

# 神经网络挖掘模型在水电评标决策支持中的应用

曲俊华 汪 萌 (华北电力大学 计算机科学与技术系 北京 102206)

**摘要:** 水电项目评标决策支持系统克服传统评标主观性和不确定性的弊端,充分利用历史评标数据,为水电工程项目中的评标环节提供决策支持。提出了基于神经网络算法的数据挖掘模型,分析和研究了模型特点及建立过程,并将其应用于水电工程项目评标决策支持系统中,结合水电项目招投标的实际业务需要,实现了一个公正合理的水电项目评标决策支持系统。

**关键词:** 神经网络;数据挖掘;招投标;决策支持;商务智能

## Application of Neutral Net Data Mining Model in Bidding Evaluation Decision Support

QU Jun-Hua, WANG Meng

(Department of Computer Science and Technology, North China Electric Power University, Beijing 102206, China)

**Abstract:** Bidding evaluation decision support system overcomes the shortcomings like subjectivity and uncertainty in traditional bidding evaluation. It takes full advantage of history data to provide decision making support in bidding evaluation. This paper presents a data mining model based on neutral net algorithm, analyzes and studies its features and establishment, then applies the model in the hydropower project bidding evaluation decision support system, with the practical business demands of hydropower project bidding, to implement a fair and rational hydropower project bidding evaluation decision support system.

**Keywords:** neutral net; data mining; bidding; decision support; business intelligence

水电工程项目的招投标管理是水电工程项目管理的重要部分,在整个工程项目管理过程中的作用举足轻重。传统的招投标决策建立在经验、主观判断和感性认识的基础上,评标结果往往带有相当成分的主观随意性和不确定性,难以使众多投标企业信服。在高度信息化的社会和 Internet 的环境中,传统的招标显然已经不能适应经济的发展。网络信息化的招投标管理是一种必然趋势。

水电项目评标决策支持系统以我国水电工程施工招标评标的实际状况为背景,以已有的水电工程项目招投标管理系统为依托,设计建立一个公正合理的评标模型和评标决策支持系统。通过该系统的实现及应用,克服传统评标主观性和不确定性的弊端,充分利用历史评标数据,为水电工程项目中的评标环节提供决策支持,从而保证水电工程项目的质量和工期,达到合理使用资源,节约成本,降低造价目的;同时也

促进建筑企业的优胜劣汰,促进投标工程项目施工企业内部结构的调整和优化。

本文提出了基于神经网络算法的数据挖掘模型,利用人工神经网络的自适应功能,良好的容错能力和非线性数据处理特点,基于历史招投标过程中的海量数据,建立一个信息电子化的招投标评标决策支持系统。

## 1 神经网络算法基本原理

### 1.1 神经网络模型特点

人工神经网络是人脑神经网络的某种简化、抽象和模拟<sup>[1]</sup>。它具有如下特点:

- (1) 人工神经网络实现了从输入状态空间到输出状态空间的非线性映射。
- (2) 人工神经网络具有相当强的容错性。
- (3) 人工神经网络具有很强的处理不确定性信息

的能力。即使输入信息不完全、不准确，神经网络仍然能够进行处理，给出正确的推理结论。

(4) 人工神经网络具有很强的自适应、自学习能力，能够根据输入信息的要求，通过网络训练和学习调整神经网络的连接权值和连接结构。

随着神经网络理论的成熟和计算机计算能力的提高，神经网络已经应用于各个领域。神经网络的众多特点，使其成为水电项目评标决策支持数据挖掘模型的首选算法。

### 1.2 神经网络模型结构

神经网络模型包含一组节点(神经元)和边，这组节点和边形成一个网络<sup>[1]</sup>。每条边都通过一个权值连接两个节点，边的方向表示预测过程的数据流。每个节点都是一个处理单元。神经网络是一个分层结构，可以分为输入节点层，隐节点层，输出节点层，如图 1 所示。

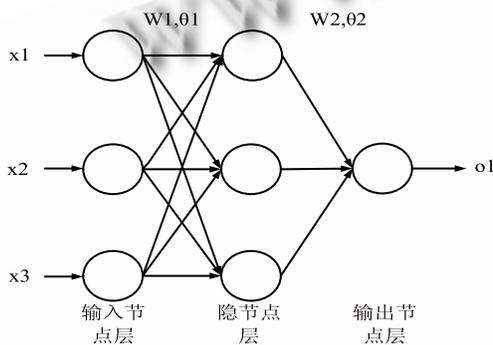


图 1 神经网络模型

$X_1, X_2, X_3$  是输入节点， $O_1$  是输出节点， $W_1$  和  $W_2$  是层与层之间的连接权重， $1$  和  $2$  是隐藏层和输出层的阈值。

每个神经元都是一个基本处理单元，有多个输入和一个输出。组合所有输入值，进行特定计算，然后触发一个输出值。它包含两个函数：输入组合函数和输出计算函数(激活函数)。一个神经元的基本结构如图 2 所示：

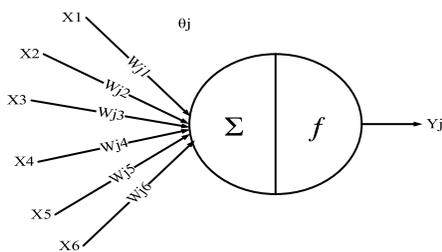


图 2 神经元基本结构

神经元有多个输入  $X_i, i=1, 2, \dots, n$ ，和一个输出  $Y_j$ 。

$$Y_j = f(w_{ji} * X_i - \theta_j)$$

其中  $w_{ji}$  为连接权值， $\theta_j$  为偏差， $f$  为激活函数。

### 1.3 神经网络模型的学习过程

处理神经网络的核心部分是反向传播。神经网络的训练过程是一个迭代过程。其过程可以被分为两个阶段：

第一阶段(正向传播)，给出输入信息，通过输入层经隐节点层逐层处理并计算输出值；第二阶段(反向传播)，若输出层未得到期望的输出值，则逐层递归地计算实际输出与期望输出之差值(即误差)，并据此差值调节权值<sup>[2]</sup>。

## 2 神经网络数据挖掘模型在水电评标决策支持系统中的应用

### 2.1 水电项目招投标过程及评标指标体系

按照《招标投标法》的规定，工程项目招标投标遵守以下的流程：确定项目策略，资格预审，招标和投标，开标，评标，授予合同<sup>[3]</sup>。评标是指在开标后，由招标单位组织的评标小组或委员会对各投标人编制和递交的投标文件进行分析比较，判断优劣，提出确定中标人的意见和建议。

评标内容随着招标项目的不同而改变，也随着工程规模和特性的不同而不同。但不论工程项目有何不同，项目评标指标体系的内容都可以归结为以下三类：

- (1) 商务评审——标价分析与比较，如分析投标单位在招标项目的总报价等。
- (2) 技术评审——工程实施方案评审，如分析投标人的施工方案，施工经验设备状况等情况。
- (3) 资质与信誉评审——承包商自身能力与信誉评审，如根据招标项目的难易程度等。

### 2.2 基于神经网络的数据挖掘水电评标模型的建立

#### 2.2.1 神经网络评标模型输入的确定及其规范化

评标体系的各项评价指标可以分为定性和定量两种。表 1 列出标书的一些基本评标条件。

对于神经网络数据挖掘评标模型来说，输入就是评标体系的各评价指标，定性和定量两种指标可以被映射为离散型和连续型两种输入。对于离散型的输入变量，最简单的方法就是要把它们映射到等价空间点上，其范围从 0 到 1。如投标人员的信誉分为 4 级：好、

良、中、差，可以分别映射到 1、0.75、0.5、0。最终得到一个位于[0, 1]区间的数值。

表 1 评标指标

ID	评标指标	指标说明	类型	取值
1	工程报价 (万元)	完成工程所需 资金	定量	200
2	设备情况	包括: 高档, 中 档, 低档	定性	中档
3	项目周期 (天)	完成项目需要 时间	定量	100
4	工程经验	大型项目, 中性 项目, 小型项 目, 无经验	定性	有中性项 目经验

神经网络需要将输入变量的值规范化为相同范围的值，否则值较大的那些变量会支配训练过程。我们采用标准化法进行规范化：

$$V = \frac{X-\mu}{\delta}$$

对于连续型变量， $\mu$  是平均值， $\delta$  是标准方差。对于离散型变量， $\mu = p$ (该状态的概率)， $\delta^2 = p*(1-p)$ 。

在神经网络中，一个连续的属性对应两个输入神经元，一个代表值，一个代表缺失状态。一个离散属性对应  $n+1$  个输入神经元， $n$  为不同状态的数目，1 代表缺失状态。下面给出一个输入映射和规范化的例子。

表 2 为训练数据的示例，其中包括连续型和离散型的输入。

表 2 训练数据

ID	报价(万元)	设备情况
1	240	低级
2	280	中级
3	300	高级
4	260	中级

表 3 为规范化之后的数据，是根据标准化法进行规范化后的输入。

表 3 规范化数据

ID	报价		设备情况			
	-	val	-	高级	中级	低级
1	0	-1.162	0	-1.732	-1.732	1.732
2	0	0.387	0	-1.0	1.0	-1.0
3	0	1.162	0	1.732	-1.732	-1.732
4	0	-0.387	0	-1.0	-1.0	1.0

### 2.2.2 神经网络模型网络结构及样本的选取

对神经网络模型进行处理前我们必须要对神经网络的拓扑结构进行确定，输入和输出的神经元数目是通过一个训练数据集来确定的<sup>[4]</sup>。在该神经网络模型中值包含一个隐含层，该层的神经元数目也影响着模型的性能。选择隐含神经元的准则是： $c*\sqrt{m*n}$ ，其中  $n, m$  分别是输入和输出神经元个数， $c$  为常量，默认为 4。

本课题选取水电工程项目管理系统的历史招投标数据作为研究对象，通过对评标体系的研究，参考有关专家打分意见，得到原始数据，然后对得到的原始数据进行归一化规范化处理，形成该神经网络模型的训练样本。这里的神经网络担任的是评标专家的角色，利用模型的学习能力，使训练后的神经网络获取了评标专家的经验、知识、判断及其对指标重要性的倾向<sup>[5]</sup>。然后输入参与评标的各项条件数据，可以根据神经网络模型得到评价结果，为水电项目招投标管理中的评标环节提供决策支持。

### 2.2.3 神经网络模型训练过程

由于神经网络训练本身有许多不确定的因素，每次训练的结果不尽相同，所以经过训练得到的神经网络并不一定具有良好的预测效果。在这种情况下，就需要利用训练得到的神经网络对未参加训练的样本进行预测，再对网络的预测效果进行检验，当网络对预测样本给出的结果不合理时，将再次重新训练，直到检验样本得到满意的答案。

神经网络训练过程算法描述：

(1) 为网络中所有权值指派初值(范围通常是 -1.0~1.0)。

(2) 对每一个训练示例，基于当前权值来计算输出。这里选择加权和作为组合函数，使用 sigmoid 函数作为激活函数。

(3) 神经网络需要一个度量标准来指出训练质量，该度量标准就是误差函数。根据误差函数计算出误差，然后为网络中的每一个输出和隐含神经元计算误差，更新权值。这里给出误差函数及权值调整函数(推导证明过程从略)。

输出节点误差函数  $Err_i = O_i(1-O_i)(T_i-O_i)$ ，其中  $O_i$  是输出节点的计算， $T_i$  是实际值。

隐节点误差函数  $Err_i = O_i(1-O_i) \sum_j Err_j W_{ij}$ ，其中  $O_i$  是隐含节点的输出，该节点有  $j$  个到下层节点的输

出。Err<sub>j</sub>是下层节点j的误差，W<sub>ij</sub>是这两个节点之间的权值。

权值调整函数  $W_{ij}=W_{ij}+l*Err_j*O_i$ 。

(4) 重复步骤(2)，直到满足条件为止。

该训练流程图如图3所示：

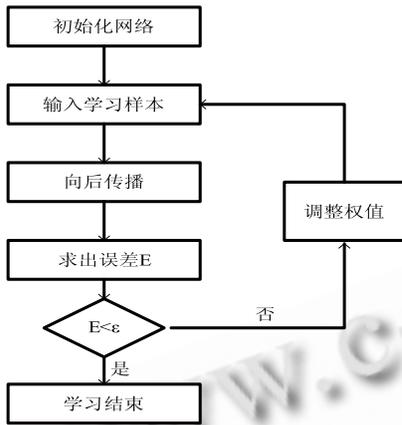


图3 训练流程图

由于系统中需要考虑的现实因素太多，在这里仅以一个简单的示例来说明神经网络模型的基本结构，并且演示神经网络训练过程。图4显示了简单的神经网络模型拓扑结构，可以看作实际模型的一个原型。

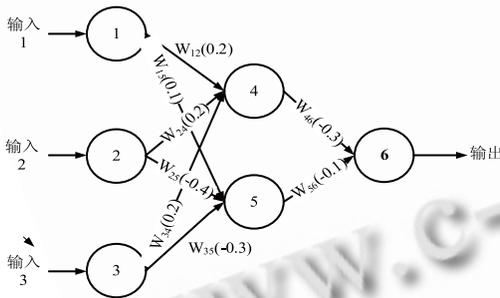


图4 神经网络模型拓扑结构

该神经网络模型中包含三个输入节点，两个隐含节点和一个输出节点，在训练过程中我们使用加权和作为组合函数，使用 sigmoid 作为激活函数。假定学习的样例为(1,1,0,1),最后的一位是输出结果。

下面表格展现出了该模型(图4)的训练过程：

对每个输入，基于当前权值，利用加权和作为组合函数，使用 sigmoid 函数作为激活函数，计算输出，其结果如表4所示。

表4 计算输出

神经元	输入	输出
4	0.2+0.2+0=0.4	1/(1+e <sup>-0.4</sup> )=0.401
5	0.1-0.4+0=-0.3	1/(1+e <sup>-0.3</sup> )=0.574
6	-0.3*0.401+(-0.1)*0.574=-0.694	1/(1+e <sup>-0.694</sup> )=0.667

对每一个输出和隐含神经元，根据给定的误差计算函数计算误差，其结果如表5所示。

表5 误差计算

神经元	误差
6	0.667*(1-0.667)*(1-0.667)=0.074
5	0.574*(1-0.574)*0.074*(-0.3)=-0.005
4	0.401*(1-0.401)*0.074*(-0.1)=-0.002

对于网络中的权值，根据计算值以及给定的权值调整函数，进行权值调整。其结果如表6所示。

表6 权值调整

权值	调整权值
W <sub>46</sub>	-0.3+0.8*0.074*0.401=-0.276
W <sub>56</sub>	-0.1+0.8*0.074*0.574=-0.066
W <sub>14</sub>	0.2+0.8*(-0.002)*1=0.198
W <sub>15</sub>	0.1+0.8*(-0.005)*1=0.096
W <sub>24</sub>	0.2+0.8*(-0.002)*1=0.198
W <sub>25</sub>	-0.4+0.8*(-0.005)*1=-0.404
W <sub>34</sub>	0.2+0.8*(-0.002)*0=0.2
W <sub>35</sub>	-0.3+0.8*(-0.005)*0=-0.3

2.2.4 神经网络模型的训练结果及评标预测

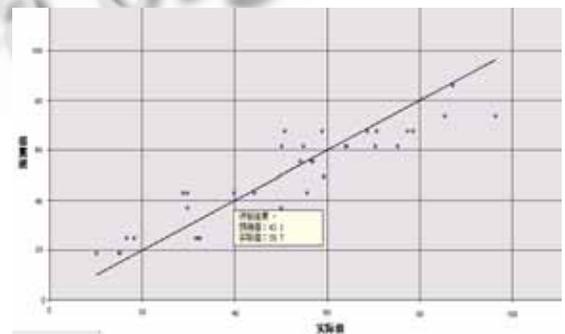


图5 散点图

模型训练完成后，需要验证模型的准确性和可用性。选取一些历史的招投标数据(最好不是训练模型时用的训练数据，否则会使得模型预测效果比实际预测效果好)，利用准确性图表来度量训练后模型的质量和

精确性,它可以针对该神经网络模型进行预测,并且将预测结果与已知结果进行比较。图 5 是该神经网络模型的散点图,它针对包含连续可预测属性的模型,以图形方式对照显示数据中的实际值与模型预测的值。沿 X 轴显示实际值,沿 Y 轴显示预测值。从图中可以看出该模型能够较准确的预测评标得分结果,散布在该直线周围的点显示实际数据偏离预期值的程度。预测值与实际值走向基本一致,说明了模型具有一定的准确型,对评标决策的预测是有效的。

确定了模型的准确性之后,可以利用这个神经网络模型来对需要参与投标的单位的评标结果进行预测。表 7 列出了预测结果。

表 7 预测结果

投标单位	评标分数	预测概率
单位 A	90.1	67.1%
单位 B	60.9	54.8%
单位 C	73.6	40.9%

其中评标分数列体现的是基于该模型得出预测的结果,预测概率列是该预测结果可信度的体现。实现这样的预测功能可以为评标提供一定参考的,评标人员可以利用预测得出的评标分数,以及体现该预测分数准确度的预测概率进行评标,能提高评标效率,避免主观随意性和不确定性,为水电项目招投标的评标

工作提供参考依据。

### 3 结语

本课题实现了神经网络数据挖掘模型在水电项目招投标的评标决策支持的应用。文中提出了神经网络评标模型,分析了它的训练处理过程,并将其应用于水电工程项目招投标系统中,为评标结果提供决策支持。实践证明,神经网络的自适应功能,良好的容错能力和非线性数据处理等特点保证了评标的准确性和客观性,同时也提高了评标的效率,具有良好的应用前景和现实意义。

### 参考文献

- 1 MacLennan J. Data Mining with SQL Server 2005.北京:清华大学出版社,2007. 77 - 78.
- 2 谢邦昌.商务智能与数据挖掘 Microsoft SQL Server 应用.北京:机械工业出版社,2008. 56 - 57.
- 3 吴广译.招投标中评标指标的分解,量化和评定.建筑经济,1998,(6):32 - 34.
- 4 Olson D,石勇.商业数据挖掘导论.北京:机械工业出版社,2007.102 - 103.
- 5 夏火松.数据仓库与数据挖掘技术.北京:科学出版社,2009.