**DECISION TREE APPLICATIONS**

Business Management

**In the past decades,** many organizations had created their own databases to enhance their customer services. Decision trees are a possible way to extract useful information from databases and they have already been employed in many applications in the domain of business and management. In particular, decision tree modelling is widely used in customer relationship management and fraud detection, which are presented in subsections below.

Customer Relationship Management

**A frequently used approach to manage customers’** relationships is to investigate how individuals access online services. Such an investigation is mainly performed by collecting and analyzing individuals’ usage data and then providing recommendations based on the extracted information. Lee et al. (2007) apply decision trees to investigate the relationships between the customers’ needs and preferences and the success of online shopping. In their study, the frequency of using online shopping is used as a label to classify users into two categories: (a) users who rarely used online shopping and (b) users who frequently used online shopping. In terms of the former, the model suggests that the time customers need to spend in a transaction and how urgent customers need to purchase a product are the most important factors which need to be considered. With respect to the latter, the created model indicates that price and the degree of human resources involved (e.g. the requirements of contacts with the employees of the company in having services) are the most important factors. The created decision trees also suggest that the success of an online shopping highly depends on the frequency of customers’ purchases and the price of the products. Findings discovered by decision trees are useful for understanding their customers’ needs and preferences.

Fraudulent Statement Detection

**Another widely used business application is the detection of Fraudulent Financial Statements (FFS).** Such an application is particularly important because the existence of FFS may result in reducing the government’s tax income (Spathis et al., 2003). A traditional way to identify FFS is to employ statistical methods. However, it is difficult to discover all hidden information due to the necessity of making a huge number of assumptions and predefining the relationships among the large number of variables in a financial statement.

**Previous research has proved that** creating a decision tree is a possible way to address this issue as it can consider all variables during the model development process. Kirkos et al. (2007) have created a decision tree model to identify and detect FFS. In their study, 76 Greek manufacturing firms have been selected and their published financial statements, including balance sheets and income statements, have been collected for modelling purposes. The created tree model shows that all non-fraud cases and 92% of the fraud cases have been correctly classified. Such a finding indicates that decision trees can make a significant contribution for the detection of FFS due to a highly accurate rate.

Engineering

The other important application domain that decision trees can support is engineering. In particular, decision trees are widely used in energy consumption and fault diagnosis, which are described in subsections below.

Energy Consumption

**Energy consumption concerns how much electricity has been used by individuals.** The investigation of energy consumption becomes an important issue as it helps utility companies identify the amount of energy needed. Although many existing methods can be used for the investigation of energy consumption, decision trees appear to be preferred. This is due to the fact that a hierarchical structure provided by decision trees is useful to present the deep level of information and insight. For instance, Tso and Yau (2007) create a decision tree model to identify the relationships between a household and its electricity consumptions in Hong Kong. Findings from their tree model illustrate that the number of household members are the most determinant factor of energy consumption in summer, and both the number of air-conditioner and the size of a flat are the second most important factors. In addition to such findings, their tree model identifies that a household with four or more members with a flat size larger than 817ft2 is the highest electricity consumption group. On the other hand, households which have less than four family members and without air-conditioners are the smallest electricity consumption group. Such findings from decision trees not only provide a deeper insight of the electricity consumptions within an area but also give guidelines to electricity companies about the right time they need to generate more electricity.

Fault Diagnosis

**Another widely used application in the engineering domain is the detection of faults,** especially in the identification of a faulty bearing in rotary machineries. This is probably because a bearing is one of the most important components that directly influences the operation of a rotary machine. To detect the existence of a faulty bearing, engineers tend to measure the vibration and acoustic emission (AE) signals emanated from the rotary machine. However, the measurement involves a number of variables, some of which may be less relevant to the investigation. Decision trees are a possible tool to remove such irrelevant variables as they can be used for the purposes of feature selection. Sugumaran and Ramachandran (2007) create a decision tree model to identify the features that may significantly affect the investigation of a faulty bearing. Through feature selection, three attributes were chosen to discriminate the faulty conditions of a bearing, i.e., the minimum value of the vibration signal, the standard deviation of the vibration signal, and kurtosis. The chosen attributes, subsequently, were used for creating another decision tree model. Evaluations from this model show that more than 95% of the testing dataset has been correctly classified. Such a highly accurate rate suggests that the removal of insignificant attributes within a dataset is another contribution of decision trees.

Healthcare Management

**As decision tree modelling can be used for making predictions,** there are an increasing number of studies that investigate to use decision trees in health-care management. For instance, Chang (2007) has developed a decision tree model on the basis of 516 pieces of data to explore the hidden knowledge located within the medical history of developmentally-delayed children. The created model identifies that the majority of illnesses will result in delays in cognitive development, language development, and motor development, of which accuracies are 77.3%, 97.8%, and 88.6% respectively. Such findings can result in assisting healthcare professional to have an early intervention on developmentally-delayed children so as to help them catch up their normal peers in their development and growth. Another example of health-care management can be found in Delen et al. (2005). In their study, a decision tree is created to predict the survivability of breast cancer patients. The classification accuracy is 93.6% in their decision tree. This classification rate indicates that the created tree is highly accurate for predicting the survivability of breast cancer patients. These studies suggest that decision tree is a useful tool to discover and explore hidden information in health-care management.