International Journal of Engineering Science Invention (IJESI)
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org || PP. 62-66

Data Mining for Customer Service Support

R.DEEPA[1], A.LAVANYA[2]
[1](Research scholar,Dept.of computer science,kamban college of arts and science for women, tiruvannamalai)
[2](Head of the department, computer science, kamban college of arts and science for women, tiruvannamalai)

***Abstract**: In traditional customer service support of a manufacturing environment, a customer service database usually stores two types of service information: (1) unstructured customer service reports record machine problems and its remedial actions and (2) structured data on sales, employees, and customers for day-to-day management operations. This paper investigates how toapply data mining techniques to extract knowledge from the database to support two kinds of customer service activities: decision support and machine fault diagnosis. A data mining process, based on the data mining tool DBMiner, was investigated to provide structured management data for decision support. In addition, a data mining technique that integrates neural network, case-based reasoning, and rule-based reasoning is proposed; it would search the unstructured customer service records for machine fault diagnosis. The proposed technique has been implemented to support intelligent fault diagnosis over the World Wide Web.*

***Keywords:** Structured data, unstructured data, decision support, machine fault diagnosis, neural network, case-based reasoning, and rule-based reasoning.*

I. Introduction

Customer service support is becoming an integral part of most multinational manufacturing companies that manufacture and market expensive machines and electronic equipment. Many companies have a customer service department that provides installation, inspection, and maintenance support for their worldwide customers. Although most of these have some engineers to handle day-to-day maintenance and small-scale troubleshooting, expert advice are often required from the manufacturing companies for more complex maintenance and repair jobs. Prompt response to a request is needed to maintain customer satisfaction. Therefore, a hot-line service centre (or help desk) is usually set up to answer frequently encountered problems from the customers.

Fig.1 shows the work low in a traditional hot-line service centre. The service centre is responsible for receiving reports on faulty machines or enquiries from customers via telephone calls. When a problem is reported, a service engineer will suggest a series of checkpoints for customers using the hot-line advisory system. Such suggestions are based on past experience. This has been extracted from a Customer Service Database, which contains previous service records that are identical or similar to the current problem. The customer can then try to solve the problem and subsequently confirm, with the Service centre,

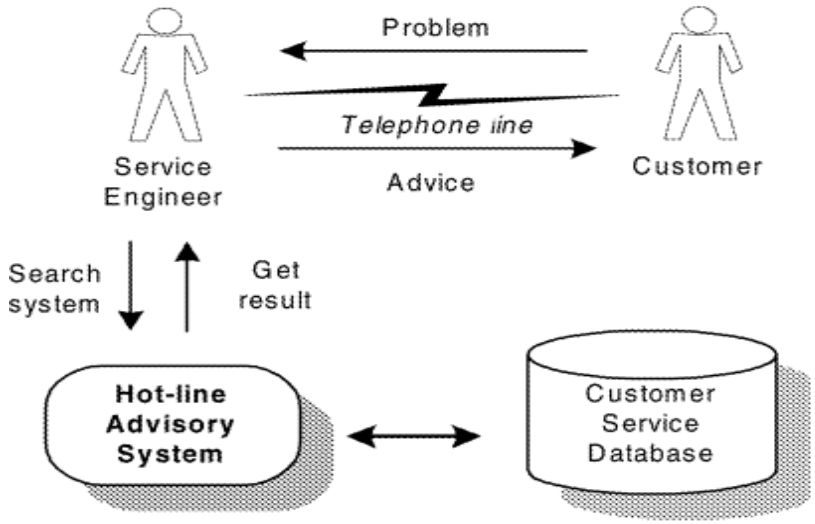

Fig. 1. Traditional hot-line service centre.

if the problem is resolved. If the problem still persists, the centre will dispatch a service engineer to the customer's premise for an on-site repair. During such trips, the service engineer will take past records of the customer's machine, related manuals, and spare parts that may be required to carry out the repair. Such a process is inconvenient.

At the end of each service cycle, a customer service report is used to record the new problem and the proposed remedies or suggestions taken to rectify it. This database is used for billing purposes, as well as for maintaining a corporate knowledgebase. The service centre stores the customer service report in the database. Apart from maintaining a knowledge base on common faults and its remedies, the customer service database also stores data on sales, employees, customers and service reports. These data are not only used for day-to-day management operations, but help the company in decision making on job assignment and promotion of service engineers, and marketing, manufacturing, and maintenance of different machine models. The customer service database serves as a repository.

II. Data Mining

Data mining, also known as knowledge discovery in databases (KDD), is a rapidly emerging field. This technology is motivated by the need of new techniques to help analyze, understand or even visualize the huge amounts of stored data gathered from business and scientic applications. It is the process of discovering interesting knowledge, such as patterns, associations, changes, anomalies and significant structures from large amounts of data stored in databases, data warehouses, or other information repositories. It can be used to help companies to make better decision to stay competitive in the marketplace. The major data mining functions that are developed in commercial and research communities include summarization, association, classification, prediction and clustering. These functions can be implemented using a variety of technologies, such as database-oriented techniques, machine learning and statistical techniques.

Recently, a number of data mining applications and prototypes have been developed for a variety of domains including marketing, banking, finance, manufacturing and health care. In addition, data mining has also been applied to other types of data such as time-series, spatial, telecommunications, web, and multimedia data. In general, the data mining process, and the data mining technique and function to be applied depend very much on the application domain and the nature of the data available.

III. Customer Service Support

Service records (or reports) are currently dened and stored in the customer service database. Each service record consists of customer account information and service details, which contain two types of information: fault-condition and checkpoint information. The former contains the service engineer's description of the machine fault, while the later indicates the suggested actions or services to be carried out to repair the machine, based on the actual fault-condition given by the customer. Checkpoint information contains checkpoint group name, and checkpoint description, with priority and an optional help. The checkpoint group name is used to specify a list of group checkpoints. Each checkpoint is associated with a priority that determines the sequence in which it can be exercised and a help that gives visual details on how to carry out the checkpoint.

3.1. Mining structured data

A list of most popular data mining tools available commercially or in public domain is given in the KDN uggets website. These tools can mine the structured data of sales, maintenance, and particulars of employees and customers in the customer service database. It is interesting to see that a number of tools support multiple approaches; i.e., more than one data mining techniques. For example, Darwin from Thinking Machine Corp. supports neural networks, regression tree (CART), k-means algorithm, and case based reasoning for classification, prediction, and clustering functions. There are also some tools that only aim at a specific data mining function. This provides edibility, the users can select different data mining tools for their problem domains to achieve the best results.

3.2. Mining unstructured data

Although DBMiner is an excellent data mining tool for large databases with structured data, it is unsuitable for extracting knowledge from the textual data of the customer service database. As the information or knowledge on common faults and their suggested remedies are stored in textual format as fault-condi-tions and checkpoints, new techniques are needed to extract knowledge from this database for machine fault diagnosis. This is known as text mining.

IV. Data Mining For Decision Support

Information, such as the best selling machines, the customers of a particular machine, a comparison of sales among different machines, and the performance of different service engineers are highly desirable for the management team.

4.1. Data mining process

Fig. 2 depicts the data mining process for extracting hidden knowledge from large databases. The process focuses on finding interesting patterns that can be interpreted as useful knowledge. It consists of seven steps.

4.1.1. Establishing the mining goals

This involves the understanding of the customer service support process, its database, and the administrative procedures of the company. A number of mining goals were identified:

- Marketing
- Customer support
- Resource management

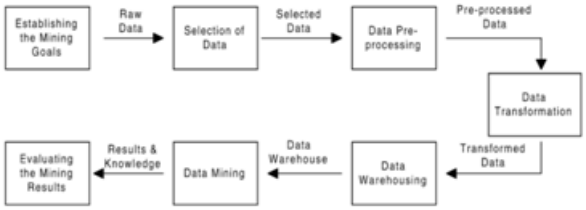

Fig. 2. Data mining process for decision support

4.1.2. Selection of data

This step identifies a subset of variables or data samples, on which mining can be performed. There are many tables in the database. However, not all are suitable for mining, since they are not sufficiently large. After an initial study, the structured data tables EMPLOYEE and CUSTOMER were found unsuitable for mining, while MACHINE and SERVICE_REPORT were considered suitable for mining.

4.1.3. Data pre-processing

This step removes the noisy, erroneous, and incomplete data. The presence of too many different categories of categorical data makes visualization of the displayed information very difficult. Hence, those categories with only a few records are eliminated. Moreover, all the records with missing values are deleted to avoid problems in visualization. cross, grill, spot, star, square and circle. Additionally shapes of palm and fingers also play very important role in such decision making for identification of diseases. Since the proportion of such records is quite small, their deletion will have little effect on the results.

4.1.4. Data transformation

The data stored in the various tables are in a specified format (defined during the construction of the database). Sometimes, it is useful to transform the data into a new format in order to mine additional information. For example, a new column `svc_repair_time' (service repair time) is created by calculating the difference, measured in number of days, between `svc_start_dt' and `svc_end_dt' in the SERVICE_REPORT table. This new attribute is useful in analyzing the performance of the service engineers.

4.1.5. Data warehousing

Data warehousing is the process of visioning, planning, building, using, managing, maintaining and enhancing databases. The data suitable for mining are collected from the various tables of the customer service database and stored in DBMiner's data warehouse. OLAP data marts are then generated from the data warehouse, which contains customized data at a higher level of summarization. Data cubes can be constructed from data marts to provide multi-dimensional views of the data. On-line analytical mining can then be performed using the multi-dimensional data cube structure for knowledge discovery.

4.1.6. Data mining

DBMiner is used to perform the data mining functions, including summarization, association, classifi- cation, prediction and clustering.

4.1.7. Evaluating the mining results

Different data mining functions have been exercised, providing data. The information obtained is next analyzed. The results are:

- Marketing
- Customer support
- Resource management

V. Data Mining For Machine Fault Diagnosis

The unstructured textual data of fault-condition and checkpoint information of the customer service database provides useful machine service information. A data mining technique based on the integration of neural network, case-based reasoning, and rule-based reasoning has been applied to the customer service database to support intelligent machine fault diagnosis.

5.1. Data mining process

Fig. 3 shows the framework of the data mining process. It consists of two major processes: the off-line knowledge extraction process and the on-line fault diagnosis process. The first extracts knowledge from the customer service database to form a knowledge base that contains the neural network models and a rule-base. The neural network models and the rulebase work within the CBR cycle to support the second, which uses the four stages of CBR cycle (retrieve, reuse, revise, and retain) to diagnose customer reported problems. It accepts user's problem description as input, maps the description into the closest fault-conditions of the faults previously stored from the knowledge base, and retrieves the corresponding checkpoint solutions for the user. The user's feedback on the fault diagnosis process is used to revise the problem and its solution. The new result is ultimately retained as knowledge for enhancing performance of future problems.

5.2. Knowledge extraction process

5.3. Fault diagnosis process

5.3.1. Pre-processing of user input

5.3.2. Neural network retrieval

5.3.3. Reuse of service records

5.3.4. Revise and retain with user feedback

5.4. Performance evaluation

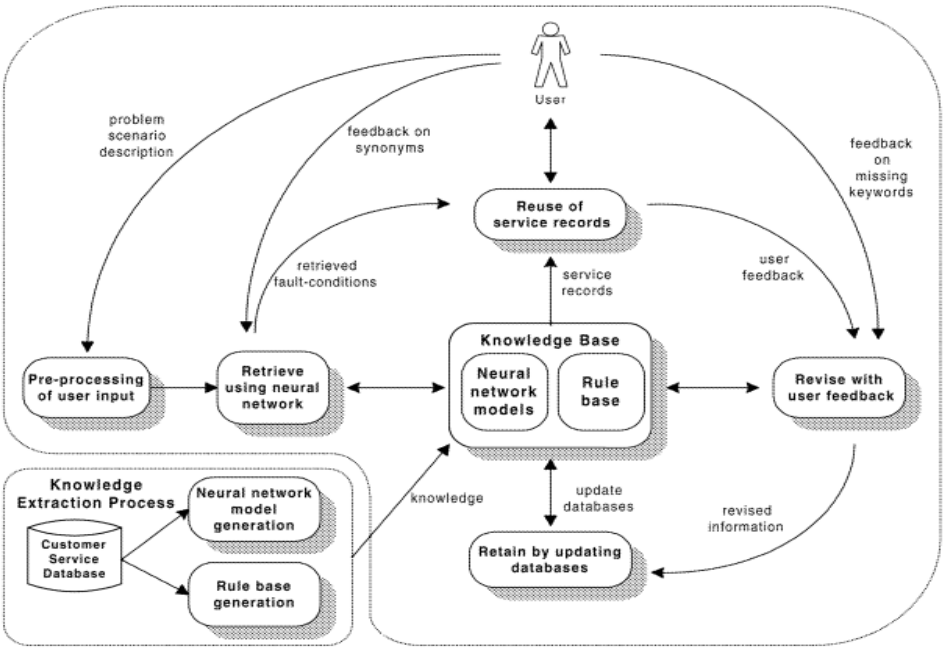

Fig. 3. Data mining process for machine fault diagnosis

VI. Conclusions

The proposed technique has been implemented to support intelligent fault diagnosis over the WWW.

As a collaborative research project with a multinational company, this research investigated the application of data mining techniques to extract knowledge from the customer service database for two kinds of

customer service activities: decision support and machine fault diagnosis. The information stored in the customer service database are classified as structured and unstructured textual data. The structured data are mined to enhance the decision making process for better management of resources and marketing of products. The unstructured data are mined to support intelligent diagnosis of machine faults over the World Wide Web.

In order to mine the structured data in the customer service database, a data mining process based on the data mining tool, DBMiner was proposed. To support machine fault diagnosis, a data mining technique based on the integration of neural network, case-based reasoning, and rule-based reasoning is incorporated. This data mining technique can operate within a system to provide efficient on-line machine fault diagnosis over the World Wide Web.

References

[1]. A.Aamodt,E.Plaza,Case-basedreasoning:foundational issues, methodological variations, and system approaches, S.C. Hui, G. Jha / Information & Management 38 Proceedings of AICom Artificial Intelligence Communications 7 (1), 1994.

[2]. H. Ahonen, O. Heinonen, M. Klemettinen, A.I. Verkamo, Applying data mining techniques for descriptive phrase extraction in digital document collections, IEEE Forum on Research and Technology Advances in Digital Libraries (ADL'98), 1998, pp. 2±11.

[3]. K. Balakrishnan, V. Honavar, Intelligent diagnosis systems, Journal of Intelligent Systems, 8(3)(1998), alsoavailableatURL:http://www.cs.iastate.edu/honavar/publist.html>.

[4]. R.J. Brachman, T. Khabaza, W. Kloesgen, G. PiatetskyShapiro, E. Simoudis, Mining business databases, Communication of the ACM 39 (11), 1996, pp. 42±48

[5]. L. Breiman, J. Friedman, R. Olshen, C. Stone, Classification of Regression Trees, Wadsworth, 1984.

[6]. S. Chaudhuri, U. Dayal, An overview of data warehousing and olaptechnology, ACM SIGMOD Record 26 (1), 1997, pp. 65±75.

[7]. M.S. Chen, J. Han, P.S. Yu, Data mining: an overview from a database perspective, IEEE Transactions on Knowledge and Data Engineering 8 (6), 1996, pp. 866±883.

[8]. R. Eustace, G. Merrington, A probabilistic neural network approach to jet engine fault diagnosis, in: Proceedings of the 8th International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems (IEA/AIE'95), Melbourne, Australia, ACM, 6±8 June 1995, pp. 67±76.

[9]. Fayyad, G. Piatetsky-Shapiro, P. Smyth, R. Uthurusamy (Eds.), Advances in Knowledge Discovery and Data Mining. AAAI Press, Cambridge, MA, 1996, pp. 1±

[10]. U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, The KDD process for extracting useful knowledge from volumes of data, Communications of the ACM 39 (11), 1996, pp. 27±34.

[11]. C. Fellbaum (Ed.), Wordnet: An Electronic Lexical Database, The MIT Press, 1998.

[12]. J. Han et al., DBMiner: a system for data mining in relational databases and data warehouses, in: Proceedings of CASCON '97: Meeting of Minds, Toronto, Canada, 1997.

[13]. J. Han, Towards on-line analytical mining in large databases, SIGMOD Record 27 (1), 1998, pp. 97±107.

[14]. Inference Corporation, CBR Content Navigator, Online document available at , 1999. [15] G. Jha, Data mining for customer service support, M.A.Sc Thesis. School of Applied Science, Nanyang Technological University, Singapore, 1999.