

# Decision Support System for Cosmetics Product using Sequential Pattern Mining based FP-Growth Algorithm

Soe Moe Myint<sup>1</sup>, Lae Wah Htun<sup>2</sup>, Moh Moh Myint Kay<sup>3</sup>, Aye Aye Cho<sup>4</sup>

<sup>1</sup>Lecturer, Faculty of Computer Systems and Technologies, University of Computer Studies, Patheingyi, Myanmar

<sup>2</sup>Assistant Lecturer, Faculty of Computer Science, Hinthada, Myanmar

<sup>3</sup>Assistant Lecturer, Faculty of Computer Science, Hinthada, Myanmar

<sup>4</sup>Associate Professor, Faculty of Computer Science, Hinthada, Myanmar

## ABSTRACT

Today information is vital in our modern technology world. Organizational decisions are largely dependent on computerized systems. In recent days, massive data are collected from the customer's purchase records using application software system in the shop, which are stored in database using relational database management systems. Manager can be discovered interesting patterns from the customer's purchase records by analyzing the sales in database, using the data mining techniques, which may reflect customer's buying patterns. Association rule mining consists of the finding frequent item set mining from which association and correlation rules can be derived.

**Keywords:** Decision Tree, Cosmetic Dataset, FP-Growth Algorithm

---

## 1. INTRODUCTION

A specific class of computerized information system supports business and organizational decision making activities. Decisions are made based on a combination of judgment and knowledge. Successful business organization to be able to react rapidly to the changing demands of the customers, must utilize a sophisticated information technology to transform vast resources of raw data into actionable business plan. Database mining is motivated by the decision support problem faced by most large retail organizations. Progress in bar-code technology has made it possible for retail organizations to collect and store massive amount of sales data, referred to as the basket data. Frequent item set is to determine what items are typically brought together. For example which items customers typically buy together in a database of computerized shopping transactions. There are many frequent item set algorithms they are Apriori, Eclat and FP-growth etc. These presents have the same input and same output, the input is transaction database and a minimum support threshold. The difference between these algorithms is to way in which they generate the output. A minimum support threshold is applied to find all frequent item set in the database. Among these methods, we will search cosmetics frequent pattern using FP-growth algorithm in computerized cosmetics shop. And, then we will find strongly rules from the frequent patterns. Finally, manager is to make decision using strongly rules with decision tree method.

## 2.BACKGROUND THEORY

We will use methods for mining the form of frequent patterns such as "Cosmetic data analysis". Generating association rules from frequent item sets. A pattern-Growth Approach for mining frequent item sets.

## 2.1 Frequent Item set mining methods

Generate association rules from the frequent items. To do consider all partitioning of the item set into rule left-hand and right-hand sides. Confidence of a candidate rule X and Y is calculated as  $\text{support}(XY)/\text{support}(X)$ . All rules that meet the confidence threshold are report as discover of the algorithm.

## 2.2 FP-Growth Algorithm

FP-Growth allows frequent item set discovery without candidate item set generation. Two step approach: Step1: Build a compact data structure called the FP-tree. FP-tree is constructed using two passes over the data-set. Pass1- Scan data and find support for each item. Discard in frequent items. Sort frequent items in decreasing order based on their support. Pass2- Nodes correspond to items and have a counter. 1. FP-growth reads one transaction at a time and maps it to a path. 2. Fixed order is used, so paths can overlap when transactions share items. 3. Pointers are maintained between nodes containing the same item, creating singly linked lists. The more paths that overlap, the higher the compression. FP-tree may fit in memory. 4. Frequent item set extracted from the FP-tree. Step2: Frequent item set Generation. 1. FP-Growth extracts frequent items from the FP-tree. 2. Bottom-up algorithm from the leaves towards the root. 3. Divide and conquer: first look for frequent item sets ending in---then---then---,etc. 4. First, extract prefix path sub-trees ending in an item (set).

**Table-1:** The original table is given below and the support count is 3.

Transaction ID	Item bought
T1	{M,O,N,K,E,Y}
T2	{D,O,N,K,E,Y}
T3	{M,A,K,E}
T4	{M,U,C,K,Y}
T5	{C,O,O,K,I,E}

**Table-2:**Steps to follow:

Pass1:

Item	Number of Transactions
M	3
O	4
N	2
K	5
E	4
Y	3
D	2
A	1
U	1
C	2
I	1

**Table-3:**Sort frequent items in decreasing order based on support count

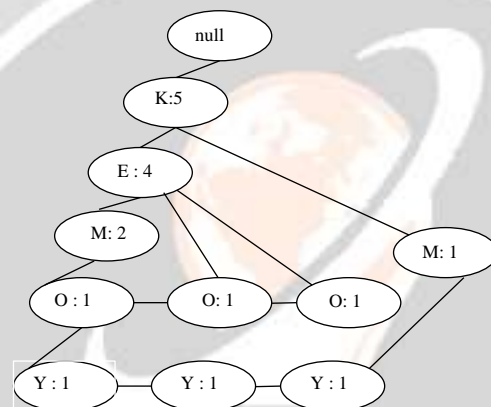
Item	Number of Transactions
K	5
E	4
M	3
O	3
Y	3

**Table-4:**Pass2: We just scan each transaction in DB in L,

Transaction ID	Item sets	Item bought
T1	{M,O,N,K,E,Y}	{K,E,O,M,Y}
T2	{D,O,N,K,E,Y}	{K,E,O,Y}
T3	{M,A,K,E}	{K,E,M}
T4	{M,U,C,K,Y}	{K,M,Y}
T5	{C,O,O,K,I,E}	{K,E,O}

**Table-5:**Construction tree

Item	Number of Transactions
K	5
E	4
M	3
O	3
Y	3



**Fig-1:**Above the FP- tree structure, we find in conditional pattern base in reverse order of frequency.

**Table-6:** Conditional pattern base

Item	Conditional pattern base
Y	{ K , E, M, O : 1 } , { K, E, O : 1 } , { K , M : 1 }
O	{ K , E, M : 1 } , { k , E : 2 }
M	{ k , E : 2 } , { k : 1 }
E	{ k : 4 }
K	null

**Table-7:**Write the conditional FP- tree based on conditional pattern base.

Item	Conditional pattern base
Y	K-3
O	K, E-3
M	K-3
E	{K:4}
K	-

Generate Frequent Pattern from Y O E M.

Y – {K, Y:3}

O – {K, O:3},{E, O:3},{O, E, K:3}

M– {M, K:3}

E – {E, K:4}

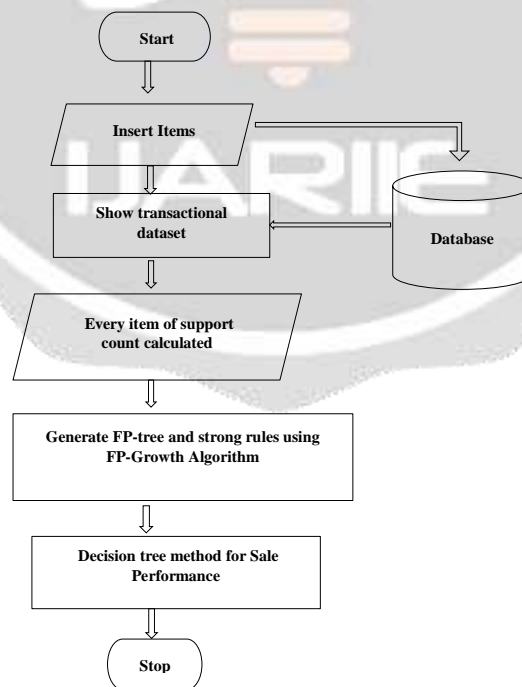
Therefore association rules found from the frequent pattern. There are two association rules ie: either k implies k. If the minimum support is provided in the question then we can find strong rules from association rules.

**2.3 Decision Tree Method**

Decision tree is a technique for representation of data in hierarchical manner. There are two approaches that decision tree follows like top down or bottom up. As name reveals, its purpose is to help in decision making process. There are various applications of decision tree are available such as education, health care, finance etc. and many software packages provide implementation of decision tree algorithms. There are various types of decision tree algorithms are available in the field of data mining but the major drawback of decision trees technique is that if behave differently with distinct types of data sets. We will find sale power of cosmetic product for manager using decision tree method.

**3.OVERVIEW OF THE SYSTM**

The overall system of the FP-Growth algorithm is discussed in this section “Item Name” “Unit Price” and “Description” for new items are entered as input to become a traditional datasets. The overview of the system of the following figure. Customers bought frequent item sets of the cosmetics shop is store in the database of the proposed system as a form of “Sales Vouchers” with Voucher numbers. There are four main parts: show transactional datasets, generate frequent item sets and strongly rule by using FP-Growth algorithm and decision making of the manager by using decision tree method. In general frequent item sets and strong rules by using FP-Growth algorithm process, firstly frequent item sets are collected from transactional database according to minimum support count value. Frequent item sets are sorted in descending order. Frequent pattern tree is constructed and frequent pattern base are created from FP-tree. Conditional frequent pattern base are generated depends on frequent pattern base. Strong rules are generated frequent pattern base with minimum support count value. This system generates frequent item sets by using minimum support value enters as an input to make the decision for manager how many frequent item sets of sale power.



**Fig-2:**Overview of the system design

#### 4.CONCLUSION

We applied the frequent pattern mining based FP-Growth method in the computerized cosmetic shop. For example, if customers have purchased the lotion in cosmetic shop, then how many greatly possibility to purchase the shampoo simultaneously by them. This merchant to decide what kind of commodity will be selected to sell and the arrangement of commodity space, this determination can increase the volume of sales.

#### 5.REFERENCES

- [1] Sumit Garg, Arvind K.Sharma, "Comparative Analysis of Data Mining Techniques on Educational Dataset," International Journal of Computer Applications(0975-8887), Volume74-NO.5, July 2013
- [2] Irfan Khan, Anoop Jain,"A Comprehensive Survey on Sequential Pattern Mining, International Journal of Engineering Research & Technology(IJERT), ISSN:2278-0181, Vol.1 Issue 4, June-2012.
- [3] Ms. Pooja Agrawal, Mr. Suresh kashyap, Mr.Vikas Chandra Pandey, Mr.Suraj Prasad Keshri, "An Analytical Study on Sequential Pattern Mining With Progressive Database," International Journal of Innovative Research in Computer and Communication Engineering, Vol.1 Issue 3, May 2013.
- [4] Aye Myint Myat, Zaw Tun, "Extraction Rules from Educational Data using FP Growth Alogorithm with Correlation,"PSC, (2062) 2010.
- [5] Evasaria M.Sipayng, Cut Fiarni, Randy Tanudjaya "DECISION SUPPORT SYSTEM FOR POTENTIAL SALES AREA OF PRODUCT MARKETING USING CLASSIFICATION AND CLUSTERING METHODS" Proceeding 8<sup>th</sup> International Seminar on Industrial Engineering and Management, ISSN: 1978-774X.
- [6] Ritu Sharma, Mr.Shiv Kumar, Mr.Rohit Maheshwari,"Comparative Analysis of Classification Techniques in Data Mining Using Different Datasets," International Journal of Computer Science and Mobile Computing, Vol.4 Issue.12, December-2015, pg.125-134.
- [7] [glitteryuki@gmail.com](mailto:glitteryuki@gmail.com), [winlei@gmail.cm](mailto:winlei@gmail.cm), "Extraction of Frequent Pattern From Diabetes Cluster," PSC 2010(2068)

