International Science and Technology Journal
المجلة الدولية للعلوم والتقنية
Volume 36 العدد
Part 2 المجلد
المجلة الدولية للعلوم والتقنية
International Science and Technology Journal
ISTJ

http://www.doi.org/10.62341/jbac3060

Received	2025/06/03	تم استلام الورقة العلمية في
Accepted	2025/06/29	تم قبول الورقة العلمية في
Published	2025/06/30	تم نشر الورقة العلمية في

A Content-Based Recommender System for News Articles

Jazya Moftah Ahmad Amshaher[1], **Basma Emhamed Ali Dihuom**[2]
Aziza Ehmaid I.Omar[3]

1-University of Sirt, Libya,
2- University of Jafara Tripoli, Libya,
3- University of Sirt, Libya
jazamoftah@su.edu.ly

Abstract:
Recommendation systems are a type of artificial intelligence that are used to predict what a user may want to watch, read, or purchase next. They are designed to personalize the user experience by suggesting items that are relevant and appealing to the individual, and are often found on websites and apps. There are several types of recommendation systems, including collaborative filtering, content-based filtering, and hybrid systems, which use different approaches to make recommendations. These systems have become an important tool for businesses and organizations in a variety of industries, and have greatly impacted the way that users interact with the internet. This article provides an overview of recommendation systems, including their types, applications, and challenges. It also discusses the potential benefits and limitations of these systems. We also explain a content-based recommendation system that helps readers to the relevant articles to them. In addition, we have implemented a content-based recommendation system for news media that helps readers find their relevant articles fast. The recommendation system is implemented in Python and used a publicly available Data Set to build and train the model.

Keywords: Recommender system, content-based, recommendation, Collaborative Filtering, Hybrid Recommendation.

International Science and Technology Journal
المجلة الدولية للعلوم والتقنية
Volume 36 العدد
Part 2 المجلد
http://www.doi.org/10.62341/jbac3060

نظام التوصية القائم على المحتوى لمقالات الأخبار

جازية مفتاح أحمد أمشهر[1]، بسمة امحمد علي ديهوم[2]، عزيزة احميد اشتيوي[3]

1 جامعة سرت، ليبيا

2 جامعة الجفارة، طرابلس، ليبيا

3 جامعة سرت، ليبيا

jazamoftah@su.edu.ly

الملخص:

أنظمة التوصيات هي نوع من الذكاء الاصطناعي يُستخدم للتنبؤ بما يرغب المستخدم في مشاهدته أو قراءته أو شرائه . صُممت هذه الأنظمة لتخصيص تجربة المستخدم من خلال اقتراح مواد ذات صلة ، وغالبًا ما توجد على مواقع الويب . هناك أنواع عديدة من أنظمة التوصيات، بما في ذلك التصفية التعاونية، والتصفية القائمة على المحتوى، والأنظمة الهجينة، والتي تستخدم أساليب مختلفة لتقديم التوصيات. أصبحت هذه الأنظمة أداة مهمة للشركات والمؤسسات في مختلف القطاعات، وقد أثرت بشكل كبير على طريقة تفاعل المستخدمين مع الإنترنت. يمكن استخدام أنظمة التوصيات في تنظيم الأخبار، مما يُسهّل القراءة والتنقل عبر الصحف. باستخدام تاريخ استهلاك المستخدمين للعناصر وملفات تعريفهم أو مصادر المعرفة الأخرى، تُمكّن هذه الأنظمة من تخصيص تجربة المستخدم، مما يُقلل من عبء المعلومات الذي نواجهه حاليًا. تُقدم هذه المقالة لمحة عامة عن أنظمة التوصية، بما في ذلك أنواعها وتطبيقاتها وتحدياتها. كما تناقش الفوائد والقيود المحتملة لهذه الأنظمة. كما نشرح نظام توصية قائم على المحتوى يُساعد القراء على الوصول إلى المقالات ذات الصلة. بالإضافة إلى ذلك، طبّقنا نظام توصية قائم على المحتوى لوسائل الإعلام الإخبارية يُساعد القراء على العثور على مقالاتهم ذات الصلة بسرعة.تم تنفيذ نظام التوصيات باستخدام بايثون، واستُخدمت مجموعة بيانات متاحة للعامة لبناء النموذج وتدريبه.

الكلمات المفتاحية: نظام التوصية، نظام التوصية القائم على المحتوى، التصفية التعاونية، التوصية الهجينة.

Introduction

Recommendation systems are a type of artificial intelligence that are used to predict what a user might like, based on their past behavior and preferences. These systems are used by a wide range of businesses, and help users select appropriate services from a wide

http://www.doi.org/10.62341/jbac3060

range of choices(Patel et al.,2023).There are several different types of recommendation systems, including collaborative filtering, content-based filtering, and hybrid systems. Collaborative filtering systems make recommendations based on the behavior of other users who have similar interests, while content-based filtering systems use the characteristics of the item being recommended to make predictions. Hybrid systems combine both of these approaches to make recommendations.

Recommendation systems can be built using a variety of algorithms, including matrix factorization, singular value decomposition, and nearest neighbor algorithms. These algorithms analyze user behavior and item characteristics to generate recommendations that are tailored to the individual user. One of the key benefits of recommendation systems is their ability to improve the user experience. By providing personalized recommendations, these systems can help users discover new products or content that they are likely to enjoy. This can lead to increased engagement and customer loyalty. Recommendation systems can be implemented in a variety of ways, including through the use of machine learning algorithms, natural language processing, and data mining techniques. These algorithms analyze user data, such as their past purchases or ratings of items, to make recommendations.

There are also several challenges to implementing recommendation systems. One challenge is the need for large amounts of data in order to make accurate recommendations. Another challenge is the potential for bias in the recommendations, which can occur if the data used to train the system is not diverse or representative of all users. Despite these challenges, recommendation systems have become an integral part of many online platforms and are likely to continue to be important in the future. They provide a valuable service to both businesses and users, and have the potential to greatly enhance the user experience on websites and apps. News recommender systems represent part of the digital media landscape (Mitova,.2023), This article gives an overview of recommendation systems, including their various types, the fields in which they are used, and the issues they may face. It also examines the potential advantages and drawbacks of these systems. Additionally, the article explains how a content-based recommendation system can assist readers in finding articles that are relevant to their interests. In particular, we have developed a content-based recommendation system for news media that enables readers to quickly find relevant

http://www.doi.org/10.62341/jbac3060

articles. The system, implemented in Python, utilized a publicly available data set to build and train the model.

Literature review

Recommender Systems (RS) play important role in enhancing user experiences by providing suggest items or content(Raza et al .,2024),(Alfaifi,2024),Some of the major works in recommendation systems research include:

- **Collaborative filtering:**

Collaborative filtering, which involves making recommendations based analysis the past behaviour of users with similar preferences(Phalle et al.,2024) Since then, there have been many advances in collaborative filtering algorithms, including matrix factorization techniques and deep learning approaches.

- **Content-based filtering,** which involves making recommendations based on the characteristics of the items being recommended, was first introduced by Gerard Salton and Chris Buckley (Salton et al.,2024). Since then, there have been many advances in content-based filtering algorithms, including the use of natural language processing and machine learning techniques.(Sina et al,2024)They proposed a system for providing instant news recommendations to users based on their previous site visits and current context. The system addresses over-specialization, uses implicit collaboration (recommending categories rather than items), and can interact with users. The system is characterized by its preference for novelty over relevance. The proposed approach has been compared with several modern alternatives and has proven successful.
- **Hybrid systems:** Hybrid recommendation systems, which combine collaborative and content-based filtering, were first introduced in the early 2000s. These systems have become increasingly popular in recent years, By hybrid recommender systems leverage multiple sources of data to enhance the quality of recommendations (Khtira et al.,2025)
- **Deep learning approaches:** In recent years, deep learning approaches have been applied to recommendation systems, with the goal of improving the accuracy and diversity of

http://www.doi.org/10.62341/jbac3060

recommendations. These approaches have shown promising results and are expected to continue to be an active area of research (Da'u et al.,2020)(Zhang et al.,2019)which is basically ready to consider both the general ratings and feature-level opinion values (as extracted from textual reviews) to perceive reviewers' inclination homogeneity. In the examination, they tried the proposed recommender algorithm with two true datasets. More notably, (Mouhiha et al.,2024) , author improved accuracy, the proposed hybrid approach combines both techniques in various ways to increase precision. While content-based recommendation based on the similarity of the attributes ,collaborative recommendations are based on the similarity of the users and implemented collaborative filtring through deep neural networks. In (Belhaouari et al.,2023), solve the problem of techniques that lead to unfair recommendations due to biased data, they proposed using content-based filtering (CBF) and collaborative filtering (CF) with semantic relationships to capture relationships for readers in a digital library.Implementing recommendation systems using Python is not new. Surprise is a Python library that is used to build as well as analyze rating prediction algorithms . In addition, OpenRec (Guo et al.,2023)in Python supports neural-network inspired algorithms, while implicit1 specializes in implicit feedback recommendation and Light FM implements a hybrid algorithm based on matrix factorization. Other popular recommendation libraries with similar functionalities from other languages include Lib Rec in Java or My Medi aLite in C# (Singh et al.,2020).

The recommendations systems for news

The recommendation of news articles is a hard task because of a highly dynamic environment, which leads several challenges, e.g. frequent changes in the set of news articles, set of users, rapid changes in user's preferences. Recommendation algorithms must be able to process continuous incoming news streams in real-time. The complex requirements of news recommendations based on some relevancy among news articles that best fulfill the user's requirements lead to many rich research scenarios and make it more interesting. Unique characteristics of news items are as follows:

http://www.doi.org/10.62341/jbac3060

- **Large volume:**
  Unlike other types of web objects, news articles tend to be in flood within a short period of time, requiring much more computation for recommendation.
- **Unstructured format:**
  The unstructured format of a news story is more difficult to analyze than other objects with structured properties.
- **Recency:**
  News items typically have short self lives. For example, few sports fans will be concerned with the past breaking scores of two days ago. In contrast, the shelf lives of products and movies extend several months or even years.
- **Entity preference:**
  Most news articles describe the occurring of specific events. Online news readers tend to be interested in the named entities of information like what, when, who and where the event occurred.
- **News selection and Ranking:**
  The interest of news articles with respect to a user is regressive. he clicks the first piece of news he/she interested in, the interest of the user on a news item may change if he chooses other news items also.
- **Scalability:**
  The scalability of news recommendation requires elegant algorithms to efficiently deal with large news corpus. Map Reduce, a programming model prototype by Google, aims to support distributed computing on large datasets with clusters of computers and has been widely used in many data mining and Machine learning tasks.

The main steps of a recommender system for news

- **Data Loading and Cleaning**

Load the data set from a file or database.
Clean the data by removing any missing or invalid values.
Normalize the data if necessary.

- **Data Analysis**

Explore the data to understand its structure and characteristics.
Identify any patterns or relationships in the data.

http://www.doi.org/10.62341/jbac3060

- **Model Building**

Select a suitable model for the recommendation system (e.g., similarity).
Train the model on the data.

- **Getting Final Recommendations**

Accept input from the user (e.g. a list of their favorite articles).
Use the trained model to generate recommendations for the user based on their input. We have used Python to implement the recommendation system for news. Figure 1 shows the main steps of the implemented recommendation system.

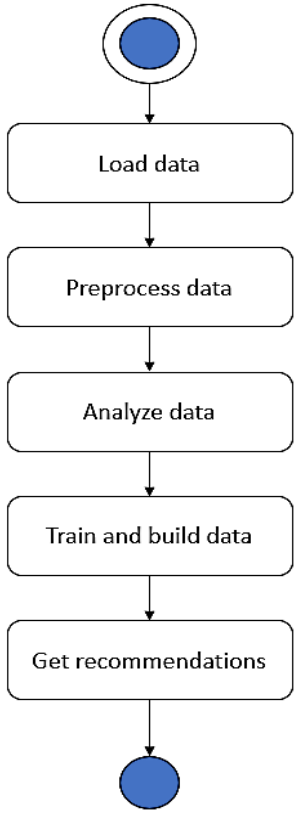

Figure 1. The main steps of implemented recommendation system

Evaluation

In this section, we evaluate the implemented recommendation system for news. In particular, we introduce the Data Set. Then, we present results for queries that we used to evaluate the implemented recommendation system.

http://www.doi.org/10.62341/jbac3060

Data Set

The data set used for building this recommender system is a news category data set which is a collection of news articles that have been labeled with one or more categories or topics. These categories might include politics, sports, entertainment, business, technology, and so on. News category datasets are often used to train machine learning models to classify news articles into different categories or to analyze trends in media coverage.

the data set used consists of 25,000+ news articles collected from a publicly available data set online, including news aggregators, media outlets, and academic institutions. Some news category datasets may be freely available, while others may require a fee or subscription to access.

The data set includes following information:

- Article text:

The full text of the article is often included in a news category data set.

- Categories:

The categories or topics that the article belongs to are typically labeled in the data set.

- Timestamps:

News category datasets often include timestamps indicating when the article was published or updated.

- Authors:

Many news category datasets include information about the authors of the articles.

- Headlines:

The headline of the article is often included in a news category data set. In table 1 we show the categories that are exist in this Data Set as well as the number of articles for each category.

TABLE 1 . The data set categories and article counts.

Category	Number	Category	Number
U.S. NEWS	1377	FOOD & DRINK	114
COMEDY	782	MEDIA	512
PARENTING	114	QUEER VOICES	608
WORLD NEWS	1917	HOME & LIVING	125
CULTURE & ARTS	44	WOMEN	407
TECH	95	BLACK VOICES	531
SPORTS	590	TRAVEL	103

http://www.doi.org/10.62341/jbac3060

ENTERTAINMENT	3373	MONEY	49
POLITICS	6932	RELIGION	110
WEIRD NEWS	383	LATINO VOICES	121
ENVIRONMENT	121	IMPACT	140
EDUCATION	58	WEDDINGS	2
CRIME	377	COLLEGE	3
SCIENCE	76	PARENTS	105
WELLNESS	118	ARTS & CULTURE	40
BUSINESS	188	STYLE	99
STYLE & BEAUTY	165	GREEN	90
TASTE	44	HEALTHY LIVING	86

Results

In this section, we show the top ranked recommendations based on the inputs that is given. In the following figure, you see the input that is given, which is in our case “Biden Says U.S. Forces Would Defend Taiwan If China Invaded”. This input is from cluster Politics. Below this input, you see the output, which is the top ranked retrieved articles. The example, where we used an input from Sport category. As we can observe that most of the retrieved articles are related to sport. The input that we used is “Carlos Alcaraz Wins U.S. Open 1st Slam Title, Top Ranking”

```
getting recommendation for: Carlos Alcaraz Wins U.S. Open For 1st Slam Title, Top Ranking
569      Rafael Nadal Wins French Open For 14th Time, D...
13576    Tennis Legend Althea Gibson To Be Honored With...
3183     Naomi Osaka Beats Jennifer Brady To Win Second...
4528     Serena Williams And Daughter Olympia Are Tenni...
12513    Trevor Noah Rips Tucker Carlson For Saying Gun...
16244    Seth Meyers Gives Donald Trump Advice About Hi...
12493    Ohio Legislator Says Students Should Be Able T...
16636    Serena Williams Covers Vogue With Her Baby Gir...
3203     Serena Williams Talks Preparing To Face 'Incre...
10226    Robbin' Season's True Criminal Is Revealed In ...
```

Figure 2. Recommendations related to Sport

Note that the reported results in Figure 2 are based on one single input. That means, if the recommendation system is integrated into website news, where the articles that have been navigated by the user are considered, we are confident that the results will even be improved.

We used a sample articles from the dataset to evaluate the effectiveness of proposed system, and measured how well the recommended articles matched the actual categories.Their

http://www.doi.org/10.62341/jbac3060

experimental analysis indicated that the achieved results ,Precision was 78%, meaning that most recommended articles were relevant and Recall was 74%, indicting the system successfully retrieved the majority of the relevant articles.

Conclusion

Recommendation systems are artificial intelligence that predict what a user is likely to want to watch, read, or purchase next. They personalize the user experience by suggesting items that are relevant and appealing to the individual, and can be found on websites and apps. There are three main types of recommendation systems: collaborative filtering, content-based filtering, and hybrid systems. These systems use different methods to make recommendations and have become crucial for businesses and organizations in various industries. They have significantly influenced how users interact with the internet.

Recommender systems open new opportunities of retrieving personalized information on the Internet. It helps to alleviate the problem of information overload which is a very common phenomenon with information retrieval systems and enables users to have access to products and services which are not readily available to users on the system .

This article offers a comprehensive review of recommendation systems, including their various types, uses, and issues. It examines the pros and cons of these systems, and also describes a content-based recommendation system that assists readers in finding articles that are relevant to their interests. In addition, we developed a content-based recommendation system for news media using Python, which helps readers quickly find relevant articles. The recommendation system was trained on a publicly available data set.

References

Alfaifi,Y,2025,Recommender Systems Applications: Data Sources, Features, and Challengesmdp,https.//www.mdpi.com/2078-2489/15/10/660.

Amal,K,Khtira,A,&Sabiri,B,2025,Hybrid Quality-Based Recommender Systems: A Systematic Literature Review,Journal of Imaging,https://www.mdpi.com/journal/jimaging.

Belhaouari,M,Fareed,A,&Hassan,S,2023,A Collaborative Filtering Recommendation Framework Utilizing Social

http://www.doi.org/10.62341/jbac3060

Networks.Machine LearningwithApplications.https://www.researchgate.net/publication/364626538_A_Collaborative_Filtering_Recommendation_Framework_Utilizing_Social_Networks.

Da'u, A, & Salim, N,2020, Recommendation system based on deep learning methods: a systematic review and new directions. Artificial Intelligence Review, 53(4),pp. 2709-2748.https://link.springer.com/article/10.1007/s10462-019-09744-1.

Guo, G, Zhang, J, Sun, Z, & Yorke,N,2023, LibRec: A Java Library forRecommenderSystems.https://www.cs.cornell.edu/~ylongqi/paper/YangBGHE18.

Raza,S,Rahman,M, Kamawal,S,& Toroghi,M,2024,A Comprehensive Review of Recommender Systems: Transitioning from Theory to Practice.Mizanur Rahman National Institute of Cancer Research andHospital.HTTP.//WWW.researcher.net/publication/382364052_A_Comprehensive_Review_of_Recommender_Systems_Transitioning_from_Theory_to_Practice.

Phalle,M,& Bhushan,S,2024,Content Based Filtering And Collaborative Filtering: A Comparative Study.Journal Of Advanced Zoology.45(S4):96100.HTTP://WWW.researchgate.net/publication/378841543_Content_Based Filtering And_Collaborative Filtering_A_Comparative_Study.

Patel,D, Patel,m & Chauhan,U,2023,Recommendation Systems: Types, Applications, and Challenges.International Journal of Computing and Digital Systems,Sys.13,No.1,https://www.researcher.net/publication/370498330_Recommendation_Systems_Types_Applications_and_Challenges.

Salton,G,& Buckley,C,2024,Improving retrieval performance by relevance feedback. Journal of the American Society for Information Science, 41,pp. 288-297.https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291097-

http://www.doi.org/10.62341/jbac3060

4571%28199006%2941%3A4%3C288%3A%3AAID-ASI8%3E3.0.

Sina,B,Eliza,M &Edina,S,2024,User Perceptions of News Recommender Systems and Trust in Media Outlets: A Five-Country Study.Zurich Open Repository and Archive.https://www.zora.uzh.ch/id/eprint/263669/9/User_Perceptions_of_News_Recommender_Systems_and_Trust_in_Media_Outlets__A_Five_Country_Study.pdf.

Singh, R,Maurya, S,Tripathi, T, Narula, T,& Srivastav, G,2020,Movie recommendation system using cosine similarity and KNN, International Journal of Engineering and Advanced Technology,9(5),pp.556559.https://www.ijeat.org/wpcontent/uploads/papers/v9i5/E9666069520.

Mitova,E,Blassnig,N,Strikovicb,S,&Urmanc,2023,News recommender systems: a programmatic research review,ANNALS OF THE INTERNATIONAL COMMUNICATION ASSOCIATIONA,, VOL. 47, NO. 1, 84113.https://pure.uva.nl/ws/files/181532079/News_recommender_systems.pdf .

Mouhiha,M,Oualhaj,O.&Mabrouk,B,2024,Combining Collaborative Filtering and Content Based Filtering for Recommendation Systems. 2024 11th International Conference on Wireless Networks and Mobile Communications(WINCOM).https://www.researchgate.net/publication/383811239_Combining_Collaborative_Filtering_and_Content_Based_Filtering_for_Recommendation_Systems .

Zhang,S,Yao,L,Sun,A,&Tay,Y,2019,Deep learning based recommender system: A survey and new perspectives. ACM Computing Surveys (CSUR), 52(1), pp.1-38.https://dl.acm.org/doi/10.1145/3285029.