Retraction

Retraction: Research on Personalized Healthy Diet Recommendation Based on Artificial Intelligence (J. Phys.: Conf. Ser. 1852 022085)

Published 9 September 2022

This article has been retracted by IOP Publishing following an allegation that raises concerns this article may have been created, manipulated, and/or sold by a commercial entity. In addition, IOP Publishing has seen no evidence that reliable peer review was conducted on this article, despite the clear standards expected of and communicated to conference organisers.

The authors of the article have been given opportunity to present evidence that they were the original and genuine creators of the work, however at the time of publication of this notice, IOP Publishing has not received any response. IOP Publishing has analysed the article and agrees there are enough indicators to cause serious doubts over the legitimacy of the work and agree this article should be retracted. The authors are encouraged to contact IOP Publishing Limited if they have any comments on this retraction.

Retraction published: 9 September 2022
Research on Personalized Healthy Diet Recommendation Based on Artificial Intelligence

Xiaoxuan Tang¹,*
¹Shandong Vocational College of Science and Technology, Weifang Shandong, 261053

*Corresponding author e-mail: glh56789@sdvcst.edu.cn

Abstract. With the development of the Internet, China's good data mining technology in e-commerce has achieved development, and the recommendation service has gradually become popular and has been adopted by the majority of citizens. In order to facilitate people to quickly find what they need from a large number of commodities and save consumers' time, the recommendation system pays more and more attention to the user experience. This paper constructs a healthy diet recommendation model based on improved Apriori. When judging the user's dietary preferences, the improved Apriori algorithm is used to recommend the diet. The original algorithm and the improved algorithm are compared to the same example. The results show that the improved algorithm has advantages in mining association rules and can better realize personalized healthy diet recommendations.

Keywords: artificial intelligence, Apriori algorithm, healthy diet recommendation system

1. Introduction

With the rapid development of my country’s economy, people’s living pace is getting faster and faster while people’s living standards are improving. People’s poor living habits have caused a series of health problems. Unreasonable eating habits and dietary structure will affect people’s health. And how to develop healthy eating habits and ways in busy work is a problem that people are concerned about. Therefore, people have gradually noticed the importance of food choices for three meals a day, and a series of nutritional products have been derived. However, due to the large population of our country and the lack of talents in the nutritionist industry, not everyone can have a professional dietitian to arrange the corresponding food intake according to personal physique. The professional dietitian is expensive and inefficient, the penetration rate is not high, so the intelligent diet recommendation designed by the computer is gradually adopted. This article proposes a diet recommendation model based on improved Apriori diet recommendation model, and analyzes its effect.

2. Apriori Algorithm
Apriori algorithm is a classical frequent itemset algorithm for mining association rules. It mainly finds out the relationship between itemsets in the database through the iterative way of layer by layer search, so as to form rules. The process is mainly composed of connection (class matrix operation) and pruning (removing the unintended intermediate results). The basic idea is: firstly, all candidate 1-itemsets C1 are found, and then frequent 1-itemsets L1 are generated according to item set C1. The frequency of occurrence of these frequent sets needs to be equal to or greater than the given minimum support, that is, association rules satisfying the minimum confidence and support are generated from frequent sets. Secondly, candidate 2-itemsets C2 are generated by L1 according to iteration rules, and frequent 2-items are generated from C2 by C2 Set L2, in turn, until the location of the frequent itemsets can no longer be generated. The circulation flow chart is as follows:

Figure 1. Apriori algorithm loop flow chart

3. Apriori Algorithm Improvement
As a classic association rule mining algorithm, Apriori algorithm has a great influence in the development and research of data mining technology. The algorithm can mine data through repeated connection operations and pruning operations. However, with the increase of the data mined, the defects of this algorithm are gradually revealed, and its main disadvantages are as follows:

(1) When the number of candidate itemsets is large, the time consumed by the algorithm will increase accordingly. When a transaction database contains equal to or greater than 100 frequent 1-itemsets, the number of candidate frequent 2-itemsets generated is 200. The calculation of such a large number of candidate sets will consume huge memory space and very long time.

(2) The database will be rescanned along with the connection operation. Multiple scans means that multiple input and output operations are required. In the long run, the efficiency of the algorithm will
decrease.

(3) Once the support of the algorithm is determined, it cannot be changed, otherwise it will affect the accuracy of data mining.

This study improves the Apriori algorithm from the aspects of reducing the generation of scanning transaction databases and frequent itemsets, so as to achieve the purpose of optimizing the Apriori algorithm and improving its computational efficiency. The specific improvements are as follows:

(1) Scan the candidate set C1 generated by the transaction database D, and calculate the support of each candidate set.

(2) Analyze the candidate item set C1 to obtain L1.

(3) Analyze L1 and find the most supported items, and list the possible candidate item sets one by one according to the obtained support item set (k-item set Ck(2,3,...)).

(4) Analyze the support of each data in the frequent K-items set Ck, and then use the properties of the Apriori algorithm to compare and calculate the obtained data one by one, and finally get the items set Cn.

4. Examples And Effects of Diet Recommendations

4.1. Instance Data

Use the Apriori algorithm to mine the frequent itemsets of the data in the system constructed in this article. The following table shows the dishes in a user's attention list in the system:

Table 1. The diet menu of a user in the system

|                         | Beauty | Strengthen the spleen | Nourish the stomach | Invigorate Qi | Nourish blood | Promote the secretion of saliva | Diuresis | Expectorant | Calm the nerves |
|-------------------------|--------|-----------------------|--------------------|--------------|--------------|-------------------------------|----------|-------------|----------------|
| Dinghu Plain Dish (T1) | 1      | 0                     | 0                  | 0            | 0            | 0                             | 1        | 1           | 1              |
| Quick fried yellow croaker with vinegar (T2) | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| Mint and coriander sauce (T3) | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 |
| Hangzhou three delicacies (T4) | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| Braised rice with sausage (T5) | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| Fried eggs with fish sauce (T6) | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| Roast pork strips (T7) | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Steamed fish | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
4.2. Mining of Association Rules before Algorithm Improvement

From the diet menu of a user in Table 1, it can be learned that there are 10 diet records related to the user, then the value of D is 10 and the steps of using Apriori algorithm to mine frequent itemsets are as follows:

(1) During the first round of the algorithm, each item that meets the standard conditions is an element in the candidate 1-item set C1. The algorithm is used to scan D in the database to determine each item in C1. The support of the element, as shown in Figure 2:

![Figure 2](image)

| Item set | Support frequency |
|----------|------------------|
| \{I_1\}  | 6                |
| \{I_2\}  | 2                |
| \{I_3\}  | 5                |
| \{I_4\}  | 1                |
| \{I_5\}  | 1                |
| \{I_6\}  | 4                |
| \{I_7\}  | 3                |
| \{I_8\}  | 1                |
| \{I_9\}  | 4                |

(2) Set the minimum support vacation \(\min_{\text{sup}} = 2/9 = 22\%\) to determine the frequent item set 1-L1, which is composed of elements in the candidate 1-item set C1.

(3) In order to find frequent itemsets 2-itemsets L2, Apriori algorithm \(L_1 \oplus L_1\) is used to obtain candidate 2-itemsets C2, where C2 includes 2-itemsets in \(\binom{|L_1|}{2}\). Then the D in the transaction database is scanned twice to obtain the element support in the candidate 2-item set C2, as shown in Figure 3:

![Figure 3](image)
(4) Therefore, the content in the frequent 2-item set \( L_2 \) is determined, and its generation is composed of all 2-item sets with support greater than the minimum support in the candidate 2-item set \( C_2 \).

(5) Use the algorithm \( L_1 \oplus L_1 \) to calculate the content in the frequent 2-item set \( L_2 \) to obtain the candidate 3-item set \( C_3 \). The calculation process is as follows:

i. Connection process:

\[
C_3 = L_2 \oplus L_2 = \{(I_1, I_1), (I_1, I_2), (I_1, I_3), (I_1, I_4), (I_1, I_5), (I_1, I_6), (I_1, I_7), (I_1, I_8), (I_1, I_9), (I_2, I_3), (I_2, I_4), (I_2, I_5), (I_2, I_6), (I_2, I_7), (I_2, I_8), (I_2, I_9), (I_3, I_4), (I_3, I_5), (I_3, I_6), (I_3, I_7), (I_3, I_8), (I_3, I_9), (I_4, I_5), (I_4, I_6), (I_4, I_7), (I_4, I_8), (I_4, I_9), (I_5, I_6), (I_5, I_7), (I_5, I_8), (I_5, I_9), (I_6, I_7), (I_6, I_8), (I_6, I_9), (I_7, I_8), (I_7, I_9), (I_8, I_9)\} \oplus \{(I_1, I_1), (I_1, I_2), (I_1, I_3), (I_1, I_4), (I_1, I_5), (I_1, I_6), (I_1, I_7), (I_1, I_8), (I_1, I_9), (I_2, I_3), (I_2, I_4), (I_2, I_5), (I_2, I_6), (I_2, I_7), (I_2, I_8), (I_2, I_9), (I_3, I_4), (I_3, I_5), (I_3, I_6), (I_3, I_7), (I_3, I_8), (I_3, I_9), (I_4, I_5), (I_4, I_6), (I_4, I_7), (I_4, I_8), (I_4, I_9), (I_5, I_6), (I_5, I_7), (I_5, I_8), (I_5, I_9), (I_6, I_7), (I_6, I_8), (I_6, I_9), (I_7, I_8), (I_7, I_9), (I_8, I_9)\}
\]

ii. Use the properties of the Apriori algorithm to perform pruning operations, and then define all the subsets in the frequent itemsets as frequent itemsets according to property 1. The following is the process of judging whether the candidate set contains infrequent itemsets:

In the first step, the subsets of \( \{I_1, I_3, I_9\} \) are \( \{I_1, I_3\} \), \( \{I_1, I_9\} \), \( \{I_3, I_9\} \) respectively, and can be found in \( L_2 \), then this item is included in \( C_3 \).

In the second step, the subsets of \( \{I_1, I_7, I_9\} \) are \( \{I_1, I_7\} \), \( \{I_1, I_9\} \), \( \{I_7, I_9\} \) respectively, and, which can be found in \( L_2 \), then this item is included in \( C_3 \).

In the third step, the subsets of \( \{I_1, I_7, I_9\} \) are \( \{I_1, I_7\} \), \( \{I_1, I_9\} \), \( \{I_7, I_9\} \) respectively, and, which can be found in \( L_2 \), then this item is included in \( C_3 \).

In the fourth step, the subsets of \( \{I_3, I_7, I_9\} \) are \( \{I_3, I_7\} \), \( \{I_3, I_9\} \), \( \{I_7, I_9\} \) respectively, and, which can be found in \( L_2 \), then this item is included in \( C_3 \).

The fifth step, if some itemsets are not found in \( C_3 \), then these itemsets are not included in \( C_3 \).

Then the candidates after the pruning operation are

\[C_3 = \{(I_1, I_3, I_9), (I_1, I_5, I_9), (I_1, I_7, I_9), (I_3, I_5, I_9)\} \]

One point that needs special attention is that since the search of the Apriori algorithm is carried out in a hierarchical cycle, only one candidate item set (k-item set \( C_k \)) is given, then it is only necessary to determine whether (k-1)-item set is Frequent itemsets are enough, search candidate 3-itemsets and
frequent 3-itemsets are shown in Figure 4.

(6) D in the transaction database is scanned to determine the content in L3, which is composed of 3-itemsets with greater than minimum support in C3.

(7) Finally, the candidate 4-itemset C4 will be obtained through the algorithm $L_3 \oplus L_3$. Then, $C_4 = \{I_1, I_3, I_7, I_9\}$, and only C4 has $\{I_1, I_3, I_7, I_9\}$, so the result of C5 cannot be obtained. Therefore, C5 is an empty set, and the Apriori algorithm ends here because it can no longer cycle and cannot obtain a new itemset.

Combined with the user data in the system, we can select the dishes that are suitable for the user's physique and have the health attributes to recommend to the users.

### 4.3. Effect of Improved Algorithm

This study still analyzes the implementation process of the improved Apriori algorithm according to the user data in the table. The procedure of obtaining candidate 1-itemset C1 and frequent 1-itemset L1 in the improved algorithm is the same as that of the original algorithm, followed by the step of obtaining candidate k-itemset CK. This step is different from the previous one. The improved algorithm obtains the result based on the analysis of L1. Through the comparison in L1, we can get the most relevant items I1, I2, I3 and I4 with the support degree of 6, 5, 4 and 3 respectively. We can know that if the support degree of i7 is 3, then the support degree of I1 is 2, so we can get the result that there is no I1 in the frequent 4-item set. If k = 4 in the candidate k-itemset Ck, then only I1, I3, i9 and i7 are possible. Therefore, the candidate k-itemset Ck is shown in Table 2.

**Table 2. Candidate k-item set C k**

| Item set | Support frequency | Item set | Support frequency |
|----------|-------------------|----------|-------------------|
| $\{I_1, I_2\}$ | 0                  | $\{I_1, I_2, I_3\}$ | 0                 |
| $\{I_1, I_3\}$ | 3                  | $\{I_1, I_2, I_7\}$ | 0                 |
| $\{I_1, I_4\}$ | 4                  | $\{I_1, I_2, I_9\}$ | 2                 |
| $\{I_1, I_9\}$ | 3                  | $\{I_1, I_3, I_9\}$ | 2                 |
| $\{I_2, I_3\}$ | 1                  | $\{I_1, I_7, I_9\}$ | 2                 |
| $\{I_2, I_7\}$ | 0                  | $\{I_2, I_3, I_7\}$ | 0                 |
| $\{I_2, I_9\}$ | 0                  | $\{I_2, I_5, I_9\}$ | 0                 |
Combining the data in the above table to compare the results of the two algorithms, the frequent itemsets obtained are all \( \{I_1, I_3, I_7, I_9\} \), so that the conclusion is consistent with the conclusion of the Apriori algorithm.

4.4. Analysis of computational efficiency before and after the improvement of Apriori algorithm

Through the above examples, it is not difficult to see that the biggest difference between the improved Apriori algorithm and the pre-improved Apriori algorithm is that the improved Apriori algorithm only needs to scan the transaction database twice, and it is analyzed from the structure of the first scan The candidate k-item set Ck (k>2) that may exist in the database is generated, thereby reducing the generation of candidate sets. In order to test the difference in the operational efficiency of the Apriori algorithm before and after the improvement, different amounts of data were used for testing, and the results are shown in Table 3:

| Number of test data | Average test time of the original Apriori algorithm | Improved Apriori algorithm test average time |
|---------------------|--------------------------------------------------|-------------------------------------------|
| 100                 | 121                                              | 78                                        |
| 200                 | 218                                              | 144                                       |
| 300                 | 353                                              | 206                                       |
| 400                 | 497                                              | 311                                       |
| 500                 | 606                                              | 383                                       |

5. Conclusion

With the development of Internet technology in China, various recommendation systems have been sought after by people. This article builds a framework for a healthy diet recommendation system based on artificial intelligence, which is an intelligent recommendation system that recommends different dishes according to different users' physical conditions and dietary preferences. In order to build a recommendation system with a better user experience, a healthy diet recommendation based on the Apriori algorithm is constructed, thereby reducing the number of database scans, achieving the purpose of reducing the generation of candidate sets and improving the computational efficiency of the algorithm.

References

[1] Ji Wenlu, Wang Hailong, Su Guibin, Liu Lin(2020). Research review of recommendation methods based on association rule algorithms. Computer Engineering and Applications, vol. 56, no. 22, pp: 33-41.
[2] Ding Li(2013). Research on user behavior data mining based on Apriori algorithm. Bulletin of Science and Technology, vol. 29, no.12, pp: 214-217.

[3] Luo Ke, He Caiwang(2006). An improved association rule extraction algorithm based on Apriori algorithm[J]. Computer and Digital Engineering, no.4, pp: 48-51.

[4] Chen Jin(2020). Improvement of Apriori algorithm and its application in book intelligent recommendation system. Journal of Chongqing University of Science and Technology (Natural Science Edition), vol. 22, no.5, pp: 91-95.

[5] Yang Jingrong, Hou Xiangning(2020). Research on data mining algorithm for positive and negative association rules[J]. Computer Technology and Development, vol. 30, no. 11, pp: 64-68.

[6] Fan Liang, Li Zhao(2019). Big data mining analysis algorithm based on neutrosophic fuzzy association rules generation. Computer Applications and Software, vol. 36, no. 10, pp: 285-292.

[7] Zong Wanli, Zhu Xijun(2020). Association rule mining of food sampling data based on Apriori algorithm[J]. Journal of Food Safety and Quality Inspection, vol. 11, no.4, pp: 1334-1337.

[8] Huang Ling, Yu Xia(2020). E-commerce commodity intelligent recommendation system based on cloud platform. Modern Electronic Technology, vol. 43, no.5, pp: 183-186.

[9] Chen Bin, Zhang Rongmei(2018). A review of research on intelligent recommendation system[J]. Journal of Hebei Academy of Sciences, vol. 35, no.3, pp: 82-92.

[10] Chen Jin(2020). Improvement of Apriori algorithm and its application in book intelligent recommendation system. Journal of Chongqing University of Science and Technology (Natural Science Edition), vol. 22, no.5, pp: 91-95.