

# Analysis of an Enormous Telemetry Database of High Complexity

Fugjher Nderw\*

Department of Intelligent Mechatronics Engineering, Sejong University, Seoul, Uk

\*Corresponding author: Fugjher Nderw, Department of Intelligent Mechatronics Engineering, Sejong University, Seoul, Uk, E-mail: fugjhernder@gmail.com

**Received date:** March 06, 2023, Manuscript No. IJIRCE-23-16841; **Editor assigned date:** March 09, 2023, PreQC No. IJIRCE-23-16841 (PQ); **Reviewed date:** March 18, 2023, QC No. IJIRCE-23-16841; **Revised date:** March 29, 2023, Manuscript No. IJIRCE-23-16841 (R); **Published date:** April 06, 2023, DOI: 10.36648/ijirce.8.02.116

**Citation:** Nderw F (2023) Analysis of an Enormous Telemetry Database of High Complexity. Int J Inn Res Compu Commun Eng Vol.8 No.02:116

## Description

This study focuses on the development of automation platform for performing data mining on a telemetry database for computer systems. It is common for computer systems to encounter failures in an unexpected manner. It is therefore valuable to have prognostics capability for computer systems to minimize the effects of unexpected system failure. Data acquisition schemes employing telemetry techniques are considered the most effective method for collection of in-service information for computer systems. Analysis of an enormous telemetry database of high complexity must be completed before useful knowledge can be extracted. In this research, an automatic data mining platform is reported for the extraction of useful knowledge from the telemetry database. This paper describes the structure and basic theories underlying the data mining of the telemetry database. Also, an automatic computer program capable of performing database management, filtering, data analysis, and reporting is described. Some useful data generated by the platform are reported for the telemetry database. In a data mining project developed on a relational database, a significant effort is required to build a data set for analysis.

## Database Transformations

The main reason is that, in general, the database has a collection of normalized tables that must be joined, aggregated and transformed in order to build the required data set. Such scenario results in many complex SQL queries that are written independently from each other, in a disorganized manner. Therefore, the database grows with many tables and views that are not present as entities in the ER model and similar SQL queries are written multiple times, creating problems in database evolution and software maintenance. In this paper, we classify potential database transformations, we extend an ER diagram with entities capturing database transformations and we introduce an algorithm which automates the creation of such extended ER model. We present a case study with a public database illustrating database transformations to build a data set to compute a typical data mining model. Mining Maximal Frequent Patterns (MFPs) in Transactional Databases (TDBs) and Dynamic Data Streams (DDSSs) is substantially important for business intelligence. MFPs, as the smallest set of patterns, help

to reveal customers' purchase rules and Market Basket Analysis (MBA). Although, numerous studies have been carried out in this area, most of them extend the main-memory based Apriori or FP-growth algorithms. Therefore, these approaches are not only unscalable but also lack parallelism. Consequently, ever increasing big data sources requirements cannot be met. In addition, mining performance in some existing approaches degrade drastically due to the presence of null transactions. We, therefore, proposed an efficient way to mining MFPs with Apache Spark to overcome these issues. For the faster computation and efficient utilization of memory, we utilized a prime number based data transformation technique, in which values of individual transaction have been preserved. After removing null transactions and infrequent items, the resulting transformed dataset becomes denser compared to the original distributions. We tested our proposed algorithms in both real static TDBs and DDSSs.

Experimental results and performance analysis show that our approach is efficient and scalable to large dataset sizes. After an enterprise builds a data warehouse, it can record information related to customer interactions using structured and unstructured data. The intention is to convert these data into useful information for decision-making to ensure business continuity. Hence, this study proposes a new Chinese text classification model for the Project Management Office (PMO) using fuzzy semantics and text mining techniques. First, content analysis is performed on the unstructured data to convert important textual information and compile it into a keyword index. Next, a classification and decision algorithm for grey situations and fuzzy (GFuzzy) is used to categorize textual data by three characteristics: maximum impact, moderate impact, and minimum impact. The purpose is to analyze consumer behaviors for the accurate classification of customers. Lastly, a more effective marketing strategy is formulated to target the various customer combinations, growth models, and the best mode of service. A company database of interactions with customers is used to construct a text mining model and to analyze the decision process of its PMO. The purpose is to test the feasibility and validity of the proposed model so that enterprises are provided with better marketing strategies and PMO processes aimed at their customers. Many different approaches of data mining have been proposed to satisfy various demands of users. Erasable pattern mining is one of the interesting areas in frequent pattern mining, which was

proposed to diagnose and solve financial problems caused in industrial fields. Since its original concept emerged, various relevant approaches have been devised.

## Data Mining

Analyzing incremental data becomes more important because interesting data are continually accumulated in various application fields including industrial areas. For this reason, an incremental method for erasable pattern mining has also been suggested in order to reflect such a trend. Since incremental data become gradually larger and more complicated with the passage of time, it is important to process such data as quickly and efficiently as possible. However, the previous method has limitations in this respect. Motivated by this challenge, we propose a new incremental erasable pattern mining algorithm including new data structures and mining techniques for efficient incremental data processing. We also demonstrate that the proposed method outperforms previous state-of-the-art approaches through extensive, empirical performance tests. The purpose of this paper is to develop a customer-oriented organisational diagnostic model, 'PARA' model, based on data mining of customer-complaint databases. The proposed 'PARA' model, which is designed to diagnose and correct service failures, takes its name from the initial letters of the four analytical stages of the model: (i) 'primary diagnosis'; (ii) 'advanced diagnosis'; (iii) 'review'; and (iv) 'action'. In the primary-diagnosis stage, the customer-complaint database is comprehensively analysed to identify themes and categories of complaints. In the advanced-diagnosis stage, a data-mining technique is employed to investigate the relationship between the categories of customer complaints and the deficiencies of

the service system. In the review stage, the identified weaknesses of the service system are reviewed and awareness of these weaknesses is enhanced among the organisation's employees.

In the action stage, a strategy of action plans for improvement is developed. An empirical case study is conducted to demonstrate the practical efficacy of the 'PARA' model. The paper concludes by summarising the advantages of the proposed model and the implications for future research. The utilisation of data mining methods has become common in many fields. In occupational accident analysis, however, these methods are still rarely exploited. This study applies methods of data mining (decision tree and association rules) to the Finnish national occupational accidents and diseases statistics database to analyse factors related to Slipping, Stumbling, and Falling (SSF) accidents at work from 2006 to 2007. SSF accidents at work constitute a large proportion (22%) of all accidents at work in Finland. In addition, they are more likely to result in longer periods of incapacity for work than other workplace accidents. The most important factor influencing whether or not an accident at work is related to SSF is the specific physical activity of movement. In addition, the risk of SSF accidents at work seems to depend on the occupation and the age of the worker. The results were in line with previous research. Hence the application of data mining methods was considered successful. The results did not reveal anything unexpected though. Nevertheless, because of the capability to illustrate a large dataset and relationships between variables easily, data mining methods were seen as a useful supplementary method in analysing occupational accident data.