International Journal of Mathematical, Engineering, Biological and Applied Computing, 2022, 1, 12-16
www.scipublications.org/journal/index.php/ijmebac
DOI: 10.31586/ijmebac.2022.291

Article

Movie Recommendation System Modeling Using Machine Learning

Raja Marappan *, **S. Bhaskaran**

School of Computing, SASTRA Deemed University, Thanjavur, India

*Correspondence: Raja Marappan (professor.m.raja@gmail.com)

Abstract: The task of recommending products to customers based on their interests is important in business. It is possible to accomplish this with machine learning. To reduce human effort by proposing movies based on the user's interests efficiently and effectively without wasting much time in pointless browsing, the movie recommendation system is designed to assist movie aficionados. This work focuses on developing a movie recommender system using a model that incorporates both cosine similarity and sentiment analysis. Cosine similarity is a standard used to determine how similar two items are to one another. An examination of the emotions expressed in a movie review can determine how excellent or negative a review is and, consequently the overall rating for a film. As a result, determining whether a review is favorable or adverse may be automated because the machine learns by training and evaluating the data. Comparing different systems based on content-based approaches will produce results that are increasingly explicit as time passes.

Keywords: Cosine Similarity, Sentiment Analysis, Machine Learning, Recommendation Systems, Content-Based Approaches

How to cite this paper: Marappan, R., & Bhaskaran, S. (2022). Movie Recommendation System Modeling Using Machine Learning. *International Journal of Mathematical, Engineering, Biological and Applied Computing*, *1*(1), 12–16. Retrieved from https://www.scipublications.com/journal/index.php/ijmebac/article/view/291

Received: March 27, 2022
Accepted: May 04, 2022
Published: May 06, 2022

1. Introduction

Technology has taken centre stage in virtually every industry, resulting in an overabundance of data and information [1]. The use of recommendation systems can resolve this issue. To deal with large amounts of data, recommender systems must first filter out irrelevant information before attempting to predict user preferences. To implement recommendation systems, the most commonly used techniques include content-based filtering, collaborative filtering (CF), and hybrid filtering. Content-based filtering is the most common technique employed to implement recommendation systems [12,13,15].

2. Review of the Literature

To deliver movie recommendations, a system has been developed that uses the information already known about the user [1]. This approach attempts to resolve the problem of unique recommendations that arise as a result of neglecting the data specific to each user. The psychological profile of the person and their viewing history and data, including movie scores from other websites, are all gathered together. They are based on an estimate of aggregate similarity between two things. The system is a hybrid model that makes use of both content-based filtering and CF techniques to achieve its goals [12,13,15].

MODREC is a movie recommendation system developed with the CF approach to make recommendations. The information provided by the user is utilized in CF. That information is analyzed, and a movie is recommended to the users in a sorted manner with the movie with the highest rating appearing first. The system also includes a feature that

allows the user to specify the features the user would like the movie to be recommended [3-5].

An analysis is performed on the standard recommender systems, namely, content-based filtering and CF. A novel approach, which combines both Bayesian networks and CF, was proposed because they both have their own set of shortcomings. The suggested system is optimized for the challenge at hand and generates probability distributions that may be used to draw helpful conclusions about the problem [6-11].

3. Model of a Movie Recommender

The movie recommender model is developed using some recommendation and machine learning (ML) strategies with [12,13,15,27]. The user preferences, history, and interests are taken into consideration, and the substance of each item when using content-based filtering to make recommendations [29]. On the other hand, CF is a method of making recommendations based on similar users, as shown in Figure 1. It is designed to simulate user-to-user recommendations. An example of a hybrid recommender system combines both content-based and CF techniques. The cosine similarity can also be used to find similarities between two vectors in an inner product space. It is determined by the cosine of the angle between two vectors and is used to determine whether or not two vectors are pointing in the same general direction. In the text analysis, it is frequently used to determine how similar two documents are to one another.

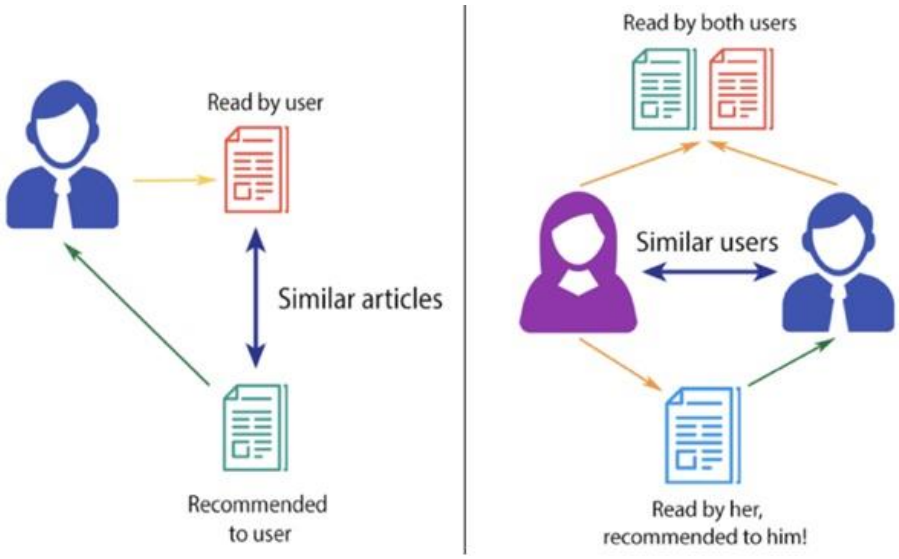
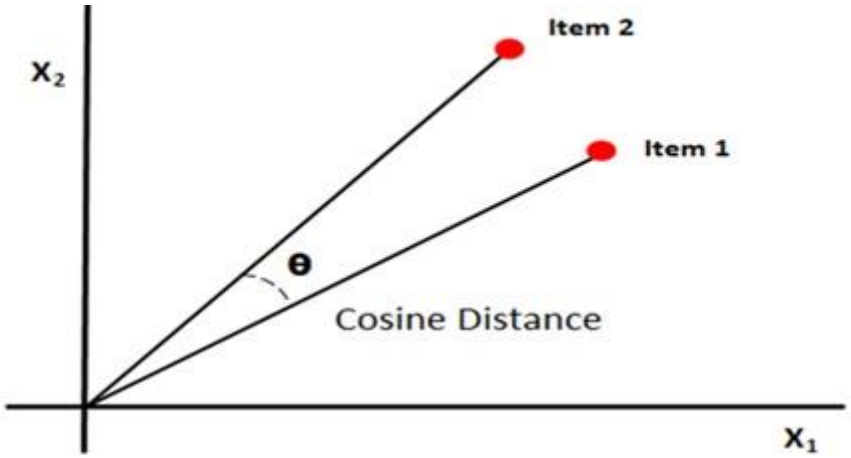


Figure 1. Content (left) versus CF (right)

The direction of a vector is determined by the angle formed between two vectors, as defined in Figure 2. As soon as the angle between the two vectors is equal to zero, the two vectors overlap and appear similar. The type of reviews that a movie receives from its audience determines its level of popularity.



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

Figure 2. Cosine similarity model

The opinions expressed in these reviews can influence the choices of other users. The users are more inclined to choose a movie that was overwhelmingly favored than a movie that was overwhelmingly despised by the general public. This can be accomplished through the use of sentiment analysis. In sentiment analysis, natural language processing (NLP) is used to extract information from a textual source and categorize the statement or document as either positive or negative as illustrated in Figure 3. The Naive Bayes (NB) classifier and the support vector machine (SVM) are two algorithms that are used in sentiment analysis [28].

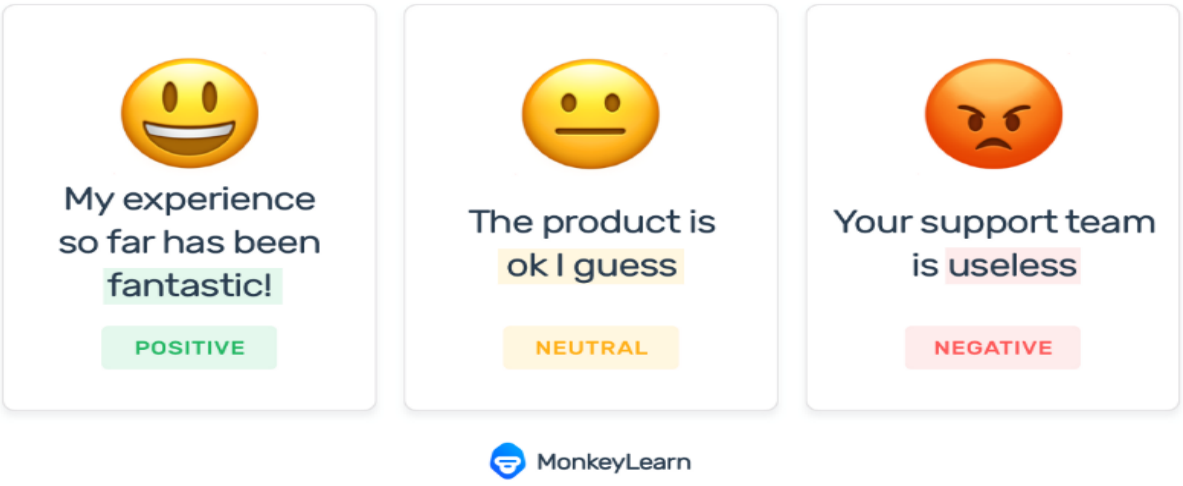


Figure 3. Sentiment analysis

4. Conclusions & Future Plan

The cosine similarity algorithm is a good fit for the movie recommendation system since it is fast and accurate. In addition to the quantum of solace, the cosine similarity algorithm was used to predict the outcomes of five more films, including Never Say Never Again, Skyfall, Thunderball, and From Russia with Love. In the case of sentiment analysis, the SVM classifier performs significantly better than the NB classifier when identifying movie reviews. In the future, different soft computing strategies can be mixed to develop a hybrid recommendation system [14, 16-26].

References

[1] Fisk, D. (1997). An application of social filtering to movie recommendation. In *Software agents and soft computing towards enhancing machine intelligence* (pp. 116-131). Berlin, Heidelberg: Springer.

[2] Chen, Q., & Aickelin, U. (2008). Movie recommendation systems using an artificial immune system. *6th international conference in adaptive computing in design and manufacture (ACDM 2004).* arXiv preprint arXiv:0801.4287\

[3] Uluyagmur, M., Cataltepe, Z., & Tayfur, E. (2012, October). Content-based movie recommendation using different feature sets. *Proceedings of the world congress on engineering and computer science* (Vol. 1, pp. 17-24). http://www.iaeng.org/publication/WCECS2012/WCECS2012_pp517-521.pdf

[4] Kumar Manoj, D.K. Yadav, Singh Ankur, Kr Vijay, "A Movie Recommender System: MOVREC", 2015 International Journal of Computer Applications, vol. 124, pp. 7-11.

[5] A. V. Dev, A. Mohan, "Recommendation system for big data applications based on set similarity of user preferences", 2016 International Conference on Next Generation Intelligent Systems (ICNGIS), pp. 1-6, 2016.

[6] Koen Verstrepen, Bart Goethals, "Unifying nearest neighbors collaborative filtering", Proceedings of the 8th ACM Conference on Recommender systems, October 06- 10, 2014.

[7] A. Jain, S. K. Vishwakarma, "Collaborative Filtering for Movie Recommendation using RapidMiner", International Journal of Computer Applications, vol. 169, no. 6, pp. 0975-8887, July 2017.

[8] Giles C.L., Bollacker K.D., and Lawrence S., "CiteSeer: An automatic citation indexing system," in Proceedings of the third ACM conference on Digital libraries, 1998, pp. 89–98.

[9] Seroussi Y., "Utilising user texts to improve recommendations," User Modeling, Adaptation, and Personalization, pp. 403–406, 2010.

[10] Beel J., Langer S., and Genzmehr M., "Mind-Map based User Modelling and Research Paper Recommendations," in work in progress, 2014.

[11] MacQueen J.. Some methods for classification and analysis of multivariate observations. In Proc. Of the 5th Berkeley Symp. On Mathematical Statistics and Probability, pages 281-297. University of California Press, 1967.

[12] Bhaskaran, S., Marappan, R. Design and analysis of an efficient machine learning based hybrid recommendation system with enhanced density-based spatial clustering for digital e-learning applications. Complex Intell. Syst. (2021). https://doi.org/10.1007/s40747-021-00509-4

[13] Bhaskaran, S.; Marappan, R.; Santhi, B. Design and Analysis of a Cluster-Based Intelligent Hybrid Recommendation System for E-Learning Applications. Mathematics 2021, 9, 197. https://doi.org/10.3390/math9020197

[14] Marappan, R., Sethumadhavan, G. Solving Graph Coloring Problem Using Divide and Conquer-Based Turbulent Particle Swarm Optimization. Arab J Sci Eng (2021). https://doi.org/10.1007/s13369-021-06323-x

[15] Bhaskaran, S.; Marappan, R.; Santhi, B. Design and Comparative Analysis of New Personalized Recommender Algorithms with Specific Features for Large Scale Datasets. Mathematics 2020, 8, 1106. https://doi.org/10.3390/math8071106

[16] Marappan, R.; Sethumadhavan, G. Complexity Analysis and Stochastic Convergence of Some Well-known Evolutionary Operators for Solving Graph Coloring Problem. Mathematics 2020, 8, 303. https://doi.org/10.3390/math8030303

[17] Marappan, R., Sethumadhavan, G. Solution to Graph Coloring Using Genetic and Tabu Search Procedures. Arab J Sci Eng 43, 525–542 (2018). https://doi.org/10.1007/s13369-017-2686-9

[18] R. Marappan and G. Sethumadhavan, "Solution to graph coloring problem using divide and conquer based genetic method," 2016 International Conference on Information Communication and Embedded Systems (ICICES), 2016, pp. 1-5, doi: 10.1109/ICICES.2016.7518911.

[19] Marappan, R., & Sethumadhavan, G. (2015). Solving graph coloring problem for large graphs. Global Journal of Pure and Applied Mathematics, 11(4), 2487-2494.

[20] Marappan, R., & Sethumadhavan, G. (2015). Solution to Graph Coloring Problem using Evolutionary Optimization through Symmetry-Breaking Approach. International Journal of Applied Engineering Research, 10(10), 26573-26580.

[21] Marappan, R., & Sethumadhavan, G. (2015). Solution to graph coloring problem using heuristics and recursive backtracking. International Journal of Applied Engineering Research, 10(10), 25939-25944.

[22] G. Sethumadhavan and R. Marappan, "A genetic algorithm for graph coloring using single parent conflict gene crossover and mutation with conflict gene removal procedure," 2013 IEEE International Conference on Computational Intelligence and Computing Research, 2013, pp. 1-6, doi: 10.1109/ICCIC.2013.6724190.
[23] R. Marappan and G. Sethumadhavan, "A New Genetic Algorithm for Graph Coloring," 2013 Fifth International Conference on Computational Intelligence, Modelling and Simulation, 2013, pp. 49-54, doi: 10.1109/CIMSim.2013.17.
[24] Raja Marappan, Gopalakrishnan Sethumadhavan, R.K. Srihari, New approximation algorithms for solving graph coloring problem – An experimental approach, Perspectives in Science, Volume 8, 2016, Pages 384-387, ISSN 2213-0209, https://doi.org/10.1016/j.pisc.2016.04.083.
[25] Raja Marappan, Gopalakrishnan Sethumadhavan, U. Harimoorthy, Solving channel allocation problem using new genetic operators – An experimental approach, Perspectives in Science, Volume 8, 2016, Pages 409-411, ISSN 2213-0209, https://doi.org/10.1016/j.pisc.2016.04.091.
[26] Raja Marappan: A New Multi-Objective Optimization in Solving Graph Coloring and Wireless Networks Channels Allocation Problems. Int. J. Advanced Networking and Applications Volume: 13 Issue: 02 Pages: 4891-4895 (2021)
[27] Raja Marappan, S. Bhaskaran, N. Aakaash, S. Mathu Mitha. (2022) Analysis of COVID-19 Prediction Models: Design & Analysis of New Machine Learning Approach. Journal of Applied Mathematics and Computation, 6(1), 121-126. DOI: http://dx.doi.org/10.26855/jamc.2022.03.013
[28] Raja Marappan, S. Bhaskaran, S. Ashwadh, H. Aathi Raj. (2022) Extraction of Drug Review Polarity Using Sentimental Analysis. Journal of Applied Mathematics and Computation, 6(2), 167-177. DOI: http://dx.doi.org/10.26855/jamc.2022.06.001
[29] Marappan, R., & Bhaskaran, S. (2022). Analysis of Network Modeling for Real-world Recommender Systems. International Journal of Mathematical, Engineering, Biological and Applied Computing, 1(1), 1–7.