

- Word Count: 3866

Plagiarism Percentage

4%

sources:

- 1 2% match (Internet from 05-Feb-2015)
<http://etd.lib.nsysu.edu.tw/ETD-db/ETD-search/getfile?URN=etd-0621110-150934&filename=etd-0621110-150934.pdf>
- 2 1% match (Internet from 29-Nov-2016)
<https://pdfs.semanticscholar.org/e27c/0b10949db4618c41e2f2245164a542935209.pdf>
- 3 1% match (Internet from 22-Mar-2010)
http://www.pha.jhu.edu/~ehoverst/obs_astro/tutorial/ehgal004.fits
- 4 1% match (Internet from 24-Sep-2010)
http://umbra.nascom.nasa.gov/eit_qkl/2007/08/04/efr20070804.070608

paper text:

Sequential Pattern Mining Application to Support Customer Care “X” Clinic Alexander Setiawan¹, Adi Wibowo², Samuel Kurniawan³ Informatics Engineering Department of Petra Christian University Siwalankerto 121-131, Surabaya – Indonesia Telp.: +62 31 8439040, 8494830-31 1alexander@petra.ac.id; 2adiw@petra.ac.id; 3m26406143@john.petra.ac.id Abstract. The X clinic that was one of the pioneers in the aesthetics clinic in Indonesia, had much experienced manpower. In a supplied manner this experience, the doctors and the nurse could to the X clinic serve consultations and could give the suggestion or the recommendation to the customer for the following maintenance. This matter gave comfort for the customer to take the decision. However, not all the customers had time that was enough to consult with the doctor. With used Free Span, one of the algorithms in the method sequential pattern mining was expected to be able to satisfy the requirement for the clinic customer of X. The use of sequential pattern mining in this recommendation system could help the doctor in increasing the recommendation, and helping the customer in taking the decision. This algorithm used the historic data the maintenance of the available customer. Results that were given in this program took the form of the pattern that in accordance with the available situation to the clinic of X. The result of the recommended selected based on existing customer categories, namely gender, priority customers in the clinic, and age range. Expected with the available category to be able to give the recommendation that agreed with the customer's available criterion.

Keywords. Sequential pattern mining, Free Span, Clinic.

1 Introduction The X clinic was known as the pioneer in Aesthetic Clinic in Indonesia. Supplied with the experience for more than 16 years, the X clinic is currently known as one of the foremost aesthetics clinics in Indonesia with 11 branches that were spread in Jakarta, Surabaya, Denpasar, Kuta, Balikpapan, Malang, Batam and Makassar. With the "Beyond Result" concept, apart from giving results of the first-rate maintenance, the X clinic also gave comfort, the security

in each one of his maintenance as well as the friendliness and privation for his customer. Now the role of the internet is felt to be increasingly important and significant. Through this internet, the X Clinic wanted to give a special service for the customers in the X clinic by giving an application web the recommendation. This recommendation could be used by the nurse or the doctor to be able to give maintenance information what could be taken by the customer, and could be used by the customer as help in taking the decision. Therefore, with the existence of the data concerning historic the customer who carried out the maintenance in this clinic could be known by the maintenance what could be taken by the further customer. From this available data, also could be known how the place or the pattern of the maintenance that often was taken by the customer with used sequential pattern mining that was used to look for and find relations between item available to some of dataset. Specifically problems that was appointed in the production of this web was [5] :

- How to change from the long database (in this case used .dbf) became the new database (the database sql).
- How the algorithm that was used in the method sequential pattern mining.
- How processed the customer's data to the X clinic with used sequential pattern mining.
- How could the application give information concerning any maintenance that could be used in order with used sequential pattern mining.

2 Sequential Pattern Mining Sequential pattern mining is a data mining technique used to identify patterns of ordered event within a database. First introduced in 1995 by Rakesh Agrawal of IBM's Almaden Research

2

Center [1]. Was the method of look for and finding relations between item available to some dataset. The program sequential pattern mining aimed at finding information item-item that was connected with each other in the form of rule. Therefore sequential pattern exact to be applied in mining sequential pattern [2]. The available problem in the data mining was to find the pattern sequential. Inputan his data was a collection sequence that was mentioned

data-sequences. Each data sequential was a list from transactions, where each transaction was a

1

collection item. Generally a transaction was distinguished according to occurrence time of this transaction. Some sequential-pattern also consisted the list from a collection item [2], [3]. Anyway, all item in one a part of a pattern sequential must be gotten in a data transaction sequential to support this pattern. For the ease could diasumsikan that there is no data-sequence had more than one transaction with same transaction time, and used transaction time as the connoisseur of this transaction. So, in this case the quantity or the amount of some item was not counted on [3].

2.1 FreeSpan Algorithm The FreeSpan Algorithm consisted of several stages, that is the first stage made sequence the database and searched f list or also could be mentioned as length-1 sequential pattern, the second stage made frequent item matrix, the third stage made the table pattern generation that consist of length-2 sequential pattern, item repeating pattern, and projected DB, and the last stage was the production sequential patterns [4]. The first stage from this algorithm was made sequence the database was based on the available data of the transaction data [4]. This data could be distinguished in accordance with the time and did not see the quantity from this transaction shown in Table 1. Table 1. The Example Table of The Transaction ID Tanggal

10	28/03/2010	10	15/04/2010	10	23/05/2010	10	22/06/2010	10	22/06/2010	20	21/03/2010	20	21/03/2010	20
----	------------	----	------------	----	------------	----	------------	----	------------	----	------------	----	------------	----

b d c b a c b f c e b f g a a ... e 3

From the example of the data to the Table 1 could be made by one sequence the data- base. The production sequence the database made the time (in the example above, take the form of the date) as the reference in grouped sequence the database like the example to the Table 2. [2], [4]. FreeSpan Algorithm considering the pattern-projection method for mining sequential patterns. It is the original approach for mining sequential pattern recursive projecting the data sequence into smaller database [5]. Table 2. Tabel Sequence Database ID SEQUENCE ITEM_PATTERN 10

(b,d),c,b,(a ,c) a,b,c,d 20 (b, f),(c, e),b,(f,g) b,c,e, f,g 30 (a, h),(b,f),a,b,f a,b, 4

f,h 40 (b,e),(c,e),d b,c,d,e 50 a,(b,d),b,c,b,(a,d,e) a,b,c,d,e From the Table 1 could be seen how the production of the Table 2 in the column se- quence, contained the list of the available transaction with used the date as the arbiter. For the transaction that happened in the same date, his writing was distinguished by signaling (). For example will be explained by the production sequence by ID 10. In the date 28/03/2010 bought the thing b and d, so as to be grouped to (b,d). In the date 15/04/2010 only bought c, so as only was written c. In the date 23/05/2010 the avail- able transaction only bought the thing b, so as only was written b. In the date 22/06/2010 bought an and c, so as to be grouped to (a,c). From the data - this data had finally been grouped to one sequence, that is (b,d) ,c,b, (a,c). After getting the table sequence, looked for the collection from the data that often emerged. This collection was also mentioned as length-1 sequential pattern. This col- lection will be put in order be based on support descending order, before looking for this collection determined previously minimal support. Minimal support for this ex- ample was 2. Length-1 sequential pattern in this example was as follows

b: 5, c: 4, a: 3, d: 3, e: 3, f: 3. 1

Length-1 sequential pattern this also could be mentioned frequent item list or more briefly f list. The second stage that is made

frequent item matrix F to count the frequency 1

from the incident every time length-2 sequence that was formed be based on f list. To f list (i1,i2,... ,im), F was matrix the F triangle

[j,k], where 1 j m. and 1 k j. F [j,j] (for 1 j m.) only had 1 counter, whereas 1

the rest of them of

F [j,k] (1 j m. and 1 k j)

1

had 3 counters:

(A,B,C), where A was the number of

1

incidents where ik was after ij (for example in the table sequence was gotten < ij,ik >), B was the incident where ik was before ij (for example in the table sequence was gotten < ik,ij >), and C was the incident where ik happened together be the same as ij (for example in the table sequence was gotten < (ij,ik) >). Was based on this example, was gotten 6 available things in f list. This matter to make the triangle matrix with the measurement 6x6, with each value = 0, could be seen in the Table 3. Table 3. Table Frequent Item Matrix Item Output length-2 sequential pattern f <bf>:2;<fb>:2;<(bf)>:2;<ff>:2 e <be>:3;<(ce)>:2 d <bd>:2;<db>:2;<(bd)>:2;<cd> :2;<dc>:2;<da>:2 a <ba>:3;<ab>:2;<ca>:2;<aa>:2 c <bc>:4;<cb>:3 b <bb>:4 Ann. On Repeating Items

{b+f+},<ff+><b+e> {b+d};< da+><aa +>;{a+b+};< ca+>{b+c}

3

<bb+> Ann. On Projected DB \emptyset <(ce)>:{b} <da>:{b,c};{cd}:{b} <ca>:{b} \emptyset \emptyset The matrix above will be used to make the stage further, that is making the table pattern generation that consist of length-2 sequential pattern, item repeating pattern, and projected DB. The step to make the table pattern generation was as follows: • Made length-2 sequential patterns: To every time counter from the matrix, if thought from counter was not smaller than minimal support, put forward frequent pattern this. • made annotation ounce item repeating patterns for each column j: for the side diagonal, if F [j,j] min support, put forward < jj+ >. For the column i j, there was something that must be paid attention to: if F [i,i] min support then was contained i+, if F [j j] min support then was contained j+. • made annotation ounce projected databases for each column j: To every time i < j, if F [i,j], F [k,j], and F [i,k] (k < i) can make pattern generating triple (with the note of each pair often emerged), k could be added in the column projected to i. After investigated all the columns in front i, the column projected could be determined. The production of the table pattern generation for the example above could be seen in the Table 4. Table 4. Table The Pattern Generation b 4 c a d e f (4,3,0) (3,2,0) (2,2,2) (3,1,1) (2,2,2) b 1 (2,1,1) (2,2,0) (1,1,2) (1,1,0) c 2 (1,2,1) (1,0,1) (1,1,0) a 1 (1,1,1) (0,0,0) d 1 (1,1,0) e 2 f 3 Analysis & System Design The X clinic that was the pioneer in Aesthetic Clinic in Indonesia has used the computer system in recording the transaction and his details base on website [6]. This system also recorded the data from this clinic customer. The customer who just the first time was present being asked for by the data himself, in part: the name, the address, gender, and the date of birth. After information was finished was recorded, the customer could choose the maintenance kind what will be taken and chose the doctor to be able to carry out consultations. For the customer who for a long time has been registered could make the promise or appointment with this clinic. This agreement contained the date and the hour for the maintenance or carried out consultations with the doctor. After carrying out the agreement, immediately will be carried out by the good action medically or non-medical. However the problem that was dealt with was several patients who came to experience the difficulty in choosing the maintenance kind,

especially for the patient who just tried the maintenance. Whereas the X clinic, still used the method of maintenance consultations in a traditional manner that is consulting directly with the doctor. From the problem that often emerged, then could be concluded that the company needed a based system the computer that could help in decision making by using the story data of the maintenance of the available customer as the source of the beginning data. The system that was needed by the clinic was:

- The system was based on the data mining, that could produce information for the customer in carrying out decision making by using the story data of the maintenance of the available customer.
- The system that used the concept sequential pattern, where could show connection relations between the maintenance from various abstraction stages, not only from treatment that was taken simultaneously, but also in order.
- The computer system that could integrate all the factors of the customer's decision making to produce good information for the customer and for the clinic. The system used the concept of Frequent Pattern-Projected Sequential Pattern Mining (FreeSpan), where the factor or this decision making variable was: gender, range the age, the category of the customer, and the price of the maintenance. The process mining that the first time was carried out was the process selection or the election. In this process was carried out by the type selection customer.

Fig. 1. Flowchart Selection Process The selection process on figure 1. was gotten by 2 process sorts more, that is generate sequence the database and generate item pattern the Election was based on gender, range the age, the category of the customer, and the price of the maintenance. The process generate Item pattern this will be used to get length-1 sequential pattern or also could be mentioned as frequent item list (f list). took the data item pattern from the table sequence that will be used to search length-1 sequential pattern or often was mentioned frequent item list (f list). Frequent Item Matrix. Matrix that was produced from this process was very important his use in searching length-2 sequential patter. After that was carried out Generate Pattern Generation. This process used Frequent Item Matrix and the table sequence in his production. Results of this process of taking the form of 3 sorts, that is length-2 sequential pattern, annotation ounce repeating items, and annotation ounce projected databases. The flowchart generate pattern generation can be seen figure 2.

Fig. 2. Generate Pattern Generation Along with this was the data Current in the system mining in the form of DFD that was the explanation from the figure 1 and figure 2. In DFD this was gotten by 6 processes that will be it was explained further that is the process selection, sorted frequentitems, frequent item matrix, generate pattern generation, generate rule, and translate rule can be seen in Figure 3.

Managerial Customer Customer Customer Id Treatment Customer Data Treatment Data 1 RecommendationelRecetipoonrt Recommendation Data Selection Data 2 Sequence Table Sequence Data Sorted Frequent Items 6 Sequence Data Translate Rule 3 Frequent Item Matrix Sequence Data Sorted Frequent Data Frequent List Frequent Item List Data Matrix Data Matrix Table Matrix Data Rule List Data 5 4 Rule Data Generate Pattern Data Generate Rule Table Rule Pattern Data Pattern Generation Table Pattern Generation

Fig.3. Process Data Flow Diagram Mining Clinic "X"

4 Implementation and Testing This explained concerning the implementation to each controller together with available functions to each controller. Controller connected view and the model. Each controller had the function that differed to every time and the available process in this recommendation system.

4.1 Testing of Data Mining

The page of the management mining was used to get the data from the recommendation be based on historic the customer and the category who were given. The category that was used in getting the recommendation to be gender, the category customer, and range the age. The Home pages mining for representative of Aesthetic Clinic can be seen figure 4. To this page was gotten by main activity access from the clinic of "X".

Fig. 4. Home Pages Mining To this page in figure 5. was gotten by the category that was used to get the recommendation data. The available category was gender, the priority in the customer, and range the age.

Fig. 5. Enter Mining Category Gender that was used as the reference was L (The Man) and P (The Woman). The priority in the customer who was used was Core (the core customer), Regular (the public's customer), and Low (the customer with the minimal priority). Range the age that was

given here was used to restrict the historic data the customer who was used. From the category that was given, the program will make the table mining to facilitate the process of the production sequential pattern by means of looking for the historic data the customer that in accordance with the category and put in order was based on the customer's data. The process was in the production of this table put in order be based on id from the customer. The column sequence was historic that was recorded from the customer with id this, and the column item pattern was the list from the maintenance that was recorded in historic this customer. The writing in the column sequence was adapted to time that was used by this customer. The mining sequence table can be seen figure 6. Fig. 6. Form Mining Sequence Table After getting the mining sequence table, the process will be followed by the Matrix production after got frequent item list or length-1 sequential pattern from the column item pattern, the form matrix table can be seen figure 7. Fig. 7. Form Matrix Table From matrix that was formed, the process mining will be continued by searching length-2 sequential pattern that will produce repeating item pattern and projected the database. In this program, the process will be continued to Projected Database in accordance with that was put forward to the figure 8. Fig. 8. Projected Database Mining From the data that was received in projected the database, could be looked for the pattern with combined projected the database. And results of this combination were used to get sequential pattern and support him can be seen figure 9. Fig. 9. Form Sequential Pattern Results of the end that was obtained were the pattern from sequential pattern that was interpreted. Results of this end were the reference in giving the recommendation to be based on the category that was determined in the beginning of the process of the management mining this. The figure 10. put forward the example produced by the end that was received after carrying out the process mining towards the data customer that was matched with the category that was chosen at the beginning of the process. Fig. 10. Form Final Result Mining

4.2 Testing of Recommended User

To get the recommendation, member will be asked for to put id member before. Id member this was used to get the data himself member and was compared with the available category to the recommendation table. The data himself who could be taken was gender, the priority, and the date of birth. This page in figure 11. put forward results of the recommendation that was taken from the database with the data himself the customer with id 11979. The data himself who was taken repeatedly was the category that was needed, and the last maintenance that was carried out. In these results was seen that the last maintenance that was carried out by the customer had "MRC CONTOURING BODY TREATMENT MCBT", because that was based on the available data the recommendation that was first that is taking "MRC RF LIFT" and could be followed by "MRC INJEKSI ACNE 1". Fig. 11. Form Recommendation Member

5 Conclusion and Advise

Was based on available discussions in chapters beforehand then could be taken by several conclusions as follows:

- From the test with the data that actually, could be produced the pattern of the pattern that could be recommended to the customer. The pattern of the pattern that was gotten in these results was made be based on FreeSpan Algorithm in the method sequential pattern mining. The source of the data that was used in this algorithm, was processed in accordance with the category that is gender, the priority in the customer, and range the age.
- Through website that was made, the Administrator could get the pattern of the pattern of the maintenance that was taken by the customer in accordance with range the age in the category that was given.
- From the process of the testing on the whole, this program was it was felt more inappropriate in the search for the maintenance recommendation to the customer. Because the taking of the maintenance that was carried out by the customer often was carried out on the basis of the recommendation from the doctor, so as the pattern of the recommendation that was given not more maximal.
- Must be tried to develop this recommendation system with more made his execution time efficient and maximized results that were received.
- Facilities in this application were still limited so as big the opportunity of increasing and increasing facilities to this site so that the achievement of this site could be more optimal.

References 1. Agrawal, Rakesh., Srikant Ramakrishnan, Mining sequential patterns. In : Proc int conf data p.3-p.14.

(1995) 2. Han, Jiawei., Kamber, Micheline. *Datamining: Concept & Techniques 3rdEdition*. America : Morgan Kaufmann Publisher is an imprint of Elsevier. (2012) 3. Wright, Aileen P., Wright, Adam T., McCoy, Allison B., & Sittig, Dean F. "The use of se-quential pattern mining to predict next prescribed medications". *Journal of biomedical in- formatics*. Elsevier Inc. (2014) 4. Han, Jiawei., Pei, Jian. *FreeSpan: Frequent pattern-projected sequential pattern mining*. Retrieved April 7, 2010, from <http://www.cs.sfu.ca/~jpei/publications/freespan.pdf>. (2010) 5. Desai, Niti, & Ganatra, Amit. "Draw Attention to Potential Customer with the Help of Subjective Measures in Sequential Pattern Mining (SPM) Approach". *Proceeding of Inter- national Conference on Recent Trends in Information, Telecommunication and Comput- ing, ITC*, (2014) 6. Myer, Thomas. "Professional CodeIgniter". Wiley Publishing. (2008)