

运用资料探勘技术于制程质量分析改善-以LCD驱动IC封装产业为例

Application Data Mining Technique to improve Manufacturing Quality -A Case Study of LCD Driver IC Packaging Industry

摘要(abstract)

近年来由于专业的分工,以LCD驱动IC封装厂来说,为提升产品良率及加速了解产品不良原因及找出对策,建置一产品质量问题分析模式,有效率,更快取得所需的信息,并提高客户对信息系统的满意度。

本研究以资料仓储的星状资料纲要作为在线分析处理的基础,用资料探勘的决策树方法及类神经网络方法来将封装厂制造过程不良原因,建构一质量分析系统,提供问题分析的界面,能够迅速掌握问题原因作出判断,达到缩短整个质量问题解决的时间。研究的结果,使用决策树(decision tree)方法及类神经网络(neural network)方法在降低内引脚(inner lead)不良颗数、树脂(resin)问题不良颗数及芯片问题不良颗数方面均获得改善, and决策树方法较类神经网络方法更适用于LCD驱动IC封装产业之质量问题分类分析。结果亦可用于计算机组装业、电子零件制造、半导体封装制造业质量问题分析。

关键词:客户关系管理、资料探勘、资料仓储、决策树

1. 简介(Introduction)

在半导体制造业讲求专业分工制造,上下游产业供应链明显,如何在产业中竞争,提升客户满意度以及维持客户忠诚度是企业竞争的关键因素。从Burgess[9]的研究结果得知,制造业竞争策略的优先级有「质量」(Quality)、「成本」(Cost)、「时间」(Time)以及「弹性」(Flexibility)等四项关键因素,因此企业如何提升产品良率、提高产品的质量、降低生产制造的成本、满足客户交货期限以及弹性制造需求等,这些都是刻不容缓的议题。

而在知识经济的时代中,企业体认到顾客关系管理的重要以及利用信息科技来提升竞争优势,近年来更有许多新的技术及观念应用出来,如人工智能、资料仓储、资料探勘等,藉由计算机化系统的运作,辅助企业进行资料处理、信息分析,并整合各来源客户信息并快速响应与解决问题,推进到主动行销与服务、协助企业作出各种合理的反应处理或决策动作[5, 10, 11] 亦即计算机化的境界可由最基础资料处理为目的的思考,提升并发展对信息的管理,进而形成决策辅助的目的,而运用知识规则达到智能决策最佳的目的[2]。

目前对于制造质量问题研究有针对半导体制程机台组合用分类式资料探勘,以

避免产生质量不良的机台组合,提高半导体制程的良率[14],以及利用决策树方法,针对半导体晶圆允收测试及相关制程资料,发展制程事故诊断资料探勘[15],及针对半导体生产制造资料,利用SOM以及决策树方法,找出生产绩效指针间的特征与规则,以做为半导体厂制造决策之依据[16]。

而针对半导体驱动IC封装产业中,目前尚无相关研究,因此充分整合及运用信息及系统功能,建构一个运用资料仓储及资料探勘的技术在分析制程异常的产生作为决策依据。透过基本及历史资料取得建立,分析过去产品质量的问题资料,能够快速提供找出产品问题原因及降低对问题处理因应时间,以提高客户的满意度。

2. 资料探勘

2.1. 知识发掘

资料探勘可视为资料库知识发掘(Knowledge Discovery in Database; KDD)的一部份,资料库知识发觉是一连串可创造知识的程序,资料探勘只是其中的一个步骤[4, 13]。Brachman 等人[6]认为:「知识探索中所有的活动及过程,是为了从资料中找到有用的样型,藉由运用资料探勘的算法,以及对于知识的事后处理或再处理,以找到解决问题的关键原因。资料探勘实为整个知识发现过程中最重要的核心步骤,而对于整个资料探勘的过程,其实是不间断地重复确认问题、分析资料、采取行动、评估结果四个步骤[7]。

Pyle [8]指出,利用资料探勘技术来发觉知识皆有其相似的地方。例如首要对欲解决的问题与需求做明确的定义,其次选择目标资料库,将对噪声资料进行排除,并且对资料进行编码来转化资料显示格式,以增加处理上的效能。

2.2. 资料探勘

在资料探勘的过程中,由于不同的问题定义会产生不同的探勘结果类型,以提供作为不同决策时使用。目前常见的问题类型如下:

1. 预测(Prediction):Berry 与Linoff[7, 12]认为预测会根据某些未来的行为来分类,或推估某变数未来可能的值以及趋势。
2. 关联规则(Association Rule):
Berson[12]等认为关联规则是从资料库探勘出满足一定条件的依赖性关系。
3. 聚类(Clustering/Segmentation):
Berson[12]等认为聚类可以根据一组特性,将资料库区隔成几个子集合。
4. 分类(Classification):分类所作的是为每一种类别找到一个合理的描述或模型,然后再利用这些分类的描述或模型,来对未知的资料进行分类[7]。

3. IC封装制程质量问题

3.1. LCD 驱动IC封装制程

随着IC制程的快速微细化，传统打线的封装技术已无法应付高电性需求，而直接以芯片之凸块(Bump)与基板連結的覆晶封装技术却能因应电子产品小型化与芯片性能高速化之趋势，使得满足轻、薄、短、小特性的芯片尺寸封装(CSP)蔚为风潮。由于以LCD驱动IC的封装为主要对象，LCD驱动IC系制造液晶显示器(LCD)零件之一，主要应用于笔记型计算机及液晶监视器大尺寸TFT-LCD面板上或行动电话、PDA、翻译机等小尺寸STN面板上。

LCD驱动IC封装制程大致可以区分成以下几道流程，其流程架构如图2所示：

建构LCD驱动IC封装制程产品质量问题分析系统，以质量的观点而言，必须找出每个工作站中可能会影响生产质量的因素，如表1所示主要为制造工作站中所包含的制程作业以及可能发生的缺陷。

3.2. 质量问题分析

质量分析系统主要是由资料库、在线分析处理、资料探勘等组件组成，如图3所示，主要分几个层级：

在资料准备阶段将进行资料汇整与筛选、资料清理与新资料产生与转换等3个资料准备的动作。因此建立问题资料格式的主要欄位必须包括当初制造制令编号、属性与问题。在资料属性的部分，大致定义成4个属性，分别为人员(man)、设备(machine)、材料(material)与方法(method)，一般俗称4M，其定义如下：

1. 人员因素：如未按照流程方法进行、漏检、参数未设定等。
2. 设备因素：乃有关设备运作的问题如压力过大、Tool头过高、Sensor无感应等。
3. 材料因素：为相关于材料本身的问题如胶过浓、内引脚过细等。
4. 方法因素：如程序设计不良、指示不明确、未使用无尘布等。

将属性变数及质量问题类别予以编号，以编号0, 1, 2, 3, ... 的方式，属性变数若为「无错误发生」则计为0，依序再按编号排列之。同样的，质量问题类别也以编号编排，即1, 2, 3, ... ，制作属性变数及质量问题编号对照表，如表2所示：

个案公司之质量问题资料搜集自2002年01月至2003年12月。资料总数为9752笔。问题类型总计15个类型，而资料探勘4个属性下的个别变数个数分别为：人员17个变数，设备13个变数，材料10个变数，方法12个变数。

4. 质量问题探勘系统设计

(1). 决策树算法

决策树是以树状资料结构为基础的分类分析方法。其详细的算法如下所示：

Step1: 设定根节点为目前节点C，此时所有对象都属于C的对象集合。

Step2: 如C中所有的对象都属于同一个类别，则设定节点C为此类别，然后停止，否

则继续执行Step3。

Step3: 计算其熵值 $E(C)$ 。若集合 C 分属于 j 个不同类别, $E(C)$ 为: $E(C) = -\sum p_j \log_2(p_j)$
其中 $p_j = (\text{属于类别}j\text{的总对象数}) / (C\text{的总物件数})$ 。

Step4: 针对所有尚未出现在从根节点到目前节点之路径的属性 A_i (称之为候选属性), 以 A_i 分别对 C 的对象集合进行分割, 并计算所导致之部分决策树的熵值 $E(A_i)$ 与信息获利 $G(A_i) = E(C) - E(A_i)$ 。选择某一属性 A_i 为决策树节点, 意味着在此节点下建立 m 个子节点。因此, 以属性 A_i 为决策树节点所产生的子决策树之熵值计算如下:

$$E(A_i) = \sum (n_k/n) * E(C_k)$$

其中 C_k 为对象集合 C 中其 A_i 属性值相同的对象子集合 k

$E(C_k)$ 为对象子集合 C_k 之熵值

n 为对象集合 C 的总对象数

n_k 对象子集合 C_k 的对象数

信息获利(Information Gain)是指因选择属性 A_i 为决策树节点而带来熵值的改变程度, 也就是原来对象集合的熵值与以 A_i 为决策树节点的子决策树之熵值间的差距, 计算如下:

$$G(A_i) = E(C) - E(A_i)$$

Step5: 计算获利比(Gain-Ratio)做为属性选择衡量方法, 其计算如下:

$GR(A_i) = G(A_i) / IV(A_i)$ 其中 $G(A_i)$ 为 A_i 属性对 C 的对象集合进行分割所得的信息获利 $IV(A_i)$ 为 A_i 属性的信息值 $= -\sum p_m \log_2 p_m$ ($m=1..k$) 其中 $p_m = (C\text{集合中属于}A_i\text{属性值为}a_m\text{的总对象数}) / (C\text{的总物件数})$

Step6: 选择获利比最大的候选属性, 将之当成节点 C 的分类属性。

Step7: 在节点 C 下建立子节点 C_1, C_2, \dots, C_m (假设所选定的分类属性有 m 个属性值), 依据分类属性值, 将 C 中的所有对象分派至适当的子节点中。

Step8: 对每个子节点 C_i 当成目前节点 C , 并分别由Step2继续执行。

C4.5 修剪决策树的标准是以预估错误率(Predicted Error Rate)的值作为判断条件。错误比率来评估其它非学习资料的错误比率。假设被分类于某一子树的所有学习资料个数为 N , 其中有 E 个学习资料分类错误, 而错误预估率就是由 E/N 来评估当有新的资料被测试时, 利用此一条件判断结果为不正常的机率值[1]。

表3为LCD驱动IC封装产业质量问题资料, 设备、材料、方法及人员为主要预测变数:

一般在决策树的每一个节点上使用信息增益度量来选择测试属性, 假设 S 是 s 个资料样本的集合。假设类别卷标属性具有 m 个不同值, 定义 m 个不同类

$C_i (i=1, \dots, m)$ 。设 S_i 是类 C_i 的样本数。对于一个给定的样本分类所需的期望信息由下列式子算出:

(a). 计算预测变数的期望信息与信息获利:

$$I(s_1, s_2, \dots, s_m) = -\sum_{i=1}^m p_i \log_2(p_i) \quad I=1, \dots, m$$

(1). 计算材料变数期望信息与信息获利：

对于材料= " 胶本身过浓 "

$$S_{11}=2 \quad S_{21}=3 \quad I(S_{11}, S_{21})=0.971$$

对于材料= " 胶本身气泡 "

$$S_{12}=4 \quad S_{22}=0 \quad I(S_{12}, S_{22})=0$$

对于材料= " Tape本身内缩 "

$$S_{13}=3 \quad S_{23}=2 \quad I(S_{13}, S_{23})=0.971$$

$$I(S_1, S_2)=I(9, 5)=0.940$$

$$\text{期望信息 } E(\text{材料})=(5/14)*I(S_{11}, S_{21})+(4/14)*I(S_{12}, S_{22})+(5/14)*$$

$$I(S_{13}, S_{23})=0.694$$

$$\text{信息获利 Gain(材料)}=I(S_1, S_2)$$

$$-E(\text{材料})=0.246$$

(2). 计算设备变数期望信息与信息获利：

对于设备= " 无 "

$$S_{11}=3 \quad S_{21}=1 \quad I(S_{11}, S_{21})=0.811$$

对于设备= " Marking压力有异 "

$$S_{12}=2 \quad S_{22}=2 \quad I(S_{12}, S_{22})=1$$

对于设备= " Curing温度有异 "

$$S_{13}=4 \quad S_{23}=2 \quad I(S_{13}, S_{23})=0.918$$

$$\text{期望信息 } E(\text{设备})=(4/14)*I(S_{11}, S_{21})+(4/14)*$$

$$I(S_{12}, S_{22})+(6/14)*I(S_{13}, S_{23})=0.911$$

$$\text{信息获利 Gain(设备)}=I(S_1, S_2)$$

$$-E(\text{设备})=0.940-0.911=0.029$$

(3). 计算方法变数期望信息与信息获利：

对于方法= " 温度有异 "

$$S_{11}=3 \quad S_{21}=4 \quad I(S_{11}, S_{21})=0.985$$

对于方法= " 无 "

$$S_{12}=6 \quad S_{22}=1 \quad I(S_{12}, S_{22})=0.592$$

$$\text{期望信息 } E(\text{方法})=(7/14)*I(S_{11}, S_{21})$$

$$+(7/14)*I(S_{12}, S_{22})=0.789$$

$$\text{信息获利 Gain(方法)}=I(S_1, S_2)$$

$$-E(\text{方法})=0.151$$

(4). 计算人员变数期望信息与信息获利：

对于人员= " 无 "

$$S_{11}=6 \quad S_{21}=2 \quad I(S_{11}, S_{21})=0.811$$

对于人员= " Marking压力有異 "
 $S_{12}=3 \quad S_{22}=3 \quad I(S_{12}, S_{22})=1$
 期望信息 $E(\text{人员})=(8/14)*I(S_{11}, S_{21})$
 $+ (6/14)* I(S_{12}, S_{22})=0.892$
 信息获利 $\text{Gain}(\text{人员})= I(S_1, S_2)$
 $-E(\text{人员})=0.048$

(b). 计算信息获利比Gain-Rate(GR) :

GR(材料)
 $=0.246/(-(5/14)*\log_2(5/14)-(4/14)*(\log_2(4/14)-5/14*\log_2(5/14)))=0.156$

GR(设备)
 $=0.029/(-(4/14)*\log_2(4/14)-(4/14)*(\log_2(4/14)-6/14*\log_2(6/14)))=0.019$

GR(方法)=
 $0.151/(-(7/14)*\log_2(7/14)-(7/14)*(\log_2(7/14)))=0.151$

GR(人员)=
 $0.048/(-(8/14)*\log_2(8/14)-(6/14)*(\log_2(6/14)))=0.048/0.986=0.049$

由于材料在属性中有最高信息获利比，它被选作测试属性，因而以此建立节点，其它则配合修剪演算或限制深度等作法达到决策树分类的目的。

(2). 類神经网络算法

類神经网络是由一组神经元节点(node)与具有权重的节点间連結(link)所构成的网路。神经网络的多层结构可分为输入层(input layer)、隐藏层(hidden layer)以及输出层(output layer)，而层与层间的神经元则以完全連結(fully connected)的方式向前连接着(forward connected)，其算法如下所示[5]：

假设類神经网络为单一隐藏层，输入层有*i*个单元，隐藏层有*j*个单元，则连接权重以 W_{ij} 表示，容忍误差以 ϵ_j 表示

Step1: 设定网路参数，包括各层神经元数目，学习率1、训练次数、容忍误差 ϵ_j 。

Step2: 设定网路的初始权值 W_{ij} ；乃随机产生之亂数值。

Step3: 将输入向量及输出向量依神经元转换函数之值域，分别将其正规化(normalize)。神经元转换函数为 $f(x)=1/(1+e^{-x})$

Step4: 计算隐藏层及输出层之输出。

第*j*单元输入向量为 $I_j = \sum W_{ij} * O_i + \theta_j$

第*j*单元输出向量 $O_j = 1/(1+e^{-I_j})$

Step5: 计算输出层及隐藏层之误差。

输出层误差为 $\text{Err}_j = (O_j(1-O_j)(T_j-O_j))$

隐藏层误差为 $\text{Err}_j = (O_j(1-O_j) \sum \text{Err}_k * W_{kj})$

其中 W_{kj} 是由下一较高层中单元*k*到单元*j*的连接权重，而 Err_k 而是单元*k*的误差，

而是j基于给定的已知类别标签的真正输出。

Step6: 计算网路权值修正量, 并修正之。

$$\text{网路权值修正量}(W_{ij} = W_{ij} + L * \text{Err}_j * O_i)$$

$$\text{误差修正量}(\text{Err}_j = \text{Err}_j + L * \text{Err}_j)$$

Step7: 跳回Step3, 重复计算直到误差量小于容忍误差或指定的训练次数为止。

图4为一多层前馈神经网络范例, 假设学习率为0.9。

(a). 该网路的初始权值和偏差值以及第一个训练样本 $X=(1, 0, 1)$, 如表4所示。

即学习率 $L=0.9$ 、容忍误差 $\epsilon=0.01$ ($j=4, 5, 6$)、初始权值 W_{ij} ($i=1, 2, 3, 4, 5; j=4, 5, 6$) 乃随机产生之乱数值。

(b). 给定第一个训练样本 X , 首先将样本提给

网路, 并计算每个单元的净输入和输出这些值表示在表5中。依转换函数为 $f(x)=1/(1+e^{-x})$ 计算隐藏层及输出层之输出。

第j单元输入向量为 $I_j = \sum W_{ij} * O_i + \theta_j$, 第j单元输出向量 $O_j = 1/(1+e^{-I_j})$

$$I_4 = \sum W_{i4} * O_i + \theta_4 = -0.7$$

$$O_4 = 1/(1+e^{-0.7}) = 0.332$$

$$I_5 = \sum W_{i5} * O_i + \theta_5 = 0.1$$

$$O_5 = 1/(1+e^{-0.1}) = 0.525$$

$$I_6 = -0.105$$

$$O_6 = 1/(1+e^{-0.105}) = 0.474$$

(c). 接着计算每个单元的误差, 并向后传播。此误差值在表6中。

输出层之误差为 $\text{Err}_j = O_j(1-O_j)(T_j-O_j)$, 隐藏层之误差为 $\text{Err}_j = (O_j(1-O_j) * \sum \text{Err}_k * W_{kj})$ 其中 W_{kj} 是由下一较高层中单元k到单元j的连接权重, 而 Err_k 而是单元k的误差, 而是j基于给定的已知类别标签的真正输出

而更新后的权重和偏差在表7中。误差修正量($\text{Err}_j = \text{Err}_j + L * \text{Err}_j$), 网路权值修正量($W_{ij} = W_{ij} + L * \text{Err}_j * O_i$)

$$\text{Err}_6 = (O_6(1-O_6)(T_6-O_6)) = 0.1311$$

$$\text{Err}_5 = (O_5(1-O_5) * \sum \text{Err}_k * W_{5k}) = -0.0065$$

$$\text{Err}_4 = (O_4(1-O_4) * \sum \text{Err}_k * W_{4k}) = -0.0087$$

(d). 而更新后的权重和偏差在表7中。误差修正量($\text{Err}_j = \text{Err}_j + L * \text{Err}_j$), 网路权值修正量($W_{ij} = W_{ij} + L * \text{Err}_j * O_i$)

$$W_{66} = W_{66} + L * \text{Err}_6 * O_6 = 0.1 + 0.9 * (0.1311) = 0.218$$

$$W_{55} = W_{55} + L * \text{Err}_5 * O_5 = 0.2 + 0.9 * (-0.0065) = 0.194$$

$$W_{44} = W_{44} + L * \text{Err}_4 * O_4 = -0.408$$

$$W_{46} = W_{46} + L * \text{Err}_6 * O_4 = -0.261$$

$$W_{56} = W_{56} + L * \text{Err}_6 * O_5 = -0.2 + 0.9 * 0.1311 * 0.525 = -0.138$$

$$W_{14} = W_{14} + L * \text{Err}_4 * O_1 = 0.2 + 0.9 * (-0.0087) * 1 = 0.192$$

$$W_{15} = -0.3 + 0.9 * (-0.0065) * 1 = -0.306$$

$$W_{24} = 0.4 + 0.9 * (-0.0087) * 0 = 0.4$$

$$W_{25} = 0.1 + 0.9 * (-0.0065) * 0 = 0.1$$

$$W_{34} = -0.5 + 0.9 * (-0.0087) * 1 = -0.508$$

$$W_{35} = 0.2 + 0.9 * (-0.0065) * 1 = 0.194$$

(e). 重复如此步骤直至学习结束为止。

5. 实作

5.1. 系统环境

产品质量问题分析系统的实作环境，分为以下几个部份：

1. MES Server: 记录每日生产交易资料。
2. ERP server: 记录每日企业各项营运流程交易资料。
3. DataMart Server: 整合ERP以及MES关于分析系统所需资料纲要及资料。
4. Data Mining Server: 作为资料探勘系统运作环境。
5. OLAP Server: 从DataMart Server撷取资料，作OLAP分析。
6. Client PC: 以可视化呈现方式，供使用者操作分析用。

而质量问题分析系统整体环境架构图如图5所示：

5.2. 资料探勘模式实作

5.2.1. 决策树方法

1. 即由已建立的资料仓储中选取质量问题的资料，并定义出属性变数栏及类别变数栏，属性变数为人员、设备、材料、方法，类别变数为质量问题。
2. 设定决策树算法的各项参数，如是否修剪决策树、决策树最高深度、训练组及测试组比例等，取得个案公司2002年1月至2003年12月计24个月资料，资料总笔数9752笔，以乱数选取方式将训练组及测试组以3:1比例作分类分析。
3. 决策树算法针对先前所设定的参数，其实际与推估列出矩阵如下，其训练组的总笔数为7314笔，训练组分类正确笔数计5851笔，因此正确率高达80%，如图6所示。
4. 最后依决策树算法针对先前所设定的参数以乱数选取测试组的方式进行预测推估验证，测试组总笔数为2438笔，实际与推估列出矩阵如下，其测试组的分类正确比数计1682笔，正确率高达69%，如图7所示。

5.2.2. 类神经网络方法

1. 将质量问题的资料建立起单一表格资料库，定义出属性变数栏及类别变数栏。

2. 设定類神经网络算法的各项参数，如学习率、隐藏层层数、训练组及测试组比例等，个案公司2002年1月至2003年12月计24个月资料，资料总笔数9752笔，以亂數选取方式将训练组及测试组以3：1比例作類神经网络分類分析。
3. 類神经网络算法针对先前所设定的参数，其训练组的模拟学习的正确率，如图8所示跑出实际与学习推估分類矩阵图，资料总笔数9752笔，其训练组的总笔数为7314笔，训练组分類正确笔数计4535笔，正确率高达62%。
4. 最后依類神经网络算法针对先前所设定的参数以亂數选取测试组的方式进行预测推估验证，其实际与推估列出矩阵如下，测试组总笔数为2438笔，其测试组的分類正确比数计1292笔，正确率高达53%，如图9所示類神经网络实际与预测分類推估矩阵图。

5.3. 效益分析

根据所建置的资料库，决策树及類神经网络的方法，所达成的效益有下列几项：

1. 降低内引脚不良颗数：

以2004年1至6月系统使用前内引脚不良颗数如表8所示，而分别以决策树及類神经网络的方法在2004年7至12月系统使用后内引脚不良颗数如表8所示，决策树方法总平均改善率为10.4%；同样以類神经网络方法总平均改善率为2.95%，而内引脚不良颗数改善前后统计曲线图如图10所示。

2. 降低树脂问题不良颗数：

以2004年1至6月系统使用前树脂问题不良颗数如表9所示，而分别决策树及類神经网络的方法在2004年7至12月系统使用后树脂问题不良颗数如表9所示，以决策树总平均改善率为9.2%；同样以類神经网络方法总平均改善率为4.4%，而树脂问题不良颗数改善前后统计曲线图如图11所示。

3. 降低芯片问题不良颗数：

以2004年1至6月系统使用前芯片问题不良颗数如表10所示，而决策树及類神经网络的方法在2004年7至12月系统使用后芯片问题不良颗数如表10所示，决策树总平均改善率为11.6%；同样以類神经网络总平均改善率为4.6%，而芯片问题不良颗数改善前后统计曲线图如图12所示。

针对个案公司以资料探勘决策树算法及類神经网络算法针对质量问题资料进行分类分析，资料总笔数为9752笔，以随机取样方式将资料分为训练组及测试组，其比例为3:1，训练组笔数7314笔，测试组2438笔，以决策树算法作分類分析其训

練組正确筆數5851，正确率达80%，测试組正确筆數1682，正确率达69%；而類神经网络算法作分類分析，其訓練組正确筆數4535，正确率达62%，测试組正确筆數1292，正确率达53%；另外分別以决策树及類神经网络的方法在2004年7至12月系统使用后，决策树在内引脚不良顆數总平均改善率为10.4%，類神经网络方法总平均改善率为2.95%；决策树在树脂问题不良顆數总平均改善率为9.2%，類神经网络方法总平均改善率为4.4%；决策树在芯片问题不良顆數总平均改善率为11.6%，類神经网络方法总平均改善率为4.6%，由此可知决策树算法较适用于质量问题分類分析，如表11所示。

6. 結論

本研究的总结主要提出适用于LCD驱动IC封装厂产品质量问题分析模式并以案例公司研究，以LCD驱动IC封装厂來說，为提升产品良率及加速了解产品不良原因及找出对策，建置一产品质量问题分析模式，可使工作更有效率，更快取得真正所需的信息，并提高客户对信息系统的满意度。

在质量问题分析的效益上，决策树在内引脚不良顆數总平均改善率为10.4%，树脂问题不良顆數总平均改善率为9.2%，芯片问题不良顆數总平均改善率为11.6%；類神经网络方法在内引脚不良顆數总平均改善率为2.95%，树脂问题不良顆數总平均改善率为4.4%，芯片问题不良顆數总平均改善率为4.6%，决策树方法较類神经网络方法适用于质量问题分類分析。

此方法可适用于计算机组装业、电子零件制造、半导体封装制造业质量问题分析。