CHAPTER - 1

INTRODUCTION
CHAPTER- 1

INTRODUCTION

To operate in this world shaped by globalization and information revolution, it is necessary to collect, analyze, understand, and act rightly upon the information which successfully leads to the target [1]. To be more perfect the highest level of data accuracy and consistency should be the two essentialities needed in the real time applications. Approaching in this way, eventually the partner of any business or management should predict on making right moves to stay competitive. The vital component of the success is the prediction and execution of a plan that supports the strategic goals and drives towards all targeted decisions.

Recent research publications in the area of Computer Science and Information Technology witness how the business events are performed by mining the hidden valuable information to improve the vital business operations like process, flow, data organization and social structures [2]. Especially in the domain of managerial decision making, the event logs of the system, popularly known as ‘transaction log’ [3] are oriented to some business activities leading to a process. The event logs are to be incorporated effectively in a planned
automation process so that an efficient system of business management is set up [4].

Pursuing this line, many system tools are developed in the market and the commercial interest of these tools towards the business domain is mounting due to the increasing awareness of the companies to incorporate effectively the same in the cross functional entities[5]. Hence there is a need for a system which maps all business functions. The design architecture of any proposed system in exploring the information has to serve all the required functions. The present work is so devised that it builds a system which acts more intelligently than the legacy system competing in the information age. This has been achieved by designing a hybrid computational model involving contemporary algorithms. The design of the system [6] is fabricated in the way that the system is not only application-intelligent but also compatible in managing the concepts of time complexity and usage memory utility.

1.1. INFORMATION SYSTEM

An Information System (IS) [1] is a collective system of data, records, persons and activities. The system processes the data to derive useful information in an organization. Usually the term is used
erroneously as a synonym for computer-based information systems, which is only the information technology component of an Information System [7]. The computer-based information systems are the specific field of study of interest in the area of Information Technology (IT).

The term information system has different meanings in different fields. As we are concerned with knowledge representation and decision making, we define an information system as one that consists of data through three levels of activity – (i) syntax level, (ii) semantic level, and (iii) knowledge level.

Data which can be automatically processed by the application system correspond to the syntax-level. In the context of an individual who interprets the data they become information, which correspond to the semantic-level. Information becomes knowledge when an individual understands and evaluates the information for a specific task.

Embossing the said concept, the present work aims at developing the hybrid intelligent information system architecture which enables the construction of a good business information system
evaluating knowledge discovery and decision making to suggest suitable business intelligence activities.

Besides mapping the various business activities, the system of monitoring these activities is also needed as illustrated by today’s buzz words such as Business Activity Monitoring (BAM), Business Operations Management (BOM), and Business Process Intelligence (BPI). The goal of building such a good business infrastructure encompassed by business processes starts with mining of information.

With regard to mining, there are three different perspectives as follows:

(i) the process perspective (‘How’), (ii) the organization perspective (‘Who’), and (iii) the case perspective (‘What’).

The process perspective focuses on the control flow of processes whereas the organizational perspective focuses on the inter relations between the performers. The case perspective focuses on properties of cases. Thus as an overview, the information system design is hybrid to encompass all the above mentioned features to act as intelligent on applications.
1.2. BUSINESS INTELLIGENCE

Business Intelligence (BI) is a business management [8] term, which refers to applications and technologies that are used to gather, provide access to, and analyze data about company operations. Business intelligence systems can help companies to have more comprehensive knowledge of the factors affecting their business, such as metrics on sales, production, internal operations, and they can help companies to make better business decisions. Business Intelligence should not be confused with competitive intelligence, which is a separate management concept.

Business intelligence applications and technologies [9] can enable organizations to make more informed business decisions, and may give a company a competitive advantage. For example, a company could use business intelligence applications or technologies to extrapolate information from indicators in the external environment and forecast the future trends in their sector. Business intelligence [10] is used to improve the timeliness and quality of information and enable managers to better understand the position of their firm in comparison with its competitors. When a BI system is well-designed and properly
integrated into a company's processes and decision-making process, the company's performance may be well set for improvement.

Managing timely and accurate information is the prime factor for a company to practice business intelligence, which will expedite decision-making and improve customer relationship. Business intelligence applications and technologies can help companies analyze the following:

(i) Changing trends in market share,
(ii) changes in customer behavior and spending patterns,
(iii) customers' preferences, and
(iv) company capabilities and market conditions.

Business intelligence can be used to help analysts and managers in determining the most likely adjustments to affect trends. Business intelligence systems can help companies develop consistent and "data-based" business decisions [11] producing better results than making decisions on "guesswork."

Proceeding on this line, the work develops a decision support information system predicting suggestions and decisions in the area of customer trading in stock market. Almost in all stock trading sectors,
the online trading [12] has come into practice. While practicing such online trading or even the conventional trading, other than the buy and sell deal of the customer the value assets of the customer investing in the trade have to be also analyzed. This analysis streamlines the customer behavior and enhances the valued added benefits of the customer. It also helps in avoiding default conditions like customer facing the risk of loss in investment and inability of making prompt payments on delivery of units or derivatives bought.

1.3. DATA MINING

Data mining, also known as “knowledge discovery in databases” [13], is the process of discovering interesting patterns in databases that are useful in decision making. Data mining is a discipline of growing interest and importance, and an application area that can provide significant competitive advantage to an organization by exploiting the potential of large data warehouses.

The task of finding patterns in business data is not new. Traditionally, it was the responsibility of business analysts, who generally used statistical techniques [14]. The scope of this activity, however, has recently changed. Widespread use of computers and
Networking technologies has created large electronic databases that store business transactions.

Retailers capture millions of sales transactions through their point-of-sale terminals. Transactions can be analyzed to identify buying patterns of individual customers as well as customer groups, and sales patterns of different stores. Intense competition is forcing companies to identify innovative ways to capture and enhance market shares at reduced cost.

Data mining methodologies has enabled companies to organize and store large volumes of business data in a form that can be analyzed. The field of “computational intelligence” has created a set of techniques of “machine learning” that are useful in automating tedious and crucial activities of discovering patterns in databases. These factors have changed the way that business data are analyzed and given rise to data mining, which integrates machine learning, statistical analysis and visualization techniques. With the intuition and knowledge of the business analyst, meaningful and interesting predictions are made in business data employing computational intelligence.
Data mining is a complex process involving multiple iterative steps. Figure 1 gives an overview of this process. The first step is the selection of data for analysis. Selection of data involves problem definition and acquisition of background knowledge. Normally, for this process historical data is used and the data set may be retrieved from a single source, such as a data warehouse, or may be extracted from several operational databases. The selected data set then undergoes cleaning and preprocessing. Lack of consistency across database creates serious problems when data is extracted from multiple
databases. The cleaning operation removes discrepancies and inconsistencies in the data and improves its quality. Examples include transformation of data from one scale to another, identification of predictive attributes in the data set, and reduction of the dimension through rearrangement of data.

The data set is analyzed next to identify patterns, i.e. models that represent relationships among data. The model is then validated with new data sets to ensure its versatility. It should be possible to translate the model into actionable business plans that are likely to help the organization achieve its goals. A model or pattern that satisfies these conditions becomes knowledge. The steps in the mining process are performed iteratively until meaningful business knowledge is extracted. Pattern discovery data mining algorithms based on machine learning techniques generate patterns that are easy to understand. They are, therefore, finding wide popularity in data mining applications. Any technique has its own strengths and weaknesses. Understanding these in the context of business data mining, the work is aimed at selecting a suitable technique for a specific application that serves the objective to build an information system.
1.4. INTELLIGENT INFORMATION SYSTEM

Computational Intelligence (CI) is the study of computational methods to automate the process of knowledge acquisition from examples [15] leading to the specified targets. This discipline which is a predecessor of the artificial intelligence field eliminates the laborious and expensive knowledge engineering process involved in developing knowledge-based systems. Computational intelligence discovers a knowledge pattern using the training data set and the pattern is used to classify and/or predict the behavior of applied examples.

For training examples with a known classification there are algorithms [16] available. The data search algorithm has to work on the required data set in the way that the design of the system aims, with intelligent and effective behavior. In such a context the architecture of the information system is designed to be a fusion system of computational utilities and is popularly termed as ‘A Hybrid Information System’.

The proposed work attempts at engineering an intelligent hybrid model of an information system by incorporating the analysis of data for knowledge representation applying the Adaptive Fuzzy Apriori
(AFA) – Tree Search algorithm on fuzzy transformation of data sets and the potential itemsets were identified. These potential itemsets were trained using a multi layer feed forward Artificial Neural Network (ANN) applying the rule based algorithm to make predictions and hence decisions.

The source data on which the AFA-Tree search algorithm is run is obtained by Rule Induction (RI) segregating the heterogeneous database on which fuzzy transformations were made creating a fuzzy database. The resultant database is structured like a spanning decision tree or a set of decision variables forming a 5 by 3 matrix, trying to scan all combinations of variables generally called the samples of the data set. Redundant combinations of similar attributes are eliminated. This is the stage where the real AFA-Tree search is applied to extract the potential combinations which are checked for a support threshold and this constitutes the ‘Knowledge Discovery Module’ of the model.

This process of attribute selection and splitting is continued until each terminal node is included in all possible clustering. Then on to the multi layer ANN the derived combinations were trained using the rule based algorithm and the new dataset is tested for results which constitutes the ‘Decision Making Module’ of the model. This hybrid
model is effectively run on real time data to predict ‘Least Risk in Investor Trading in Indian Stock Market’. For convenience of manipulation, the Indian perspective is taken and hence the money value is in terms of the ‘Rupee’.