Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 3, June 2021:62- 68

Research Article

Fındıng Mıssed Product and Loss Predıctıon Usıng Market Basket Analysıs (MBA)

A.Gokilavani[1], S.Srimathi[2], Meenalochini.M[3], Dr. P.Megaladevi[4]

Abstract

Market Basket Analysis (MBA), it is the process of modeling technique, if the customer buys the certain group of item, at the next period of purchase they may like to buy the same items. MBA, it has determination and prediction of customer's behavior based on the pattern of previous customer. MBA is applied not only in retail shop. In the existing, Market basket prediction, i.e., according to the customer's purchase, it will provide the shopping list for the next purchase. The same approaches are not used for all the time it may differ based on their decision. The following four approaches are used to identify the individual customer's behavior. They are co-occurrence, sequentiality, periodicity and recurrence. This project defines a Temporal Annotated Recurring Sequence (TARS) to identify the individual purchasing behavior and TARS based predictor to solve the basket prediction problem (i.e., suggestion for next purchase).

MBA based on multidimensional model is used to conduct a study of Market basket analysis, to make a choice of purchasing and sale of stocks in an equity market. In this thesis using MBA for improving methods of arranging products on shelves are identified. The proposed project is mainly focuses on association rule for recommendation system and cosine model for transaction similarity. The best sales of the market are only predicted by using profit and loss. MBA it also provides profit and loss the market.

Keywords: *Market basket analysis (MBA), Temporal Annotated Recurring Sequence (TARS), TARS based prediction.*

---

[1] Associate Professor, Dept. of CSE, Jai shriram Engineering college, Tirupur, Tamilnadu, India.,
[2]PG Scholar, Dept. of CSE, Jai shriram Engineering college, Tirupur, Tamilnadu, India.,
[3]Assistant Professor, Dept. of CSE, Jai shriram Engineering college, Tirupur, Tamilnadu, India.
[4]Professor, Dept. of Management studies, Jai shriram Engineering college, Tirupur, Tamilnadu, India.
[1]gokilavani@jayshriram.edu.in ,[2]srimathi257@gmail.com ,[3] meenalochini@jayshriram.edu.in ,
[4] deanjsrsms@gmail.com
Received Accepted

Introduction

Now days, detecting purchase habits and its evolution in time are a crucial challenge for the marketing policies. In this context, retail market offer their customers a most promising facilities is basket prediction i.e., automated forecasting of the next basket based on customer purchase. The solution of the basket prediction it must be adapt to the evolution of customers behavior, recurrence of their purchase and their periodic changes. It contains the Temporal Annotated Recurring Sequence (TARS) it helps to identify the individual's purchasing behavior by using four characteristics they are: co-occurrence, sequentiality, and periodicity, recurrence. In this process, future basket can identified based on the basket already they purchased. But the customer habits may affect both endogenous and personal factors [2]. In this periodicity is a crucial characteristic for basket prediction. Then exploit the TARS to build TARS based prediction (TBP) it solves basket prediction problem. It provides the recommendation and the list of items has suggested for next purchase. TBP it predicts up to next 20 baskets and the quality of TARS is perfect only after 36 weeks of purchase. Both TARS and TBP are user-centric approaches, it provides only individual data of customer to predict their next basket [7][8][9].

Related Work

TARS extraction procedure:

TARS models two aspects: First is customer's recurrence and its sequentially of the customers purchases, it is used to identify the items that is purchased together and second is recurrence it means sequential purchase, i.e., how many such list of purchase occurred in previous purchase . To get perfect TARS here it uses FP-Growth algorithm, it pursue the result with the help of purchase history [36].FP-Growth produces results that are easily interpretable, and it builds an FP-Tree structure. In this algorithm, each node and branch is represented as item and different association. FP-Growth it attached additional information to FP-Tree.

In the TARS Extraction procedure two steps are used they are Data-Driven parameters estimation and sequence filtering. The first processor have a parameters as ^{max}, q^{min}, p^{min} . These parameters are not only for the individual customer's [20]. The data-driven contains the set of the parameters { s^{max} },{qs^{min}},{ps^{min}}. Similarly, obtain {ps^{min}} by following two ways (i) identify the sum of the number of occurrences of base sequence

Group the base sequence with similar items.

The similarity also can be identify by using cosine similarity

$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}},$$

The second is the sequence filtering; it selects the base recurring sequences. This filtering is used to reduce the search space it builds the TARS-tree and the TARS extraction. It removes too many TARS which are used in prediction and it considers too many valid base sequences and not prunes enough the search space.

TARS Based Prediction:

TARS Based Prediction it is markedly personalized and user-centric [7] the prediction is performed only based on the customer's purchase history. TARS simultaneously embed complex items for interactions such as co-occurrence, sequential relationship, periodicity, re-purchase. These factors are used to measure the customer's recent purchase history and active patterns. This active pattern is used to provide the items that are needed for the next purchase.

At first, it sorts the purchase history using recent basket to the oldest one, then it loops the pair of consecutive baskets. When it finds a potentially active TARS then it consider two cases. If the sequence is encounter for first time then the algorithm adds to the set of active TARS.

It identifies the next basket of purchase for the customer based on the previous purchase history.

Existing Method:

In the existing system, TARS AND TBP both the process is used for the customer's purchase method. TARS based method is used to identify the individual purchase behavior by using following characteristic. They are co-occurrence: a customer's systematically set of items purchased together. Sequentiality: a customer systematically purchases a set of items one another. Periodicity: it is the sequential purchase in specific periods of the year. Because, it may change environment and personal reason of customer.

Recurrency: sequential purchase during each period. These factors are used to identify the individual behavior of the customer.

If a customer typically purchases basket with few items it is useless to predict basket with large items or if the customer purchase basket with the large items then the prediction of small basket if not covered with mostly purchased items. Predictions is measured by using fixed length each and every customer have the fixed size as $k=k^*_c$,k_c^* indicates the aerage basket length of customer c.

For each customer, the precision and recall is measured for the prediction process:

Precision $(b,b^*) = |b \cap b^*| / |b^*|$

Recall$(b,b^*) = |b \cap b^*| / |b^*|$

Hit ratio, it is the ratio of the customers who received atleast one correct prediction.

Hit – Ratio$(b,b^*) = 1$

if $b \cap b^* \neq 0$ otherwise

First, it have to deal with the retail transaction then the customer should provide the rating for the each and every purchased item. Second, individual purchase frequency because, the user recommendation the item which is most frequently purchased rather that item are easily forgettable. Ranking measures are used for the items are more useful when appearing earlier in result list.

The following approaches build the predictive model of the customer purchase data. LST: the next basket prediction is the last basket purchased by customer $b_{t\,n+1}=b_{t\,n}$. TOP: it predicts top-k most frequent items. MC: This prediction is based on the last purchase $b_{t\,n}$ and Markov chain. CLF: each item purchased by customer builds a classifier on temporal extracted from customer's purchase history it has two classes: item i purchased yes/no. This is used to predict next basket using TARS based prediction based on purchase history.

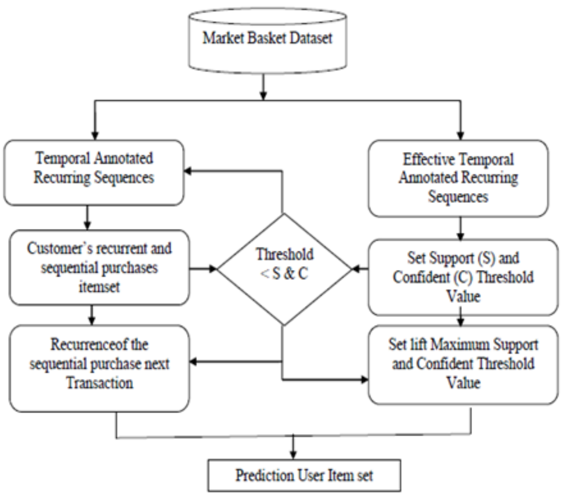
Here, it implements four state-of-the-art methods that are not used in user- centric, i.e., they require the purchase data of all the customers

NMF (Non-negativity matrix factorization), it is a collaboration filtering method which is non negativity matrix it provides purchase history of all customers. Then FMC (Factorizing personalized makov chain), HRM(Hierarchical Representation Model)these also used for predict purchase behavior.

Finally market basket prediction, when setting the length it predicts basket equals to average basket length for each prediction of each individual customer, $k=k_c^*$.

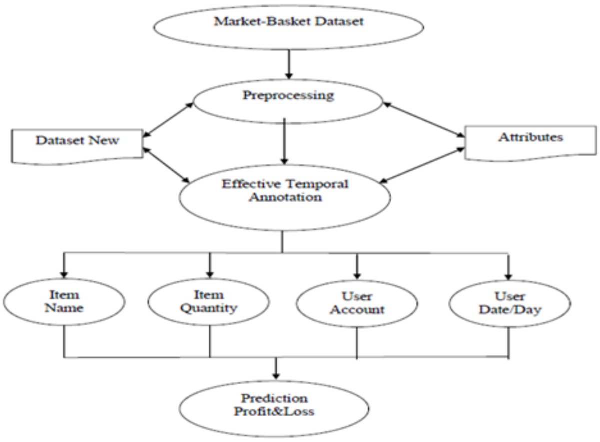
The positive gap between TBP and the competitors increases : for the customers which TBP it correctly predicted by TBP are more accurate and it covers large items then the baskets predicted by other methods.

Proposed System:



This is the architecture for the market basket analysis (MBA), it has two process that is Temporal Annotated Recurring Sequences (TARS) and effective Temporal Annotated Recurring Sequences (Effective TARS). TARS it is used to identify the customer's individual purchase behavior and sequentially purchased item , here it also has the rating for each product they purchased then the step is it also provides the recommendation for the next basket list based on previous history.

The other side is effective TARS that is used the identify the support and confidence value for each and every product. The support and confidence of the product is used to analyze the customer purchase for the next basket.



Threshold, it compares with support and confidence with sequentially purchased item to predict for next basket. It will predict the user's item set. The above diagram is used to predict loss and profit by using user purchase history.

Drawback of Existing System:

- Only one to one collaboration it filters and recommends are based on similarity of purchase history.
- Common groups of rules by using cauterization methods and filtering of non-conforming parts of data.
- Do not find missing data of customer Transaction data
- Do not predicated missing item with loss of profit by customer purchase transaction dataset.

Advantage of Proposed System:

- The quality of TBP's predictions stabilizes only after 36 weeks of purchase.
- Find missing date of customer Transaction data.
- Predicated missing item with loss of profit by customer purchase transaction dataset.

Conclusion:

In this work, data driven, user-centric method is used for market basket prediction method. Here it also defines the Temporal Annotated Recurring Sequence and TARS based prediction for basket prediction. This two approaches are mainly used for market basket prediction. TARS used four characteristic to identify the individual purchase habit and TBP is used to provide the suggestion for the next basket it based on the previous purchase history. It also has the recommendation system as FP- Growth method and cosine similarity for similarity measurement. This method is possible only after the 36 weeks of the individual purchase history. This market basket analysis is used for retail shops for the purpose of customer's satisfaction and the profit of the shop. The proposed as it provides the profit and loss for shop keeper based on the customer purchase and their ratings.

References

[1] B. Mittal and W. M. Lassar, "The role of personalization in serviceencounters," J. Retailing, vol. 72, no. 1, pp. 95–109, 1996.

[2] A. Dagher, "Shopping centers in the brain," Neuron, vol. 53, no. 1,pp. 7–8, 2007.

[3] B. Knutson, et al., "Neural predictors of purchases," Neuron,vol. 53, no. 1, pp. 147–156, 2007.

[4] R. Guidotti, M. Coscia, et al., "Behavioral entropy and profitability in retail," in Proc. IEEE Int. Conf. Data Sci. Adv. Analytics, 2015,pp. 1–10.

[5] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, et al., "A survey ofmethods for explaining black box models," ACM Comput. Surv.,vol. 51, no. 5, 2018, Art. no. 93.

[6] R. Guidotti, J. Soldani, et al., "Explaining successful dockerimages using pattern mining analysis," in Software Technologies:Applications and Foundations, Toulouse, France: Springer, 2018.

[7] A. Pentland, et al., "Personal data: The emergence of a new assetclass," in An Initiative of the World Economic Forum, 2011.

[8] C. Kalapesi, "Unlocking the value of personal data: From collectionto usage," in World Economic Forum Technical Report, 2013.

[9] R. Guidotti, "Personal data analytics: Capturing human behaviorto improve self-awareness and personal services through individualand collective knowledge," PhD Thesis, 2017.

[10] R. Guidotti, A. Monreale, M. Nanni, F. Giannotti, and D. Pedreschi,"Clustering

individual transactional data for masses of users,” inProc. 23rd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining,2017, pp. 195–204.

[11] Y.-A. de Montjoye, E. Shmueli, S. S. Wang, and A. S. Pentland,“openPDS: Protecting the privacy of metadata through safeanswers,”PloS One, vol. 9, no. 7, 2014, Art. no. e98790.

[12] M. Vescovi, C. Perentis, C. Leonardi, B. Lepri, and C. Moiso, “Mydata store: Toward user awareness and control on personal data,”in Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.: AdjunctPublication, 2014, pp. 179–182.

[13] R. Guidotti, G. Rossetti, L. Pappalardo, et al., “Market basket predictionusing user-centric temporal annotated recurring sequences,” inProc. IEEE Int. Conf. Data Mining, 2017, pp. 895–900.

[14] A. Mild and T. Reutterer, “An improved collaborative filteringapproach for predicting cross-category purchases based on binarymarket basket data,” J. Retailing Consum. Serv., vol. 10, no. 3,pp. 123–133, 2003.

[15] K. Christidis, et al., “Exploring customer preferences with probabilistictopic models,” in Proc. Eur. Conf. Mach. Learn. Knowl. Discoveryin Databases, 2010, pp. 1–13.

[16] G. Linden, et al.,“Amazon. com recommendations: Item-to-itemcollaborative filtering,” IEEE Internet Comput., vol. 7, no. 1, pp. 76–80, Jan./Feb. 2003.

[17] Y. Hu, Y. Koren, and C. Volinsky, “Collaborative filtering forimplicit feedback datasets,” in Proc. 8th IEEE Int. Conf. Data Mining,2008, pp. 263–272.

[18] X. Su and T. M. Khoshgoftaar, “A survey of collaborative filteringtechniques,” Advances Artif. Intell., vol. 2009, 2009, Art. no. 4.

[19] G. Adomavicius and A. Tuzhilin, “Toward the next generation of recommender systems: A survey of the state-of-the-art and possibleextensions,” IEEE Trans. Knowl. Data Eng., vol. 17, no. 6,pp. 734–749, Jun. 2005.

[20] Meenalochini.M and A.Gokilavani “Enhanced trust based secure routing for manet”,Journal of huazhong university of science and technology, vol 50, no.3, 2021.

[21] C. Chand, A. Thakkar, and A. Ganatra, “Sequential pattern mining:Survey and current research challenges,” Int. J. Soft Comput.Eng., vol. 2, no. 1, pp. 185–193, 2012.