Chapter XI
Application of Decision Tree as a Data Mining Tool in a Manufacturing System

S. A. Oke
University of Lagos, Nigeria

ABSTRACT

This work demonstrates the application of decision tree, a data mining tool, in the manufacturing system. Data mining has the capability for classification, prediction, estimation, and pattern recognition by using manufacturing databases. Databases of manufacturing systems contain significant information for decision making, which could be properly revealed with the application of appropriate data mining techniques. Decision trees are employed for identifying valuable information in manufacturing databases. Practically, industrial managers would be able to make better use of manufacturing data at little or no extra investment in data manipulation cost. The work shows that it is valuable for managers to mine data for better and more effective decision making. This work is therefore new in that it is the first time that proper documentation would be made in the direction of the current research activity.

INTRODUCTION

General Overview

In today's digital economy, knowledge is regarded as an asset, and the implementation of knowledge management supports a company to developing

innovative products and making critical management strategic decisions (Su, Chen, & Sha, 2005). This digital economy has caused a tremendous explosion in the amount of data that *manufacturing organizations* generate, collect, and store, in order to maintain a competitive edge in the global business (Sugumaran & Bose, 1999). With global competition, it is crucial for organizations to be able to integrate and employ intelligence knowledge in order to survive under the new *business environment*. This phenomenon has been demonstrated in a number of studies, which include the employment of artificial neural network and decision tree to derive knowledge about the job attitudes of "Generation Xers" (Tung, Huang, Chen, & Shih, 2005). The paper by Tung et al. (2005) exploits the ART2 neural model using the collected data as inputs. Performance classes are formed according to the similarities of a sample frame consisting of 1000 index of Taiwan manufacturing industries and service firms. While there is a plethora of *data mining techniques and tools* available, they present inherent problems for end-users such as complexity, required technical expertise, lack of flexibility, and interoperability, and so on. (Sugumaran & Bose, 1999). Although in the past, most data mining has been performed using symbolic artificial intelligence *data mining algorithms* such as C4.5, C5 (a fast variant of C4.5 with higher predictive accuracy) and CART (Browne, Hudson, Whitley, Ford, & Picton, 2004), the motivation to use *decision tree* in this work comes from the findings of Zhang, Valentine, & Kemp, (2005). The authors claim that *decision tree* has been widely used as a modelling approach and has shown better predictive ability than traditional approaches (e.g., regression). This is consistent with the literature by considering the earlier study by Sorensen and Janssens (2003). The authors conduct an exploratory study that

Figure 1. Data generated in a modern manufacturing system