Decision Support System using Data Mining

Prof. P. M. Daflapurkar
University of Pune
MMIT,Lohegaon Pune

Himalay S Joshi
University of Pune
MMIT,Lohegaon Pune

Shraddha S Yashwante
University of Pune
MMIT,Lohegaon Pune

Vikas V Chavan
University of Pune
MMIT,Lohegaon Pune

Abstract - CRM Systems are developed for the business organizations to manage their customer's data effectively and to improve their profit which helps to increase the organization's business. It will maintain customer's detail and their data. Business organizations can generate reports in various formats and these reports help business organizations for various types of analysis. They can use these reports in decision making processes and also for customer satisfaction by analyzing their needs and problems. The proposed Decision Support System, is an application which takes the database, processes it with data mining methods
and gives the output reports in the form of graphs, pie-charts, pdf , CSV formats and excel file. These reports are then read by the CRM systems to undergo their functions easily. The system is based on predictive analysis of the customers on the basis of the data. The system will be able to predict the potential customers for the organisation, and thus it will increase its productivity by saving their time, money and energy. Thus, this system will be a effective tool for decision making process of the organisation.

Keywords: Decision support system (DSS), Data Mining, predictive analysis CRM.

1. INTRODUCTION

CRM Systems are developed for the business organizations to manage their customer's data effectively and to improve their profit which helps to increase the organization's business. It will not just maintain customer's detail and their data but it also provides some features that helps business organizations to increase communication with their customers using some facilities of the system such as Email and SMS facilities. It also provides the customization facility to user so that user can customize this system as per his/her organizational needs. Business organizations can manage their inventory management and support tasks through this system. They can also manage their sales orders, invoices,

quotes, assets, complaints and much more through this system. Business organizations can also generate reports in various formats and these reports help business organizations for various types of analysis. They can use these reports in decision making processes and also for customer satisfaction by analyzing their needs and problems. So, CRM systems plays a very important role for business organizations. A Decision Support System, is a application which takes the database, processes it with data mining methods and gives the output reports in the form of graphs, pie-charts, CSV formats and excel file. This reports and then read by the CRM systems to undergo their functions easily. The system is based on predictive analysis of the customers on the basis of the data. The system will be able to predict the potential customers for the organisation. So, every CRM systems needs a DSS which Processes the data and gives out the reports.

2. LITERATURE SURVEY

In many organizations the customer's data is managed manually so it is very tedious and crucial task to manage huge data of customers manually. Some organizations use the CRM systems to manage their data but it is a e-CRM and it uses internet and centralized data so availability of data is 24X7. And it also provides more functionalities than CRM. Some
organizations use CRM but it's not provide the facilities of Customization. In the existing
system, Huge Data is handled for system and business evolution without prior processing.
Analyzing and prediction of customer behavior is randomly done. Descriptive analyses
of the information leads to inaccurate results are obtained. Pattern extraction was never known.
Problems faced in the existing system :

1.Getting and keeping satisfied customers.
2. There is ever increasing competition and businesses are finding things difficult. There is rise in demanding and knowledgeable customers and a host of new competitors flooding the market.

3. Challenges to identify the market segment they want to target. No proper means for sustainable follow-up efforts to build strong pipeline and convert the footfalls into customers.

4. Keeping important information up-to-date and make it accessible to employees so that they could provide better services to customers.

5. Inability to create a marketing calendar integrated with financials, providing a centralized view to manage and schedule all relevant enterprise marketing plans and campaigns.

6. Unable to scale segmentation for targeting new prospects.

7. Unstructured sales processes and de-centralized prospect communication.

To overcome all these problems we use some data mining methodologies. We are following systematic implementation of SDLC (system development life cycle) approach for our product development. During our System Analysis and Study phase data mining models like CHAID and Neural models were studied . We can also use other data mining models like classification, clustering and regression.

3. PROPOSED WORK

The CRM application is used to give the reports, which can be easily read by the managers of the company. These applications support the decision making process of the manager in a most accurate way.CRM systems needs the pre-processed data for the generation of usefull reports. Decision support system provides the best input to the CRM system, to manage all the data about customers, campaigns, leads etc. in an efficient way. In this system by using Classification, Clustering and Regression module, the database can be analyzed and managed is the best possible ways. Whereas by using the Report module, The detailed reports in form of graphs, pie-charts, CVS format and Excel format can be generated.

Scope of the Project
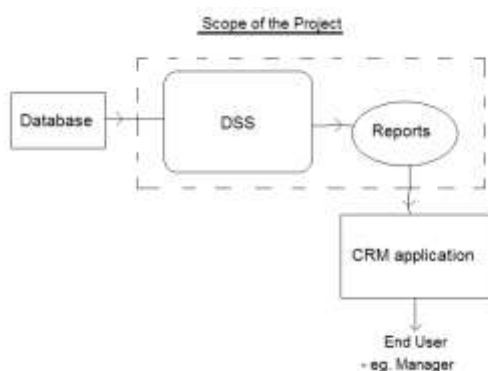
The system Consists of 4 modules :
- Classification.
- Clustering.
- Regression.
- Reports.

The figure shows the overall structure of the CRM system and our area of the project. The database sample will be provided by the company. The data samples are tested. This is done to ensure the authentisity of the data sample. These data samples are then passed to our DSS application. Here the data sample is thoroughly analyzed under the above modules. The reports are then generated. The reports in CSV format and excel sheet will be imported by the CRM application. Reports in DOC and pdf _le are for the users to analyse data manually if required. It provides the fully integrated decision support system for CRM systems. The reports

generated by this system can be used as input to various CRM systems. Thus the CRM system gets the preprocessed database, which supports CRM applications to generate re-ports, and take decisions about the organizations. The output of our system is read by the CRM systems, and the output of CRM systems is read by the Managers of the Organizations. Advantages over existing system :

- Data mining Approach is used.
- Data mining is a technology used in different disciplines to search the significant relationships among variables or components in large data sets.
- Data mining enables organizations to use their current reporting capabilities to un-cover and understand hidden patterns in vast databases.
- The use of pattern recognition logic is to id entity trends within a sample data set and extrapolate this information against the larger data pool. (I.e. a sample of few records is considered for analyzing purpose).
- In particular, Naive Bayesian classifier is used for detailed analysis.
- Classification, clustering and regression are the Data Mining methods which helps in giving accurate results.
- The reports will be genarated in graphs, pie diagrams, CSV format and hardcore excellent. The reports goes to CRM and then to the user of the application. eg. Manager of the organisation.

4. MATHEMATICAL MODEL

RS = d,c,r,s,r
where,
RS = Result Set
d = Data sample as an input
c = Classification module
r = Regression module
s = Segmentation module
r = Reporting module

A.1.1 Classification Module
Naive Bayes Theorem
$p(H|X) = p(X|H) \times p(H)/p(X)$
$p(response|custid) = p(response|interest) \times p(response|interest) \times$
$p(response|houshold - size) \times p(response|cust - income) \times$
$p(response|creditlimit) \times (response|age - of - head) \times$
$p(response|cust - lastinvestment) \times p(response)/p(custid)$
(A.1)
22
A.1.2 Segmentation module
K-means algorithm

$$J= \sum_{j=1}^{k} \sum_{i=1}^{n} \left| x_i^j - c_j \right|^2$$

5. CONCLUSION

Our project is restricted to the DSS and the refine data generated in form of reports. The database contents will be provided by the company. The DSS application will process the data using different data mining methods. The reports are generated in form of CSV format, pie charts graphs etc. These reports are thus given to the CRM system and finally viewed by the end user i.e. a senior manager. This in turn helps the senior manager to take better decision for profit maximization.

REFERENCES

1. Norjansalika Janom Nor Hapiza Mohd Ariffin, Abd Razak Hamdan Khairuddin Omar. Customer relationship management (crm) implementation: A soft issue in knowledge management scenario. *IEEE*, 2012.
2. Wu Danhong Shao Jihong. Views on crm under e-business environment. *IEEE*, 2011.
3. Wang Zhen-hua Shao Yi-xing, Wang Wei. Crm measurement indicator system of logistics enterprises for data-mining. *IEEE*, 2010.
4. Robert Ian Whitfield Iain M. Boyle Shaofeng Liu, Alex H. B. Duffy. Integration of decision support systems to improve decision support performance. *Springer*, 2009.
5. Nong Ye Senior Member IEEE Xiangyang Li, Member IEEE. A supervised clustering and classification algorithm for mining data withmixed variables. *IEEE*, 36(2), 2006.
6. Hung-Ju Huang and IEEE Chun-Nan Hsu, Member. Bayesian classification for data from the same unknown class. *IEEE*, 32(2), April 2002.
7. IEEE Jerry B. Weinberg Gautam Biswas, Senior Member and Douglas H. Fisher. Iterate: A conceptual clustering algorithm for data mining. *IEEE*, 28(2), 1998.
8. Jian Pei Jiawei Han, Micheline Kamber. *Data Mining Concepts and Techniques*. Morgan Kaufmann Publishers, second edition edition, 2012.
9. M. Turk and A. Pentland. Face recognition using eigenfaces. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, volume 591, 1991.
10. Pedro Domingos Daniel Lowd. Naive bayes models for probability estimation.
11. Nello Cristianini Colin Campbell. Simple learning algorithms for training support vector machines. 20
12. Harry Zhang. The optimality of naive bayes.
13. I. Rish. An empirical study of the naive bayes classifier.