

申请编号 : 151600018

广州市科技计划项目 申报书

项目名称：基于数据挖掘技术的财政收入分析预测模型研究

申报单位：广州番禺职业技术学院

组织单位：广州番禺职业技术学院

计划类别：软科学研究专项

申报日期：2014/08/05



广州市科技和信息化局

填写说明

- 1.本申报书需一式两份，每份需加盖公章。
- 2.申报书及有关材料、附件均采用 A4 纸双面打印。
- 3.本申报书的各项内容必须如实详细填写，无此项内容则填“无”。
- 4.本申报书不够书写的，可另行加页。
- 5.申请单位名称需填写全称，地址及电话等联系方式要准确无误。

目 录

一、项目基本信息	4
二、申请单位和主要参加单位的基本情况	4
三、项目情况	5
四、经费预算	21
五、项目组人员情况	22
六、项目负责人及主要研究人员承担的省、市科技计划项目情况	22
七、申请单位意见	23
八、参加单位意见	23
九、推荐部门意见	23
附件目录	24

一、项目基本信息

项目名称	基于数据挖掘技术的财政收入分析预测模型研究					
项目主要内容摘要	构建广州市历史财政收入数据与同期社会、经济发展相关数据的数据库；开展业务现状调研，梳理影响财政收入关联指标的有关数据，分析、识别影响财政收入的关键影响因素；分析研究各影响因素与财政收入的相关性，精选财政收入评价指标，并研究基于数据挖掘技术的广州市财政收入分析预测的参考模型。					
项目负责人	姓 名		性别		出生年月	
	最高学历		电话		手 机	
			E-mail			
	证件名称		证件号码			
新投入经费（万元）	合 计		申请科信局经费		自筹资金	
	30		15		15	
项目起止时间	2015-01-01 至 2016-12-30					

二、申请单位和主要参加单位的基本情况

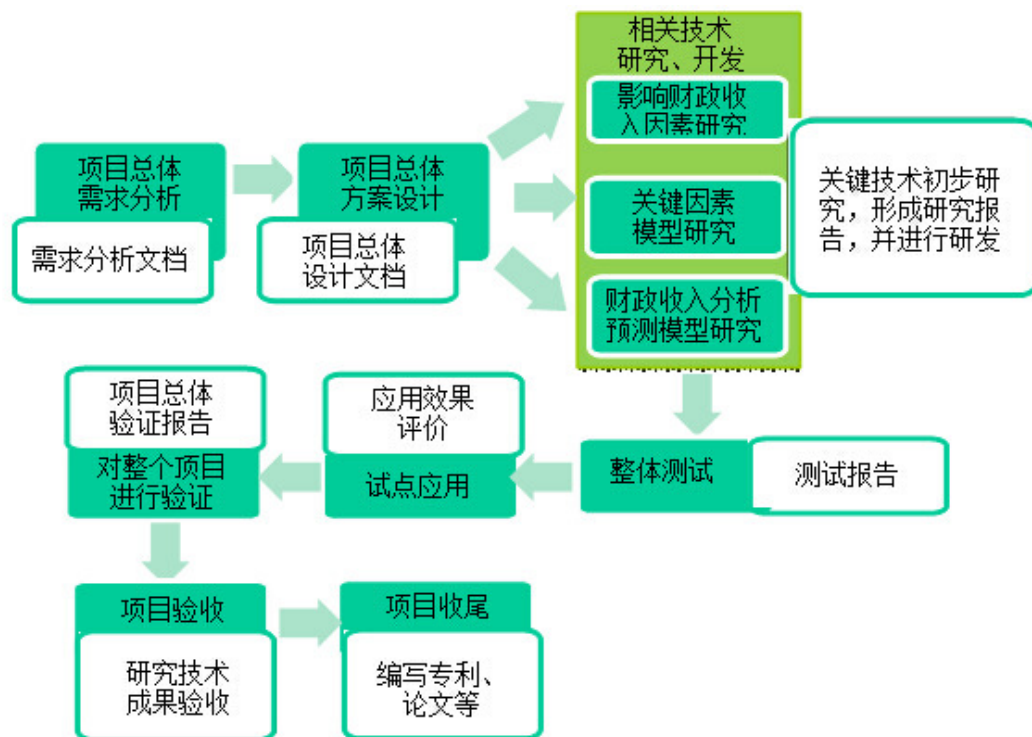
第一申请单位名称	广州番禺职业技术学院					
法定代表人					电话	
通讯地址	广州市番禺区沙湾镇市良路 1342 号				邮编	
联系人		电话	020-34832750	手机		
传真				E-mail		
申请单位职工总数		其中技术人员	396 人	中、高级职称		
主要参加单位	1、广州泰迪智能科技有限公司					
	2、					
	3、					

三、项目情况

(一) 研究方案

1、研究总路线图

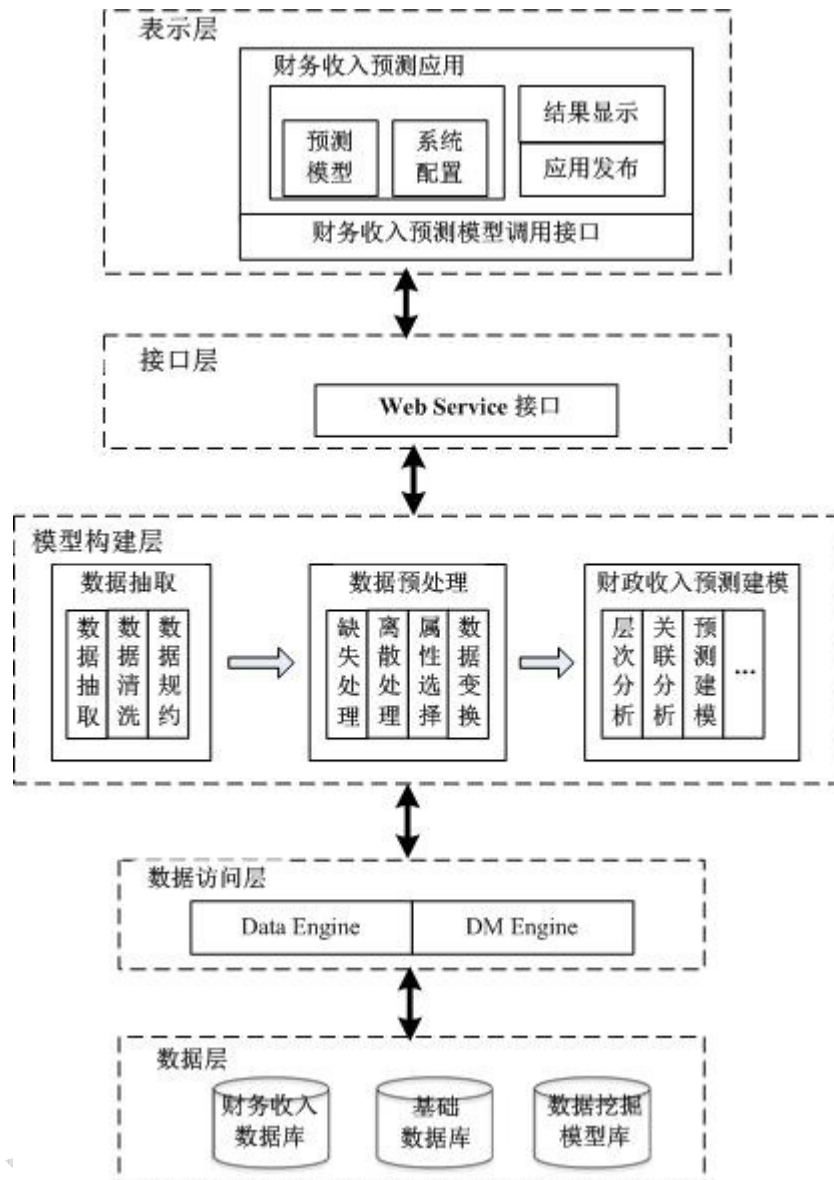
项目研究技术实现路线如下图所示：



从图中可以看出，本项目将从项目总体分析、方案设计、相关技术研究/开发、整体测试、试点应用、整个项目验证、项目验收等几个阶段开展工作。

2、技术路线

平台采用多层体系结构，分为“表示层”、“接口层”、“模型构建层”、“数据访问层”及“数据层”。其中“接口层”、“模型构建层”、“数据访问层”及“数据层”构成的平台的服务器端，“表示层”为客户端，客户端通过 SOAP 调用服务器端功能，在平台服务器端的支撑下用户可定制客户端或将 SDM 集成于第三方软件中，平台体系结构如下：



数据层功能说明如下:

根据财政收入数据与同期社会、经济发展等相关因素,分析会影响财政收入的各种对象以及所需要的数据,构建广州市历史财政收入数据与同期社会、经济发展相关数据的数据库。

地方财政收入包括地方财政预算收入和预算外收入:

(1) 地方财政预算收入的内容包括:

① 主要是地方所属企业收入

② 各项税收收入包括营业税、地方企业所得税、个人所得税、城镇土地使用税、固定资产投资方向调节税、土地增值税、城镇维护建设税、房产税、车船使用税、印

花税、农牧业税、农业特产税、耕地占用税、契税，增值税、证券交易税（印花税）的 25%部分和海洋石油资源税以外的其他资源税。

③ 中央财政的调剂收入

④ 补贴拨款收入

⑤ 其它收入

(2) 地方财政预算外收入的内容主要有

① 各项税收附加：（正税以外的其他税收）

② 城市公用事业收入

③ 文化、体育、卫生及农、林、牧、水等事业单位的事业收入

④ 市场管理收入

⑤ 物资变价收入等

另外，财政收入与国民收入、工业总产值、农业总产值、总人口、就业人口、固定资产投资等因素有关。根据这些因素，也可预测财政收入问题，因此这些数据也需要采集入库。

模型构建层功能说明如下：

a) 数据抽取：主要用于完成数据挖掘建模所需的数据提取，即从企业大量数据中抽取出一个与预测目标相关的样本数据子集。抽取数据的标准，一是相关性，二是可靠性，三是最新性。

b) 数据预处理：拿到了一个样本数据集后，它是否达到我们原来设想的要求；其中有没有什么明显的规律和趋势；因素之间有什么相关性；它们可区分成怎样一些类别……，这些都是要数据探索的主要内容。

由于采样数据中常常包含许多含有噪声、不完整、甚至是不一致的数据。显然对数据挖掘所涉及的数据对象必须进行预处理，消除对挖掘结果产生重要负面影响，达到改善数据质量的目的。数据预处理主要包括：数据清洗、数据集成、数据转换、数据消减数、坏数据处理、据离散化、降维处理等。

c) 财务收入预测建模：由于财政收入预测模型是多种多样的，而且要求能适用不同结构的样本数据，因此，系统的主体构成要素应包括数据探索、预处理、挖掘建模、模型评估、查询统计，挖掘系统可以实现灵活调整。

运用层次分析法，可将地方财政收入分析，地方财政收入包括地方财政预算收入和

预算外收入。首先，本方案利用主成分分析法得出这些影响财政收入的主要因素。着重对主要因素进行分析，进行深入的数据挖掘、建立数学模型；对次要因素建立合理的、科学的数学模型；此处假设地方财政收入主要是财政预算内收入。其次，继续根据现行财政体制划归预算内收入的主要组成进行层次分析，并根据分析结果分别建立数学模型，此处假设预算内收入主要是税收收入。在广州市财政收入的税收收入方面，可针对主要税收税种建立预测模型。其中主要税收税种包括：企业所得税、营业税、增值税和个人所得税等。再次，将主要的税收收入税种——企业所得税所，对征收的企业类进行分解为几类，然后对各分类企业分别分析研究对象、分析所需数据、通过数据挖掘找出影响企业所得税收入的因素，建立数学模型。最后，综合所有分析内容与建立的数学模型，对模型予以验证。整理出具有相当可信度的广州市财政收入预测模型，并生成分析报告。

模型构建完成后，需要对模型进行优化，即对所要建立模型的思路进行阐述，对所得的结果进行数学上的分析、检查。将模型分析结果与实际情形进行比较，以此来验证模型的准确性、合理性和适用性。如果模型与实际吻合较差，则应该修改假设，再次重复建模过程。

3、关键技术

1) 数据预处理技术

本项目中用到的数据预处理技术主要有：主成分因子分析、特征提取和属性重要性选择。

特征选择与提取是为了求出一组对分类最有效的特征，这就需要有一个定量的准则（判据）来衡量特征对分类的有效性。把一个高维空间变换为低维空间的映射有很多，哪种映射对分类最有利，需要一个比较标准。

主成分因子分析就是把维数比较多的空间最终压缩到维数比较少的空间，为下一步的特征筛选而服务的过程，是把原特征空间众多特征进行某种组合（通常是线性组合）而减少特征，降低维数的。

特征选择就是从一组特征里挑出一些最有效的特征以降低特征空间的维数的过程。选择有两种途径：首先是从原始特征中去挑选出一些最有代表性的特征，再者就是用变换的方法把原始特征变换为较少的新特征。

属性重要性选择是数据预处理的一部分，因采集的数据中的每一个属性对于整个数据的挖掘结果的作用不是完全对等的，一些属性对结果的影响占主导地位，一些属性对结果的影响不大，甚至没有影响，采用相应的算法，对数据的属性值进行评估，如去掉某个属性后对挖掘结果无影响，从而减少后续挖掘算法的运行时间，同时也能有效地去除数据中含有的噪声数据。

如果建模数据集的维度较高，或输入属性与输出属性的相关性不明确时，对其进行属性选择是必要的步骤。综合考虑应用实现的复杂性，可使用标准属性选择方法，用一个评估标准对属性的有用性进行度量，并选择其最有用的一部分属性作为下一部分算法的输入。为了确定属性选择的标准，可选用多种方法进行属性评价，选出 5-10 个不同的属性，然后对于处理后的数据集进行测试，综合评价维度约减前后的预测模型的性能及效果。

属性重要性选择是用于对属性进行筛选，搜索数据集中全部属性的所有可能组合，找出预测效果最好的那一组属性，一般在预处理阶段进行。为实现这一目标，必须设定两个东西：属性评估器和搜索策略。属性重要程序评估是对属性/属性子集进行评估确定，决定了怎样给一组属性安排一个表示它们好坏的值。搜索策略确定搜索算法，决定了要怎样进行搜索。

2) 数据挖掘技术

数据挖掘是为了发现事先未知的规则和联系而对大量数据进行选择探索和建模的过程。目的在于得到对数据库的拥有者来说清晰而有用的结果。其实质是在数据预处理的基础上，选择合适算法,提取规则，进行评价和解释，从而获取隐藏在数据背后的知识。数据挖掘把人们对数据的应用从低层次的简单查询提升到从数据这挖掘知识，提供决策支持。在这种需求牵引下，其技术涉及到了多个领域的相关知识，如数据库技术、数理统计、模式识别、人工智能、机器学习、可视化技术等。如何挖掘和利用这些海量电力数据，使之转换为有价值的信息和知识，用于解决各种各样的科学和应用问题，已经成为新世纪信息技术发展的重大挑战。

1、层次分析法

层次分析法（Analytic Hierarchy Process）简称（APH），最早是由美国运筹学专家沙特（T.L.Saaty）于 20 世纪 70 年代末提出的。它通过建立所谓判断矩阵的过程，逐步分层将众多的复杂因素和决策者的个人偏好综合起来，进行逻辑思维，然后用定量的形

式表示出来，从而使决策者能对复杂问题做到“心中有数”，进而作出正确的决策。技术评估本身是一种复杂的决策问题。因此，层次分析法在技术评估中得到广泛的应用。

本方法是先将复杂问题中各种因素通过划分相互联系的有序层次使之条理化，然后根据对一定客观现实的判断就每一层次的相对重要性给予定量表示，再利用数学方法确定表达每一层次的全部元素相对重要性次序的权值，最后通过排序结果分析，解决问题。

层次分析法的步骤。第一步，明确问题，建立层次结构。对于所要解决的问题，首先进行系统分析，明确问题的范围、所包含的因素以及因素之间的定性关系等，然后根据这些初步分析，将各因素分层分组，建立层次结构。第二步，构造判断矩阵。对于所建立的层次结构，可以构造一系列的判断矩阵。首先可以构造最高层目标下属有联系各元素的判断矩阵，然后依次由上而下地建立上层某元素与下一层有关元素之间的判断矩阵。第三步，层次单排序。也即对各判断矩阵进行求解，计算出反映上层某元素和下层与之有联系的元素重要性次序的权重来。与此同时，还要对各判断矩阵进行一致性检验。第四步，结果分析。在一致性基本问题满足的前提下，我们就可以根据层次单排和层次总排序结果对问题进行定量的分析，从而作出决策。

2、模糊综合评价法

模糊综合评价法是一种基于模糊数学的综合评标方法。该综合评价法根据模糊数学的隶属度理论把定性评价转化为定量评价，即用模糊数学对受到多种因素制约的事物或对象做出一个总体的评价。它具有结果清晰,系统性强的特点，能较好地解决模糊的、难以量化的问题，适合各种非确定性问题的解决。

模糊集合理论(fuzzy sets)的概念于 1965 年由美国自动控制专家查德（L.A. Zadeh）教授提出，用以表达事物的不确定性。

模糊综合评价法的最显著特点是：

1) 相互比较

以最优的评价因素值为基准，其评价值为 1；其余欠优的评价因素依据欠优的程度得到相应的评价值。

2) 函数关系

可以依据各类评价因素的特征，确定评价值与评价因素值之间的函数关系（即：隶

属度函数)。确定这种函数关系(隶属度函数)有很多种方法,例如,F统计方法,各种类型的F分布等。当然,也可以请有经验的评标专家进行评价,直接给出评价值。

模糊综合评价法的一般步骤:

1) 模糊综合评价指标的构建

模糊综合评价指标体系是进行综合评价的基础,评价指标的选取是否适宜,将直接影响综合评价的准确性。进行评价指标的构建应广泛涉猎与该评价指标系统行业资料或者相关的法律法规。

2) 采用构建好权重向量

通过专家经验法或者 AHP 层次分析法构建好权重向量。

3) 构建评价矩阵

建立适合的隶属函数从而构建好评价矩阵。

4) 评价矩阵和权重的合成

采用适合的合成因子对其进行合成,并对结果向量进行解释。

3、人工神经网络

人工神经网络是一种应用类似于大脑神经突触联接的结构进行信息处理的数学模型。在工程与学术界也常直接简称为神经网络或类神经网络。神经网络是一种运算模型,由大量的节点(或称神经元)之间相互联接构成。每个节点代表一种特定的输出函数,称为激励函数(activation function)。每两个节点间的连接都代表一个对于通过该连接信号的加权值,称之为权重,这相当于人工神经网络的记忆。网络的输出则依网络的连接方式,权重值和激励函数的不同而不同。而网络自身通常都是对自然界某种算法或者函数的逼近,也可能是对一种逻辑策略的表达。

人工神经网络是由大量处理单元互联组成的非线性、自适应信息处理系统。它是在现代神经科学研究成果的基础上提出的,试图通过模拟大脑神经网络处理、记忆信息的方式进行信息处理。人工神经网络具有四个基本特征:

1) 非线性非线性关系是自然界的普遍特性。大脑的智慧就是一种非线性现象。人工神经元处于激活或抑制二种不同的状态,这种行为在数学上表现为一种非线性关系。具有阈值的神经元构成的网络具有更好的性能,可以提高容错性和存储容量。

2) 非局限性一个神经网络通常由多个神经元广泛连接而成。一个系统的整体行为不仅取决于单个神经元的特征,而且可能主要由单元之间的相互作用、相互连接所决

定。通过单元之间的大量连接模拟大脑的非局限性。联想记忆是非局限性的典型例子。

3) 非常定性人工神经网络具有自适应、自组织、自学习能力。神经网络不但处理的信息可以有各种变化,而且在处理信息的同时,非线性动力系统本身也在不断变化。经常采用迭代过程描写动力系统的演化过程。

4) 非凸性一个系统的演化方向,在一定条件下将取决于某个特定的状态函数。例如能量函数,它的极值相应于系统比较稳定的状态。非凸性是指这种函数有多个极值,故系统具有多个较稳定的平衡态,这将导致系统演化的多样性。

人工神经网络中,神经元处理单元可表示不同的对象,例如特征、字母、概念,或者一些有意义的抽象模式。网络中处理单元的类型分为三类:输入单元、输出单元和隐单元。输入单元接受外部世界的信号与数据;输出单元实现系统处理结果的输出;隐单元是处在输入和输出单元之间,不能由系统外部观察的单元。神经元间的连接权值反映了单元间的连接强度,信息的表示和处理体现在网络处理单元的连接关系中。人工神经网络是一种非程序化、适应性、大脑风格的信息处理,其本质是通过网络的变换和动力学行为得到一种并行分布式的信息处理功能,并在不同程度和层次上模仿人脑神经系统的信息处理功能。它是涉及神经科学、思维科学、人工智能、计算机科学等多个领域的交叉学科。

人工神经网络是并行分布式系统,采用了与传统人工智能和信息处理技术完全不同的机理,克服了传统的基于逻辑符号的人工智能在处理直觉、非结构化信息方面的缺陷,具有自适应、自组织和实时学习的特点。

4、模糊神经网络算法

模糊神经网络(Fuzzy Neural Network, FNN)是具有模糊权系数或者输入信号是模糊量的神经网络,是模糊系统与神经网络相结合的产物,它汇聚了神经网络与模糊系统的优点,集联想、识别、自适应及模糊信息处理于一体。模糊神经网络的主要特点是利用神经网络调整模糊逻辑推理系统的隶属度函数和推理规则,利用模糊推理规则的形式构造神经网络结构,从而充分发挥各自的优点。模糊规则经过神经网络的学习,以“权值”的形式体现出来,这样规则的生成和修改则转化为权值的确定和修改。模糊神经网络的结构决定了模糊规则条数以及隶属度函数个数,而参数决定了每条规则的具体表达及隶属度函数的形状。

模糊神经网络的本质就是在常规的神经网络中输入‘模糊输入信号’和‘模糊权值’。其结构像神经网络,功能上是模糊系统,这是目前研究和应用最多的一类模糊

神经网络。

该模糊神经网络是多层网络，除了输入结点层和输出结点层外，还有多个隐含层，相邻两层之间都有连接，且每个连接都对应于一个权值。该模糊神经网络共有五层：输入层、模糊化层、模糊推理层、去模糊层和输出层。

模糊神经网络的优点：

- 1) 并行分布式处理：神经网络具有高度的并行结构和并行实现能力，具有高速寻找优化解的能力，能够发挥计算机的高速运算能力，可能很快找到优化解；
- 2) 非线性处理：人脑的思维是非线性的，故神经网络模拟人的思维也应是非线性的。这一特性有助于处理非线性问题；
- 3) 具有自学习功能：通过对过去的历史数据的学习，训练出一个具有归纳全部数据的特定的神经网络，自学习功能对于预测有特别重要的意义。

5、ARIMA 模型

ARIMA 模型全称为差分自回归移动平均模型，是由博克思和詹金斯于 70 年代初提出的一著名时间序列预测方法，博克思-詹金斯法。其中 ARIMA(p,d,q)称为差分自回归移动平均模型，AR 是自回归，p 为自回归项；MA 为移动平均，q 为移动平均项数，d 为时间序列成为平稳时所做的差分次数。

ARIMA 模型的基本思想是：将预测对象随时间推移而形成的数据序列视为一个随机序列，用一定的数学模型来近似描述这个序列。这个模型一旦被识别后就可以从时间序列的过去值及现在值来预测未来值。

(二) 预期提供的研究成果及形式

- 1、基于数据挖掘技术的财政收入分析预测模型 2~3 个
- 2、基于数据挖掘技术的财政收入分析预测模型研究报告 1 份
- 3、发表核心期刊论文 2 篇
- 4、申报发明专利 1 个
- 5、财政收入分析预测软件（含源代码）1 套

(三) 研究基础

- 1、申报单位及承担单位概况（人员、资产、业务与管理状况）：

广州番禺职业技术学院财经学院现开设金融管理与实务、会计电算化、国际金融、国际商务、投资与理财、审计六个专业。其中，国家级示范性专业一个，省级示范性专业一个，市级示范性建设专业二个。全系约 50 个教学班，共有在校生 2100 余人，现有教职工 50 余人，其中，教授 2 人，副教授 9 人，中级职称教师 30 余人；博士 4 人，硕士 30 余人；高级经济师、高级会计师、律师、注册会计师、注册税务师 10 余人。现有国家级精品课程 2 门、省级精品课程 2 门、市级精品课程 3 门、校级精品课程 13 门；配备了银行基本技能实训室、银行综合柜台业务实训室、信贷业务实训室、证券业务实训室、证券资讯中心、保险业务实训室、账务处理与展示实训室、电算化会计与审计实训室、会计中介业务实训室、国际金融业务实训室、金融项目中心等校内实训设施，建设了三十多个校外实训基地。近三年的就业率均为 100%，毕业生受到了用人单位的普遍欢迎。

广州泰迪是广州经济技术开发区内的一家专业从事数据挖掘软件基础研究、培训、咨询服务及数据挖掘应用开发的高新技术企业。公司开始运作于 2005 年，是国内目前具有较强实力的数据挖掘系统研发及咨询服务商之一，注册资金 500 万元，在职员工 37 名，研究生以上学历 8 人，85% 为本科以上学历。公司通过一流的咨询服务，为企业进行资源预测、决策优化、故障检测、模式分类和企业管理服务；公司已成为广州金发科技、格力电器、美的集团、广东省电信规划设计院、交通运输部公路科学研究院、珠江水产研究所、广东电网公司等供应商，为他们成功提供了数据挖掘咨询和预测建模的服务。

2、本项目现有的研究工作基础（包括现有科研装备条件）：

本项目由广州番禺职业技术学院与广州泰迪智能科技有限公司联合研发，两大力量强强联合，既保障了项目组成员充分熟悉本领域业务知识，又保障了项目的研发力量和二次开发条件。

2010 年以来，广州番禺职业技术学院与广州泰迪智能科技有限公司开始致力于数据挖掘技术与金融领域的合作。2013 年财金系采购了由泰迪公司研发的 TipDM 数据挖掘平台（教研版），并定制了针对金融领域的数据挖掘教学案例。同年聘请泰迪公司高级工程师来我校举办针对全系师生的“面向金融领域的科研与教学的数据挖掘培训”实践周活动，大大增强了财金系教师的科研能力和水平。

广州泰迪智能科技有限公司具备软硬件集成的经验和市场推广能力，在与广州番

禺职业技术学院合作的基础上，形成了一系列的市场化产品，

先后研发了企业智能预测开发平台、大数据挖掘基础平台、数据挖掘辅助教学套件等产品，并成功应用到政府、电信、电力、银行、教育等行业。

顶尖数据挖掘平台（TipDM）是广州泰迪智能科技有限公司花费数年时间自主研发的一个数据挖掘平台，基于云计算和 SOA 架构，使用 JAVA 语言开发，能从各种数据源获取数据，建立各种不同的数据挖掘模型。系统支持数据挖掘流程所需的主要过程，并提供开放的应用接口和常用算法，能够满足各种复杂的应用需求。2010 年初“顶尖数据挖掘平台（TipDM）”通过了由广州赛宝软件评测中心的功能和性能测试。现产品在广东省电信规划设计院、珠江水产研究所、华南师范大学、广东中医药大学、武汉理工大学、交通运输部公路科学研究院等单位成功试用，受到用户的赞许与肯定。TipDM 以智能预测技术为核心，并提供开放的应用接口，用户也可以嵌入其它自己开发的任何算法，平台支持 CRISP-DM 流程所需的主要过程，完成包括对数据进行预处理，包括空值处理、降维处理、离散处理，因子分析、主成分分析、抽样、过滤等，创建、训练、评估模型，预测、修改模型参数，误差分析等一系列功能。平台主要界面如下：





The screenshot displays the LM Neural Network software interface. The main window is titled '分类与回归 (LM神经网络)' (Classification and Regression (LM Neural Network)). It features a '训练数据' (Training Data) table with columns for index, DO, PH, NH3-N, NO2-N, SD, and TYPE. The table lists 13 data points (index 241 to 254). To the right of the table, there are input and output properties for the model. The input properties include DO, PH, NH3-N, NO2-N, and SD. The output property is TYPE. The model parameters section on the right lists the following values: 显示间隔次数: 25.0, 最大循环次数: 100.0, 目标均方误差: 0.02, 动量常数: 0.0010, and 学习速度: 0.0010. The bottom of the interface shows a status bar with the current position '分类与回归-LM神经网络' and a '操作流程' (Operation Process) button.

公司拥有多年数据挖掘应用开发和实施经验，公司拥有一批熟悉数据挖掘技术的专家。数据挖掘相关项目案例如下：

单位：万元

项目名称	项目规模	服务客户	项目周期
电力生产数据挖掘分析平台软件		广东电科院	2014.6~2015.5
并行计算平台及营销应用模型研发项目		广西电科院	2013.5~2014.12
Hadoop 平台大数据挖掘算法框架开发		广东联通	2013.10~2014.10
水产养殖区环境综合评价专家系统		黄海水产研究所	2013.8~2013.12
决策支持系统营销域客服分析应用开发项目		广州供电局	2013.8~2014.8
计量自动化项目 Tivoli ITCAM 服务		广东电网公司	2013.10~2014.10
计量自动化项目 Tivoli ITCAM 服务		广州供电局	2013.10~2014.10
营销稽查数据挖掘与智能分析项目		广州供电局	2013.7~2014.12
服务器智能分析系统		广州市灵通新技术有限公司	2013.10~2014.10
IT 集中监控系统数据挖掘分析		同方鼎欣信息技术有限公司	2014.1~2015.7
TipDM(教研版)销售合同		甘肃萃英信息科技有限公司	2013.8~2014.7
数据挖掘培训服务合同		珠海格力电器股份有限公司	2014.4.17~19

公司拥有 10 台 IBM 服务器的大数据挖掘环境及公司组织编写的图书专著。





目前项目相关的开发及测试环境都已经基本具备。这些条件包括研发办公区 300 多平方米，办公位 50 个、企业和高校专业实验室共 3 个、交换机 3 台、数据挖掘服务器 10 台、SIP 服务器 2 台、版本控制服务器 1 台、文件服务器 2 台、PC 机 10 台、笔记本电脑 35 台等。

3、已取得与本项目相关的研究成果（包括专著、论文、专利等）：

广州泰迪公司已经取得的专利及软件著作权有：大菱鲆养殖水质评价的模糊神经网络专家系统的构建方法（专利号：201010011863.0）、一种自适应的防窃漏电诊断方法（专利号：201210287477.3）、一种基于海量计量数据的电力客户在线分群方法（专利号：201210484712.6）、一种电力行业中基于客户分群的电力负荷预测方法（专利号：201210484721.5）、一种智能防窃电监测装置（专利号：201320552008.X）、一种缺陷图像的自动分割方法（专利号：201310556840.1）；TipDM 商标（商标号：10828082）；基于云计算模式的餐饮智能服务平台软件（注册号：2014SR069119）、基于 hadoop 的电力用户智能分析与服务系统（注册号：2013SR114284）。

项目组主要研究人员相关专著及论文成果：

- 邓华丽.保险实务.中国财政经济出版社,2009.5
- 邓华丽.保险营销.中国人民大学出版社,2012.3
- 杨则文,邓华丽.国家税收（财政部规划教材）.中国财政经济出版社,2010.05
- 杨则文,邓华丽.国家税收学习指导与练习（财政部规划教材）.中国财政经济出

版社,2010.06

- 邓华丽.论促进西部开发的财税政策问题.经济研究导刊,2007.7
- 邓华丽,李修全.基于混沌时间序列分析的股票价格拐点预测方法.统计与决策,2007.5
- 邓华丽.银行风险控制及全口径风险管理系统实现.经济研究导刊,2007.6
- 邓华丽,张良均. 基于公司内在价值评价的证券策略投资研究,中国证券期货,2013 年第 05 期
- 邓华丽,张良均.模糊神经网络在企业信用评级中的应用. 统计与决策 2007 年第 07 期
- 邓华丽,张良均.基于人工神经网络的个人信用评估模型.番禺职业技术学院学报 2006 年 第 01 期
- 张良均,曹晶,蒋世忠.神经网络实用教程[M].北京:机械工业出版社,2008.2
- 张良均,陈俊德,刘名军,陈荣.数据挖掘:实用案例分析[M].北京:机械工业出版社,2013.7
- Xiaohui Hu, Liangjun Zhang.Spoken Arabic Digits Recognition Based on Wavelet Neural Networks, 978-1-4577-0653-0/11 ©2011 IEEE
- 张文阳,张良均,李娜,周红艳. 多元回归和 BP 人工神经网络在预测混合厌氧消化产气量过程中的应用比较,环境工程学报,2013 年第 2 期
- 宋晖,薛云,张良均. 基于 SVM 分类问题的核函数选择仿真研究,计算机与现代化,2011 年第 08 期
- 宋晖,薛云,胡晓晖,张良均. 基于人工神经网络的大气质量智能评价预警系统的设计与应用,现代计算机(专业版),2011 年 08 期

(四) 研究计划及进度安排

时间进度	阶段目标主要内容及成果
2015/1/1 至 2015/2/30	深入理解问题的财政收入的实际背景, 明确其实际意义, 掌握财政收入的各种信息。
2015/3/1 至 2015/5/30	分析对象和所需数据, 构建相关数据库
2015/6/1 至 2015/7/30	数据挖掘与财政收入相关的指标
2015/8/1 至 2015/10/30	分析建立数学模型以及对模型进行优化
2015/11/1 至 2016/4/30	编写模型实验程序



7415612255428

2016/5/1 至 2016/8/30	专利及论文材料整理
2016/9/1 至 2016/12/20	资料整理，生成准确分析报告准备验收

广州市科技计划项目

四、经费预算

单位：万元

经费开支科目	新增经费	其中 1：自 筹经费	2、市科信局 经费	使用说明
一、人员费	21	10	11	人员薪酬及奖励
二、设备费	2	1	1	用于购置设备及试制费
三、能源材料费	2	1	1	办公场地及材料费
四、外协费	0	0		用于检测分析费用
五、会议差旅费	1	1		用于会议及差旅费用
六、管理费	2	1	1	项目日常管理费用
七、其他费用	2	1	1	验目验收等费用
合计	30	15	15	

2. 申请单位自筹资金情况说明：

特此承诺！

3. 主管部门配套资金情况说明（如主管部门有配套经费支出请填写）：

主管部门（盖章）

年 月 日

五、项目组人员情况

1、项目负责人								
序号	姓名	身份证号	年龄	学历	职称/职务	分工	单位名称	签名
1								
2、项目组主要成员								
序号	姓名	身份证号	年龄	学历	职称/职务	分工	单位名称	签名
1								
2								
3								
4								
5								
6								

六、项目负责人及主要研究人员承担的省、市科技计划项目情况

姓名	项目编号	项目名称	起止日期	负责或参加	进展情况

七、申请单位意见

户名：广州番禺职业技术学院
账 号： 44074301040002340

开户银行： 中国农业银行广州番禺沙湾支行
(公章)

年 月 日

八、参加单位意见

参加单位 1：广州泰迪智能科技有限公司

负责人：(签名)

单位公章：

年 月 日

参加单位 2：

负责人：(签名)

单位公章：

年 月 日

九、推荐部门意见

负责人：(签名)

单位公章：

年 月 日

附件目录

序号	附件名称	附件类型
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		