

基于主成分分析法的高校 EHR 系统

邢德海, 董旭源

(天津大学管理学院, 天津 300072)

摘 要: 主成分分析法是系统分析中一种新型的数学方法, 将它应用于 EHR 系统的多个环节, 使人力资源管理的决策实现定性分析与定量分析的有机结合, 从而使许多因素问题得到相对科学的解决。天津大学的实践表明, 系统的应用实现了与高校人员相关的数据、过程和资源的集成化管理, 提高了管理效率。

关键词: 主成分分析; 人力资源管理; EHR

University EHR System Based on Main Component Analysis

XING Dehai, DONG Xuyuan

(School of Management, Tianjin University, Tianjin 300072)

【Abstract】 Main component analysis is a new mathematical method in system analysis. The application of main component analysis in EHR system induces effectively to combine qualitative and quantitative analysis in decision makers of human resource departments, which makes many multiple-factor problems be solved scientifically. The practicality of Tianjin University indicates that the application of EHR system integrating management of data, processes and resources, which are relative to personnel, improves the efficiency of human resource management, and can be improved.

【Key words】 Main component analysis; Human resource management; EHR

高校电子人力资源管理 (Electronic Human Resource, EHR) 系统用信息技术来改变传统人力资源管理模式, 以管理信息化促进教育现代化, 顺应了教育发展的必然趋势。主成分分析法是一种客观赋值法, 将其应用到教育评价领域中能够比较科学地解决权重的确定问题。例如高校教师的招聘与甄选问题就涉及到多种因素和标准, 因而用主成分分析法可以发挥定性与定量分析的双重优势, 使高校教师评聘问题走向科学化。

1 高校 EHR 系统的需求分析

1.1 基于 Web 的 EHR 的整体框架

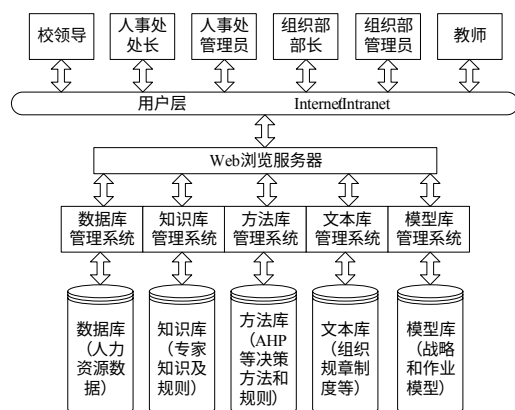


图1 基于WEB的高校EHR系统框架

基于 Web 的 EHR 系统是利用网络环境的人力资源管理系统, 同以往的人事管理系统存在本质的差异, 是对以往人事系统的合理修正和提升, 具体表现在如下几个方面: (1)EHR 系统融合了专家知识系统、系统分析与评价、知识管理、模式识别及系统集成等技术和理论; (2)EHR 系统具备了

较好的决策支持功能; (3)EHR 系统能根本地改变组织的人力资源管理模式, 提供教职工参与管理的平台, 组成学校高层、管理人员、教师的三维管理网, 提高组织的人力资源管理效率。

基于 Web 的高校 EHR 系统的用户层分为校领导、人事处处长、人事处管理人员、组织部部长、组织部管理人员、其它部处领导、教师等。高校 EHR 系统由数据库及其管理系统、模型库及其管理系统、文本库系统、知识库系统、方法库系统等 5 个子库组成的系统^[1]。EHR 系统框架结构见图 1。

1.2 高校 EHR 系统的体系结构和功能

EHR 系统能为学校在不同时期提供不同的管理策略和功能, 主要涉及到学校的人力资源规划、招聘与甄选策略、组织机构变革、绩效管理、薪资模式及用人政策规章制度、人事档案管理、素质测评、人力资本、知识管理、质量体系标准等方面。高校 EHR 系统体系结构根据其功能可以大致划分为如下模块^[2,3], 其体系结构如图 2 所示。

(1) 人力资源规划。根据高校的发展战略规划、组织需求以及职务分析的要求, 进行人力资源管理的规划, 充分发挥所有教职员工的潜力, 使他们在现有的岗位或今后可能调动的岗位上得到发展, 这是人力资源管理为学校的远景发展规划所做的基础工作。

(2) 招聘和甄选。依据人力资源规划和职务分析的要求, 实施招聘工作。招聘一般分为外部招聘和内部竞聘。招聘计划确定后, 依据职务说明书, 通过各种途径发布招聘信息, 吸引优秀的人才来应聘, 最终在考核评价的基础上, 对应聘人员进行甄选和录用。

基金项目 国家“985”工程重点建设天津大学基金资助项目 (S30201)

作者简介: 邢德海 (1979—), 男, 博士生、研究实习员, 主研方向: 信息管理与信息系统; 董旭源, 副研究员

收稿日期: 2006-01-20 **E-mail:** xingdehai@yahoo.com.cn

(3)培训与开发。教职工培训一般分为由高校相关部门对相关岗位的人员进行的内部培训和外派深造的外部培训。

(4)组织机构设计。根据输入的部门名称及其所处的层次、岗位名称及其所处的层次和所属关系,自动生成层次状的组织机构图、岗位图,并能从文本库中调用相应的部门职能和岗位工作说明书。根据岗位图自动生成人力资源需求总量,提供人机交互的修改功能。

(5)绩效评估与考核。根据科学、合理的绩效考评,可以对学校教职工实行动态管理,如晋升、调动等。

(6)薪酬与福利。薪酬与福利指通过工资薪酬、福利措施给教职工为学校做出的贡献予以报酬,并激励其工作积极性的过程。

(7)人力资本管理。自动从系统各模块和系统外采集人力成本数据,生成人力成本和人力资源产出数据表及各自历年的曲线,其中生成曲线所需的数值分析方法从方法库中调用,所需的模型从模型库中调用。模块自动分析和评价学校的人力资本指数等级。

(8)文本管理。提供较静态的文本数据管理功能,包括系统其他模块的文本数据、学校的规章制度及质量体系的标准等。

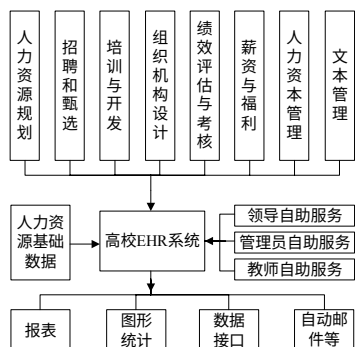


图2 高校EHR系统的体系结构

2 主成分分析法的数学模型

主成分分析法是一种把多个指标转换为少数几个综合指标的一种多元统计分析方法,它在社会经济、企业管理等各个领域中都存在着广泛的应用。在多指标(即变量)的研究中,往往由于变量个数太多,并且彼此之间存在着一定的相关性,因此使得所采集的数据在一定程度上反映的信息有所重叠,而且当变量较多时,在高维空间中研究样本的分布规律比较困难。主成分分析正是把这种情况简化,即采取一种降维的方法,找出几个综合因子来代表原来众多的变量,使这些综合因子能尽可能地反映原来变量的信息量,且彼此之间不相关。用主成分分析法进行多指标综合评价的基本步骤如下^[4]:

(1)原始指标数据的无量纲化

数据的无量纲化处理有许多方法,其中 Z-Score 法比较适合被评价对象个数较多的情况,更符合本系统的要求。

设有 n 个样本, P 项指标,可得数据矩阵 $X = (X_{ij})_{n \times p}$, $i=1,2,\dots,n$, $j=1,2,\dots,p$, X_{ij} 表示第 i 个样本的第 j 项指标值。用 Z-Score 法对数据进行标准化变换,变换公式为

$$Z_{ij} = \frac{X_{ij} - \bar{X}_j}{S_j}$$

其中 $\bar{X}_j = \frac{1}{n} \sum_{i=1}^n X_{ij}$, $S_j = \left[\frac{1}{n-1} \sum_{i=1}^n (X_{ij} - \bar{X}_j)^2 \right]^{\frac{1}{2}}$, $i=1,2,\dots,n$; $j=1,2,\dots,p$ 。

(2)指标权数的确定

为使由多个指标合成的综合价值更能准确地反映高校教职工的真实情况,从而保证综合评价的科学性,还必须对经过无量纲化变换后的指标赋予不同的权数。本系统采用主成分分析法进行综合评价时,各指标的权值是由主成分分析

法自动确定和保证的,无需另外的方法为各指标设定权值大小。

(3)确定指标数据的相关矩阵

指标数据的相关矩阵 $R = (r_{jk})_{p \times p}$, 其中 $j=1,2,\dots,p$, $k=1,2,\dots,p$ 。 r_{jk} 为指标 j 与指标 k 的相关系数。

$$r_{jk} = \frac{1}{n-1} \sum_{i=1}^n \left[\frac{(X_{ij} - \bar{X}_j)^2}{S_j} \right] \left[\frac{(X_{ik} - \bar{X}_k)^2}{S_k} \right], \quad r_{jk} = \frac{1}{n-1} \sum_{i=1}^n Z_{ij} Z_{ik},$$

有 $r_{ij} = 1$, $r_{jk} = r_{kj}$, 其中 $i=1,2,\dots,n$, $j=1,2,\dots,p$, $k=1,2,\dots,p$ 。

(4)求相关矩阵 R 的特征根和特征向量,确定主成分

由特征方程式 $|\lambda I_p - R| = 0$, 可以求得 p 个特征根 λ_g ($g=1,2,\dots,p$), 将其按大小顺序排列为 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$, λ_g 是主成分的方差,它的大小描述了各个主成分在描述被评价对象上所起作用的大小。特征方程的每一个特征根对应一个特征向量 L_g , $L_g = L_{g1}, L_{g2}, \dots, L_{gp}$, $g=1,2,\dots,p$ 。

将标准化后的指标变量转换为主成分:

$F_g = L_{g1} Z_1 + L_{g2} Z_2 + \dots + L_{gp} Z_p$, $g=1,2,\dots,p$ 。 F_1 称为第 1 主成分, F_2 称为第 2 主成分, ..., F_p 称为第 p 主成分。

(5)求方差贡献率,确定主成分个数

一般地,主成分个数等于原始指标个数。如果原始指标个数较多,进行综合评价时就比较麻烦。主成分分析法就是选取尽量少的 k 个主成分 ($k < p$) 来进行综合评价,同时还要使损失的信息量尽可能少。 k 值由方差贡献率 $\sum_{g=1}^k \lambda_g / \sum_{g=1}^p \lambda_g \geq 85\%$ 决定。

(6)对 k 个主成分进行综合评价

先求每一个主成分的线性加权值, $F_g = L_{g1} Z_1 + L_{g2} Z_2 + \dots + L_{gp} Z_p$, $g=1,2,\dots,k$, 再对 k 个主成分进行加权求和,即得最终评价价值,权数为每个主成分的方差贡献率 $\lambda_g / \sum_{g=1}^p \lambda_g$, 最终评价价值 $F = \sum_{g=1}^k \left(\lambda_g / \sum_{g=1}^p \lambda_g \right) F_g$ 。

3 主成分分析法在系统中的应用实例

3.1 评价实例

主成分分析法在天津 EHR 系统中有广泛的应用,下面就以高校教师的招聘与甄选为例加以说明。结合学校当前发展阶段的目标,为教师的招聘与甄选建立如下指标体系^[5]: (1)该教师承担的国家级科研项目 and 获得的国家级科研成果奖项 (x_1); (2)该教师的语言表达能力和教学方法的适当运用情况 (x_2); (3)该教师参与学术活动和校系活动的情况 (x_3); (4)该教师在国内外权威杂志上发表的论文数和论文的影响程度 (x_4); (5)该教师的教龄和教学上的获奖情况 (x_5); (6)该教师在科研和教学过程中表现的创造性和进取心 (x_6); (7)该教师的身体健康状况 (x_7)。

(1)原始指标数据的标准化

结合上述评价指标对 2004 年参与天津大学化工学院某岗位竞聘的 11 位教师进行评价,其中 3 位教师的各项指标的得分情况如表 1 所示。

表 1 竞聘教师原始指标得分

姓名	x_1	x_2	x_3	x_4	x_5	x_6	x_7
教师冯	81.51	62.98	65.10	68.17	60.00	63.00	72.95
教师杜	89.47	23.04	28.00	74.89	84.01	84.03	87.95
教师付	89.82	21.01	24.93	75.20	84.00	72.19	88.01

(2)相关系数矩阵

各项评价指标的相关矩阵 R 的结构如表 2 所示。

表 2 各项评价指标的相关矩阵表

	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇
X ₁	1	-0.792	-0.882	0.999 8	0.740	0.712	0.962
X ₂		1	0.993	-0.789	-0.765	-0.776	-0.720
X ₃			1	-0.879	-0.787	-0.76	-0.848
X ₄				1	0.739	0.713	0.96
X ₅					1	0.897	0.61
X ₆						1	0.534
X ₇							1

(3)求相关矩阵 R 的特征根和特征向量,确定主成分

相关矩阵 R 是实对称矩阵,由特征方程 $|\lambda I - R| = 0$,求得 7 个特征根 $(\lambda_1, \dots, \lambda_7)$ 及其贡献率,具体数值参照表 3。

表 3 相关矩阵的特征根及其贡献率

	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7
特征根	6.949	1.996	0.578	0.256	0.098	0.082	0.028
贡献率(%)	69.4	20.1	5.8	2.6	1.0	0.8	0.3
累计贡献率(%)	69.4	89.5	95.3	97.9	98.9	99.7	100.0

因为前两项特征根的累计贡献率为 $89.5\% > 85\%$,所以可以用第 1 主成分和第 2 主成分作为评价的综合指标,评价的可信度为 89.5% 。

对应 λ_1 的特征向量为

$$L_1 = (0.369, -0.336, -0.358, 0.368, 0.321, 0.314, 0.342)$$

对应 λ_2 的特征向量为

$$L_2 = (-0.135, -0.082, 0.022, -0.132, 0.274, 0.349, -0.301)$$

(4)求应聘教师综合评价并排序

根据上述特征向量可知,

第 1 主成分

$$F_1 = 0.369Z_1' - 0.336Z_2' - 0.358Z_3' + 0.368Z_4' + 0.321Z_5' + 0.314Z_6' + 0.342Z_7'$$

第 2 主成分

$$F_2 = -0.135Z_1' - 0.082Z_2' + 0.022Z_3' - 0.132Z_4' + 0.274Z_5' + 0.349Z_6' - 0.301Z_7'$$

$Z_1, \dots, Z_7$ 是各位教师原始数据做完标准化变换后的数据。

将综合评价值 $F = 0.694F_1 + 0.2F_2$ 代入标准化变换后的数据,可得各位应聘教师的综合评价值和排序情况,如表 4。

表 4 教师的综合评价值和排序

姓名	综合评价值	排序
教师付	3.061 5	1
教师杜	2.500 9	2
教师冯	-2.339 1	11

3.2 评价结果分析

从表 1 中各位应聘教师的 7 个评价指标的原始得分很难

直接来评价出每人的情况并做出排名,而使用主成分分析法则可以把这些变量概括为 2 个或 3 个综合变量(即主成分),以便帮助我们较为简便地分析这些数据。

从表 3 可以看出,前两个主成分的累计贡献率达到了 89.5% ,已经大于通常筛选主成分个数的经验百分比 85% ,因此可以考虑只取前两个主成分,它们能够很好地概括这些组数据。

由于第 1 主成分中 x_1, x_4 两个指标所对应的权值最大,这两个指标又分别代表着科研、出版科技论著,因此第 1 主成分主要是反映教师的科研和发表科技论著的情况,可认为教师具有较强的科研能力和较好的知识水平。第 2 主成分中 x_6 这一指标所对应的权值最大,而这一指标代表着教师在科研和教学过程中的创造精神和进取心,因此第 2 主成分是反映教师从事科研和教学的敬业情况,可解释为教师如具有创造精神和足够的进取心就能够更好的胜任科研和教学工作。

通过上述对综合评价结果的分析,给我们提供了两方面的提示信息:(1)承担过国家级科研项目并获得过国家级科研成果奖项的教师具有较强的科研能力;(2)教师要有敬业精神,在科研和教学活动中要具有创造精神和足够的进取心,只有这样的教师才能达到高校科研和教学工作的要求。

4 系统的实施效果

基于主成分分析法的天津大学 EHR 系统经过 1 年的适用期,2005 年正式在天津大学推广使用,实现对以教师为核心的高校人员相关的数据、过程和资源的集成化管理,极大地提高了高校人力资源管理的效率,保证了与高校人员相关的数据准确性和有效性。

以全校性的工资调整为例。在系统实施前,一次工资调整要用约 1 个月的时间,还会因为数据不够准确,使约 100 人的工资调整出现错误;在系统实施后,一次全校性的工资调整只用了 3 天时间,没有一个人的工资调整出现错误。

参考文献

- 1 黄梯云. 智能决策支持系统[M]. 北京: 电子工业出版社, 2001.
- 2 Kappel T A, Rubenstein A H. Creativity in Design: The Contribution of Information Technology[J]. IEEE Transactions on Engineering Management, 1999, 46(2): 132-143.
- 3 Yassine A A, Chelst K R, Falkenburg D R. A Decision Analytic Framework for Evaluating Concurrent Engineering[J]. IEEE Transactions on Engineering Management, 1999, 46(2): 144-157.
- 4 胡永宏, 贺思辉. 综合评价方法[M]. 北京: 科学出版社, 2000.
- 5 李成友, 王学周, 王 芳等. 主成分分析法在网络教学评价系统中的应用[J]. 计算机应用研究, 2003, 20(12): 128-133.

(上接第 56 页)

参考文献

- 1 杜明义, 郭达志. 空间数据仓库技术与模型研究[J]. 计算机工程与应用, 1999, 35 (12): 16-18.
- 2 曹学宝, 冯艳芬. 空间数据集成研究——以空间数据仓库的构建为例[J]. 广州大学学报(自然科学版), 2004, 3(3): 227-230.

- 3 Microsoft 公司. Microsoft SQL Server2000, Analysis Manager 概念与教程[R]. 2000-07.
- 4 Atluri V, Mazzoleni P. A Uniform Indexing Scheme for Geo-spatial Data and Authorizations[C]. Proc. of the 16th IFIP WG11.3 Working Conference on Database Security, 2002-07.