

# Implementation of Data Mining Approach for Building Automated Decision Support Systems

<sup>1</sup>Farrukh Saleem

<sup>1</sup>Computer Science Department  
Faculty of Computer Science and  
Information Systems  
Universiti Teknologi Malaysia,  
Malaysia.

<sup>1</sup>[farrukh800@yahoo.com](mailto:farrukh800@yahoo.com)

<sup>2</sup>Abdullah Saad Al-Malaise

<sup>1&2</sup>Information System Department  
Faculty of Computing and  
Information Technology  
King Abdulaziz University, Jeddah  
Kingdom of Saudi Arabia.

<sup>2</sup>[abdmalaise@gmail.com](mailto:abdmalaise@gmail.com)

## Abstract

*Decision support systems (DSS) has remarkably increased in today's competitive business environment. Organizations are aggressively emphasizing on computerized support for building comprehensive automated DSS. The main purpose behind this approach is to build intelligent business and reduces the business pressure from the competitors. In this paper, we proposed a model for building efficient DSS with connection of data mining view. The presented model guide the business leaders to get extra support from data mining abstract to create effective DSS and compete the business world with more appropriate manner. Model represent the combination of DSS components and data mining tasks for the generation of better decisions, results, rules, and patterns from operational databases.*

### Key Words

*Decision Support Systems, Data Mining Tasks, DSS Model.*

## 1. Introduction

The existence of large databases is the major components for smooth running of every business and organization. These databases have more complex data for the implementation of any task because of its multi dimensional attributes. The different type of format and attributes of data create more difficulties includes text, images, videos, graphs etc.

This paper presented a model for reducing some pressures from the decision makers by building a DSS with the combination of data mining applications. In the methodology sections we defined it in detail. Before moving to the

model, presented small introduction of DSS and data mining approaches in the succeeding sections for more clear understanding of the model.

### 1.1 Decision Support System

Nowadays, the environment in which organizations are working is becoming more and more complex. Continuous business pressures, competitor's attacks, and knowledge of current affairs play significant role for building healthy business. Therefore, computerized support has become the major components for every organization to deal with the customers and competitors. In this environment the best decision can lead the organization on top.

In this era DSS is one of the main tool which can apply for all type business environmental factors such as; markets, customer demand, technology, and societal factors for generating best decision from the list of alternatives.

A DSS may have several phases to analyze and take the decision. As Turban described that a DSS must consider four phases for complete decision making process such as; Intelligent, Design, Choice and Implementation [2].

Organizations own DSS for presenting themselves proactive and anticipative. The development of DSS can be based on data or model. According to Turban, the major classification of DSS based on (i) Data Oriented DSS, (ii) Model Oriented DSS [2]. Although other scholars also categorize DSS into individual and group DSS, spreadsheet DSS, solver oriented DSS, etc.

Fernando, sketched the broad view of DSS components in figure-1 [3]. Although in this figure the part of DSS contains all the sub-parts of DSS as described above. DSS tools may build by using any programming language which is

specially designed for back end working like data processing and finding list of alternatives. This figure illustrates the large view of DSS, also considering figure-1 we proposed an enhanced version of this model in the methodology section.

In the subsequent section discussion of data mining approaches and working will clear more understanding of the purpose of data mining tasks inside DSS tool.

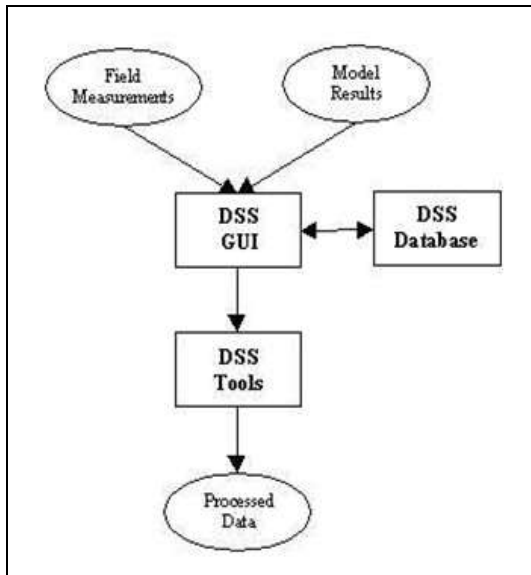


Figure-1 - Components of DSS [3]

## 1.2 Data Mining Approaches

Data Mining, also popularly known as Knowledge Discovery in Databases (KDD), refers to the nontrivial extraction of implicit, previously unknown and potentially useful information from data in databases. While data mining and knowledge discovery in databases (or KDD) are frequently treated as synonyms, data mining is actually part of the knowledge discovery process [5, 8]. Furthermore, Abdullah et. Al described the data mining in the sense of decision support systems(DSS) that, in decision support management terminology, data mining can be consider as a decision support process in which decision maker is searching to generate rule for the help in decision making [9, 7].

Mainly, data mining tasks has been divided into descriptive and predictive methods. Classification, clustering and rule association mining are most common techniques use for predictive and descriptive analysis [10].

Therefore, mainly scholars describe data mining in three major tasks. As Zaine [5] stated in his book chapter about major techniques of data mining as follows:

**Classification** – Classification analysis is the organization of data in given classes. Also known as supervised classification, the classification uses given class labels to order the objects in the data collection.

Classification consider as an important task of data mining. Using this approach data must be already defined a class label (target) attribute. Firstly we divide the classified data into two sets; training and testing data [11]. Where each datasets contains others attributes also but one of the attributed must be defined as class lable attribute. Jiawei Han [11] described classification task in two steps process; first is model construction and the second is model usage. The main target of this task is to build the model by using training dataset and then assign unseen records into a class by using the trained model as accurately as possible. While training data set is use to build the model on the other hand testing data set is use to validate the model [10].

**Clustering** – Similar to classification, clustering is the organization of data in classes. However, unlike classification, in clustering, class labels are unknown and it is up to the clustering algorithm to discover acceptable classes. Clustering is also called unsupervised classification.

Clustering is one of the major task has been applying for data mining, work on unsupervised data (no predefined classes) [12]. Clustering is a collection of data objects, clustered by taking similar object to one another within the same cluster, and dissimilar to the objects related in other clusters. Cluster differentiate by using similarities between data according to the characteristics found in the data and grouping similar data objects into clusters [11].

**Association** – Association analysis is the discovery of what are commonly called association rules. It studies the frequency of items occurring together in transactional databases, and based on a threshold called support, identifies the frequent item sets.

Data can be use to find association between several attributes, generate rules from data sets, this task is known as association rule mining

[12]. Given a set of transactions, find rules that will predict the occurrence of an item based on the occurrences of other items in the transaction. The goal of association rule mining is to find all rules having support  $\geq$  minsup (minimum support) threshold and confidence  $\geq$  minconf (minimum confidence) threshold [10].

Moreover, association rule mining can be viewed as a two-step process, first, find all frequent itemsets: items satisfying minimum support. Second, generate strong association rules from the frequent itemsets: these rules must satisfy minimum support and minimum confidence [11].

## 2. Methodology

The conceptual model of virtuous data mining cycle presented in figure-2, presented by [4]. Represent that data mining approach start with the identification of the problem followed by the data selection and data preprocessing. Finally any of the data mining task can work for finding rules/patterns and prediction of the future data. Figure-2 showed the basic steps involved in data mining [6]. Abdullah [1] described accordingly for this figure as a complete data mining process which can also consider a life cycle of data mining process may have several other iterative steps. According to four boxes in the figures representing the four basic steps, may includes some other sub-tasks in each of the steps. For example, in the second step data transformation consider comprehensive steps must include in the data mining cycle.

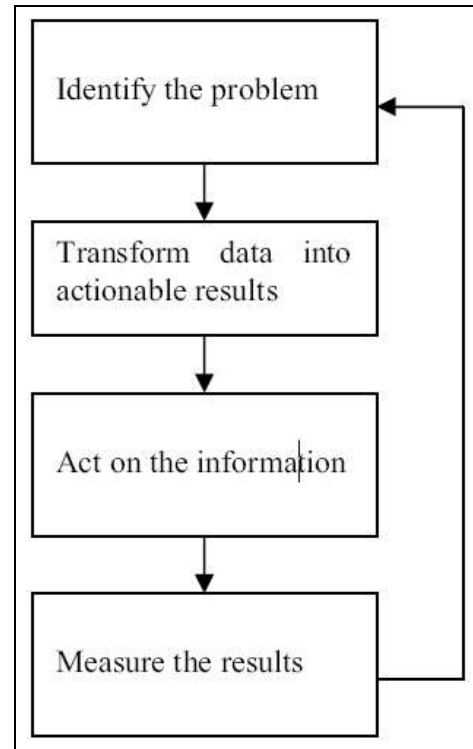


Figure-2 - Virtuous Cycle of Data Mining [4]

Reference to the figure-1; list of components include in a DSS and figure-2 basic data mining cycle with more than one iteration, in the next part of methodology we proposed an enhanced model with the combination of DSS and Data mining abstract. Implementation of this will guide the decision maker more concrete and reliable decision which is extracting by using large databases followed by all phases of DSS.

### 2.1 Explanation of the Model – Figure-3

The model presented in figure-3 is the major part of this paper. A new approach of building DSS by using data mining abstract for best result and analysis from the large databases. Where discovering of the new rules, patterns, models and finally choose the best decision is the main purpose of this model. Proposed model has two parts to be described in detail; DSS Abstract and Data Mining Abstract.

In DSS abstract we took the model sketched by [3] having several basic components of DSS. DSS graphical user interface (GUI) is the front end interface for the user to ask and find the required query. DSS tool is the major component in this model. We can build DSS tool for the particular problem by using selective DSS

technique such as analytical hierarchal process (AHP). DSS database has all the operational data of the organization. Moreover, results/models generate by the data mining abstract and DSS tool will also be saved in the DSS database for future queries.

DSS GUI always first looks at the local database and knowledge management database to find out the best solution. If not succeeded and does not find the better solution, then a new request will be generated and forwarded to the DSS tool. Now DSS tool is the common tool which can apply any of the DSS technique such as AHP for finding the best alternative from the possible more than one outcome. It always need to process a selected data either coming from the local database or may require to forward this query towards the data mining abstract. For both conditions the result/model will be saved in the DSS database and finally the result will show to the user against the asked query.

The enhanced part in the figure-3 is encapsulating the “process data” part with the data mining abstract. Data mining abstract will invoke when this step will need any extra information to be extracted from the selective data.

For the query if possible results are not available in the local DSS. Then, query will forward to the data mining abstract with particular data. Data mining module start with the selective data by applying data pre processing such as; data discrization and data transformation.

After data preprocessing, suitable data mining task such as; Association, Classification and Clustering, will use for the selected data. The technique will depend on the query and data. If the query is about prediction then classification and association technique can be use. If the query is particular about finding the association between different attribute then association mining is the best technique we can apply.

Finally, generated rules and pattern will be then forward into the knowledge management. And for current purpose or in future DSS GUI can access the knowledge management for further correspondence. The model showed the implementation of data mining abstract in the DSS tool. The importance of the model is to combine both phases of DSS and data mining. Implementation of this model using experienced data is the future part of this paper.

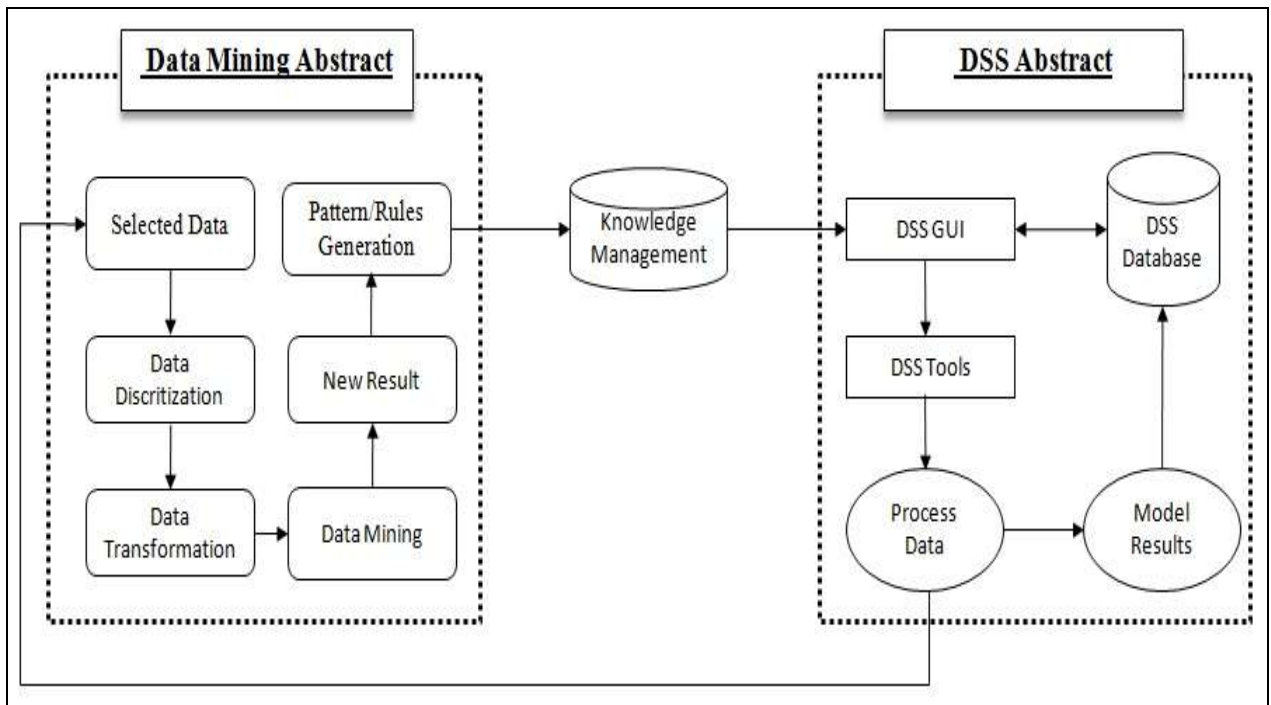


Figure-3 – Proposed Model of DSS using Data Mining Abstract

### 3. Future Work

The model is applicable in future by using any case study of the organization. Where different

data mining tasks can provide several outputs with appropriate DSS techniques.

In future we will validate this model by presenting some experienced data to show the complete process of choosing best decision with the help of data mining abstract. Selection of DSS and data mining techniques can lead towards the best possible results by using the same model.

### 4. Conclusion

In today's business environment the size of the database is increasing more and more. There is a need to analyze those data by using appropriate techniques. On the one side data mining is there to extract new patterns and rules from the vast amount of data. Other hand DSS is providing the facility to select the best alternative from the list of alternatives. In this paper we combined these two approached and proposed the model with the combination of data mining and DSS. Proper use and enhancement of the model will be highly motivating factor to improve this research field.

### References

- [1] Abdullah Saad ALMALAISE ALGHAMDI, "Efficient Implementation of FP Growth Algorithm Data Mining on Medical Data", International Journal of Computer Science and Network Security (IJCSNS)-2011.
- [2] Efraim Turban, Ramesh Sharda, Dursun Delen, Decision Support and Business Intelligence, 9<sup>th</sup> Edition, published by Pearson Education, Prentice Hall-2011.
- [3] Fernando Toro and Roberto Mayerle, A Decision Support System for Enhancing Model Development and Application, Esri International Conference-2001.
- [4] Hai Wang, Shouhong Wang, Medical Knowledge Acquisition through Data Mining, Proceedings of 2008 IEEE International Symposium on IT in Medicine and Education.

- [5] Osmar R. Zaiane, "Chapter I: Introduction to Data Mining", CMPUT690 Principles of Knowledge Discovery in Databases, 1999.
- [6] Abdullah Al- Mudimigh, Farrukh Saleem, Zahid Ullah, The Effects Of Data Mining In ERP-CRM Model – A Case Study Of MADAR, WSEAS Transaction, 2009
- [7] Dr. Abdullah Al- Mudimigh, Farrukh Saleem, Zahid Ullah, The Role Of Data Mining In ERP-CRM Model", International Conference on Applied Computer & Applied Computational Science (ACACOS '09), Hangzhou, China, 2009.
- [8] Data Mining for Customer Queries in ERP Model- A Case Study of MADAR", IEEE / ICICT-2009, Third International Conference on Information & Communication Technologies, Karachi, Pakistan, 2009.
- [9] Dr. Abdullah Al- Mudimigh, Farrukh Saleem, Zahid Ullah, Efficient Implementation of Data Mining: Improve Customer's Behavior", The 7<sup>th</sup> IEEE/ACS, International Conference on Computer Systems and Applications, Rabat, Morocco, to be held on May, 10-13, 2009.
- [10] Pang-Ning Tan, Michael Steinbach & Vipin Kumar, "Introduction to Data Mining", Addison Wesley, 2005, ISBN 0321321367
- [11] Jiawei Han, Micheline Kamber, Jian Pei, "Data Mining: Concepts and Techniques", 2nd edition, 2005, Morgan Kaufmann, ISBN 1558609016
- [12] Farrukh Saleem, Areej Malibari, DATA MINING COURSE IN INFORMATION SYSTEM DEPARTMENT–CASE STUDY OF KING ABDULAZIZ UNIVERSITY, IEEE, ICEED- 3<sup>rd</sup> International Congress on Engineering Education, 7<sup>th</sup> – 8<sup>th</sup> December, 2011.