Chapter 15
Intelligent Information Processing

ABSTRACT

The data stored in a database exhibits a certain pattern at different degrees of abstraction although by nature it is random. The availability of patterns in the data would be useful in classifying the same with a tag and clustering the same. Query over such a data cluster would provide a quick response. In a huge database, it is difficult to come out with exact patterns. The result of classification and clustering is often probabilistic. As a result, the estimation would turn statistical calling for the merger of the responses. In this chapter, the clustering and classifier algorithms are explained along with the decision process to merge the results.

INTRODUCTION

In the previous chapter, the application of the pattern matching in the data for the retrieval of the required content is discussed. To make it happen, the data needs to be organized and grouped in to clusters based on the similar patterns. Here, the methods used for the identification of patterns in the data are explored.

Pattern refers to the ability to depict a certain known features in the observed data or information. The algorithms and applications making use of pattern matching work on the assumption that the patterns remain invariant when they undergo transformations. Specifically, human observers take it for granted. The typical transformations include shift, rotation and scaling in a plane. In spite of these transformations, a human observer is able to recognize patterns in the handwritten characters, faces, voices etc, with different size, position and orientation.

DOI: 10.4018/978-1-60566-888-8.ch015
Like biological systems, neural networks have the ability to learn and recognize patterns in the data. They are capable of generalizing the results (Arun K. Pujari, 2000). I.e., they are able to find out the right class of the patterns, to which they have not been exposed in the training phase. The output provided is invariant to the distortions produced at the input. For a complex task such as face recognition (J. Dahmen, D. Keysers, H. Ney, and M. O. G"uld, 2001) a large number of neural networks are required to be used collectively. It is a tough task to collate their results and inferences, calling for the use of an intelligent element.

By providing historical data to a neural network, it is possible to train the network towards an ideal estimator. It happens by imparting feedback of various degree where by the network will learn adequately to generalize the inferences or observations.

Pattern recognition (Ripley B. D, 1996) is the science and art for looking for patterns in the data. It makes use of features extracted from the data, data clustering and classification, comparison and error analysis. It is used in automation applications including character recognition, image processing, speech analysis and speaker recognition, person identification etc. Statistical approach and the artificial neural networks are popularly used for pattern recognition application. (Bhagat, P. M, 2005). A variety of models are supported depending up on the complexity. If the data and the patterns are simple, simple models are sufficient. The conclusions derived and the decisions made are largely dependent on the model. (A. Krause, 2002). Complex models are required to learn complex patterns. The complexity of the model may be increased by adding more variables. It helps in capturing the contribution from each of them and increases the accuracy. Depending up on the model, various techniques are employed to improve the scalability and complexity. For example, in a neural network model, the number of hidden layers controls the complexity, while in a regression model, the number of polynomial terms and the interactions take the role. The complexity of a tree based model depends up on the number of branches. All these models are highly scalable.

Today, databases used for supply chain management are too large and grow at a faster rate making it difficult for the algorithms to analyze the data completely. In this chapter the tools, techniques and algorithms are introduced to find the patterns of interest in large databases and data streams.

For some of the applications, the requisite background knowledge can be provided through a Bayesian network. This background knowledge is integrated with the discovery process to get a new set of patterns in addition to the one known already. A neural network with a good background knowledge acquired through historical data is able to handle arbitrarily large databases and Bayesian networks that are too large for the exact inference to be feasible.

This chapter is concerned with the simple optimization problem of the applications around clustering. It is expected to provide insight on pattern recognition techniques making use of a feedback neural network.

**BACKGROUND**

Today, the different players generate information in heterogeneous format including video, audio, speech and machine generated data formats all along the supply chain. It is very difficult to synchronize and organize this information. The help of automated programs and machines intelligence are sought for the same.