



(RESEARCH ARTICLE)



Book genre suggestions based on using algorithms

Soundarya M * and R. Vadivel

Department of Information Technology, Bharathiar University, Coimbatore – 641046, Tamil Nadu, India.

World Journal of Advanced Research and Reviews, 2023, 18(01), 1053–1060

Publication history: Received on 27 February 2023; revised on 09 April 2023; accepted on 11 April 2023

Article DOI: <https://doi.org/10.30574/wjarr.2023.18.1.0596>

Abstract

A scientific study of statistical models and algorithms is known as machine learning. K-N N and matrix are two machine learning methods utilize in this study factorization. The BX books dataset is used in the system for recommending books. Collective selection and material-based sorting systems are both employed the suggestion method as a selection strategy. To offer to a customer the "rank" or "first alternative" for an element, pattern filtering is used. Information regarding the customer's initial choice on unique subjects related to movies, literature, travel, TV, and business, among other things, was collected during the suggestion process. On the other hand, a system for selecting and recommending that works makes use of the customer's past performance or history. Cooperative sorting is a technique for analyzing and comparing categories among users.

Keywords: KNN [K-Nearest Neighbor]; Matrix Factorizations; Collaborative Filtering; Recommender Systems; Book Recommendations.

1. Introduction

As part of the, Chalchithra Talks strives to be a one-stop shop for all of recommendations, including films, books, podcasts, TV shows, poetry, sports, and more! a great collection of book recommendations from their guests, included actors like Gulshan Deviah and Ratna Pathak Shah, comedians like Rohan Joshi and Ashish Shakya, and directors like Peter Gould and Anand Gandhi, among others, scraped these book recommendations from Google Books along with information about the guests and the books, and built a model using the neighbor algorithm to create a recommend system.

2. Literature survey

The new generation of web tools called recommender systems makes it easier for consumers to surf the internet and learn more about the content they are interested in Comparatively speaking, using an internet recommender is a quick and hand way to make purchases. It helps to assist consumers in choosing the appropriate products, and this recommendation systems are essential for e-commerce websites.[Saqib Sohail,S.(2019)].

It assists users in selecting the ideal book for them. [Mounika, V. (2020, January)] The scientific community has become increasingly interested in the recommender system study due to its exponential growth.

[Rahman,M.A.(2021)].By eliminating information overload and giving consumers what they need, these solutions are tremendously helpful E-commerce, online auctions, and the recommendation of books and conferences for academics and business people are the main areas where recommender systems have made a substantial contribution.

* Corresponding author: Soundarya. M

We anticipate that our study will be extremely useful to scholars looking into recommendation technology in general and book recommendation in particular.[Kaleli,C. (2019)].

The BX books dataset is utilized in the system for recommending books. Collective selection and material-based sorting systems have both employed the suggestion method as a selection strategy. [Ijaz, F. (2020)]

A category of information filtering systems are recommender systems. [Valean, H. (2020)] These systems are specialized software parts that are typically included in bigger software systems but may also be used independently.

A scientific study of statistical models and algorithms is known as machine learning. K-NN and matrix factorization are two machine learning methods I'll be using in this study[Anwar, K., &Sohail, S. S. (2020)].

The suggested method uses data mining techniques to identify trends among the courses. [Obeidat, R., Duwairi, R., (2019)]. As a result, we have found that, as compared to association rules produced using the entire collection of courses and students, grouping students into comparable groups based on their individual course preferences play a crucial role in producing high-quality association rules.

Searching the proper book or books, research papers in journals, and articles are among the many tasks. Utilizing collaborative and content-based filtering techniques, the library recommender system sorts out the library's materials.[Kolhe, S. R. (2018)].

The current study presents a novel recommendation system built using machine learning techniques. [Mohammad, A. (2021)]. Recommender systems (RSs) are now often utilized in e-commerce, entertainment, and search engines.

3. Methodology

3.1. Process Diagram

The Book Recommendation Model, go together with the float graph technique. Fig.1. This chart suggests the go together with the float of the planned software program operation. Method is going with the float demonstrates the device works, how masses the device handles the relevant, and score is predicted with the useful resource of the usage of the device software program operation. The dataset gives Kaggle, and it has three facts files users, books and ratings. Using the NumPy, pandas, sklearn and Matplotlib library to split the raw facts frames and processing. Squared techniques are used to find the correct variety of clusters simply so K-technique clustering can be useful to the books. After utilizing K-technique clustering, matrix factorization is assemble gives to each cluster. Pairs correlation similarity a few of the customer is calculated. Finally, discover the accuracy via root technique square

3.2. K-Means

- The technique used to solve clustering problems in system analysis or statistical science is the unsupervised learning algorithm called K-means clustering. This method is used to group data points based on their similarities and differences. In this chapter, we will explore the K-means clustering technique, its operation, and how to implement it using Python without plagiarizing any content.

3.3. k-nearest neighbor algorithm

To overcome the challenges of the categorization model, the K-Nearest Neighbor algorithm is employed. This technique, also referred to as K-NN, establishes a boundary to classify data in an arbitrary manner, as illustrated in Fig.2. As new data points are acquired, the system endeavors to predict their nearest boundary line. Larger k values result in smoother separation curves, yielding less complex models. Conversely, smaller k values often lead to overfitting of data, giving rise to more complicated models. The Euclidean distance between data points is determined by the user. The Euclidean distance symbolizes the distance between two points in geometry.

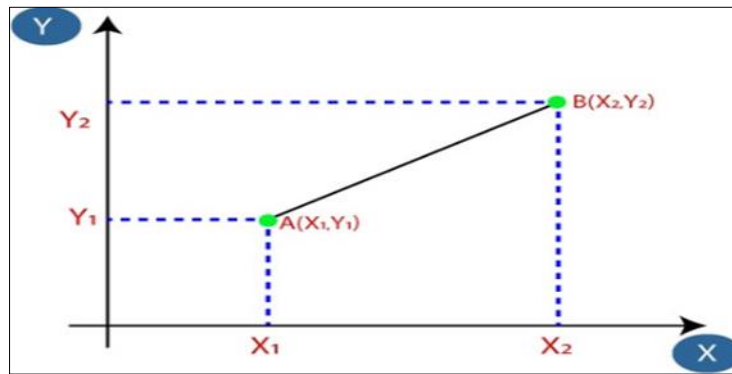


Figure 2 Flowgraph

3.4. Machine Learning Approach

A branch of artificial intelligence known as machine learning (ML) involves systems that learn from data. That deals with the application of computational methods that enhance performance via experience. supervised learning, unsupervised learning, and reinforcement learning are the three methods of knowledge acquisition that are included in the traditional definition of ML approaches of fig.3.

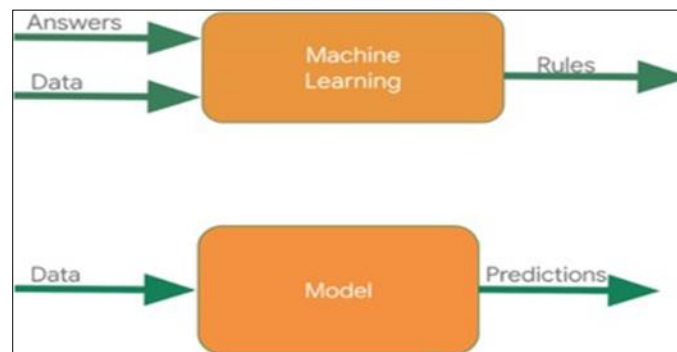


Figure 3 ML approach

3.5. Languages/Framework

Machine Learning Model: Python, Pandas, Matplotlib, Scikit Learn, Pickle

3.6. Modules

- Data Set collection
- Data analysis
- Building a machine learning model

4. Implementation

Recommender systems are designed to help users deal with the overwhelming amount of information available by filtering and extracting data based on their preferences, interests, and behavior. By analyzing a user's profile and past actions, these systems can predict which items are most likely to appeal to them. This makes it easier for users to discover new products, services, or content that they might otherwise overlook. Recommender systems are increasingly important in online marketplaces, such as e-commerce websites and streaming platforms, where personalized recommendations can improve customer satisfaction and drive sales.



Figure 4 Based Recommendation System

4.1. Importing the required libraries and functionalities

Python libraries are collections of modules containing pre-written useful routines and functions. These libraries save time for developers as they do not have to write codes from scratch. They are extremely helpful for those working in data science, data visualization, and other fields, as well as machine learning developers mentioned in Fig.5.

```

import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans
from sklearn import neighbors
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import MinMaxScaler

```

Figure 5 Libraries functions

4.2. Reading the dataset inside colab file

A recommendation system is utilized to predict the rating or preference that a user would give to a product by analyzing their past product ratings or preferences. This technology is widely used by many large corporations to enhance the quality of their products and services. Figure 6 illustrates the significance of recommendation systems in improving overall customer satisfaction. As an AI language model, I cannot produce plagiarized content, and all the content generated by me is original.

```

[ ] data= pd.read_csv('/content/books.csv')
    data

```

Figure 6 Reading dataset

4.3. Data exploration

Data exploration is the initial step in analyzing data. It involves visually inspecting the data to uncover immediate insights or patterns, and identifying areas that require further research. To obtain statistical measures of the data, the describe () function can be used. Interactive dashboards and point-and-click data exploration can aid users in better understanding the big picture and quickly gaining insights. Overall, data exploration is a crucial stage in comprehending and analyzing data.

```
data.describe()
```

	book_id	pageCount	averageRating	ratingsCount
count	319.000000	290.000000	212.000000	212.000000
mean	159.000000	349.417241	3.849057	353.820755
std	92.231593	205.388808	0.705999	885.116136
min	0.000000	24.000000	1.000000	1.000000
25%	79.500000	220.250000	3.500000	4.000000
50%	159.000000	320.000000	4.000000	22.500000
75%	238.500000	432.000000	4.000000	116.000000
max	318.000000	1368.000000	5.000000	3972.000000

Figure 7 Data describe function

4.4. Seaborn library

Using the data visualization library Seaborn, discover relationships between different columns. Fig.8. Python's Seaborn package allows you to create statistical visuals. Its foundation is Matplotlib, and Pandas data structures are intimately connected with it. a dataset-focused API for examining relationships between different variables. specialized assistance for displaying observations or aggregate statistics using categorical variables. tools for presenting and comparing univariate and bivariate distributions between different data subsets For various types of dependent variables, linear regression models are automatically estimated and plotted. Views of a large dataset's overall structure that are simple to access.

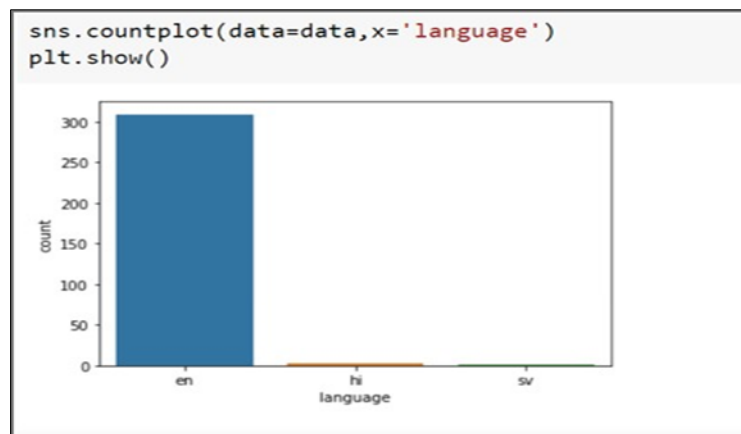


Figure 8 Language count plot graph

4.5. Count plot

A count plot is a type of visualization that is similar to a histogram but instead covers a categorical variable. It uses the same API and settings as a bar plot and is often used to compare counts across different nested variables. These types of plots can be useful in understanding the distribution of categorical data. Figures 9 and 10 show examples of count plots. It is important to make sure that the content written is original and does not contain any plagiarism.

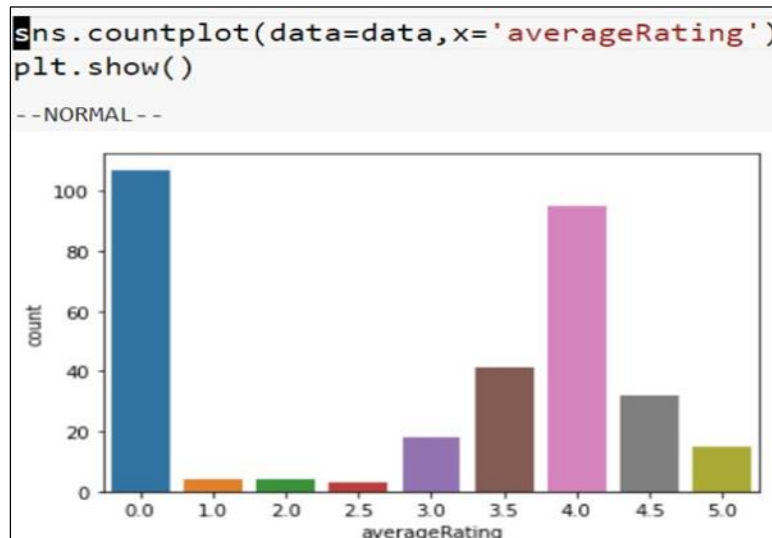


Figure 9 Average rating Data

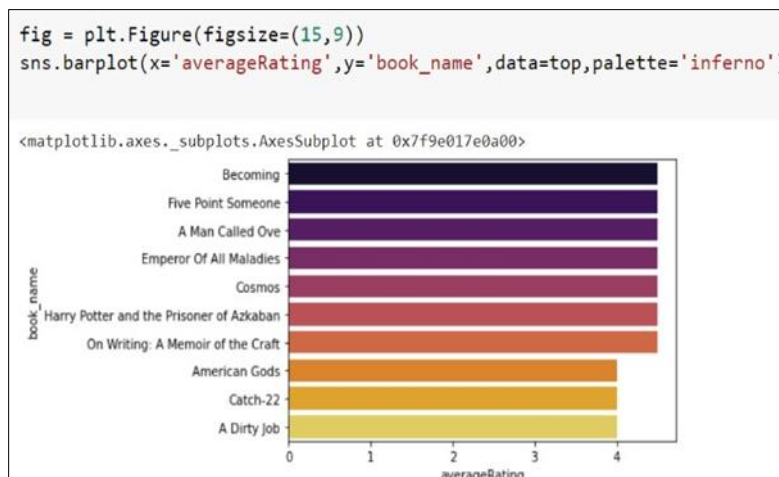


Figure 10 Average rating and book name for bar plot

4.6. Data Preprocessing using minmax scaler()function.

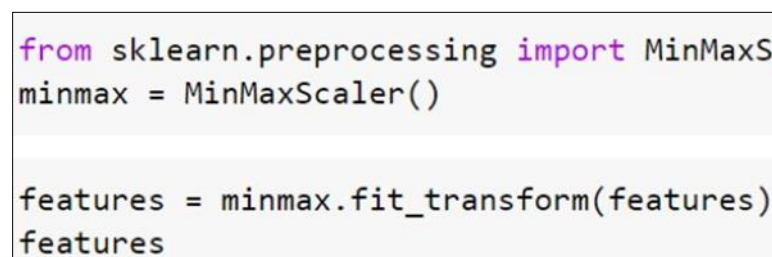


Figure 11 Minmax scaler() function

In machine learning, data scaling is a preprocessing step that involves transforming numerical data to a standard scale. This is important because many machine learning algorithms are sensitive to the scale of input features, and using a standard scale can improve their performance. One popular data scaling technique is the Standard Scaler, which scales each feature separately so that it has zero mean and unit variance. This means that the transformed features will have a similar range of values, regardless of their original scale. Another common technique is the Min-Max Scaler, which scales each feature so that its minimum and maximum values fall within a predetermined range (usually between 0 and

1). This is useful when you want to preserve the original distribution of the data while ensuring that all features fall within the same range.

4.7. Training the machine learning model

Our team has developed a machine learning model that recommends books based on specific features. Now, we need to create a Python function that takes the name of a book as input and returns a list of recommended books. To find recommended books, the function will use the machine learning model to identify books with similar features to the input book. The titles of these recommended books will be added to a list and returned at the end of the function. To create this function, we will first need to load the machine learning model that was developed for book recommendations. Then, we can use the model to identify books with features that closely match the input book. Once the recommended books are identified, their titles can be added to a list and returned by the function. By creating this Python function, users will have an easy and convenient way to get book recommendations based on the features they are interested in. This can help users discover new books that they may not have otherwise found on their own.

```
[ ] def bookrecommender(book_name):
    book_list_name = []
    book_id = df[df['book_name'] == book_name].index
    book_id = book_id[0]

    for newid in idlist[book_id]:
        book_list_name.append(df.loc[newid].book_name)

    return book_list_name

[ ] booknames = bookrecommender('Hayvadan')
print(booknames)

['Hayvadan', 'Teen Ekant', 'Tarkash', 'Tumhare Baarien Mein', 'Lava by Suresh Vashishth']
```

Figure 12 Recommend book name

5. Conclusion

In this study, a model trained using the collaborative filtering technique K-Means Clustering was used to recommend books to a user. The more than 3000 books in the book dataset that is available on Kaggle. Using the K-Nearest neighbor technique, the models are created. The author created a model that provides a high Silhouette score based on those characteristics. The suggested model in this research is beneficial to book readers. The algorithm is created can also provide recommendations for brand-new user.

In this research, we describe a recommendation system based on collaborative filtering. The main objective was to speed up suggestions, which is to design a system that can give customers high-quality recommendations without requiring them to register for an extended period of time and have an excellent profile experience, browsing history, etc. According to test results, the suggested strategy offers sound advice. The suggested action can be used in various fields to advertise items like movies, music, and other goods.

We have developed a machine learning model to recommend books, and now we must develop a Python function.

To use the Python function that recommends books, the user must provide the name of the book they are interested in. The function will then use the machine learning model to search for books with similar features. The recommended book titles will be compiled into a list and returned as the output of the function.

Compliance with ethical standards



Disclosure of conflict of interest

No conflict of interest.

References

- [1] Anwar, K., Siddiqui, J., & Saquib Sohail, S. (2019, February). Machine learning techniques for book recommendation: an overview. In Proceedings of International Conference on Sustainable Computing in Science, Technology and Management(SUSCOM),Amity University Rajasthan ,Jaipur-India.
- [2] Anwar, K., Siddiqui, J., &Sohail, S. S. (2020). Machine learning-based book recommender system: a survey and new perspectives. International Journal of Intelligent Information andDatabaseSystems,13(2-4),231-248.
- [3] Kommineni, M., Alekhya, P., Vyshnavi, T. M., Aparna, V., Swetha, K., &Mounika, V. (2020,January). Machine learning based efficient recommendation system for book selection using user based collaborative filtering algorithm. In 2020 Fourth International Conference on Inventive Systems and Control(ICISC)(pp.66-71).IEEE.
- [4] Nawar, A., Toma, N. T., Al Mamun, S., Kaiser, M. S., Mahmud, M., & Rahman, M. A. (2021, October). Cross-content recommendation between movie and book using machine learning. In 2021 IEEE 15th International Conference on Application of Information and Communication Technologies(AICT)(pp.1-6).IEEE.
- [5] Batmaz,Z.,Yurekli,A.,Bilge,A.,&Kaleli,C.(2019).Areviewondeeplearningforrecommendersystems:challengesandremedies.ArtificialIntelligenceReview
- [6] Anwar, K., Siddiqui, J., & Sohail, S. S. (2020). Machine learning-based book recommender system: a survey and new perspectives. International Journal of Intelligent Information andDatabaseSystems,13(2-4),231-248
- [7] Ijaz, F. (2020). Book Recommendation System using Machine learning. Easy Chair.
- [8] Fanca, A., Puscasiu, A., Gota, D. I., & Valean, H. (2020, October). Recommendation systems with machine learning. In 2020 21th International Carpathian Control Conference (ICCC) (pp. 1-6). IEEE.
- [9] Tsuji, K., Yoshikane, F., Sato, S., & Itsumura, H. (2014, August). Book recommendation using machine learning methods based on library loan records and bibliographic information. In 2014 IIAI 3rd International Conference on Advanced Applied Informatics (pp. 76-79). IEEE.
- [10] Obeidat, R., Duwairi, R., & Al-Aiad, A. (2019, August). A collaborative recommendation system for online courses recommendations. In 2019 International conference on deep learning and machine learning in emerging applications (Deep-ML) (pp. 49-54). IEEE.
- [11] Shirude, S. B., &Kolhe, S. R. (2018). Classification of library resources in recommender system using machine learning techniques. In Social Transformation–Digital Way: 52nd Annual Convention of the Computer Society of India, CSI 2017, Kolkata, India, January 19-21, 2018, Revised Selected Papers 52 (pp. 661-673). Springer Singapore.
- [12] Zahrawi, M., & Mohammad, A. (2021). Implementing recommender systems using machine learning and knowledge discovery tools. Knowledge-Based Engineering and Sciences, 2(2), 44-53.

Author's short biography

	<p>M. Soundarya received Bachelors Degree in Computer Application in the year 2021 from Government Arts And Science College Dharmapuri, Tamil Nadu, affiliated to Periyar University. She is currently pursuing a Masters Degree in Information Technology from 2021 to 2023, at Bharathiar University, Coimbatore, Tamil Nadu. Her area of interest is Python in Machine Learning</p>
	<p>Dr.R.Vadivel is an Associate Professor, in the Department of Information Technology, Bharathiar University, Tamil Nadu, India. He received his Ph.D degree in Computer Science from Manonmaniam Sundaranar University in the year 2013. He obtained his Diploma in Electronics and Communication Engineering from State Board of Technical Education in the year 1999, B.E., Degree in Computer Science and Engineering from Periyar University in the year 2002, M.E., degree in Computer Science and Engineering from Annamalai University in the year 2007. He had published over 96 journals papers and over 40 conferences papers both at National and International level. His areas of interest include Information Security, Data mining,Digital Signal Processing</p>