.

Comparison of Borda and NRF (Normalized Rating Frequency) in Recommender System

Taufiq Abidin[1], Slamet Wiyono[2]
Politeknik Harapan Bersama, Tegal, Indonesia

Article Info	ABSTRACT
	The Collaborative Filtering method is a popular method in making recommender systems. Although CF is a popular method, it has major problems, namely cold start and sparsity . Several studies have been conducted to treat cold starts and sparsity. One way to overcome cold start and sparsity is the Borda calculation method. Research using the Borda method has been carried out a lot but has not utilized the rating optimally. The NRF method is a new method offered to maximize the use of ratings. By using dummy test data, the NRF method is more effective than Borda in calculating recommendation scores.

Corresponding Author:

Slamet Wiyono,
Department of Engineering informatics,
Politeknik Harapan Bersama, Tegal, Indonesia.
Email: oc_slametwiyono@poltektegal.ac.id

1. INTRODUCTION

Recommender systems are increasingly being used in online media that offer products/services. This is because of the increasing amount of data causes the complexity of user data. For example, online trading platforms such as Tokopedia, Bukalapak, Lazada are getting more and more users. Not only that, movie or music streaming services on demand such as Netflix and Amazon Prime Video also continue to grow from the number of users.

Recommender system work by gathering information about user preferences for the product either explicitly or implicitly [1]. The recommender system is different from decision support systems, such as those used to select the best airline [2] or select a project manager [3]. In general, recommendation systems are grouped into three categories: content-based, collaborative filtering (CF), and hybrid [4]. The content-based method generates recommendations based on user-profiles and the similarity of product descriptions, while CF generates recommendations for users based on the transaction history (assessment) of other users [5]. However, there is no known good algorithm for making recommender systems. Some of the research conducted only focuses on the application of the method. Some of these studies; collaborative filtering recommendation system using the ALS method [6], making a recommender system using the Apriori algorithm [7], making a recommender system for UMKM (micro, small, and medium enterprises) [8], making a recommender system using the K-Means algorithm [9], making a recommender system using the association method [10], making a recommender system using Apache Mahout [11], making a recommender system using content-based filtering method for food crops [12], making a recommender system using Naïve Bayes method [13], making a recommender system using TOPSIS algorithm [14], and using Weight Product algorithm [15]. There have not been many studies comparing algorithms, such as the comparison of KNN, SVM, and Decision Tree [16].

The CF method is a popular method in making recommender systems [5]. Although CF is a popular method, it has major problems, namely cold start and sparsity [17]. Research to overcome the problem of cold start in CF has been carried out by Lika et al. [18] and Uyangoda et al. [19]. Tang and Tong conducted research to address sparsity in CF using a ranking-based approach [20], and other studies using a ranking-based approach were carried out using the Copeland Score aggregation method [21]

Rank-based research that has been conducted using the Borda and Copeland methods has not utilized rating data with maxima. One of the studies that tried to use rating data to the maximum has been carried out by Lestari by offering the NRF (normalized rating frequency) method [22]. However, the methods offered have not yet compared the effectiveness of their use. This study aims to compare the effectiveness of the Borda and NRF methods in a collaborative filtering-based recommender system.

2.   RESEARCH METHOD

The research was conducted by conducting several stages, starting from; literature review, data collection, data preparation, testing, analysis, and conclusion. The research procedure carried out is shown in Figure 1.
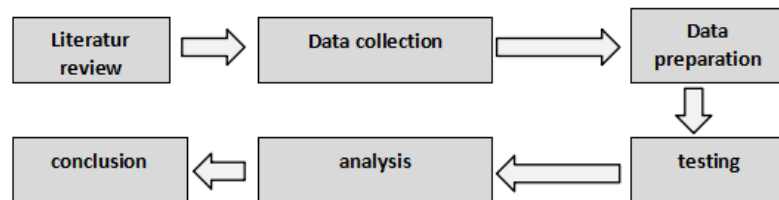

Figure 1. Research procedure

Library sources come from national/international journals, books, papers, and the internet. Literature study continues to be carried outside by side with other research stages until the end of the research. This is done so that if in the next stage other reference sources are found that support the research, these reference sources can be used as literature to help complete the research that will be carried out. The research will be conducted using dummy data. To determine what data to use for research, literature studies are carried out through journals, books, and the internet related to the research to be carried out. Data preparation is done by cleaning the data so that it can be processed easily and correctly. In addition, at this stage, data descriptions and data summaries are also carried out so that the characteristics of the data can be known.

The test was carried out by comparing the Borda and NRF methods in making rankings. The analysis is carried out by testing the model obtained. The Borda method calculation is shown in equation 1, while the NRF method is shown in equation 2.

$$V_{(i)} = \sum_i W_{(i,j)} - \sum_i L_{(i,j)} \tag{1}$$

$$NRF_{(Ph)} = \sum_r NR_{(ri,Ph)} \tag{2}$$

The test results were analyzed by analyzing the test result score. The last one is to conclude whether the objectives of the research have been achieved or not.

3.   RESULTS AND DISCUSSION

The film rating data used has obtained from Movie Lends data. The data contains 80000 rows and 4 columns. Table 1 shows the top 6 data, consisting of; user id, item id, rating, and timestamp.

Table 1. Data rating

User_id	Item_id	Rating	Timestamp
1	1	5	874965758
1	2	3	876893171
1	3	4	878542960
1	4	3	876893119
1	5	3	889751712
1	6	4	875071561

By using a sample of 10 film items, the calculation results of the Borda and NRF methods can be seen in Figure 2 and Table 2. Table 1 shows the sample data used for calculations. The data used is a film rating of 943 users. The first user gives a rating of 5 for film with id item 1, rating 3 for with id item 2, rating 4 for with id item 3, rating 3 for with id item 4, rating 3 for with id item 5, and rating 4 for with id item 6.

Table 2. Calculation results

Order of Recommendations	Borda		NRF	
	Item_id	Skor	Item_id	Skor
1	2	9	2	4.9
2	1	8	1	3.83

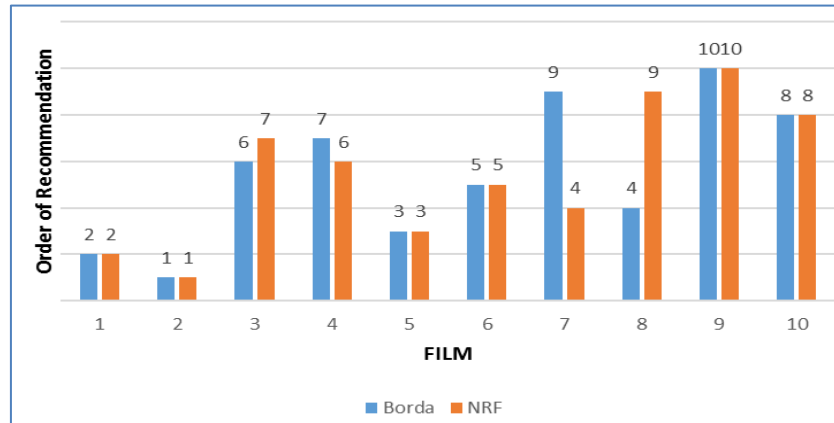| Order of Recommendations	Borda		NRF	
	Item_id	Skor	Item_id	Skor
3	6	8	7	3.67
4	7	7,9	6	3,5
5	3	7,7	3	3,47
6	5	7	5	3,10
7	9	6	4	3
8	4	6,3	9	2,95
9	10	6,3	10	2.93
10	8	5	8	2,8

Figure 2. Comparison of order recommendation

Figure 2 and Tabel 2 show a comparison of the recommendation order. The Borda and NRF methods have similarities for the films with items id 1, 2, 3, 5, 8, and 10. However, both have different sequences for films with items id 4, 6, 7, and 9. The Borda method gives the third recommendation for a film with item id 6, while the NRF gives the third recommendation for a film with item id 7. Likewise, for the film with item id 7, Borda gives the ninth-order and the NRF gives the fourth-order. The Borda method produces the same score for the second and third recommendation order, a score of 8. While the NRF method produces different scores for each recommendation sequence. This shows that the Borda method is less effective than the NRF method. Table 3 shows a very clear difference in determining the order based on the scores obtained. The Borda method still produces the same score so it is confusing to determine the order, in contrast to the NRF method which does not produce the same score so it is easy to sort.

Table 3. The difference in score between Borda and NRF methods

Order of Recommendations	Skor	
	NRF	Borda
2	3.83	8
3	3.67	8
8	2,95	6,3
9	2.93	6,3

Both NRF and Borda use the ranking method in determining recommendations. However, the NRF method yields different / detailed ranking scores (see Table 3). Borda produces the same score (score 8) for 2nd and 3rd order, even though 2nd and 3rd are different films. Thus, the correct order could be 3 then 2. It is different from the NRF method which produces a different score so that it clearly shows the 2nd and 3rd order. Borda produces the same score so it is confusing in providing recommendations, in contrast to the NRF which gives a different score so that it is not confusing. The calculation method results in less effective calculations. This is due to several things. First, the use of rating data is only for compiling the user preference profile. Second, point determination is only based on products in the preference list, not considering other ratings such as the same number of ratings and index.

4.   CONCLUSION

The Borda and NRF methods have the same method of determining recommendations, namely based on ratings. However, it is not yet known which method is the best in calculating the rating score. After the comparison of the score calculations using the film rating data, the NRF method is more effective in providing recommendations because it produces different calculations for each recommendation sequence. The NRF method can calculate more detail than Borda in calculating the order of recommendation. NRF produces scores that are always different, while Borda produces the same score

calculations for 2nd and 3rd (score 8) and 8th and 9th (score 6.3). For further research, it is necessary to compare the two methods by conducting user testing or usability testing.

ACKNOWLEDGMENTS

REFERENSI

[1]   S. Lestari, T. B. Adji, and A. E. Permanasari, "Performance Comparison of Rank Aggregation Using Borda and Copeland in Recommender System," 2018 Int. Work. Big Data Inf. Secur., pp. 69–74, 2018.

[2]   H. T. Hidayat and N. G. A. P. H. Saptarini, "Sistem Pendukung Keputusan Pemilihan Maskapai Terbaik Pada Rute Domestik Dengan Menggunakan Metode Multi Factor Evaluation Process (MFEP)," MATRIX, vol. 4, no. 3, 2014.

[3]   P. I. Ciptayani and N. G. A. P. H. Saptarini, "Sistem Pendukung Keputusan Pemilihan Manajer Proyek Menggunakan Analytical Hierarchy Process," MATRIX, vol. 4, no. 2, 2014.

[4]   S. K. Gorakala and M. Usuelli, Building a Recommnebder System with R. Birmingham-Mumbai: Packt Publishing Ltd., 2015.

[5]   S. Natarajan, S. Vairavasundaram, S. Natarajan, and A. H. Gandomi, "Resolving Data Sparsity and Cold Start Problem in Collaborative Filtering Recommender System Using Linked Open Data," Expert Syst. Appl., 2020.

[6]   A. S. Girsang, B. AL Faruq, H. R. Herlianto, and S. Simbolon, "Collaborative Recommendation System in Users of Anime Films," in Journal of Physics: Conference Series, 2020, pp. 1–7.

[7]   R. Hidayat, "Membangun Sistem Rekomendasi Berbasis Algoritma Apriori Pada Website Penjualan UD Rahmat," STMIK Widya Cipta Dharma, 2019.

[8]   W. A. Prakoso, "Pembangunan Sistem Rekomendasi Dalam Strategi Bisnis Umkm Berdasarkan Lokasi Di Kota Bandung Pada Platform Android," Universitas Komputer Indonesia, 2019.

[9]   A. Wiyanto, "Sistem Rekomendasi Kuliner Dengan Menggunakan Metode K-Means Untuk Mendukung Strategi Penjualan Makanan," Universitas Pembangunan Nasional Veteran Jakarta, 2019.

[10]   N. P. Sari, C. A. Putra, and F. Muttaqin, "Rancang Bangun Sistem Rekomendasi Pencarian Produk Properti Dengan Menggunakan Metode Association Rule Di PT.Brighton," J. Inform. dan Sist. Inf., vol. 01, no. 1, 2020.

[11]   K. R. Sari, W. Suharso, and Y. Azhar, "Pembuatan Sistem Rekomendasi Film dengan Menggunakan Metode Item Based Collaborative Filtering pada Apache Mahout," REPOSITOR, vol. 2, no. 6, 2020.

[12]   P. Nastiti, "Penerapan Metode Content Based Filtering Dalam Implementasi Sistem Rekomendasi Tanaman Pangan," TEKNIKA, vol. 8, no. 1, 2019.

[13]   A. S. Guntara, "Implementasi Metode Naïve Bayes pada Sistem Rekomendasi Pembelian Mobil pada PT. Honda Prospect Motor," REMIK, vol. 4, no. 1, 2019.

[14]   I. T. Sevandri, R. K. Dewi, and M. T. Ananta, "Implementasi Algoritma Topsis Pada Sistem Rekomendasi Pencarian Lokasi Gym Berbasis Android (Studi Kasus: Kota Malang)," JPTIIK, vol. 3, no. 4, 2019.

[15]   T. W. Harjanti and L. Lipindi, "Rancang Bangun Sistem Rekomendasi BIrotravel pada Fitur Custom Trip untuk Aplikasi Travel Vaithme Menggunakan Algoritma Weighted Product Berbasis Android," in SEMNASTEK 2019, 2019.

[16]   S. Wiyono and T. Abidin, "COMPARATIVE STUDY OF MACHINE LEARNING KNN , SVM , AND DECISION TREE ALGORITHM TO PREDICT STUDENT ' S PERFORMANCE," IJRG, vol. 7, no. January, pp. 190–196, 2019.

[17]   maryam K. Najafabadi, A. Mohamed, and C. W. Onn, "An impact of time and item influencer in collaborative filtering recommendations using graph-based model," Inf. Process. Manag., vol. 56, no. 3, pp. 526–540, 2019.

[18]   B. Lika, K. Kolomvatsos, and S. Hadjiefthymiades, "Expert Systems with Applications Facing the cold start problem in recommender systems," Expert Syst. Appl., vol. 41, no. 4, pp. 2065–2073, 2014.

[19]   L. Uyangoda, S. Ahangama, and T. Ranasinghe, "User Profile Feature-Based Approach to Address the Cold Start Problem in Collaborative Filtering for Personalized Movie Recommendation," in Thirteenth International Conference on Digital Information Management (ICDIM), 2018, pp. 24–28.

[20]   Y. Tang and Q. Tong, "BordaRank : A Ranking Aggregation Based Approach to Collaborative Filtering," in International Conference on Computer and Information Science (ICIS), 2016, pp. 1–6.

[21]   G. Al-sharrah, "Ranking Using the Copeland Score : A Comparison with the Hasse Diagram," J. Chem. Inf. Model., pp. 785–791, 2010.

[22]   S. Lestari, T. B. Adji, and A. E. Permanasari, "NRF : Normalized Rating Frequency for Collaborative Filtering Paper," in International Conference on Applied Information Technology and Innovation (ICAITI), 2018, pp. 19–25.