前的就业形势下,准确地获取大学生就业信心指数及其变化趋势,可及时调整学生工作重心,确保大学生对就业问题的信心.从图 3 可知,德州市的 3 所高校根据大学生就业信心指数,及时调整、制定了学生工作计划,将就业信心教育提升到大学生思想政治教育的层次上,同时多次举办各种职业规划和实践活动,解读国家和各级政府就业政策,将大学生的就业信心指数稳定在了上升趋势中.

3 结 论

单个预测模型在大学生就业信心指数预测中,表现不同,BP 模型具有较强的拟合能力,TAR 模型具有较强的泛化能力,GM 模型在就业信心指数预测中表现一般.

CF-AHP 组合预测模型,综合利用了单个预测模型的预测信息,极大地提升了模型的预测性能.CF-AHP 在拟合期的表现与 BP 模型接近,而在预测期则获得了远小于其他 3 种模型的预测误差,显示出强大的泛化预测能力.

大学生就业信心指数的预测结果,可作为高校制定学生工作计划的依据.根据预测结果,及时调整工作方向,可进一步地稳定大学生的就业信心.

参考文献:

- [1] 易忠, 张明平. 西部大学生就业模式研究[J]. 重庆工商大学学报:自然科学版, 2013, 30(6):29-31
- [2] 裴菁. 上海市大学生就业区域流向的实证研究[J]. 上海理工大学学报:社会科学版, 2014 (1):90-95
- [3] 周红霞. 大学生就业信心状况调查与对策分析[J]. 东北师大学报:哲学社会科学版, 2011(3):214-217
- [4] 严春红. 大学生就业信心指数的设计与分析[D]. 金华: 浙江师范大学, 2007
- [5] 杨光军. 灰色神经网络在大学生就业信心指数预测中的应用[J]. 计算机系统应用, 2013, 22(8): 190-193