International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

IJIEMR Transactions, online available on 12th May 2020. Link :http://www.ijiemr.org/downloads.php?vol=Volume-09&issue=ISSUE-05

Title: **MECHANIZED HEALTH ALERT SYSTEM USING HUI MINER**

Volume 09, Issue 05, Pages: 14-24

Paper Authors

P. NARENDRA BABU, K. KALYANI, V. VARSHITHA, B. HIMA BINDU

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

MECHANIZED HEALTH ALERT SYSTEM USING HUI MINER

P. NARENDRA BABU[1], K. KALYANI[2], V. VARSHITHA[3], B. HIMA BINDU[4]
[1]Associate Professor, NRI Institute of Technology
[2, 3, 4] Scholars, NRI Institute of Technology

ABSTRACT

The care setting is sometimes perceived as being 'information rich' all the same 'knowledge poor'. There is a wealth of data in the market within the care systems. However, there is insufficiency of effective analysis tools to induce hidden relationships and trends in information. Knowledge discovery and processing have found varied applications in business and scientific domain. Valuable knowledge is also discovered kind application of data mining techniques in care system. The care business collects Brobdingnagian amounts of care information that, sadly, do not appear to be "mined" to induce hidden knowledge. For information pre-processing and effective deciding Naïve man of science classifier is utilized. It's associate extension of Naïve man of science to general possibilities that aim at delivering sturdy classifications collectively once managing small or incomplete information sets. The HUI manual labourer is utilized to hunt out the high utility item sets from info. Discovery of hidden patterns and relationships generally gets untapped. Exploitation medical profiles like age, sex, sign and glucose it'll predict the likelihood of patients getting a heart condition. It permits vital knowledge, e.g. patterns, relationships between medical factors related to heart condition, to be established.

Keywords: *Behavioural bio-markers, eldercare monitoring, health alerts, in-home sensing.*

I.INTRODUCTION

High Utility Item set Mining is also in vogue processing task for locating useful patterns in shopper human action databases. It consists of discovering item sets that yield a high utility (e.g. high profit), that is High Utility Item sets. Besides shopper human action analysis, HUIM to boot has applications in different domains like click stream analysis and biomedicine. HUIM are going to be viewed as associate extension of the matter of Frequent Item set Mining (FIM), where a weight (e.g. unit profit) is additionally assigned to each item, and where purchase quantities of things in transactions are not restricted to binary values. HUIM is usually observed as a difficult downside, as a result of the utility live utilized in HUIM is neither monotonic nor anti-monotonic, in distinction to the support sleep in FIM that is, the utility of associate item set is additionally larger, smaller or up to the utility of its subsets. For this reason, economical search space pruning techniques developed in FIM can't be utilized in HUIM.

II.RELATED WORK

The field of high utility pattern mining is gaining additional importance within the recent overdue to the rise in information generation and therefore, they have to be compelled to get unidentified patterns from the notable information sets. Many analysis works are planned to satisfy the problems of high utility pattern mining. The varied planned algorithms for mining high utility patterns square measure delineate as follows.

Liu et al., have planned pseudo projection formula that is basically completely different from those planned within the past. This formula uses 2 completely different structures like array based mostly and tree-based to represent projected dealings subsets and heuristically decides to create unfiltered pseudo projection to create a filtered copy per options of the subsets. This work builds tree-based pseudo projections and array-based unfiltered projections has been build for projected dealings subsets that make formula each central processor time economical and memory saving. This formula grows the frequent item set tree by depth 1st search, wherever as breadth 1st search is employed to create the higher portion of the tree if necessary. This formula isn't solely economical on distributed and dense databases in any respect levels of support threshold and additionally extremely ascendible to terribly giant databases. The disadvantage of this formula is, it solely supports minimum description code length with tiny variety of patterns.

Han et al., have projected a frequent pattern growth (FP-Growth) algorithmic rule for mining frequent pattern with constraints. During this work the frequent pattern tree (FP-tree) structure that is associate extended prefix tree structure developed for storing crucial info concerning frequent patterns. The pattern fragment growth mines the whole set of frequent patterns exploitation the FP-growth. This algorithmic rule constructs an extremely compact FP-tree and applies a pattern growths methodology for info scans that are sometimes considerably smaller than the first info by that pricey info scans are saved within the frequent mining processes. The disadvantage of this algorithmic rule is it reduces multi-pass candidate generation method within the 1st section by discarding isolated things to cut back the quantity of candidates. Additionally, this work shrink the info scanned in every pass, and it takes a lot of computations time.

Liu et al., have projected a two-phase algorithmic program to seek out high utility item sets. This algorithmic program expeditiously prunes down the amount of candidates and obtains the entire set of high utility item sets. This work has developed with 2 phases. Section one uses dealing-weighted downward closure property that is applied to feature high transaction weighted utilization sets throughout the amount wise search. In

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

section 2, any over calculable low utility item sets square measure filtered exploitation an additional info scan. This algorithmic program needs fewer info scans, less memory area, and less procedure price for big info and performs o.k. in terms of speed and memory price on each artificial and real database. The most disadvantage of this algorithmic program is, the poor frequent counts for continual candidate item set which might lose attention-grabbing patterns.

Li et al., have planned an associate degree isolated things discarding strategy (IIDS) formula for utility mining. This formula discovered high utility item set with less variety of candidates that improve the performance of the pattern mining. This formula shows that thing set share mining downside may be directly born-again to utility mining downside by replacement the frequent values of every item in very dealings by its total profit, i.e., multiplying the frequency price by its unit profit. During this work the share frequent set mining scans the information to calculate the share price of every item set and removes all useless candidate item sets and remaining candidate to come up with. The direct condition generation may be a level wise methodology associate degree it maintains an array for every candidate throughout each pass. It provides associate degree economical thanks to design vital operations by victimization dealings.

III.PROBLEM DEFINITION

Anticipating patient's future behavior on the given history is one among the vital applications of knowledge mining techniques which will be utilized in health care management. A serious challenge facing tending organizations (hospitals, medical centers) is the provision of quality services at cheap prices. Quality service implies identification patients properly and administering treatments that square measure effective. Poor clinical selections will result in fateful consequences that square measure thus unacceptable. Hospitals should conjointly minimize the price of clinical tests. They will attain this result by using applicable computer-based info and/or call support systems. Health care knowledge is very large. It includes patient central knowledge, resource management knowledge and reworked knowledge. The availability of integrated info via the large patient repositories, there's a shift within the perception of clinicians, patients and payers from quantitative assessment of information of knowledge with the supporting of all clinical and imaging data.

IV.ALGORITHM

HUI-Miner uses a novel structure, called utility-list, to store both the utility information about a data and the heuristic information for pruning the search space of HUI-Miner.

HUI-MINER ALGORITHM:

1. Input: P.UL, the utility-list of item set P, initially
2. empty;
3. ULs, the set of utility-lists of all P's

International Journal for Innovative
Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

4. 1-extensions;
5. minutil, the minimum utility threshold.
6. Output: all the high utility item sets with P as prefix.
7. foreach utility-list X in ULs do
8. if SUM(X.iutils)≥minutil then
9. output the extension associated with X;
10. end
11. if SUM(X.iutils)+SUM(X.rutils)≥ minutil then
12. exULs = NULL;
13. foreach utility-list Y after X in ULs do
14. exULs = exULs+Construct(P.UL, X,Y);
15. end
16. HUI-Miner(X, exULs, minutil);
17. End
18. End

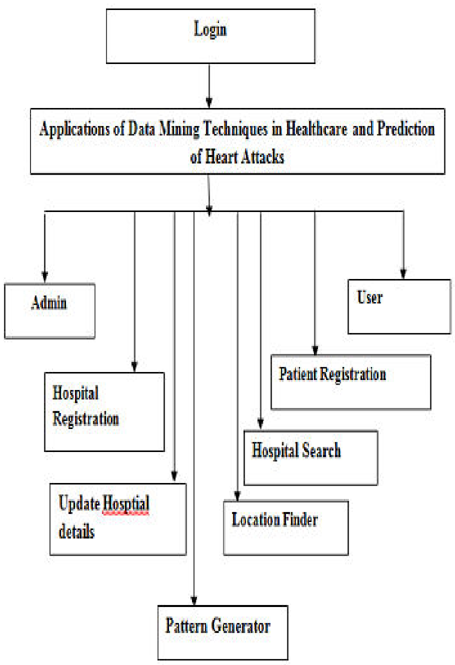

Fig. : Process Diagram for Automated Health Alert System

V.ARCHITECTURE

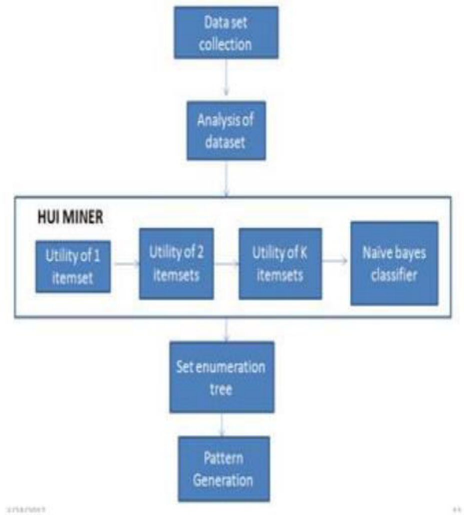

Fig. Architecture Diagram for Automated Health Alert System

In a database, firstly, all 1-itemsets are candidate high utility item sets. After scanning the database, the algorithms eliminate unpromising 1-itemsets and generate 2-itemsets from the remaining 1-itemsets as candidate high item sets.

After the second scan over the database, unpromising 2-itemsets are eliminated and 3-itemsets as candidates are generated from the remaining 2-itemsets... The procedure is performed repeatedly until there is no generated candidate item set.

VI.MODULES ADMIN MODULE

The hospitals to be registered desires to contact the administrator. The administrator provides the login ID for the doctors operating within the hospitals that have registered. The doctors United Nations agency wish to be a member of the website, however operating within the non-registered hospitals can also produce their account directly. This doesn’t mean there's no correct authentication, however paves the thanks to offer an improved data discovery. Those outside doctors will register their details and has got to send the patient details to the administrator. The administrator then checks out the small print and at last when the verification is over, admin sends the ID and positive identification to the revered doctor's email. If the other drawback arises then straightaway the admin is named, the internal operations are properly handled by the admin solely.

USER MODULE

The user can login and view the hospital details and the generated analysis.

HOSPITAL MODULE

The doctors will login with their user name and that they are provided to look at the patient record. The patient's record are updated and inserted by the doctor and build an entry within the information. Through this the patient's record is maintained utterly. The hospitals that are registered are provided with the registration ID, this provides the users to look at the registered hospitals, doctors list and patient details.

REPORT GENERATION MODULE

The patient's record maintained is then created use in generating the pattern. The rate of death because of heart diseases and their risk factors area unit shown within the pattern and effectively generated. This pattern creates the user to administer right call through effective mining of data's from a definite range of hospitals and therefore, you look after causes.

EXTRACTING DATA ITEMS

In the first module, the user can choose the dataset to get high utility item sets, after choosing the dataset, the data items should be extracted from the transactions dataset, the data items are individual items which occur in the transaction.

CALCULATING TU (TRANSACTION UTILITY):

After constructing break down tree for the XML information. The dealing with Utility (TU) of every item is calculated. The dealing with Utility is calculated by exploitation the worth of profit and its amount. Dealing with Utility=Profit*quantity.

Finding TWU (Transaction Weighted Utility):

After calculating Transaction Utility (TU) of each item from the parse tree. We need to find Transaction Weighted Utility(TWU) for each items from the calculated Transaction Utility(TU).The TWU of item set whose value is less than the given threshold value are discarded by pruning items .The discarded item sets are called unpromising item set they don't yield more utility to the user. The item set whose TWU value is less than minimum utility is called unpromising item set, otherwise promising item set.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

TNO	Items	Transaction utility	Item utilities for this transaction
t1	3 5 1 2 4 6	30	1 3 5 1 0 6 5
t2	3 5 2 4	20	3 3 8 6
t3	3 1 4	8	1 5 2
t4	3 5 1 7	27	6 6 1 0 5
t5	3 5 2 7	11	2 3 4 2

Mining frequent patterns based on M-LP-tree

Mining frequent patterns based on M-LP-tree (LP-growth) LP-growth searches LP-tree and create a conditional LP-tree for mining frequent patterns. To do that, our algorithm first selects the bottom item from the header list and traverses nodes connected to corresponding node links. Then, supports of the visited nodes are stored, and nodes from each linked node to a root are searched. Each node can be accessed directly if the search is conducted within one LPN. In other words, given a current node, the algorithm immediately accesses N (k1) to approach a parent node of. Iterating the traversal regarding one LPN, the algorithm reaches a header of the LPN, where the header refers to its parent node, i.e. the other LPN.

Mining High-Utility Item sets in a Transaction Database using the IHUP Algorithm

How to run this example?

- If you're mistreatment the graphical interface, (1) select the "IHUP" rule, (2) choose the input data "DB_utility.txt", (3) set the computer file name (e.g. "output.txt") (4) set the minimum utility to thirty and (5) click "Run algorithm".
- If you wish to execute this instance from the instruction, then execute this command: java -jar spmf.jar run IHUP DB_utility.txt output.txt thirty during a folder containing spmf.jar and also the example input data DB_utility.txt.
- If you're mistreatment the ASC II text file version of SPMF, launch the file "MainTestIHUP.java" within the package ca.pfv.SPMF.tests.

What is IHUP?

IHUP (Ahmed et al., TKDE 2009) is associate degree formula for locating high-utility item sets during dealing information containing utility data.

Note that the initial IHUP formula is intended to be progressive. During this implementation of IHUP will solely be run in batch mode.

Also note that additional economical formula is recently planned like FHM (2014) and HUI-Miner (2012). These latter algorithms outperform IHUP by over associate degree order of magnitude, and are offered in SPMF.

What is the input?

IHUP takes as input dealing information with utility data and a minimum utility threshold min_utility (a positive integer). Let's take into account the subsequent information consisting of five transactions (t1, t2...t5) and seven things (1, 2, 3, 4, 5, 6, and 7). This information is provided within the document "DB_utility.txt" in the conglomeration ca.pfv.spmf.tests of the SPMF distribution.

Each line of the info is: a set of things (the 1st column of the table),

The add of the utilities (e.g. profit) of those things during this group action (the second column of the table),

The utility of every item for this group action (e.g. profit generated by this item for this transaction) (the third column of the table).

Note that the worth within the second column for every line is that the add of the values within the third column.

What are real-life examples of such a database? There are several applications in real life. One application is a customer transaction database. Imagine that each transaction represents the items purchased by a customer. The first customer named "**t1**" bought items 3, 5, 1, 2, 4 and 6. The amount of money spent for each item is respectively 1 $, 3 $, 5 $, 10 $, 6 $ and 5 $. The total amount of money spent in this transaction is 1 + 3 + 5 + 10 + 6 + 5 = 30 $.

What is the output?

The outputs of IHUP is the set of high utility item sets having a utility no less than a min_utility threshold (a positive integer) set by the user. To explain what a high utility item set is, it is necessary to review some definitions. An item set is an unordered set of distinct items. The utility of an item set in a transaction is the sum of the utility of its items in the transaction. For instance, the utility of the item set {1 4} in transaction t1 is 5 + 6 = 11 and the utility of {1 4} in transaction t3 is 5 + 2 = 7. The utility of an item set in a database is the sum of its utility in all transactions where it appears. For instance, the utility of {1 4} in the database is the utility of {1 4} in t1 plus the utility of {1 4} in t3, for a total of 11 + 7 = 18. A high utility item set is an item set such that its utility is no less than min_utility For instance, if we run IHUP with a minimum utility of 30, we obtain 8 high-utility item sets:

Itemsets	utility	support
{2 4}	30	40 % (2 transactions)
{2 5}	31	60 % (3 transactions)
{1 3 5}	31	40 % (2 transactions)
{2 3 4}	34	40 % (2 transactions)
{2 3 5}	37	60 % (3 transactions)
{2 4 5}	36	40 % (2 transactions)
{2 3 4 5}	40	40 % (2 transactions)
{1 2 3 4 5 6}	30	20 % (1 transactions)

If the database is a transaction database from a store, we could interpret these results as all the groups of items bought together that generated a profit of 30 $ or more.

Input file format

The input file format of IHUP is defined as follows. It is a text file. Each lines represents a transaction. Each line is composed of three sections, as follows.

- First, the items contained in the transaction are listed. An item is represented by a positive integer. Each item is separated from the next item by a single space. It is assumed that every one thing among a same group action (line) area unit sorted in line with a complete order (e.g. ascending order) which no item will seem doubly among a similar group action.
- Second, the image: seems and is followed by the group action utility (an integer).
- Third, the image: seems and is followed by the utility of every item during this group action (an integer), separated by single areas.

For example, for the previous example, the input file is defined as follows:

3 5 1 2 4 6 : 30 : 1 3 5 10 6 5
3 5 2 4 : 20 : 3 3 8 6
3 1 4 : 8 : 1 5 2
3 5 1 7 : 27 : 6 6 10 5
3 5 2 7:11:2 3 4 2

For example, for the previous example, the input file is defined as follows: Consider the first line. It means that the transaction {3, 5, 1, 2, 4, 6} has a total utility of 30 and that items 3, 5, 1, 2, 4 and 6 individually have a utility of 1, 3, 5, 10, 6, 5 in this transaction. The following lines follow the same format.

Output file format

The output file format of IHUP is described as follows. It is a text file, where each line represents a high utility item set. On each line, the items of the item set are first recorded. Each item is represented by an integer, accompanied by a single space. After, all the items, the keyword " #UTIL: " appears and is accompanied by the utility of the item set. For example, we show below the items, the keyword " #UTIL: " seems and is followed by the utility of the item set. For instance, we have a tendency to show below the computer file for this instance.

2 4 #UTIL: 30
2 5 #UTIL: 31
1 3 5 #UTIL: 31
2 3 4 #UTIL: 34
2 3 5 #UTIL: 37
2 4 5 #UTIL: 36
2 3 4 5 #UTIL: 40
1 2 3 4 5 6 #UTIL: 30

For example, the first line indicates that the item set {2, 4} has a utility of 30. The following lines follow the same format.

Performance

High utility item set mining is a

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

more difficult problem than frequent item set mining. Therefore, high-utility item set mining algorithms are generally slower than frequent item set mining algorithms. The IHUP (2009) formula was the fastest formula for high-utility item set mining in 2009.However, additional economical rule are recently planned. UP Growth (2010) is AN ameliorated version of IHUP. The HUI-Miner (2012) rule outperforms UP Growth (2009) by over AN order of magnitude, and additional recently the FHM rule (2014) was shown to be up to 6 times quicker than HUI-Miner. Additional recently, the EFIM rule (2015) was planned and was shown to surpass IHUP, and different recent algorithms like FHM (2014), HUI-Miner (2012), HUP-Miner (2014). Of these algorithms square measure offered in SPMF (see "performance" page of this website).

Data Set Link:-

http://archive.ics.uci.edu/ml/machine-learning-databases/00319/

****

Health dataset

****

Authors: OrestiBaños, Rafael Garcia, Alejandro Saez

Date: 22/10/2013

Institution: University of Granada (UGR)

Department: Department of Computer Architecture and Computer Technology.

Contact: oresti@ugr.es (oresti.bl@gmail.com)

NOTE: if you use this dataset please cite the following work

Banos, O., Garcia, R., Holgado-Terriza, J.A., Damas, M., mobile health applications.In: Proceedings of the sixth International Work-conference on close motor-assisted Living an energetic Ageing (IWAAL 2014), Belfast, uk, December 2-5 (2014).

VII.CONCLUSION

In this paper, we've got a bent to gift studies designed to analysis embedded health assessment. A forward search was initial accustomed severally investigate the feature house of embedded in-home sensors. We've got a bent to boot delineates a prospective study exploitation 1-D health alerts. Clinical ratings on the health alerts were provided by clinicians and conversant in train and check multi-D classifiers. Solely the sole 6-D performance was achieved by associate FPT supported domain data solely, though the SVM (trained on labelled employment data) had a homogeneous performance. To boost this performance, we've got a bent to face live getting to investigate online learning exploitation the alert ratings as feedback. The work given here shows that domain data is employed for initial classification to create up enough data to support online learning ways in which during which at intervals which. Finally, supported the study results and our experience exploitation health alerts prospectively, we've got a bent to line up a model for investigation health decline with in-home sensors. A randomized management study exploitation this model with the hydraulic bed device, motion sensors, and in-home gait is current to a good deal of checks the potential of embedded health assessment.

VIII.FUTURE WORK

As technology changes the new requirements unit of measurement expected by the user, to spice up the usefulness of the merchandise may would like new versions to be introduced. Though the system is complete and dealing with efficiency, new changes that enhance the system usefulness are going to be superimposed with none major changes to the entire system. In future if we've got a bent to use FP Growth algorithm in place of HUI working person algorithm as a results of It permits frequent item set discovery whereas not candidate generation. It builds a compact arrangement referred to as FP tree with a pair of passes over the information. It takes out frequent item sets straight from the FP tree and passes over through the FP tree. Therefore then the abnormal browse of patients square measure about to be recognized extra accurately and fastly.

IX. REFERENCES

[1] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules," inProc. Int. Conf. Very Large Data Bases, 1994, pp. 487–499.

[2] C. Ahmed, S. Tanbeer, B. Jeong, and Y. Lee, "Efficient tree structures for high-utility pattern mining in incremental databases," IEEE Trans. Knowl. Data Eng., vol. 21, no. 12, pp. 1708–1721, Dec.2009.

[3] K. Chuang, J. Huang, and M. Chen, "Mining top-kfrequent patterns in the presence of the memory constraint,"VLDB J., vol. 17,pp. 1321–1344, 2008.

[4] R. Chan, Q. Yang, and Y. Shen, "Mining high-utility itemsets," in Proc. IEEE Int. Conf. Data Mining, 2003, pp. 19–26.

[5] P. Fournier-Viger and V. S. Tseng, "Mining top-k sequential rules," inProc. Int. Conf. Adv. Data Mining Appl., 2011, pp. 180–194.

[6] P. Fournier-Viger, C. Wu, and V. S. Tseng, "Mining top-kassociation rules," inProc. Int. Conf. Can. Conf. Adv. Artif. Intell., 2012, pp. 61–73.

[7] P. Fournier-Viger, C. Wu, and V. S. Tseng, "Novel concise representations of high utility itemsets using generator patterns," in Proc. Int. Conf. Adv. Data Mining Appl. Lecture Notes Comput. Sci., 2014, vol. 8933, pp. 30–43.

[8] J. Han, J. Pei, and Y. Yin, "Mining frequent patterns without candidate generation," inProc. ACM SIGMOD Int. Conf. Manag. Data, 2000, pp. 1–12.

[9] J. Han, J. Wang, Y. Lu, and P. Tzvetkov, "Mining top- kfrequent closed patterns without minimum support," in Proc. IEEE Int. Conf. Data

Mining, 2002, pp. 211–218.

[10] S. Krishnamoorthy, “Pruning strategies for mining high utility itemsets,”Expert Syst. Appl., vol. 42, no. 5, pp. 2371–2381, 2015.

[11] Mathews, Sherin M., Luisa F. Polanía, and Kenneth E. Barner. "Leveraging a discriminative dictionary learning algorithm for single-lead ECG classification." Biomedical Engineering Conference (NEBEC), 2015 41st Annual Northeast. IEEE, 2015.

[12] Mathews, Sherin M., Chandra Kambhamettu, and Kenneth E. Barner. "Maximum correntropy based dictionary learning framework for physical activity recognition using wearable sensors." International Symposium on Visual Computing. Springer International Publishing, 2016.

[13] Mathews, Sherin M., Chandra Kambhamettu, and Kenneth E. Barner. "Centralized class specific dictionary learning for wearable sensors based physical activity recognition." Information Sciences and Systems (CISS), 2017 51st Annual Conference on. IEEE, 2017.

[14] Mathews, Sherin Mary. Dictionary and deep learning algorithms with applications to remote health
monitoring systems.