Customer Segmentation and Personalized Marketing Using K-Means and APRIORI Algorithm

Gayathri K, Arunodhaya R

{*kgi.mca@psgtech.ac.in, arunaravi0228@gmail.com}

Assistant Professor, Department of Computer Applications, PSG College of Technology, Tamilnadu, Coimbatore-641004, India[1],Student, Department of Computer Applications, PSG College of Technology, Tamilnadu, Coimbatore-641004, India[2]

Abstract- In today's business environment regardless of what type of industry we are in or what kinds of products and services that is sold, customers are the most important part of a business. Without the customer, sales doesn't happen. If the customers' views are not taken into an account, then it's likely the campaigns will not be successful. Hence classifying the right customers also matters a lot to make the products to be bought frequently. Companies follow different strategies to segment the customers. In this paper, RFM and K-means clustering is used to segment the customers. It also provides a combo offer recommendation feature which can be implemented in any commercial websites using ECLAT and Apriori algorithm. This helps in analyzing the performance of the products and also about the customers whom can be focused more for selling the products.

Keywords- Segmentation, Marketing, Offer recommendation, RFM, clustering, K-means, Apriori, ECLAT

1 Introduction

Customer Segmentation [4] and Personalized Marketing [13], provides a solution for marketing the available products to particular group of customers and also display the possible offers which can be provided to them that will yield profit to the company, retain their customers and also increase the sales. This is done by segmenting the customers into group of individuals who are similar in terms of gender, spending behavior, age and demography. Hence through achieving customer segmentation companies would be able to target the customers who has the specific needs and which helps in marketing the products to right group/customers. Customer segmentation is done by using RFM technique and K-means Clustering[8]. While implementing in the real- time application, The manager or admin of the company can log in the website and type the product name available in the store and as a result they are provided with a list of customers and also an offer recommendation which consist of related products that particular customer can buy. This in turn increase the sale of the product and the companies' profit.

Customer segmentation is the practice of dividing the customer into groups of individuals that are similar in specific ways relevant to marketing such as age, gender, interest and spending habits. There are different types of customer segmentation. They are:

1) Geographic- based on a customer's location
2) Demographics- based on age, income, occupation and family size
3) Behavioral- based on the purchasing habits.
4) Psychological- based on customer's beliefs and values.

The data used for this classification depends upon the business decision. By differentiating their customer base, businesses can better target individuals and maximize sales, sell their products appropriately

ICCAP 2021, December 07-08, Chennai, India

DOI 10.4108/eai.7-12-2021.2314561

and provide more tailored shopping experiences. There are many benefits of Customer segmentation. They are:

1) Target the customer: By identifying the right customers for the product and selling to them by marketing to similar group of customers who would buy the products instead of spending in marketing to irrelevant groups.

2) Increase sales: By identifying similar groups and selling products to them gradually increase the sales for the company. This will help to have a long-term revenue from the customer.

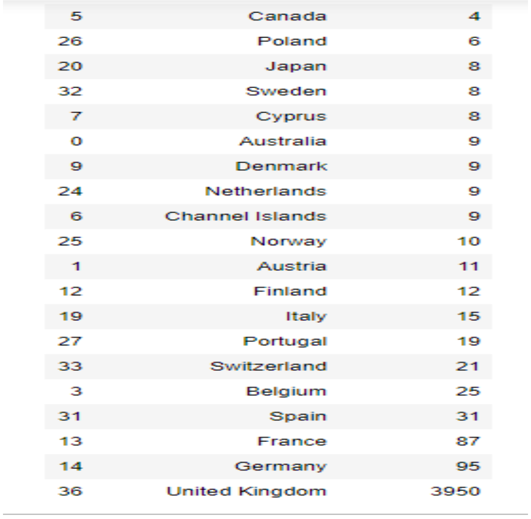
3) Improve customer satisfaction and retention of the customers: By identifying and recommending the associated products bought by each group will result in customer segmentation as their needs where satisfied continuously so it helps to retain the customers.

4) Decrease the marketing cost: By classifying the customers into groups, the products could be marketed to specific group alone who has the probability of buying the products.

Personalized Marketing is a marketing technique used for one-to-one marketing or individual marketing with digital technology to deliver individualized messages and product offering to current or prospective customers. Personalized combo offer recommendation is used for personalized marketing. It is a recommendation of an associated product that a user would like to buy along with a product that they tend to buy usually. This is taken from the past purchase history of a customer and use appropriate algorithm to predict its association product and recommend as a combo offer to that particular customer.

2 Data Preprocessing

The dataset is on the details of online retail store taken from Kaggle. The data describes about the purchase made by the customers from different countries all over the word. The attributes in the dataset are: Invoice no., Stock code, Description, Quantity, Invoice data, Unit price, Customer ID and Country name. To see the customer distribution country wise, the customers are grouped according to the countries.



5	Canada	4
26	Poland	6
20	Japan	8
32	Sweden	8
7	Cyprus	8
0	Australia	9
9	Denmark	9
24	Netherlands	9
6	Channel Islands	9
25	Norway	10
1	Austria	11
12	Finland	12
19	Italy	15
27	Portugal	19
33	Switzerland	21
3	Belgium	25
31	Spain	31
13	France	87
14	Germany	95
36	United Kingdom	3950

Fig. 1 Count of customers country wise

From the fig. 1, it is clear that the majority of the customers are from United Kingdom. Hence the customer segmentation can be performed for United Kingdom. A data mining technique that involves transformation of raw data into an understandable format is called data pre-processing. Real world data is likely to cause false conclusions, since, it may contain incomplete, inconsistent and/ or may lack in certain behavior and trends. Preparation of raw data for further processing is done by data pre-processing. If data pre-processing is not done, the results will be misleading. The process of connecting and removing corrupt and inaccurate data is done by a process called data cleaning, also known as data cleansing. This process includes smoothing the noisy data, filling the missing values or resolving the inconsistencies in the data. The data cleaning is done by removing the missing values and negative values in the dataset.

3 Methodology

In this section we describe the two techniques used for customer segmentation.

A. Recency Frequency Monetary Analysis

It is a marketing technique which determines the best customers by analyzing the recency, frequency and monetary values called Recency Frequency Monetary (RFM) analysis.

- Recency : How much time has elapsed since a customer's last activity or transaction
- Frequency : How often has a customer made transaction
- Monetary : How much a customer has spent for the product.

Customer segmentation [9] can be done with the help of RFM analysis by assigning RFM scores to each customer. RFM factors illustrate these facts:

1) The more recent the purchase, the more responsive the customer is to promotions.
2) The more frequently the customer buys, the more engaged and satisfied they are.
3) Monetary value differentiates heavy spenders from low- value purchasers.

The benefits of RFM analysis are to determine the high valued customers of the business, to target the customers who has the high chance for buying the products and to know the one- time customers of the business. There are three major groups of customers determined by RFM analysis. They are:

I. High RFM Customers:

They are the customers with high monetary, frequency and recency values. They add value to the business more and they are the promising customers where the company can market their products more.

II. Medium RFM customers:

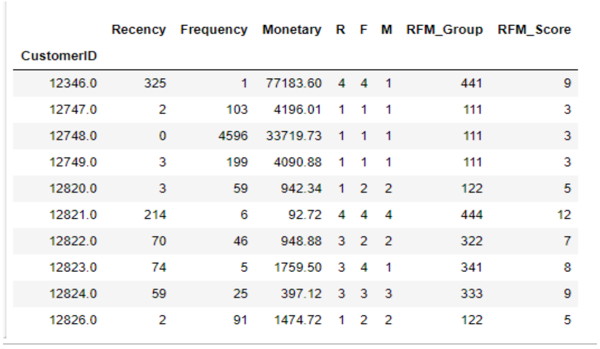
They are the customers with high recency, low frequency and monetary values. These groups of customers can be made to high RFM scored customers by using some marketing techniques.

III. Low RFM Customers:

They are the group of customers who are least engaged. They have low recency, frequency and monetary values. They could be brought to medium RFM level by providing new offers to them.

Calculation of RFM scores

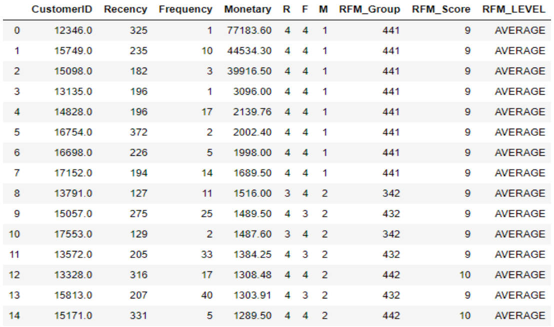
The details of each customer like Customer ID, Quantity and Unit price would be required to calculate the RFM scores. The day since the last purchase is used to calculate the recency, the total number of transactions is used for findling frequency and total money each customer has spent is calculated for monetary value. The next step is to create quartile such as 0.25, 0.50, 0.75. So that we can sub divide the set into 4 groups based on R, F and M values. In order to create segments with values 1, 2, 3 and 4, Rscoring and FM scoring is created. In Rscoring function, we assign value 1 for the lowest value of recency, because lower the value indicated the most recently visited. In FMscoring function, we assign value 1 for the highest value of frequency and monetary, because higher the value of frequency and monetary are tend to be promising customers. After applying both the functions, segmented RFM values are obtained. The next step is to calculate and add RFM-Group value column showing the combined concatenated score of RFM and also RFM-score value column to show the total sum of RFM_Group values. Fig. 2 illustrates the RFM-Group and RFM- score.



	Recency	Frequency	Monetary	R	F	M	RFM_Group	RFM_Score
CustomerID								
12346.0	325	1	77183.60	4	4	1	441	9
12747.0	2	103	4196.01	1	1	1	111	3
12748.0	0	4596	33719.73	1	1	1	111	3
12749.0	3	199	4090.88	1	1	1	111	3
12820.0	3	59	942.34	1	2	2	122	5
12821.0	214	6	92.72	4	4	4	444	12
12822.0	70	46	948.88	3	2	2	322	7
12823.0	74	5	1759.50	3	4	1	341	8
12824.0	59	25	397.12	3	3	3	333	9
12826.0	2	91	1474.72	1	2	2	122	5

Fig. 2 RFM- Group and RFM-Score

The next step is to assign loyalty levels to the customers such as Very-High, High, Average and Low. After assigning, filtering can be done for each group and it can be stored separately for retrieval purpose. Fig. 3 shows the filtering of Average valued customers.



	CustomerID	Recency	Frequency	Monetary	R	F	M	RFM_Group	RFM_Score	RFM_LEVEL
0	12346.0	325	1	77183.60	4	4	1	441	9	AVERAGE
1	15749.0	235	10	44534.30	4	4	1	441	9	AVERAGE
2	15098.0	182	3	39916.50	4	4	1	441	9	AVERAGE
3	13135.0	196	1	3096.00	4	4	1	441	9	AVERAGE
4	14828.0	196	17	2139.76	4	4	1	441	9	AVERAGE
5	16754.0	372	2	2002.40	4	4	1	441	9	AVERAGE
6	16698.0	226	5	1998.00	4	4	1	441	9	AVERAGE
7	17152.0	194	14	1689.50	4	4	1	441	9	AVERAGE
8	13791.0	127	11	1516.00	3	4	2	342	9	AVERAGE
9	15057.0	275	25	1489.50	4	3	2	432	9	AVERAGE
10	17553.0	129	2	1487.60	3	4	2	342	9	AVERAGE
11	13572.0	205	33	1384.25	4	3	2	432	9	AVERAGE
12	13328.0	316	17	1308.48	4	4	2	442	10	AVERAGE
13	15813.0	207	40	1303.91	4	3	2	432	9	AVERAGE
14	15171.0	331	5	1289.50	4	4	2	442	10	AVERAGE

Fig. 3 Filtering of average valued customers

B. K-Means Clustering

K-Means clustering is an unsupervised learning technique used when the data is unlabeled in nature.

It finds the groups based on the patterns present in the data and K is the variable that represents the number of clusters/ groups. Each data point is assigned iteratively to one of the K groups. This division is based on the features that are provided. Clustering happens based on feature similarity.

Algorithm

Step 1: Initialize k points randomly and these points are called means.

Step2: Categorize each item to the closest mean by using Euclidean Distance measure.

Step 3: Update the mean's co-ordinates by calculating the average of the items categorized in the product menu.

Step 4: Repeat the steps for some iterations till the desired number of clusters are formed.

The number of clusters cannot be formed randomly. There are two main methods to determine the optimal number of clusters. They are

I. The Elbow Method

A naïve and commonly used method to determine the number of clusters is the Elbow method. The elbow method runs k-means clustering on the dataset for a range of values for k and then for each value of k computes an average score for all clusters. By default, the score is computed, the sum of square distances from each point to its assigned center. Fig. 4 is a graph depicting the optimal number of clusters that can be formed using elbow method. Here the elbow is at k=3, hence 3 is chosen as the number of clusters.

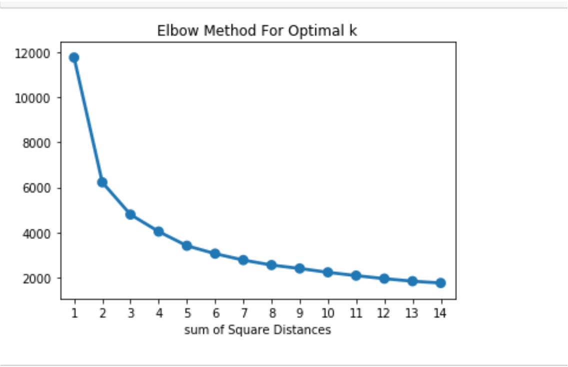


Fig.4 Elbow graph determining the number of clusters

II. The Silhouette Method

The quality of the cluster is measured by the silhouette method. A high value of the silhouette score indicates good clustering. It computes the silhouette score for all observations for different K values.

Normalization:

Normalization is used to scale the data of an attribute so that it falls in a smaller range, it is applied when the dataset has more deviation in it. Normalization is done using log transformation. Fig. 5, 6 and 7 is a graph depicting the recency, frequency and monetary values respectively after normalization.

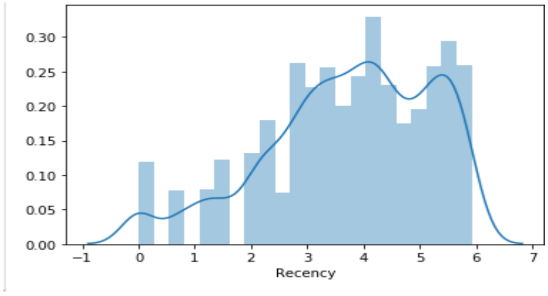


Fig. 5 Recency values after normalization

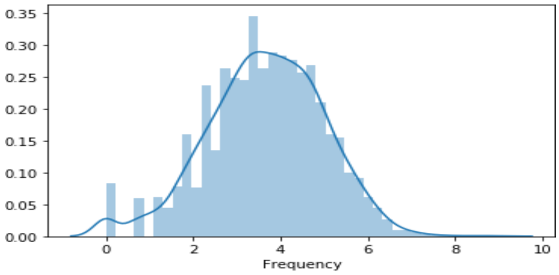
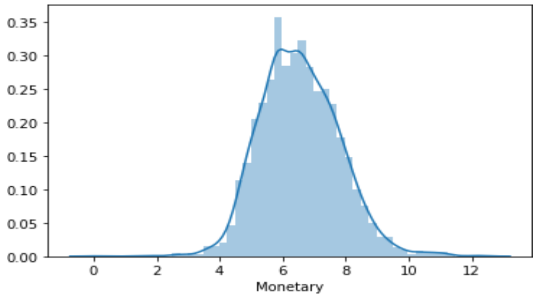


Fig.6 Frequency values after normalization



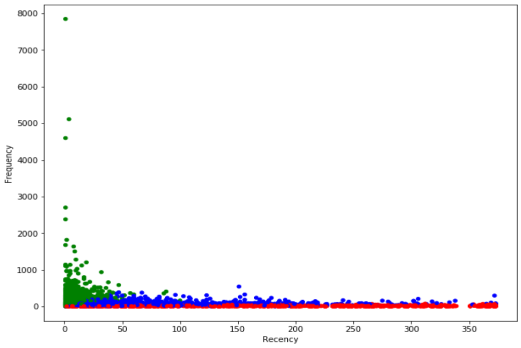


Fig. 7 Monetary values after normalization

Fig. 8 Plotted graph of the clusters

After normalization of the data, the k-means clustering is performed by having k=3 as the number of clusters. There are three clusters formed high, average and low with the color indications as green blue and red respectively. The clusters are formed after k-means clustering is performed. Fig. 8 is a plotted graph of the clusters.

4 Implementation

The benefits of personalized combo offer recommendation are: increased sales, retain customers, reduce marketing cost and customer satisfaction. Market basket analysis is the technique used for personalized marketing. It is a data mining technique used by retailers to increase sales by better understanding customer purchasing patterns, it is basically a filtering system that helps to predict and show the items that a user would like to purchase along with the product already purchased. It involves analyzing large data sets, such as purchase history, to reveal product groupings, as well as products that are likely to be purchased together. It works by looking for combinations of items that occur together frequently in transactions. This is used for combo offer recommendation. Including product combo offer recommendations as part of the user experience can increase the average order value. A timely product recommendation can lead shoppers to choose one product over another. There are two types of market basket analysis and they are:

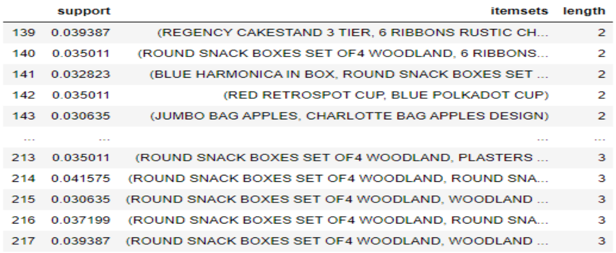
I. Predictive market basket analysis: Items purchased are considered to determine cross-sell (suggesting similar or complementary products or services).

II. Differential market basket analysis: Considers data across different stores, as well as purchases from different customer groups during different times of the day, month or year. These insights can lead to new product offers that drive higher sales.

In market basket analysis, association rules are used to predict the likelihood of products being purchased together. Association rules[14] count the frequency of items that occur together, seeking to find associations that occur far more than expected. Algorithms that use association rules include AIS, SETM, FP growth, ECLAT, Apriori etc. The Apriori algorithm is commonly used algorithm and is used to identify the frequent items in the database, then evaluate their frequency as the datasets are expanded to larger sizes. In market basket analysis Apriori, FP growth and ECLAT algorithms were the most commonly used algorithms to find the associated products. Among them Apriori and FP growth follows the Breadth- First search pattern to find the frequent item sets, whereas ECLAT algorithm uses Depth- First search pattern to find the frequent item sets. Hence Apriori algorithm is compared with ECLAT algorithm and the better can be chosen.

A. Market Basket Analysis using ECLAT Algorithm

The ECLAT algorithm stands for Equivalence Class Clustering and bottom-up Lattice Traversal. It is an algorithm for finding frequent itemsets in a transaction or database. It is more efficient and scalable version of apriori algorithm. ECLAT algorithm works in a vertical manner like Depth-First Search of a graph. According to the ECLAT algorithm, the first step is to transform the data into vertical format, so that it could execute in a depth-first search manner. This transformation is to faster the execution speed. Next is to calculate the support of each item with respect to invoice, since ECLAT algorithm uses only support as its metrics to predict the frequent item sets. The calculation of support is done for each item with minimum support 0.03 and fig. 9 illustrates the filtering of frequent item sets which has more than 2 items occurring frequently in the transactions with minimum support of 0.03.

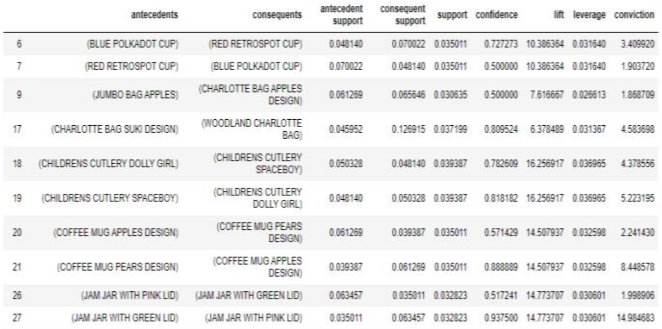


	support	itemsets	length
139	0.039387	(REGENCY CAKESTAND 3 TIER, 6 RIBBONS RUSTIC CH...	2
140	0.035011	(ROUND SNACK BOXES SET OF4 WOODLAND, 6 RIBBONS...	2
141	0.032823	(BLUE HARMONICA IN BOX, ROUND SNACK BOXES SET ...	2
142	0.035011	(RED RETROSPOT CUP, BLUE POLKADOT CUP)	2
143	0.030635	(JUMBO BAG APPLES, CHARLOTTE BAG APPLES DESIGN)	2
...	...	...	...
213	0.035011	(ROUND SNACK BOXES SET OF4 WOODLAND, PLASTERS ...	3
214	0.041575	(ROUND SNACK BOXES SET OF4 WOODLAND, ROUND SNA...	3
215	0.030635	(ROUND SNACK BOXES SET OF4 WOODLAND, WOODLAND ...	3
216	0.037199	(ROUND SNACK BOXES SET OF4 WOODLAND, ROUND SNA...	3
217	0.039387	(ROUND SNACK BOXES SET OF4 WOODLAND, WOODLAND ...	3

Fig. 9 Filtering of frequent item set

B. Market Basket Analysis using Association Rule Mining and Apriori Algorithm

The data pre-processing is performed for removing spaces from the beginning and end of the product description, removing duplicate invoices, converting invoice number to string data type and remove credit transactions. On performing market basket analysis country wise, Germany is taken for doing market basket analysis. The transactions of Germany is taken and grouped by invoice number and description and it takes the sum of the quantity, as a result it gives a basket of transaction. It has the invoice number as key and products as column. If the value is 0, then the product is not present in that invoice. If the value is greater than 0, then it is present in the transaction that many number of times. Now training the model for the basket, by applying apriori algorithm for the basket with minimum support as 0.03 and with metric as lift for predicting the associated products. The rules are generated using association rule. Finally setting up a minimum threshold value for lift and confidence for filtering which is used for offer recommendation that has a good lift and confident values. Fig.10 illustrates the final filtering for offer recommendation.

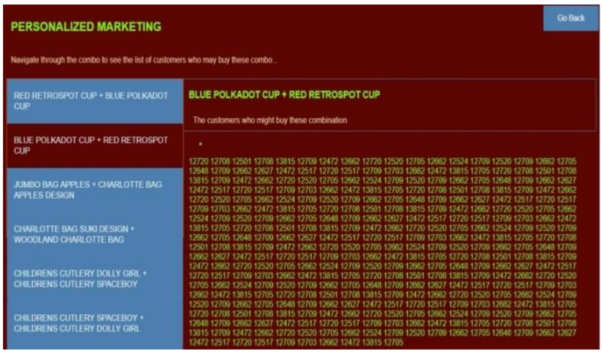


	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
6	(BLUE POLKADOT CUP)	(RED RETROSPOT CUP)	0.048140	0.070022	0.035011	0.727273	10.386364	0.031640	3.409920
7	(RED RETROSPOT CUP)	(BLUE POLKADOT CUP)	0.070022	0.048140	0.035011	0.500000	10.386364	0.031640	1.903720
9	(JUMBO BAG APPLES)	(CHARLOTTE BAG APPLES DESIGN)	0.061269	0.065646	0.030635	0.500000	7.616667	0.026613	1.868709
17	(CHARLOTTE BAG SUKI DESIGN)	(WOODLAND CHARLOTTE BAG)	0.045952	0.126915	0.037199	0.809524	6.378489	0.031367	4.583698
18	(CHILDRENS CUTLERY DOLLY GIRL)	(CHILDRENS CUTLERY SPACEBOY)	0.050328	0.048140	0.039387	0.782609	16.256917	0.036965	4.378556
19	(CHILDRENS CUTLERY SPACEBOY)	(CHILDRENS CUTLERY DOLLY GIRL)	0.048140	0.050328	0.039387	0.818182	16.256917	0.036965	5.223195
20	(COFFEE MUG APPLES DESIGN)	(COFFEE MUG PEARS DESIGN)	0.061269	0.039387	0.035011	0.571429	14.507937	0.032598	2.241430
21	(COFFEE MUG PEARS DESIGN)	(COFFEE MUG APPLES DESIGN)	0.039387	0.061269	0.035011	0.888889	14.507937	0.032598	8.448578
26	(JAM JAR WITH PINK LID)	(JAM JAR WITH GREEN LID)	0.063457	0.035011	0.032823	0.517241	14.773707	0.030601	1.998906
27	(JAM JAR WITH GREEN LID)	(JAM JAR WITH PINK LID)	0.035011	0.063457	0.032823	0.937500	14.773707	0.030601	14.984683

Fig. 10 Final filtering for offer recommendation

5 Predictions And Result

The ECLAT algorithm gives 72 combo recommendation and the Apriori algorithm gives 31 combo recommendation. Even though the ECLAT algorithm gives more combo recommendation, it uses only support as its metrics to find the associated products, sometimes it will not be accurate. ECLAT algorithm is best suited for small and medium dataset, whereas Apriori is suited for large dataset. ECLAT algorithm scans the currently generated dataset, whereas the Apriori scans the original dataset. Hence, the better choice could be apriori algorithm, since it calculates support, confidence and lift as its metrics to predict the associated products and it handles large dataset which is fixed dataset. When implemented in the real world, fig. 11 shows the result where personalized marketing is performed for each combo. It lists down allthe customer ID who can buy these combo



PERSONALIZED MARKETING

Go Back

Navigate through the combo to see the list of customers who may buy these combo...

- RED RETROSPOT CUP + BLUE POLKADOT CUP
- BLUE POLKADOT CUP + RED RETROSPOT CUP
- JUMBO BAG APPLES + CHARLOTTE BAG APPLES DESIGN
- CHARLOTTE BAG SUKI DESIGN + WOODLAND CHARLOTTE BAG
- CHILDRENS CUTLERY DOLLY GIRL + CHILDRENS CUTLERY SPACEBOY
- CHILDRENS CUTLERY SPACEBOY + CHILDRENS CUTLERY DOLLY GIRL

BLUE POLKADOT CUP + RED RETROSPOT CUP

The customers who might buy these combination

```
•
12720 12708 12501 12708 13815 12709 12472 12662 12720 12520 12705 12662 12524 12709 12520 12709 12662 12705
12648 12709 12662 12627 12472 12517 12720 12517 12709 12703 12662 12472 13815 12705 12720 12708 12501 12708
13815 12709 12472 12662 12720 12520 12705 12662 12524 12709 12520 12709 12662 12705 12648 12709 12662 12627
12472 12517 12720 12517 12709 12703 12662 12472 13815 12705 12720 12708 12501 12708 13815 12709 12472 12662
12720 12520 12705 12662 12524 12709 12520 12709 12662 12705 12648 12709 12662 12627 12472 12517 12720 12517
12709 12703 12662 12472 13815 12705 12720 12708 12501 12708 13815 12709 12472 12662 12720 12520 12705 12662
12524 12709 12520 12709 12662 12705 12648 12709 12662 12627 12472 12517 12720 12517 12709 12703 12662 12472
13815 12705 12720 12708 12501 12708 13815 12709 12472 12662 12720 12520 12705 12662 12524 12709 12520 12709
12662 12705 12648 12709 12662 12627 12472 12517 12720 12517 12709 12703 12662 12472 13815 12705 12720 12708
12501 12708 13815 12709 12472 12662 12720 12520 12705 12662 12524 12709 12520 12709 12662 12705 12648 12709
12662 12627 12472 12517 12720 12517 12709 12703 12662 12472 13815 12705 12720 12708 12501 12708 13815 12709
12472 12662 12720 12520 12705 12662 12524 12709 12520 12709 12662 12705 12648 12709 12662 12627 12472 12517
12720 12517 12709 12703 12662 12472 13815 12705 12720 12708 12501 12708 13815 12709 12472 12662 12720 12520
12705 12662 12524 12709 12520 12709 12662 12705 12648 12709 12662 12627 12472 12517 12720 12517 12709 12703
12662 12472 13815 12705 12720 12708 12501 12708 13815 12709 12472 12662 12720 12520 12705 12662 12524 12709
12520 12709 12662 12705 12648 12709 12662 12627 12472 12517 12720 12517 12709 12703 12662 12472 13815 12705
12720 12708 12501 12708 13815 12709 12472 12662 12720 12520 12705 12662 12524 12709 12520 12709 12662 12705
12648 12709 12662 12627 12472 12517 12720 12517 12709 12703 12662 12472 13815 12705 12720 12708 12501 12708
13815 12709 12472 12662 12720 12520 12705 12662 12524 12709 12520 12709 12662 12705 12648 12709 12662 12627
12472 12517 12720 12517 12709 12703 12662 12472 13815 12705
```

Fig. 11 list of customers for each combo offer

.

CONCLUSION

Customer segmentation and personalized marketing segments or groups the customers and provide services for them according to their need, since product and services needs of individual customers differs and also would help the organization to increase sales and retain customers. The techniques used are K-means clustering which is used to segment the customers and market basket analysis which is used to find the associated products to give a combo offer recommendations are sufficient to meet the solution for the company’s marketing problem by performing a targeted marketing. The future enhancement could be done by implementing with the dynamic dataset provided by the organization. It would add the real-time data and includes the dynamic recommendation in the website.

REFERENCES

[1] P. P. Pramono, I. Surjandari and E. Laoh, "Estimating Customer Segmentation based on Customer Lifetime Value Using Two-Stage Clustering Method," 2019 16th International Conference on Service Systems and Service Management (ICSSSM), 2019, pp. 1-5, doi: 10.1109/ICSSSM.2019.8887704.

[2] B. G. Muchardie, A. Gunawan and B. Aditya, "E-Commerce Market Segmentation Based On The Antecedents Of Customer Satisfaction and Customer Retention," 2019 International Conference on Information Management and Technology (ICIMTech), 2019, pp. 103-108, doi: 10.1109/ICIMTech.2019.8843792.

[3] E. Y. L. Nandapala and K. P. N. Jayasena, "The practical approach in Customers segmentation by using the K-Means Algorithm," 2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS), 2020, pp. 344-349, doi: 10.1109/ICIIS51140.2020.9342639.

[4] Ş. Ozan, "A Case Study on Customer Segmentation by using Machine Learning Methods," 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), 2018, pp. 1-6, doi: 10.1109/IDAP.2018.8620892.

[5] A. G. Aggarwal and S. Yadav, "Customer Segmentation Using Fuzzy-AHP and RFM Model," 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2020, pp. 77-80, doi: 10.1109/ICRITO48877.2020.9197903.

[6] N. R. Maulina, I. Surjandari and A. M. M. Rus, "Data Mining Approach for Customer Segmentation in B2B Settings using Centroid-Based Clustering," 2019 16th International Conference on Service Systems and Service Management (ICSSSM), 2019, pp. 1-6, doi: 10.1109/ICSSSM.2019.8887739.

[7] T. Kansal, S. Bahuguna, V. Singh and T. Choudhury, "Customer Segmentation using K-means Clustering," 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), 2018, pp. 135-139, doi: 10.1109/CTEMS.2018.8769171.

[8] Y. Huang, M. Zhang and Y. He, "Research on improved RFM customer segmentation model based on K-Means algorithm," 2020 5th International Conference on Computational Intelligence and Applications (ICCIA), 2020, pp. 24-27, doi: 10.1109/ICCIA49625.2020.00012.

[9] X. Pu, C. Song and J. Huang, "Research on Optimization of Customer Value Segmentation Based on Improved K-Means Clustering Algorithm," 2020 IEEE 3rd International Conference on Information Systems and Computer Aided Education (ICISCAE), 2020, pp. 538-542, doi: 10.1109/ICISCAE51034.2020.9236867.

[10] S. Koul and T. M. Philip, "Customer Segmentation Techniques on E-Commerce," 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 2021, pp. 135-138, doi: 10.1109/ICACITE51222.2021.9404659.

[11] A. S. M. S. Hossain, "Customer segmentation using centroid based and density based clustering algorithms," 2017 3rd International Conference on Electrical Information and Communication Technology (EICT), 2017, pp. 1-6, doi: 10.1109/EICT.2017.8275249.

[12] S. Guney, S. Peker and C. Turhan, "A Combined Approach for Customer Profiling in Video on Demand Services Using Clustering and Association Rule Mining," in IEEE Access, vol. 8, pp. 84326-84335, 2020, doi: 10.1109/ACCESS.2020.2992064.

[13] R. Ferrera, J. M. Pittman, M. Zapryanov, O. Schaer and S. Adams, "Retailer’s Dilemma: Personalized Product Marketing to Maximize Revenue," 2020 Systems and Information Engineering Design Symposium (SIEDS), 2020, pp. 1-6, doi: 10.1109/SIEDS49339.2020.9106672.

[14] T. Xu and X. Dong, "Mining frequent patterns with multiple minimum supports using basic Apriori," 2013 Ninth International Conference on Natural Computation (ICNC), 2013, pp. 957-961, doi: 10.1109/ICNC.2013.6818114.