International Journal of Advances in
Soft Computing and Intelligent Systems (IJASCIS)
2025, Vol 04, Issue 02, 119-127
ISSN: 3048-4987

Original Paper

A REVIEW DEVELOPMENT OF RECOMMENDATION SYSTEM USING MACHINE LEARNING

Akshay Deshpande[a*], Bharat Naiknaware[b], Vishnu Dabhade[c]

Dr. G.Y. Pathrikar College of Computer Science & Information Technology,
MGM University, Chhatrapati Sambhajinagar, India

*Email ids: [a]adeshpande341@gmail.com, [b]bnaiknaware@mgmu.ac.in,
[c]vishnudabhade60@gmail.com

ABSTRACT

E-commerce and entertainment to social media and online education, automated suggestions help to improve customer experience throughout the system. Although the earlier recommendation methods such as collaborative filtering or content-based filtering have set the precedent, these methods have been overtaken by more effective ones. The methods have been applied with deep learning, machine learning, reinforcement learning, and graph-based algorithms, which improves the prediction accuracy, reduces bias, and enhances the satisfaction of the users. The addition of contextual information, user modelling, and feedback loop considerably improves the relevance and accuracy of the recommendations is analysed. The recommendation systems try to fulfil the users with information, products, or content that resonates with the personal interests. Improvements within the real time recommendations system explainability and fairness also tries to address the ever-present bias and opacity within the systems. The additional integration of text, images, and user behavior data also helps in addressing the improvement of the system biases within the algorithms. The more intelligent and user-cantered systems that this recommendation algorithms provide will almost predict the future suggesting the systems with intelligent reasoning will be the most helpful for people. The recommendation system finds itself in a system of prediction where complexity and user reasoning are the determinants of its performance. The information system tries to predict user behaviour, and complexity of the information is based on the behaviour prediction. The presence of these recommendation heuristics, will try to predict systems that are emerging based on complexity and user content. The more the recommendation systems are used, the more intelligent they will be. The more

intelligent, the more they will save the user time and effort. The aim is to save the user excessive content. Predict systems that will keep users of recommendation systems engaged highly. Recommender System helps in increase product sales, improve user experience, improve business decision making etc. Machine learning algorithms such as K-Nearest Neighbours (KNN), Convolutional Neural Networks (CNN), SVM (Support Vector Machine) analyses interactions like user data, product features to detect patterns and make suggestions for product accordingly also improve the accuracy of recommendations.

KEYWORDS

Recommendation System, Content- based filtering, Collaborative Filtering, Hybrid Filtering, SVM, ANN, Supervised Learning.

1. INTRODUCTION

Recommender systems are the systems that are programmed to make recommendations of items to the user based on a large number of different parameters. Some advantages of such systems are e-commerce (Amazon, Flipkart), streaming services (Netflix, Spotify), social media sites (YouTube, Instagram), etc., which make predictions based on user past data, user similarities, demographic profile, search history, product features, and reviews. The behaviors patterns exhibited by user is how relevance is determined by machine learning systems. These systems analyses user behavior to understand preferences so personalized suggestions can be tailored to user needs. Also personalized suggestions are more easily attainable. Personalized suggestions also enhance the user experience and the business profits as the customers the products more. New users with fewer interactions pose recommendation challenges. Recommender systems also have the cold start problem. Scalability is also an issue if number users and number of items increase.

Automated recommendation systems have an undeniable impact on the overall individual experience of the user as it provides each user with tailored suggestions. As the relevance of data is increased to the recommendation the data is prompt central in the field of research in data science and the artificial intelligence.

2. LITERATURE SURVEY

In the article, the author Dewen Seng used the recommendation systems that is personalized with user and items user and item is the focus of the prediction. A recommendation model accuracy is improved with the combination of Neural Graph based collaborative filtering technique with item temporal relationship. The datasets such as LastFM, Douban, and Ciao are compared using a sliding window strategy for a substring range to be produced. This model improves problem of historical data recommendation as well as system does not have any prior data for recommendation. Here NGCF-ITS model produced more accurate result as compared to traditional models [1].

Xie feixiang proposed article author provide recommendation for online education due to overcome problem of information overload and appropriate resource availability. Hybrid recommendation model used which combine content based and collaborative based filtering. Meclass dataset with a total of 281 users and 1268 resources was used for implementation of algorithm. For measure the performance utility and rule model was used after experimentation F1 value for hybrid model is stable as compared to utility and rule model [2].

Ananta Kumar Das focuses on uncertainty of available bikes to solve this problem recommend the active bike presence using machine learning techniques. Real world Divvy Bike Sharing Dataset used which contains 1,013,659 trip records from 600 stations features like Trip ID and Bike ID are consider k-means unsupervised machine learning algorithm used to find clustering patterns. Elbow method used for find out optimal value for K. neural network approach, random forest and KNN classifier used for bike classification. Finally, after train and test dataset on different parameters KNN classifier shows best results value [3].

Geon-woo kim addressed video metadata to text speech data for recommendation. YouTube videos are utilized for video content analysis. The MovieLens dataset is used for keyword mapping. Popular set of object detection models called YOLO and STT used for convert voice information into contextualized information. For keyword classification TextRank algorithm is used. Movielens dataset mainly consist of movies and ratings which identifies two attributes movie ID, title and genre also user Id for identify unique user timestamp used for record user interactions. This technique enhances the quality of video content and keyword extraction helps to accuracy of recommendations [4].

Mohammad moradi mainly focus on enhance learning experience of children's age between 8-10 years using mobile technology based educational recommender system which monitoring students' activities and suggest relevant educational material. Focusing on a data set of 100 children, or "turtles" as they are called, they assigned educational material to a subset. Then, they used a Bayesian networks mathematical model to assess the effectiveness of the educational material recommender system. For data dissemination, cascade and threshold models were used. Bayesian networks are used for classification for the training and test data. In the end, educational progress of each child was tracked in real-time, making it possible to recommend educational material tailored to children's immediate learning needs [5].

In Sasmita Subhadarsinee Choudhury's paper, author discusses a multi-modal/trust-based recommender system, which incurs the cold start problem, data sparsity, and malicious attacks within the realm of film recommendations. In the system, trust and collaborative filtering techniques are used to find user similarities as it improves recommender system accuracy. The input dataset is a user-item rating matrix, which in this use case, was made of 610 users, 9,724 items, and 100,837 total ratings.

The expected user ratings of the matrix factorization were completed. For the BPNN model, users and items were randomly assigned and a deep neural network (DNN) was also used to predict ratings of an item for an active user and sign the item after a model comparison. The models were BPNN and SVD., Deep Neural Network (DNN), and DNN with Trust. BPNN achieved an accuracy of 41%, SVD reached 69%, and DNN obtained 78% hence DNN model shows highest accuracy and best suited for movie recombination [6].

Mahmoud Y. Shams proposed article author introduce crop recommendation model which mainly used for improve the crop yield performance also includes soil qualities, historical crop performance, and prevailing weather patterns for personalized recommendations system utilizes machine learning and eXplainable Artificial Intelligence (XAI) techniques to enhance agricultural decision-making. Dataset includes attributes such as location, season, production per square kilometer, area, and crop, which are essential for predicting the accurate crop to be cultivated in a specific region based on historical data. XAI-CROP compares with several prominent machine learning models, including Gradient Boosting, Decision Tree, Random Forest, Gaussian Naïve Bayes, and Multimodal Naïve Bayes. R2 value for XAICROP was 0.94152, indicating that the model explains 94.15% of the variability in the data comparatively Decision Tree model with an MSE of 1.1785, MAE of 1.0002, and R2 of 0.8942 Gaussian Naı¨ve Bayes models with an MSE of 1.2487 and 1.4123, respectively, and an MAE of 1.0015 and 1.0098, respectively Multimodal Naı¨ve Bayes model had the highest R2 value among all models, with a value of 0.77521 after comparison with all above models XAI-CROP model showed best performance for predicting yield [7].

Masayoshi Takeda’s current article recommend the furniture item by using Siamese network also classify the ambiguity between furniture styles. Rakuten dataset, which contains users' rating data for furniture items sold between 2015 and 2019 include 7,644 users and 127,819 items, totaling 225,095 ratings, with rating values ranging from 0 to 5. Matrix factorization method used for provide item preferences and Siamese network provide compatibility between furniture style. Multi-layered Siamese network improves the style similarity of multiple images. Finally, matrix factorization (MF) model achieved an average validation RMSE of 0.822 which is best suited for preference and multi-layered Siamese network improve the accuracy of recommend accurate furniture item [8].

Ning liu author focus on limitations of collaborative filtering by using sentiment analysis and matrix factorization on information reviews mainly two databases are used from amazon.com food dataset, which contains user reviews and user ratings about food and drinks and Clothing dataset, consisting of user reviews and user ratings for clothing, shoes, and jewelry source from kaggle. Latent Dirichlet allocation model used for classification of reviews. BERT model extracts the semantic information from reviews text. Finally, sentiment analysis and matrix factorization (SAMF) shows best performance then traditional algorithms LFM, SVD, and MFFR [9].

According to article author Stefanovic Pavel, based on the used social network, travel recommendations are given for different countries. To improve the accuracy of the model, it uses different ways like object detection, classification, similarity measures, and data clustering. The dataset contains 12460 records and 4 metadata records for each picture, which include the country, the complete address, the latitude, and the longitude. There are also 4679 data points with 169 different countries, with most of the photos containing most of the countries from the United States, Italy, and Spain. Object detection method used to identify user photos. Random Forest and Support Vector Machines algorithms are applied for identified visited countries by the user also Self-Organizing Maps (SOM) implemented for group similar photos and enhance the accuracy of recommendation model accuracy of 63% shows when user had visited based on their photos whereas model shows 96% accuracy when on considering additional countries visited by users that were not included in the input photos [10].

Ho yin kan author proposed flight recommendation for passengers to improve user experience in which some factors are considered such as individual preferences, flight quality, and the possibility of flight cancellations or delays. Dataset include airline ticket sales company in china comprising information from 131 passengers who purchased a total of 1273 domestic and international flight tickets. K-Means algorithm used for clustering dataset of user profile information. Deep learning techniques are used such as link prediction strategy user profiles and preferences and Convolutional Neural Network (CNN) model used for cancellation and delay information of flights for improve recommendation model. After experiment CNN model shows 95.13% average accuracy showing at least a 2.4% compared to other methods. Proposed recommender model shows accuracy value 72.31% which is better than other that is 15.6% indicating favourable performance of accuracy of recommender model [11].

Farhan ullah in this article author provide image-based recommendation for online shopping. System operates in two parts first part include feature extraction of products using random forest classifier and JPEG coefficient and in second part system identifies similar products. Amazon product image dataset, which comprises 3.5 million products categorized into 20 distinct categories. The integration of RF with Deep Learning enhances performance. The exciting image-based product recommendation. Magebased product recommendation system achieved a 75% accuracy in Phase 1 using the Random Forest (RF) classifier for product classification. When integrated with Deep Learning (DL), the accuracy improved to 84% Phase 2, the recommendation system demonstrated a remarkable 98% correct recommendations [12].

Po-han chiang proposed article author provide personalized recommendation for lifestyle mainly focus on blood pressure management using wearable data collected from the Galaxy Watch, which provided various lifestyle metrics including heart rate (HR), number of steps, walking/running speed, floors climbed, sleep duration, and sleep stages of the users. There are two methods are used for feature selection called F with Shapley-Value-based Feature Selection (RFSV) mainly focus on identify and

remove redundant and irrelevant features for BP prediction. Autoregressive Integrated Moving Average (ARIMA) model used to improve time series data [13].

3. PROPOSED MODEL

In this proposed methodology (Figure 1), various steps are including data collection, data pre-processing, feature creation, data training, data testing, result interpretation, and data visualization. There are different sources of data collection organization data, text data, and numerical data sources.

3.1. Data Collection

Data collected mainly from two sources first web scraping used for extracting data from websites. Scraping a web page involves fetching it and extracting from it. Fetching is the downloading of a page (which a browser does when a user views a page) and second source is from Kaggle dataset includes data on mobile phones from the top five most popular brands in India.

3.2. Data Pre-processing

In this pre-processing gathering Information like RAM, ROM, Display Size, Product Price etc. of using this pre-processing, identifying and correcting errors or inconsistencies in the data, such as missing values, outliers, and duplicates, handling data with different formats, structures, and semantics also some techniques are used like normalization, standardization, and discretization.

3.3. Feature Extraction

In this Feature extraction technique, we are finding the relevant information from the database. This technique generates the new features with retain information of original database and helps to improve the performance of data pre-processing. Feature extraction process plays important role in over fitting prevention, noisy data, and data redundancy.

3.4. Machine Learning Classifier

In this technique, we use machine learning classifier mainly used to solve regression problems. These classifiers predict the recommendation based on user behaviour pattern of user.

3.5. Training & Testing

Training data useful for improve the accuracy of model. Model give more effective predictions if training data amount in large scale. Completion of training data machine learning model test data. Testing data required unknown information for evaluate the model. These types of data important for reliability and accuracy of recommendations.

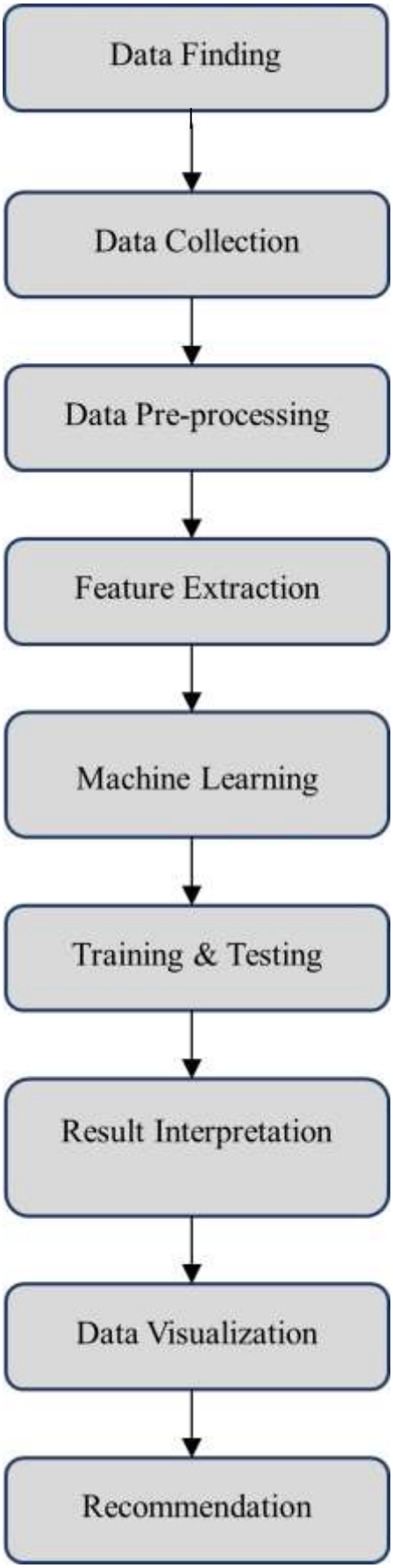

Fig. 1. Proposed Methodology.

3.6. Data Visualization

This proposed data visualization method is most important in product recommendation. The graphical representation of data helps understand, analyse, and interpret trends and insights within the data.

4. CONCLUSIONS

This review discusses thirteen recent developments in recommendation systems with machine learning across different fields, such as e-commerce, education, agriculture, healthcare, and entertainment. Incorporating metaphors will be critical, for example: today, when we need to hop on a train, we don't need a physical or printed ticket anymore, concierge in a nutshell, its smooth's out the user experience. It allows for tailored recommendations that enhance the user experience and can also help in capturing the excitement of various target audience and helping in optimizing the sales. There are various techniques used for recommendations in academia such as curated textbooks that discuss supervision, collaborations, and separation of fields of artificial intelligence. Very core techniques used for recommendations in academia are rooted in deep neural networks and transformers for better efficiency and accuracy of recommendations. In addition to multimedia content recommendations, we can also target and personalize based on wearables, and utilize systems that are adroit in contextual awareness to enhance the recommendations. There are few challenges such as the cold start challenges, slim scalability, and data isolation and bias challenges. There is also a need for advanced systems that are lightweight, and don't consume much energy, and this also extends to widgets that help gather data for federated and Artificial Intelligence systems. There are many routes of contribution for better datasets, for better hybrid routes, more calculative systems, and machines. End of the day, it rests in the hands, society to contribute and design supple interfaces that machine learning allows to incorporate. There are also machine learning systems that can help in more engaging user experiences by providing targets optimally and aiding in satisfaction of the target audience. Businesses can leverage these sophisticated technologies to increase sales and retention by tailoring their offerings to customers' preferences and actions. Nonetheless, there are still some unexplored concerns pertaining to ethics, predictive algorithm opacity, and how to formulate suggestions. Balancing concerns about bias in the training datasets and ensuring equitable possibilities to customized content are foundational to trustworthiness.

Future work needs to focus on designing structures that integrate ethical principles to promote the social utility of machine learning systems, and to render these systems just and equitable in their impact. Such structures incorporate principles of data availability, data collection, and user consent, whereby end-users are informed about the data practices in use.

Acknowledgements

Authors would like to thank to the unknown reviewers to review and provide the feedback.

Authors Contributions (Compulsory):

All authors have equally contributed.

REFERENCES

1. Seng, D., Li, M., Zhang, X., & Wang, J. (2022). Research on neural graph collaborative filtering recommendation model fused with item temporal sequence relationships. *IEEE Access*, *10*, 116972-116981, https://doi.org/10.1109/ACCESS.2022.3215161.
2. Feixiang, X. (2024). Intelligent personalized recommendation method based on optimized collaborative filtering algorithm in primary and secondary education resource system. *IEEE Access*, *12*, 28860-28872, https://doi.org/10.1109/ACCESS.2024.3365549.
3. Das, A. K., Joshi, A. M., & Dhal, S. (2020, December). A machine learning based bike recommendation system catering to user's travel needs. In *2020 IEEE 17th India Council International Conference (INDICON)* (pp. 1-6). IEEE, https://doi.org/10.1109/INDICON49873.2020.9342509.
4. Kim, G. W., Kim, W. H., Chung, K., & Kim, J. C. (2024). Extraction of meta-data for recommendation using keyword mapping. *IEEE Access*, *12*, 103647-103659, https://doi.org/10.1109/ACCESS.2024.3430375.
5. Moradi, M., Fard, K. R., & Akhlaqi, M. Y. (2022). A Recommender System Method for Children's Education Using Mobile Technology. *IEEE Access*, *10*, 123679-123696, https://doi.org/10.1109/ACCESS.2022.3223663.
6. Choudhury, S. S., Mohanty, S. N., & Jagadev, A. K. (2021). Multimodal trust based recommender system with machine learning approaches for movie recommendation. *International Journal of Information Technology*, *13*(2), 475-482, https://doi.org/10.1007/s41870-020-00553-2.
7. Shams, M. Y., Gamel, S. A., & Talaat, F. M. (2024). Enhancing crop recommendation systems with explainable artificial intelligence: a study on agricultural decision-making. *Neural Computing and Applications*, *36*(11), 5695-5714, https://doi.org/10.1007/s00521-023-09391-2.
8. Takeda, M., Ono, K., & Taisho, A. (2024). Furniture recommendations based on user propensity and furniture style compatibility. *IEEE Access*, *12*, 21737-21744, https://doi.org/10.1109/ACCESS.2024.3363459.
9. Liu, N., & Zhao, J. (2023). Recommendation system based on deep sentiment analysis and matrix factorization. *IEEE Access*, *11*, 16994-17001, https://doi.org/10.1109/ACCESS.2023.3246060.
10. Stefanovič, P., & Ramanauskaitė, S. (2023). Travel direction recommendation model based on photos of user social network profile. *IEEE Access*, *11*, 28252-28262, https://doi.org/10.1109/ACCESS.2023.3260103.
11. Kan, H. Y., Wong, D., & Chau, K. (2024). A personalized flight recommender system based on link prediction in aviation data. *IEEE Access*, *12*, 36961-36973, https://doi.org/10.1109/ACCESS.2024.3369487.
12. Ullah, F., Zhang, B., & Khan, R. U. (2019). Image-based service recommendation system: A JPEG-coefficient RFs approach. *IEEE access*, *8*, 3308-3318, https://doi.org/10.1109/ACCESS.2019.2962315.
13. Chiang, P. H., Wong, M., & Dey, S. (2021). Using wearables and machine learning to enable personalized lifestyle recommendations to improve blood pressure. *IEEE Journal of Translational Engineering in Health and Medicine*, *9*, 1-13, https://doi.org/10.1109/JTEHM.2021.3098173.