



IJIRCCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 7, July 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.542



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

High Utility Itemset Mining over Nutrition Data

Palli Jagadeeswara Rao¹, Nagalla Prasanth kumar², ³Muppalla Akhil Krishna,

⁴Namburu Naveen, C.M.Suvarna Varma⁵

U G Students, Dept. of I.T., Vasireddy Venkatadri Institute of Technology, Guntur, Andhra Pradesh, India ^{1,2,3,4}

Assistant Professor, Dept. of I.T., Vasireddy Venkatadri Institute of Technology, Guntur, Andhra Pradesh, India⁵

ABSTRACT: Nutrition is important to everyone starting from a diseased patient, children or even a healthy person to maintain survival, growth and lead a longevity life. Research towards this perspective is very much required and our project aims at this application using data mining techniques like high utility itemset mining to find the relations and effects of malnutrition by extracting the most useful itemsets over dynamic distributed databases using cloud era. Thereby, giving patterns that can actually provide proper nutrition to every kind of personality. Unlike dieticians, who plan a standard chart for all types of bodies, our project helps in giving the useful combinations of nutrition to every individual based on their health history.

KEYWORDS: HUIM(High utility itemset mining)

I. INTRODUCTION

The project entitled as “High Utility Itemset Mining over Nutrition Data” is developed using ASP.NET as front end, C# as coding language and SQL Server as back end. JavaScript used for validation purposes. The main objective of this project is to mine the frequent item set and maximum threshold signature of the Health life food suggestion. The frequent item set deals with the whole database of the meal wise food and predicts the health condition. It contains various transactions like user data, BMI data, food data and calorie details etc. Here enhanced EHAUPM (Enhanced Efficient High Average-Utility Pattern Mining) – Item set has been implemented which gives more accuracy and performance than TKU (mining Top-K Utility item sets) and TKO (mining Top-K utility item sets in One phase), which are implemented in the existing methods for mining such item sets without the consideration of entire database. This may cause inaccurate results and improper output. These methods may use an assumption purpose only.

II. PRIMARY OBJECTIVE

The main objective of this project is to mine the frequent item set and maximum threshold signature of the Health suggestion. The suggestion will be generated through the food consumed and calorie information. Here EHAUPM - Enhanced Efficient High Average-Utility Pattern Mining for efficient data retrieval. A new concept called data engineering is used in the system to find associations with the calories used and generate some health tips.

III. EXISTING SYSTEM

- ☒ The problem of frequent itemset mining is prominent. In any case, it has some critical constraints with regards to investigating client exchanges.
- ☒ A vital constraint is that buy amounts are not considered. In this manner, a thing may just show up once or zero time in an exchange. In this manner, if a client has purchased five breads, ten breads or twenty breads, it is seen as the same.
- ☒ A moment imperative restriction is that all things are seen as having a similar significance, utility of weight.

Disadvantages of existing system:

- Frequent pattern mining in K mine is the patterns which appear frequently in database and it takes more time for execution.
- For example a set of items, such as milk and bread, that appear frequently together in a transaction data set is a frequent itemset. A subsequence, such as buying first a PC, then a digital camera, and then a memory card, if it

occurs frequently in a shopping history database, is a (frequent) sequential pattern. If Pattern occurs frequently, it is called a frequent pattern. Finding such frequent patterns is more difficult in the existing system

- Frequent pattern mining has become an existing data mining task and it only focused theme in data mining research so there is no wide range of analysis is available. A typical example of frequent itemset mining is market basket analysis.
- This process analyzes customer buying habits by finding associations between the different items that customers place in their “shopping baskets”.

IV. PROPOSED SYSTEM

K – Mining is introducing in Data mining as an important part of knowledge discovery process that we can analyze an enormous set of data and get hidden and useful knowledge. Data mining is applied effectively not only in the business environment but also in other fields such as weather forecast, medicine, transportation, healthcare, insurance, government... etc.

Here the data engineering has been applied for health prediction through food consumed. Data mining has a lot of advantages when using in a specific industry. Besides those advantages, data mining also has its own disadvantages e.g., privacy, security and misuse of information. We will examine those advantages and disadvantages of data mining in different industries in a greater detail.

Here Data mining used for the process of selecting, exploring and modelling large amounts of food data. This process has become an increasingly pervasive activity in all areas of medical science research. Data mining has resulted in the discovery of useful hidden patterns from massive databases. Data mining problems are often solved using different approaches from both computer sciences, such as multi-dimensional databases, machine learning, soft computing and data visualization; and statistics, including hypothesis testing, clustering, classification, and regression techniques.

The mining model that an algorithm creates from your data can take various forms, including:

- A set of clusters that describe how the cases in a dataset are related.
- A decision tree that predicts an outcome, and describes how different criteria affect that outcome.
- A mathematical model that forecasts sales.
- A set of rules that describe how products are grouped together in a transaction, and the probabilities that products are purchased together.

Advantages of proposed system:

- **Major Techniques in data mining**

A few noteworthy information mining methods have been produced and utilized as a part of information mining ventures as of late including affiliation, order, grouping, forecast and consecutive examples. We will quickly analyze the entire example to have a decent look up with them.

- **Association**

Association is a standout amongst other known information mining system. In affiliation, an example is found in light of a relationship of a specific thing on different things in a similar exchange. For design, the affiliation strategy is utilized as a part of reservation frameworks examination to recognize in which territory clients as often as possible reserve a spot. In view of this information organizations can set up comparing reservation counters here to offer more tickets and influence more to benefit

- **Classification**

Classification is the popular process which mostly depends on machine learning approach. Arrangement strategy makes utilization of numerical strategies, for example, choice trees, straight programming, neural system and measurements. Fundamentally characterization is utilized to order everything in an arrangement of information into one of predefined set of classes or groups. For design, we can apply order in application that "given every single past record of workers who left the organization, anticipate which current representatives are most likely to leave later on." For this situation, we partition the representative's records into two gatherings that are "leave" and "remain".

V. MODULES

There are 5 types of modules we have:

- Pre processing
- Calculating BMI and work out calculation
- Implementing K mine
- Mining utility itemset

- Statistical representation

Pre Processing: Pre process is the initial process, where data format will be processed for understandable data, which can be stored in the database for further use.

Calculating BMI and work out calculation: This module contains two process BMI and Workout process calculation. Both processes will work individually as a user interface model. Initially workout time, sleeping time, rest time and sitting will be fetched from the user. Using the Fmod function the average work out time of the user will be calculated.

Implementing K mine process: This module is the implementation of the core content. According to the K mining process the "bottom up" approach has been used. Here all the data training will be done internally.

Mining Utility Itemset: Mining high utility itemsets from a transactional database refers to the discovery of item sets with high utility of various statistical information. A number of relevant approaches have been used in this module for more consideration in the analysis part.

Statistical Representation: Statistics is a special subject that deals with large (usually) numerical data. The statistical data can be represented graphically. In fact, the graphical representation of statistical data is an essential step during statistical analysis.

VI. WORKING

The main objective of this project is to mine the frequent item set and maximum threshold signature of the Health life food suggestion. The frequent item set deals with the whole database of the meal wise food and predicts the health condition. It contains various transactions like user data, BMI data, food data and calorie details etc. Here enhanced EHAUPM (Enhanced Efficient High Average-Utility Pattern Mining) – Item set has been implemented which gives more accuracy and performance than TKU (mining Top-K Utility item sets) and TKO (mining Top-K utility item sets in One phase), which are implemented in the existing methods for mining such item sets without the consideration of entire database. This may cause inaccurate results and improper output. These methods may use an assumption purpose only. EHAUPM -Item set Algorithm is used to analyze the food items which are consumed by the customers, from the consumed details the calorie details of the user will be calculated. The decisions can be made on the result of analysis, so that the item set can be identified. Normally an input given by the client to the food item in the project is taken as it is and service is provided without analyzing the input. This leads to wastage of time in decision making and also the delay in finding the frequent mining.

VII. EXPERIMENTAL RESULTS AND ANALYSIS

In this system we worked on different test cases with different data such as when we came to results we got different readings. and coming to calories and vitamins we get sometimes healthy and sometimes we get unhealthy. Coming to the calorie chart, we got a diet chart suggestion.

VIII. CONCLUSION

Towards the conclusion, all the output has been verified as per the commitment. The HUIAR performed very well according to the expectations. All the output has been generated according to the given input. From the pre-process itself the input dataset generated more information according to the interface. The High Utility Item set generated more appropriate results; in addition with AR more hidden information has been revealed. The output has been verified with three different data types in various conditions. Thus this project has been executed successfully and the output has been verified. All obtained outputs are committed in the abstract. Initially more problems occurred during the architecture creation. As mentioned above architecture has been implemented successfully. All networks are working perfectly in the architecture. And these networks will work on independent processes too. These features will make this project more successful and efficient. Displaying results will be more relevant. As mentioned in the abstract all the output has been verified.

IX. FUTURE WORK

Even though K-Mine has been implemented successfully, in future some enhancements can be done in this method. The system can be further enhanced by adding new features and facilities. Now the system is platform dependent and it can be made as platform independent software. If it is developed in an oracle environment, it is made

as platform independent software, it can be used by any intranet user of the shop. The system can also be added with other data mining techniques such as K-nearest neighbours search, neural network and genetic algorithm to find some interesting patterns in the database.

Some of the enhancements are follows:

Combining algorithm: The current system is working well, But we can combine K-Mine with the A-priori algorithm for efficient data processing. Also features can be enhanced.

Web data analysis: The current system is working under a single window platform. This can be enhanced to a web data analysis model. By using this method data analysts can fetch the data from online and K-Mine can be used anywhere at any time through the internet.

Structure Improvement: Basically the current system works on decision trees, high utility itemsets and top-k itemset. In addition some architectures like hybrid or grid architecture have been implemented.

REFERENCES

1. A. Erwin, R. P. Gopalan, and N. R. Achuthan. Efficient mining of high utility itemsets from large datasets. In Proc. of PAKDD 2008, LNAI 5012, pp. 554--561
2. C. F. Ahmed, S. K. Tanbeer, B.-S. Jeong, and Y.-K. Lee. Efficient tree structures for high utility pattern mining in incremental databases. In IEEE Transactions on Knowledge and Data Engineering, Vol. 21, Issue 12, pp. 1708--1721, 2009.
3. S.-J. Yen and Y.-S. Lee. Mining high utility quantitative association rules. In Proc. of 9th Int'l Conf. on Data Warehousing and Knowledge Discovery, Lecture Notes in Computer Science 4654, pp. 283--292, Sep., 2007
4. Y. Liu, W. Liao, and A. Choudhary. A fast high utility itemsets mining algorithm. In Proc. of the Utility-Based Data Mining Workshop, 2005.



INNO  **SPACE**
SJIF Scientific Journal Impact Factor
Impact Factor: 7.542



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 **9940 572 462**  **6381 907 438**  **ijircce@gmail.com**



www.ijircce.com

Scan to save the contact details