Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

A Web-Based Academic Article Recommendation System: Survey

Mohammed Jabardi, Ayat Abbas Fadhil*

University of Kufa, Faculty of Education, Department of Computer science/ Iraq, E-mail: mohammed.naji@uokufa.edu.iq

*Corresponding author E-mail: aayat.alkhazali@gmail.com

https://doi.org/10.46649/fjiece.v3.2.22a.1.6.2024

Abstract. *The rise of the internet and intelligent gadgets has increased platform traffic, collecting data to detect user preferences. Researchers need help finding relevant material due to the rapid proliferation of academic papers across fields. Academic article recommendation systems (ARSs) help solve this problem by suggesting articles based on research interests and needs. Academic article recommendation algorithms help navigate the vast scholarly literature. Researchers, students, and academics utilize them to find articles that match their interests and study emphasis in enormous databases. This review covers recommendation algorithms, data sources, assessment metrics, and user interfaces in ARS research. This survey also examines new trends and research directions, such as advanced machine learning and semantic analysis.*

Keywords: *recommendation system, natural language processing, Content-Based filtering, Collaborative Filtering, text mining, deep learning, machine learning*

1. Introduction

The rise of the internet and smart gadgets has increased traffic on platforms, collecting data to detect user preferences[1]. This data is used in recommendation systems, which can use click data and implicit visit information to represent user behaviour patterns. Cognitive-based recommendation systems are gaining popularity for quickly reflecting user preferences.[2] Many online platforms implement recommendation systems to generate tailored recommendations depending on previous behaviour, preferences, and similar users based on past behaviour. [3] [4]. User experience should be smooth and personalized, making it easy to access relevant content and goods and improving platform interaction Academic article recommendation algorithms help navigate the vast scholarly literature.[5] Researchers, students, and academics utilize them to find articles that match their interests and study emphasis in enormous databases. [6] The system recognizes data patterns. These patterns can be user-driven, such as preferences towards a topic, author, or type of article, or content-driven, like text similarity or article context [7]. Machine learning and AI have improved these systems' suggestion accuracy and sophistication.

1.1 Recommendation system (RS) model

Information filtering recommendation systems give users personalized item recommendations in a service environment that can store or gather data[8]. In recommendation systems, information filtering is personalized to the user's tastes or suggests only valuable items. There are three model to recommendation system.

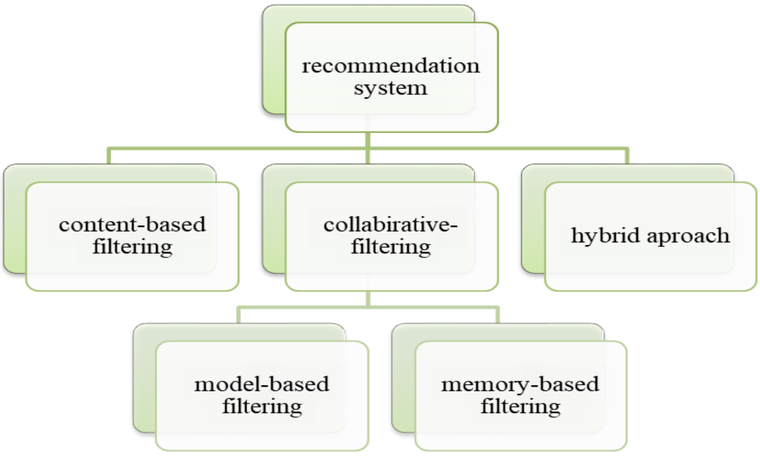


Fig.1 Recommendation system model

1.1.1 Content-Based (CB) Filtering

The Content-Based filtering technique limits recommendations to closely related items, limiting its ability to suggest new items [9]. Its drawback is that users cannot access varied stuff. Due to these constraints, this strategy is mostly utilized in services that recommend things or text data that are easy to recommend based on item and user profile information.

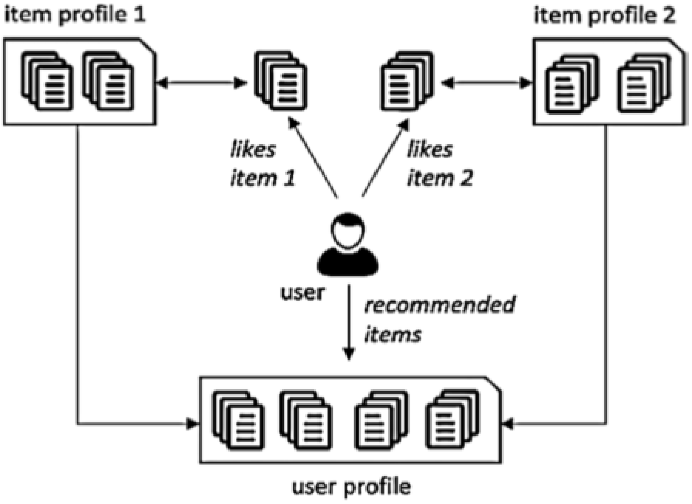


Fig.2 content-based recommender system [10]

This work introduces a Content-Similarity Author Recommendation (C-SAR) model designed to identify related academic papers based on the similarity of their titles. The model employs a Gated Recurrent Unit (GRU) for assessing document similarity alongside the Apriori algorithm for mining frequently occurring document sets within similar groups. It utilizes the AAN, integrating advanced deep learning techniques. However, it faces limitations related to memory and time constraints during implementation, which could be mitigated by leveraging cloud computing resources and high-performance computing systems.[11]
Predict latent components from text data in multimedia resources using a convolutional neural network (CNN)-based content-based approach provided here. The model uses the latent factor model for output; it employs the language model for input. The model is solved via the split Bregman iteration approach. Cai-Nicolas Ziegler of the Book-Crossing virtual book community gathered the data set in 2004. Using MAE

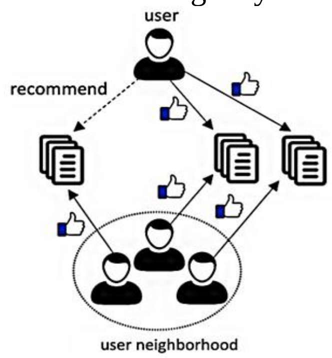
and RMSE measures, the model's accuracy is evaluated as 2.6032 for MAE and 3.3841 for RMSE. CBCNN solves start-cold difficulties rather remarkably effectively.[12]

The algorithm prioritizes appropriate conferences or journals using a web crawler to update the training set and learning model. It uses a hybrid approach based on chi-square feature selection and softmax regression, achieving an accuracy of 61.37% in 5 seconds, and F-measure is 0.23. The described strategy is used to recommend Computer science papers. Consider that only paper abstracts are recommended. The limitation here is that Accuracy and F-measure are insufficient but can be increased.[13]

1.1.2 Collaborative Filtering (CF)

Collaborative Filtering uses evaluation data to build a user's preference database and recommend things that match their taste.[9]

The collaborative techniques are divided into Memory-based and model-based; again, the memory-based collaborative filtering may be user- or item-based.[10]



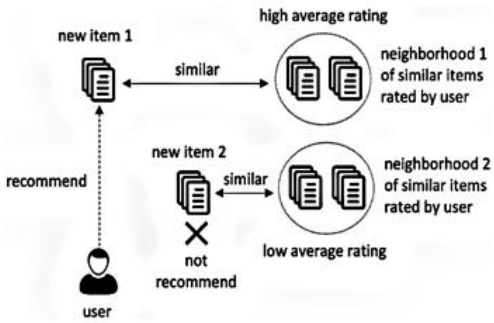


Fig.3 user-based collaborative filtering **Fig.4 item-based collaborative filtering [10]**

They propose four indirect research article evaluation methods: keyword uniqueness, text complexity, citation analysis, and scientific quality. Originality, clarity, relevance, and value are assessed to improve article suggestions. It describes how to locate, extract, and measure locating, extracting, and measuring articles better than standard methods. The dataset that is used is Scopus. Experimental results demonstrate that this feature extraction speeds article filtering. These traits match user preferences, which is vital for collaborative filtering. These evaluation approaches are proven useful by accuracy, RMSE, Precision, Recall, and F1-score (96.95, 3.04, 95.93, 96.03, 95.98 [14]

The study models observed discussions, ranks prospective quotations by conversation-quotation similarity, and optimizes recommendation performance using BERT-based hierarchical conversation modeling and neural topic models. For semantic-level conversation representation, a hierarchical architecture based on a pre-trained language model is used, and for topic-level representation, a neural topic model is used. The model outperforms leading approaches on two large datasets (Weibo and Reddit). Information retrieval measures evaluate the model's performance and stability, including MAP, P@1, P@3, and NG@5 (36.6, 32.7, 37.6,36.5), (38.0,30.0, 40.7, 38.3). The model dramatically outperforms previous works. Due to the data sets' single references, each debate context has only one ground truth citation. [15]

It creates an encoder-decoder actor-critic network using A long short-term memory (LSTM) to handle order-grabbing conflict between riders. Three rider sequencing criteria are also suggested to match LSTM unit time steps with riders. The actor-critic network can use prior riders' order lists to determine current riders' order recommendation policies, reducing order-grabbing conflict. Using Meituan delivery platform data test the proposed strategy. TIS, CR, TNGO, TNOGC, and TIRLTD metrics evaluate the method and its result (14.070, 15.844, 12.036,1.049, and 5.187), respectively. The results show that the reinforcement learning-based order suggestion technique can significantly increase grabbed orders and reduce order-grabbing disputes, improving platform and rider delivery efficiency and experience. [16]

A. *Model-based filtering*

An unsupervised learning technique based on collaborative filtering creates and constructs an autoencoder model for an efficient product recommendation system. Following the widely used CRISP-DM data mining methodology, TensorFlow 2.0.0 was used for model building and training. They compare the autoencoder approach to Singular Value Decomposition (SVD), a common recommendation system technique, in the study. SVD ran faster, but the autoencoder performed better with lower Root Mean Square Error (RMSE) values, achieving 0.029 and 0.010 on two MovieLens datasets.[17]

B. *Memory –based filtering*

The authors propose the Multi-Objective Next-Basket Recommendation (MONBR) approach to enhance the quality of recommendations. The item's temporal relevance is assessed based on factors such as utility, popularity, stability, frequency, occupancy, and novelty. The integration of MONBR with item-based collaborative filtering algorithms yields more accurate and reasonable recommendations. Empirical investigations conducted on datasets derived from real-world sources demonstrate that the MONBR-CF technique surpasses other algorithms in the Foodmart, E-commerce, TaFeng, and Dunnhumby datasets by a minimum of 15.15%, 88%, 9.43%, and 52.38%, respectively.[18]

RTN-GNNR is a Graph Neural Network Recommendation model that combines Review Text and Node features to improve item recommendations. It uses Bi-GRU text analysis, BERT, attention, GNN, FM, and MLP for feature extraction and inner- products for prediction. To evaluate the model, use RMSE and MSE. The model outperforms traditional personalized recommendation systems in Amazon's Automotive, Baby, Sports & Outdoors, Video_Games, and Toys_and_Games datasets((0.968,0.936),(0.953,0.984)(,0.967,0.933),(0.874,0.885),(0.885,0.984)) respectively, demonstrating its effectiveness in enhancing recommendation impact. The disadvantage of the model is the low speed of recommendations. [19]

This study introduces generative session-based recommendations using a generator to replicate user sequential behaviors. It considers rationality and informativeness for training and creates a doubly adversarial network. The model generates samples via reinforcement learning, and the study uses real-world datasets like ML-1M3, Amazon4, and Diginetica5. They use NDCG, Recall, and MRR-evaluated metrics. The model shows performance improvements of 4.69%, 4.19%, and 4.67% respectively across all datasets and target models.[20]

Table.1 Comparisons between two types of recommendation system [21]

Technique	**Advantage**	**Disadvantage**
Content-Based Filtering	Discover each document to calculate similarity Results reflect user preferences	Word relevancy quality is uncertain New user issue
Collaborative Filtering	The results of recommendations may be serendipitous, but the quality is guaranteed.	Cold start and Sparsity issue

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

1.1.3 Hybrid System

Both content-based and collaborative filtering techniques have limitations because they employ metadata and rating data from the user's item[22]. To increase recommendation performance and overcome the constraints of both recommendation filtering techniques, a hybrid recommendation model was presented. [23][24].

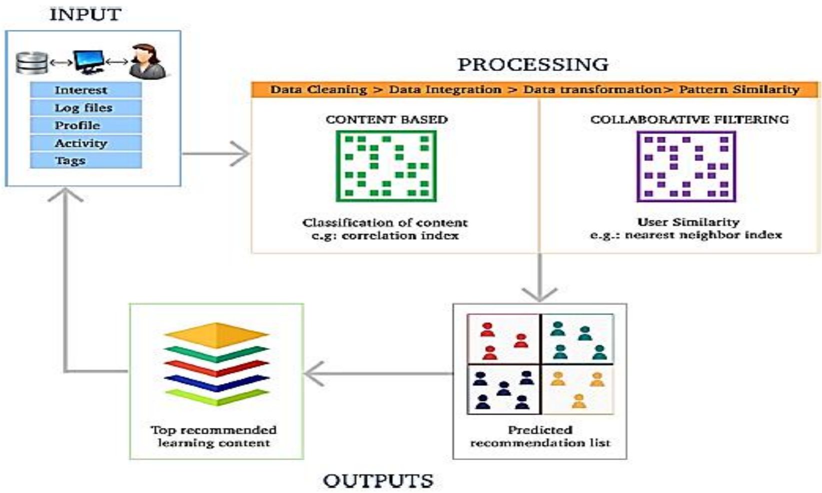


Fig.5 System architecture of hybrid recommendation system [24]

Four elements constitute the proposed system: 1- Ad Recommender, 2- Ad Profile Generator, 3- User Profile Generator, and 4- Interest Ontology. The system analyses advertising content using natural language processing methods and highlights objects connected to interest ontologies. Analysing the similarities of ads and considering their frequency of occurrence helps the Ad Recommender to generate recommendations. This traditional paradigm helps one find notable relationships between ads and involved users. Thorough validation of the system produced a mean average precision at 3 (MAP@3) of 85.6% and an aggregated f-measure of 79.2%. Using a network of linked people participating on social media networking sites (SNS).[25]

This study proposes Predictory, a unified hybrid recommender system incorporating a collaborative filtering module utilizing the SVD algorithm, a content-based system, and a fuzzy expert system. The proposed system recommends movies. The algorithm uses the user's preferred and less preferred genres to provide movie recommendations, while a fuzzy expert system evaluates their significance. The expert system considers the mean ratings of movies, the quantity of ratings received, and the similarity between movies. Therefore, the system surpasses the performance of collaborative filtering, content-based, and weighted hybrid systems. The system underwent testing using the MovieLens dataset and was compared to other traditional recommender systems. The system verification results, based on set criteria such as precision, recall, and F1-measure, indicate that the system achieved an overall accuracy of above 80% (81% precision, 83% recall, and 82% F1-measure).[26]

This research employs word embedding techniques to examine the movie's narrative and measure the degree of similarity in plot material. Additionally, it utilizes movie names, genres, directors, and stars to enhance its assessment of movie similarities. The measurement consists of two distinct stages. The first step was to categorize all movie features. The Word2Vec embedding technique was explicitly employed to generate a semantic vector representing the content of movie plots. In the second step, each feature computed movie similarities using vector-based soft cosine similarity. Using their crowdsourcing platform, OMS, they gathered the movie dataset specifically for the trials. The experimental results demonstrate that the suggested method can enhance collaborative filtering recommendation systems, with an accuracy of 82.636, precision of 83.433, recall of 77.301, and an F1-score of 80.251 [23].

1.2 Recommendation Techniques

1.2.1 Text mining

Text mining is a technique for discovering useful text information by extracting text related information from data.[27] it used to reduce the information in big data.[28]

This study presents an algorithm for recommending doctors for online pre-diagnosis based on the qualities of ontology and illness text. The algorithm considers the patients' geographical location and utilizes a dataset consisting of 20,000 patient consultation questions. Executed Python and MySQL algorithms to replace the retrieval of semantic dictionaries and word frequency statistics. Utilized word vectors to quantify the similarity between patients' prediagnosis and doctors' specialties. Developed a recommendation framework for medical departments or doctors based on sentence similarity. Generating suggestions by analysing the similarity between sentences. The methodology enhances patients' experience with online consultations and increases the convenience of offline therapy, hence augmenting the value of online prediagnosis data. The accuracy rate is 74%, and the recall rate is 78%.[29]

1.2.2 Matrix Factorization (MF)

After inferring elements from user evaluation data for things and storing them as vectors, Matrix Factorization is a technique to classify objects and user data. [30]

The project utilizes the Amazon Electronics, Yelp, CIKM-Yelp, and CIKM-Douban datasets to create collaborative filtering (CF) suggestion scenarios using heterogeneous information networks (HIN). Metagraphs constructed from a Heterogeneous Information Network (HIN) may effectively represent intricate semantic similarities between users and objects. This enables the provision of two fusion frameworks. The first framework is the "MF + FM" framework. For every metagraph, create a user-item matrix and using Matrix Factorization (MF) and nearest neighbour regression (NNR) to extract unsupervised latent properties of users and items. Subsequently, they employ a Group Lasso regularized Factorization Machine (FM) to combine latent variables from separate metagraphs, predict user-item connections, and recommend items to users. The second model is HAF, a deep learning model that extracts and combines latent features based on meta-graphs in a seamless manner. The experimental results demonstrate that these frameworks surpass current methods in terms of suggestion performance. Evaluate the dataset using the Root Mean Square Error (RMSE) values: 1.1905±0.0002, 1.2335±0.0007, 1.1139±0.0010, and 0.6975±0.0021, respectively. [31] .

This paper proposes a collaborative Matrix Factorization recommender system for Riyadh eateries. The technology predicts user preferences and recommends places based on reviews and ratings. Three methods, NMF, SVD, and SVD++, are used in the system. A comprehensive dataset of Riyadh restaurants was scraped from Foursquare.com. Results demonstrated that SVD and NMF are excellent recommendation systems, with SVD performing slightly better in RMSE (0.6439) and NMF in MAE (0.4558). The collaborative approach using Matrix Factorization methods captures complicated user-restaurant relationships. The MF methods recommend relevant eateries based on user reviews and ratings, improving customer satisfaction and restaurant income. [32]

1.2.3 Neural Networks (NN)

While the implementation and utilization of neural networks in the recommendation system domain may be limited compared to other sectors, it is now a prominent topic of focus in recommendation system research. [33][34]

The research introduces a highly effective SSR framework that utilizes a heterogeneous knowledge graph (KG) consisting of social network data and historical user behaviour data. People and objects in a knowledge graph (KG) are represented as embeddings. These embeddings are learned using a heterogeneous graph neural network (GNN) that captures information from social interactions, user-item interactions, and cross-session item transitions. A semantic role (SR) model can utilize knowledge graph (KG) embeddings to enhance the precision of recommendations by leveraging knowledge graph data. The evaluation of three public real-world datasets commonly used in SR literature (Gowalla, Delicious, Foursquare) involved the use of HR@K (Hit Rate at K) and MRR@K (Mean Reciprocal Rank at K) metrics. The results for each dataset are as follows: Gowalla (53.72, 25.67), Delicious (49.53, 21.98), and Foursquare (70.05, 34.62). They end by providing both conceptual and practical evidence to demonstrate the effectiveness of SEFrame and conducting thorough trials to verify the efficacy of SERec. The drawback of this framework arises when there are frequent changes in user and item dependencies, making the present framework's fixed heterogeneous knowledge less feasible. [35] .

The paper presents Knowledge-aware Graph Neural Networks with Label Smoothness regularization (KGNN-LS) to create customized item embeddings for users. The approach turns the knowledge graph into a customized weighted graph and detects important links inside a knowledge network using a trainable function. The approach outperforms current techniques in useful contexts including MovieLens-20M, Book-Crossing, Last.FM, and Dianping-Food datasets recommendations. The method gets, accordingly, an Area Under the Curve (AUC) of 5.1%, 6.9%, 8.3%, and 4.3% greater than the baselines. [36]

2 Datasets

Table. 2 Most dataset that used in recommendation system

Dataset name	*Used by*	*Size & Type*	*Description*	*URL source*
the Book-Crossing	[36] [37][38] [39] [40]	26MB Numeric	Cai-Nicolas Ziegler captured data from the 2004 Book-Crossing virtual book community, including 278,858 members and 1,157,112 implicit and explicit ratings.	http://www.bookcrossing.com
Foodmart	[18]	44KB Numeric	contains 1560 product and 8842 user market baskets	(https://recsyswiki.com/wiki/Grocery_shopping_datasets)
E-commerce	[18]	8MB Mixed	UK-based non-store online retailer transactions between 01/12/2010 and 09/12/2011	https://www.kaggle.com/datasets/carrie1/ecommerce-data
TaFeng	[18]	14MB Numeric	Chinese grocery store transactions from November 2000 to February 2001.	https://www.kaggle.com/chiranjivdas09/ta-feng-grocery-dataset
Dunnhumby	[18]	40.7 GB	contains two years of retailer-market household transactions.	https://www.dunnhumby.com/source-files/
MovieLens	[36] [38] [41]	190 MB numeri	Rate movies on MovieLens. The benchmark dataset is stable and collected throughout time, depending	https://grouplens.org/datasets/movielens

	[42] [43]	c	on the size of the set. Before using these data sets	
Last.FM	[36]	630KB Mixed	92,800 artist listening records from 1892 users.	https://grouplens.org/datasets/hetrec-2011/
Dianping-Food	[36]	34 KB Mixed	Dianping.com features about 10 million interactions (clicking, buying, and adding to favorites) between 2 million individuals and 1,000 restaurants. This KG has 28,115 entities, 160,519 edges, and 7 relation-types.	https://www.dianping.com/
Weibo	[15]	19 MB Text	(in Chinese, a popular Chinese microblog platform)	https://github.com/Lingzhi-WANG/Datasets-for-Quotation-Recommendation
Reddit	[15]	71.8 MB	37k English comments with sentiment labeling	https://github.com/Lingzhi-WANG/Datasets-for-Quotation-Recommendation.
The Amazon dataset	[19] [20] [44] [31] [2]	2MB Mixed	The Amazon dataset has 24 categories, including food, books, and music. Each data category includes item metadata, historical statistics, and user reviews.	https://jmcauley.ucsd.edu/data/amazon/
SemEval.	[7]	8MB Text	Semantic Evaluation is an annual semantics workshop. Its datasets are commonly utilized for text mining.	https://www.kaggle.com/datasets/azzouza2018/semevaldatadets
Inspec.	[7]	Text	2000 English abstracts from Inspec database. Computer science and information technology journal articles from 1998 to 2002 are abstracted.	https://paperswithcode.com/dataset/inspec
Yelp	[31]		Users can rate and review local businesses on this website. Ratings range from 1 to 5.	https://www.yelp.com/dataset
Epinions	[45]	4MB mixed	It included of 18,088 users and of 261,649 items	www.epinions.com https://www.kaggle.com/datasets/chrystalii/epi

				nions/data
The Wechat-Video dataset	[46]	Mixed	is a public dataset of 7.3 million Wechat Channels' Recommendation System user interaction samples from 20,000 users. Since no dataset covers CVR, staytime, and CVR prediction,	https://algo.weixin.qq.com/2021/problem-description
Gowalla	[35]	numeric	social media platform where users check-in to share their locations. 196,591 nodes (users) and 950,327 edges (friendships) make up the dataset, which includes 6,442,890 user check-ins from February 2009 to October 2010.	https://snap.stanford.edu/data/loc-Gowalla.html
Delicious	[35]	Mixed	Tag, store, organize, and share web pages with this social bookmarking service. The dataset has 1867 users, 69226 bookmarks, 487037 tags, and 437593 tag assignments. User-bookmark-tag assignments and user-user contacts are likewise time stamped.	https://grouplens.org/datasets/hetrec-2011/
Foursquare	[35]	Mixed	This location-based social network lets users check-in and evaluate establishments. The dataset has 1143092 people, 3440944 venues, 1021970 tips, and 27149 categories. The data includes venue ratings, tips, and images.	https://sites.google.com/site/yangdingqi/home/foursquare-dataset
CiteULike	[9]	Mixed	You may tag, store, and distribute scientific publications from various fields. The dataset includes abstracts, titles, tags, and citations.	http://www.citeulike.org/ https://github.com/js05212/citeulike-a
TMDB	[41]	9MB Mixed	A popular online media rating and information portal. The dataset includes movie and show titles, genres, cast, crew, budget, revenue, popularity, and user ratings.	https://www.kaggle.com/datasets/tmdb/tmdb-movie-metadata
Yahoo! Webscope R4	[47]	23MB Mixed	The Yahoo! Movies list includes user ratings and descriptions. There are 7,642 movies, 211,231 users, 8,197,596 ratings, cast, crew, synopsis, genre, awards, etc. The data is appropriate for recommender systems, collaborative filtering, relational learning, and data mining.	https://webscope.sandbox.yahoo.com/catalog.php?datatype=r&did=4

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

DBLP	[48]	Mixed	features author profiles, citation graphs, and semantic search for over 7 million publications from 15000 venues.	https://dblp.org/
Serendipity	[43]	182MB Mixed	customer reviews of popular Chinese e-commerce site product recommendations. User evaluations, preferences, personality traits, and curiosity levels are included for each recommendation.	https://github.com/greenblue96/Taobao-Serendipity-Dataset

3 Recommendation systems Machine Learning based

Machine Learning (ML) boosts recommendation system efficiency and accuracy[49]. ML's ability to learn from data and improve performance without programming is its strength. [50] This requires constantly improving recommendation systems' ability to match scholarly publications to users' interests and requirements. Search tendencies, frequently read articles, and searched topics can be detected by ML systems [51][24].

A broad article repository can become a personalized academic library with tailored recommendations using predictive modeling. Advanced recommendation algorithms can predict future interests based on reading habits. In recommendation systems, supervised and unsupervised ML are used. Supervisory learning algorithms derive conclusions from labelled input data and forecast outcomes using past data. [52]

It helps the recommendation system produce valid suggestions based on user history.[53] [54]However, unsupervised learning lets users "discover" new interests and improve their academic and learning experiences by recognizing similar structures or patterns in unlabelled data[55][56]. ML handles big data sets in the recommendation system. This is significant since recommendation systems use vast, sophisticated databases. ML efficiently maintains, evaluates, and infers such data, providing faster, more accurate recommendations.[57]

In addition to efficiency and precision, ML provides hybrid recommendation systems with collaborative and content-based filtering. This technique makes more complete suggestions using user behaviour and item characteristics.[58]

Machine learning in recommendation systems improves academic content access and engagement. Its implementation issues aside, it can considerably improve user experience, making it a desirable subject for computer science study.[59]

3.1 content-based in ML

they implemented an intelligent wearable gadget combining an IoT cloud platform and machine learning to recognize activity for specific users. Adding a recommender system to data output lets them suggest user activities. Their dataset came from the UCI machine learning repository. They examined different algorithms, but the accuracy of LR (logistic regression) was the best at 96.19 %.[60]

An automated machine learning model presents HR with appropriate application resumes based on job descriptions. The approach consists of two phases: first, it categorizes the resume. Second, it suggests resumes depending on the job description similarity index. Data came in via Kaggle and online portals. Using a linear SVM classifier accuracy of 78.53%, the proposed approach catches semantics and resumes insights. Deep learning models can improve model performance, such as convolutional neural networks, recurrent neural networks, short-term memory, and others. [61]

3.2 Collaborative-filtering in ML

This study presents an innovative CF recommendation strategy that can manage heterogeneous data, including textual reviews, ranks, votes, and video views in an extensive data Hadoop environment using the Cassandra database to increase recommendation response time. A hotel feature matrix is extracted using opinion-based sentiment analysis and saved in a database in the suggested system. Lexical, syntactic, and semantic analysis are used to understand hotel feature sentiment. The NLTK library determines textual review polarity. The system uses fuzzy heuristics to classify hotels by guest type. Euclidean distance calculates item similarities and makes correct recommendations based on guest type. Additionally, precision, recall, and F-measure have been calculated (0.80, 0.69, 0.74), respectively, and the findings show enhanced accuracy and response time compared to older methods. [62]

This paper proposes a model that is based on machine learning for product rating using 0, 2, 4. Negative 0, neutral 2, positive 4. This will complement the existing review system that handles user reviews and comments without conflict. Keras, Pandas, and Sci-kit Learning were used to implement this model internally. The proposed approach enhanced prediction with 79% Accuracy, 21% MAE, 79% Precision, 80% Recall, and 79% F1-Score for Yelp datasets of companies in 11 metropolitan areas in four countries. This experiment shows that the similarity recommendation algorithm outperforms the conventional approach and improves suggestion accuracy [63]

FCM clustering is used in an item-based collaborative filtering framework to improve prediction accuracy on public datasets. The algorithm constructs user and item membership degree vectors, adopts scoring matrix data representation, and linearly combines subjective and objective membership degree feature vectors. Several MovieLens dataset trials demonstrate the suggested system's efficiency. MAE is 0.7490, and RMSE is 0.9504. Finally, the experiment indicated that the method operates best when the combination factor is 0.1 rather than 1.0 to 0.5[64]

The effective Stochastic Gradient Descent-based approaches predicting item rate, biassed matrix factorization, and regular Matrix Factorization are compared in this work. They tested two actual public datasets: Movie Lens 100 K and Book Crossing. The study revealed that Biassed Matrix Factorization employing SGD enhanced rating prediction accuracy in both datasets. Biassed Matrix Factorization lowered RMSE by 25.78%, MAE by 14.08% for the Movie Lens 100 K dataset and RMSE by 25.78%, MAE by 19.69%, and RMSE19.69% for Book Crossing, from standard Matrix Factorization. [40]

3.3 Hybrid filtering in ML

Even though hybrid approaches are less prevalent than the other four, they are more competitive.[24]

They provide a hybrid recommendation system to recommend course content/curriculum to e-learning system users. Changing user profiles as learners go via course content and generalizing across content sources (department-taught courses) and types is easy using recommendations. Considering a document-based e-learning platform, compute a 'content vector' for each item in the database D to model document content and retrieve/recommend it. The 'bag of words' model uses extracted content vectors for similarity comparison. Then, cluster these documents using soft/fuzzy K-means++, which computes starting seeds and finds K data divisions. The system is tested on 111 transdisciplinary students from an e-learning platform. The results demonstrate that the hybrid recommendation system outperforms standard filtering by over 30%. [65]

4 Recommendation system Deep learning based.

Recommendation systems have been significantly impacted by DL, leading to more engaging and personalized user experiences. Future recommendation systems from DL should be much more advanced and potent as they develop.[66]

The way article recommendation systems operate has been entirely transformed by deep learning. We can

ATU-FJIECE, Volume: 3, Issue: 2, Jun 1, 2024, © 2020 FJIECE, All Rights Reserved

provide consumers with more precise and personalized recommendations by utilizing deep learning models to enhance their overall system experience. Future recommendation systems should become increasingly more potent and successful as deep learning continues to advance.[67][68]

They evaluate deep learning-based recommendation systems in this article to help new researchers learn about the field. [69] Analyse accumulated studies on deep learning models used in recommender systems, solutions to recommender system difficulties, awareness and prevalence over recommendation domains, and purposive features .It is also quantitatively evaluate field articles and discuss findings and future research.[70].

4.1 Content-based in DL

A scientific article recommendation (C-SAR) model based on deep learning was suggested for this work. This approach checks papers for title similarity. They used a Gated Recurrent Unit approach to detect document similarity and an association rule mining Apriori process to filter the most common documents. The model was created using the ACL Anthology Network (AAN) 2014 dataset. The model should outperform simple K-Means Clustering and user representation models. Implementation memory and time constraints limit this model. Cloud services and super configuration machines boost performance. [11]

This study proposes a novel movie recommendation system (RS). By utilizing RS output and comparable movie ratings and voting data, they created a novel collection of features and proposed a CNN deep learning (DL) model for predicting the popularity of movies across many classes. This analysis utilizes publicly available IMDb and TMDb data. It retrieves film information and inherent characteristics from TMDb and employs a content-based film recommendation algorithm to identify similar movies. Next, they will utilize IMDb beta movies and vote data to forecast movie popularity using deep learning techniques. Ultimately, it employs fuzzy c-means algorithms to predict the target audience. In a multiclass classification model, they achieved a remarkable accuracy of 96.8%, surpassing all benchmark models. This study demonstrates the utility of predictive and prescriptive data analytics in information systems for informing industry decision-making.[71]

4.2 collaborative-filtering in DL

The work presents a deep learning approach for analysing health-related medical datasets, employing machine and deep learning techniques such as logistic regression, naive Bayes, RNN, MLP, GRU, and LSTM. The hospital-collected medical dataset contains 30 patient records with 13 disease attributes and 1000 items. The LSTM deep learning model achieved an accuracy of 97.74%. The accepted class has a precision of 98%, a recall of 99%, and an F1-measure of 99%. On the other hand, the not-allowed class has a precision of 89%, a recall of 73%, and an F1 measure of 80%. [49]

They offer a deep learning system that simulates intelligent recommendations by understanding users and items. Initially, low-dimensional vectors of users and items are learned individually to encode semantic information about user-user and item-item connections. The proposed method is tested on two benchmark datasets (MovieLens 1M and MovieLens 10M) and outperforms feed-forward neural network-based methods by a large margin RSME 0.830 and 0.776 for datasets, respectively.[72]

4.3 Hybrid system in DL

In DMFL, they apply deep learning and advanced machine learning models to study user-item interactions from many angles. The DMFL model has two parts: feature learning and preference creation. The feature learning section uses two paralleled deep neural networks to extract static item and dynamic user latent feature vectors. The preference-generating section has three submodules: SDAE-FM, deep neural network, and metric learning. The three modules simultaneously receive the acquired latent feature vectors for users

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

and items in the feature learning stage, and the findings of the three submodules are merged to forecast item user preferences. Extensive trials on real-world datasets (Book Crossing, MovieLens-1M, and MovieLens-20M) demonstrate the model's usefulness and efficiency compared with another six models (Wide& Deep, CML, DeepRec, CDL2, CDL1, FM). Use AUC and recall to evaluate(AUC@100(0.7902,0.8323,0.8621), Recall@100(0.1902,0.5984,0.6102))for data sets respectively. [38]

This paper introduces an improved method for suggesting dialogue options by taking into account the time-series characteristics of the discussion setting. The system use mutual information maximization to align the semantic space by integrating external knowledge graphs. Additionally, it leverages a timing network model to deliver individualized recommendations depending on the sequencing of dialogue material. The system was assessed on the REDIAL and INSPIRED datasets, and it demonstrated excellent automatic diversity with Dist-4 values of 6.83 and 5.340, as well as strong recall with Recall@50 scores of 39.0 and 27.8. The human evaluation also demonstrated a high level of fluency (1.57) and informativeness (1.55) in the created dialogues. The researchers assert that the suggested methodology attains exceptional results in both discussion and recommendation assignments.[73]

5 Recommendation system Natural Language Processing based

Natural Language Processing (NLP), an area of AI, helps computers understand human language.[74][75] Language's complexity comes from its variety.[76] Since most processing data is unstructured text, is expected to have diverse vocabulary, misspellings, typing errors, ambiguous words, natural language understanding is crucial. Natural Language Processing (NLP) employs rule-based and machine-learning approaches to comprehend and manipulate human language. Rule-based approaches depend on pre-established rules and patterns, such as grammar, syntax, and semantics, and are efficient for activities with a clear structure. Machine learning algorithms utilize statistical models trained on extensive datasets to acquire knowledge of the fundamental patterns and structures present in the language.[77] Hybrid approaches leverage the advantages of both methods, employing rule-based techniques for clearly defined tasks and machine learning for intricate and uncertain jobs. The selection of the natural language processing (NLP) technique is contingent upon the specific task at hand, the availability of data, and the desired levels of accuracy and robustness.

NLP matters in recommendation systems. A system must understand scholarly papers to recommend them. Most academic publications are text-based, so the system needs NLP to interpret and analyse material. NLP can extract subjects, keywords, named entities, and sentiment from academic articles. [78] It can be used to generate a paper's content profile for recommendation. The recommendation system can suggest NLP-identified publications to a user interested in a topic.

Similar publications can be located using paper NLP data. NLP methods like embedding can group relevant papers by encoding words or phrases as vectors in high-dimensional space. This is needed for recommendation algorithms to suggest related papers.[79]

A common NLP algorithms and techniques:

- Text embedding: Methods like Word2Vec or GloVe transform words and sentences into dense vectors. These embeddings, with their practical application, supplement downstream activities with semantics, empowering you to navigate complex linguistic landscapes.
- NER: Named Entity Recognition NER meticulously classifies textual entities—people, organizations, and places.[77]
- By analysing article sentiment, we can delve into user preferences and emotional nuances, enriching our understanding of scholarly conversations with their depth and richness.
- Topic modeling: LDA and NMF algorithms reveal article topics.
- Summarize complex research articles into concise and engaging abstracts. They convey the

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

fundamental nature while upholding scholarly rigor.

- Extracting Keywords: Extracting important phrases improves content-based filtering, like a gemmologist sorting through large reserves. These keywords reveal knowledge riches.
- Named Entity Clarity: This is achieved by disambiguating related notions. It provides precise references, like interpreting old texts.
- The similarity of articles is quantified using either cosine similarity or the Jaccard index. Similar to the practice of scholarly cross-referencing, collaborative filtering greatly benefits from these measures.
- GPT and BERT serve as pre-trained language models that bridge the gap between human comprehension and machine analysis. They enhance scholarly discussions by providing a contextual framework for the subject matter.[80]

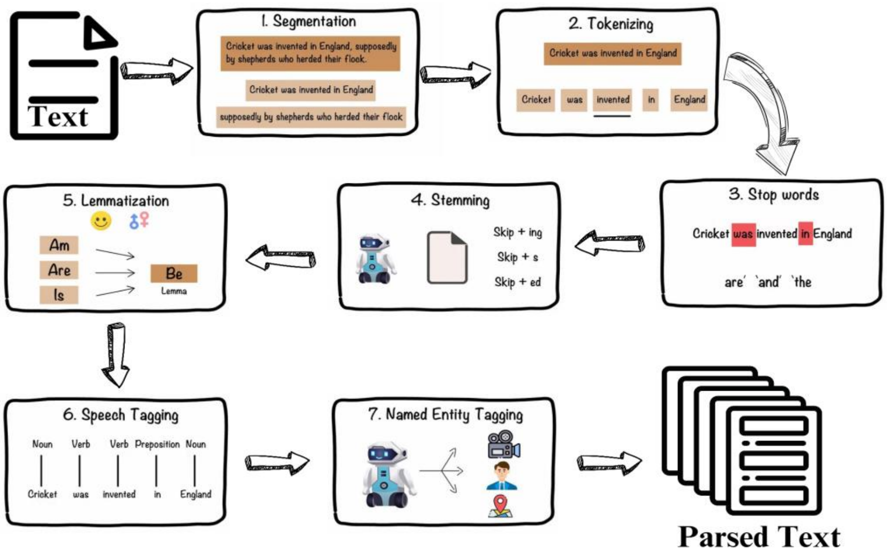


Fig.6 NLP Technique Works

5.1 Content-based in NLP

This work presents a content-based modern art recommendation system using artwork photos and artist metadata. Visual neural network designs like ResNet extracted visual information and classified artworks using attributes. Word embeddings and a proprietary BERT model were used to examine artist content and infer contextual factors. Visual and contextual information were used to embed artworks and establish an artist-artwork closeness graph. Graph analysis recommends artworks based on visual and contextual information from artists and artworks. A graph-based recommendation method provided personalized artwork recommendations. They earn an average final rating of 75% of meaningful artworks compared to professional evaluations . [81]

This system makes personalized movie recommendations based on the user's previous interest, ratings, and system engagement. The system analyses movie plot descriptions, reviews, etc., using NLP. A Count Vectorizer counts word frequency to build a vector representation of movie text. Cosine similarity is then used to compare movie vectors and recommend movies most like the ones the user has loved. Flask lets users enter a movie title and receive a list of the top 10 most comparable movies. The system tells the user

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

if the movie is not in the database. Attained an accuracy of 80%. A movie recommender system predicts better than a collaborative filtering algorithm-based recommendation system. The content-based filtering movie recommendation algorithm is more accurate and doesn't indicate a new user/item problem when new ones are added.[82]

5.2 Collaborative filter in NLP

This paper proposes a novel employment recommendation system combining ML models, NLP for talent and experience matching, and CFR approaches. The study proposes a comprehensive job suggestion system that automates and improves recruiter applicant screening using machine learning. The JobFit recommendation algorithm considers key applicant traits for recruitment. The method uses job requirements and candidate profiles to calculate a JobFit score, showing how well each applicant fits the job. The JobFit technology helps HR managers shorten the process of screening many applicants to find the best prospects. Using distinct datasets, the Similar Skills, National Longitudinal Survey of Youth 1997, and IBM HR Analytics Employee Attrition and Performance datasets. The linear regression model was chosen for the final model with an Accuracy of 0.95.[83]

This research introduces a context-aware personalized news recommendation system that utilizes contextual information to customize news suggestions. The system uses content-based and collaborative filtering techniques to offer users precise and pertinent news recommendations. 22,657 news stories were gathered from 19 widely-read online news sources. They created four systems: three for content-based recommender systems (TF-IDF, Bag-of-Words, and Word2Vec) and one for collaborative filtering based on click activity. The collaborative filtering model based on click behaviour achieved the highest performance, with a mean MAE (Mean Absolute Error) of 0.0252 and a mean RMSE (Root Mean Square Error) of 0.0364. [84]

5.3 NLP hybrid filter or system

The study develops and tests a content-based movie recommendation system using machine learning techniques and similarity measurements. The primary purpose is to propose movies based on overview, cast, crew, keywords, and genres by using the MovieLens dataset. Movie plots are converted to numerical vectors using TF-IDF, and the cosine similarity method (sigmoid and linear kernels) is used for movie vector similarity. NormalPredictor, SVD, KNNBasic, KNNBasic with similarity options, and NMF are tested. K-fold Cross Validator validates the machine learning model Error and Mean Absolute Error. Final validations are three-fold for each validator. The Best-performing SVD method has the best RMSE (90%) and MAE (69%), followed by KNNBasic with similar settings [41]

The article presents a hybrid course recommendation system for an online learning platform (GCF). They get data from two sources: First, Google Analytics logs contain implicit user interaction data (page visits, sessions, etc.). Second, Course and lesson content is stored in the MongoDB Database. The collaborative filtering recommender leverages implicit user ratings estimated from Google Analytics data to recommend courses watched by comparable individuals, while the content-based recommender uses text embedding to suggest courses with content comparable to what the user has seen. The collaborative filtering model has the best P@5 (Positive Outcome) value of 0.4, recommending two out of five courses of interest to users. The content-based strategies yield little results but help mitigate the cold start problem. [85]

6 Limitations and advantages of recommendation system in the article recommendation field

While they have limits, academic article recommendation algorithms have made great progress in improving research accessibility. Suboptimal user experiences can result from the cold-start problem, data

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

sparsity, and natural biases in recommendation systems. However, These systems also have some benefits, including scalability, tailored experiences, and better discoverability of pertinent, high-quality research materials. They could help researchers push the envelope, enable unanticipated discoveries, and increase their work output. Customized experiences help consumers to remain current with the most recent advancements, investigate uncharted subjects, and work with colleagues. Moreover, automated recommendation systems let users negotiate the research terrain with more success by offering a scalable solution to the information overload issue .Researchers and practitioners may solve these issues and raise these systems' dependability, objectivity, and openness, enabling the full potential of academic article recommendations to support scientific discovery and information exchange.

7 The challenge

building effective recommendation systems poses significant challenges, such as handling large volumes of data, dealing with sparsity and cold-start problem The sparsity problem emerges when suggestion data is scarce. The cold start problem occurs where there is no evaluation data. Comparative experiments on Amazon's product dataset demonstrated the system's accuracy and robustness, addressing cold start, data sparsity, and overspecialization.[44]. KGNN-LS also achieves strong performance in cold-start scenarios where user-item interactions are sparse.[36]

For the cold start problem of Traditional Hybrid Clustering and distance clustering to Supervised Learning, this research suggests Hybrid Content and Fuzzy C-means clustering. This approach employs fuzzy C-means clustering to solve the user and item membership degree matrix, construct the feature vector, and tackle the problem of data sparseness, preventing the traditional distance-based membership degree.[64]. The strategy overcomes data sparsity and cold start issues in recommender systems, providing users with clear explanations[86]. Content-based strategies did not produce significant results but mitigated the cold start problem and validated a hybrid strategy.[85]

8 Conclusion

Online content filtering and personalized suggestions require recommender systems. Building successful recommendation systems is challenging and needs ethical considerations. Examining current recommender system research identifies gaps and challenges to help future researchers build an effective system. The complex environment was explored and interesting tendencies were found in the survey .

A brief description of the findings:

Diverse Landscape: Recommended systems are used in e-commerce, entertainment, healthcare, learning, academics, and more. Researchers and practitioners must comprehend this variability.

Challenges Persist: Despite considerable progress, challenges persist. Scalability, cold-start issues, and data sparsity plague recommendation systems. The first step to overcoming these obstacles is acknowledging them.

Practical Implications: the practicalities are illuminated by this survey. The researcher can use these findings to improve user experiences, personalize, and reduce information overload.

List of Abbreviations

ARSs: Academic article Recommendation Systems
AI: Artificial Intelligence
RS: Recommendation System
CB: Content-Based
GRU : Gated Recurrent Unit
AAN :Attention-Augmented Network

CNN : Convolutional Neural Network
MAE: Mean Absolute Error
RMSE: Root Mean Square Error
CF: Collaborative Filtering
MAP: Mean Average Precision
LSTM: Long Short –Term Memory
TIS :Total Idle Seconds
CR : Completion Rate
TNGO :Total Number of Grabbed Orders
TNOGC :Total Number of Order-Grabbing Conflicts
TIRLTD : Total Increase in Riders' Last Time Delivery
SVD: Singular Value Decomposition
ML10,**20**: MovieLens dataset
NDCG :Normalized Discounted Cumulative Gain
MRR : Mean Reciprocal Rank
NMF :Non-negative Matrix Factorization
ML: Machine Learning
SVM: Support Vector Machine
DL: Deep Learning
RNN :Recurrent Neural Network
MLP :Multilayer Perceptron
AUC :Area Under the Curve
NLP : natural language processing
LDA : Latent Dirichlet Allocation
GPT :Generative Pre-trained Transformer
BERT: Bidirectional Encoder Representations from Transformers
TF-IDF: Term Frequency-Inverse Document Frequency

REFERENCES

[1] D. Yuxing, H. Noah, A. Gunes, A. Frank, and L. Nick, "IoT Inspector : Crowdsourcing Labeled Network Traffic from Smart Home Devices at Scale". Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2020, 4.2: 1-21.
[2] A. Beheshti, S. Yakhchi, S. Mousaeirad, and S. M. Ghafari, "Towards Cognitive Recommender Systems," 2020, doi: 10.3390/a13080176.
[3] H. El Khoury, "Ontology-based Recommender System in Higher Education," vol. 2, pp. 1031–1034, 2018.
[4] S. Yin and H. Wang, "A MOOC Courses Recommendation System Based on Learning Behaviours," 2003.
[5] M. Reba and K. C. Seto, "A systematic review and assessment of algorithms to detect, characterize, and monitor urban land change," *Remote Sens. Environ.*, vol. 242, no. May 2019, p. 111739, 2020, doi: 10.1016/j.rse.2020.111739.
[6] M. Gusenbauer and N. R. Haddaway, "Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources," *Res. Synth. Methods*, vol. 11, no. 2, pp. 181–217, 2020, doi: 10.1002/jrsm.1378.
[7] Z. Nasar, S. W. Jaffry, and M. K. Malik, "Textual keyword extraction and summarization : State-of-the-art," vol. 56, no. June, 2019, doi: 10.1016/j.ipm.2019.102088.
[8] S. Khodabandehlou, "Designing an e-commerce recommender system based on collaborative

filtering using a data mining approach," *Int. J. Bus. Inf. Syst.*, vol. 31, no. 4, pp. 455–478, 2019, doi: 10.1504/IJBIS.2019.101582.

[9] R. Glauber and A. Loula, "Collaborative Filtering vs. Content-Based Filtering: differences and similarities," 2019, [Online]. Available: http://arxiv.org/abs/1912.08932

[10] D. Roy and M. Dutta, "A systematic review and research perspective on recommender systems," *J. Big Data*, 2022, doi: 10.1186/s40537-022-00592-5.

[11] A. M. Nair, O. Benny, and J. George, "Content Based Scientific Article Recommendation System Using Deep Learning Technique," *Lect. Notes Networks Syst.*, vol. 204 LNNS, pp. 965–977, 2021, doi: 10.1007/978-981-16-1395-1_70.

[12] J. Shu, X. Shen, H. Liu, B. Yi, and Z. Zhang, "A content-based recommendation algorithm for learning resources," *Multimed. Syst.*, vol. 24, no. 2, pp. 163–173, 2018, doi: 10.1007/s00530-017-0539-8.

[13] D. Wang, Y. Liang, D. Xu, X. Feng, and R. Guan, "Knowledge-Based Systems A content-based recommender system for computer science publications," *Knowledge-Based Syst.*, vol. 157, no. February, pp. 1–9, 2018, doi: 10.1016/j.knosys.2018.05.001.

[14] A. Chaudhuri, N. Sinhababu, M. Sarma, and D. Samanta, "Hidden features identification for designing an efficient research article recommendation system," *Int. J. Digit. Libr.*, vol. 22, no. 2, pp. 233–249, 2021, doi: 10.1007/s00799-021-00301-2.

[15] K. Wong, "Quotation Recommendation for Multi-party Online," vol. 41, no. 4, 2023, doi: 10.1145/3594633.

[16] X. Wang, L. Wang, C. Dong, H. Ren, and K. Xing, "Reinforcement Learning-Based Dynamic Order Recommendation for On-Demand Food Delivery," *Tsinghua Sci. Technol.*, vol. 29, no. 2, pp. 356–367, 2024, doi: 10.26599/TST.2023.9010041.

[17] D. Ferreira, S. Silva, A. Abelha, and J. Machado, "Recommendation system using autoencoders," *Appl. Sci.*, vol. 10, no. 16, pp. 1–17, 2020, doi: 10.3390/app10165510.

[18] M. A. Fouad, W. Hussein, S. Rady, P. S. Yu, and T. F. Gharib, "An Efficient Approach for Rational Next-Basket Recommendation," vol. 10, no. June, pp. 75657–75671, 2022.

[19] B. Xiao, X. Xie, and C. Yang, "RTN-GNNR : Fusing Review Text Features and Node Features for Graph Neural Network Recommendation," *IEEE Access*, vol. 10, no. November, pp. 114165–114177, 2022, doi: 10.1109/ACCESS.2022.3218882.

[20] Z. Wang *et al.*, "Generative Session-based Recommendation," pp. 2227–2235, doi: 10.1145/3485447.3512095.

[21] X. Bai *et al.*, "Scientific Paper Recommendation : A Survey," *IEEE Access*, vol. 7, pp. 9324–9339, 2019, doi: 10.1109/ACCESS.2018.2890388.

[22] M. Goyani and N. Chaurasiya, "A Review of Movie Recommendation System: Limitations, Survey and Challenges," *Electron. Lett. Comput. Vis. Image Anal.*, vol. 19, no. 3, pp. 18–37, 2020, doi: 10.5565/rev/elcvia.1232.

[23] L. Vuong Nguyen, T. H. Nguyen, J. J. Jung, and D. Camacho, "Extending collaborative filtering recommendation using word embedding: A hybrid approach," *Concurr. Comput. Pract. Exp.*, vol. 35, no. 16, 2023, doi: 10.1002/cpe.6232.

[24] S. S. Khanal, P. W. C. Prasad, A. Alsadoon, and A. Maag, "A systematic review: machine learning based recommendation systems for e-learning," *Educ. Inf. Technol.*, vol. 25, no. 4, pp. 2635–2664, 2020, doi: 10.1007/s10639-019-10063-9.

[25] F. García-sánchez, R. Colomo-palacios, and R. Valencia-garcía, "A social-semantic recommender system for advertisements," *Inf. Process. Manag.*, vol. 57, no. 2, p. 102153, 2020, doi: 10.1016/j.ipm.2019.102153.

[26] B. Walek and V. Fojtik, "A hybrid recommender system for recommending relevant movies using an expert system," *Expert Syst. Appl.*, vol. 158, p. 113452, 2020, doi: 10.1016/j.eswa.2020.113452.

[27] S. A. Salloum, M. Al-Emran, A. A. Monem, and K. Shaalan, “Using text mining techniques for extracting information from research articles,” *Stud. Comput. Intell.*, vol. 740, pp. 373–397, 2018, doi: 10.1007/978-3-319-67056-0_18.
[28] Y. Betancourt and S. Ilarri, “Use of text mining techniques for recommender systems,” *ICEIS 2020 - Proc. 22nd Int. Conf. Enterp. Inf. Syst.*, vol. 1, no. Iceis, pp. 780–787, 2020, doi: 10.5220/0009576507800787.
[29] C. Ju and S. Zhang, “Doctor Recommendation Model Based on Ontology Characteristics and Disease Text Mining Perspective,” *Biomed Res. Int.*, vol. 2021, 2021, doi: 10.1155/2021/7431199.
[30] S. Zhao, G. Pan, J. Tao, Z. Luo, S. Li, and Z. Wu, “Understanding Smartphone Users From Installed App Lists Using Boolean Matrix Factorization,” *IEEE Trans. Cybern.*, vol. 52, no. 1, pp. 384–397, 2022, doi: 10.1109/TCYB.2020.2967644.
[31] H. Zhao, Q. Yao, Y. Song, J. T. Kwok, and D. L. Lee, “Side Information Fusion for Recommender Systems over Heterogeneous Information Network,” *ACM Trans. Knowl. Discov. Data*, vol. 15, no. 4, 2021, doi: 10.1145/3441446.
[32] R. Alabduljabbar, “Matrix Factorization Collaborative-Based Recommender System for Riyadh Restaurants: Leveraging Machine Learning to Enhance Consumer Choice,” *Appl. Sci.*, vol. 13, no. 17, 2023, doi: 10.3390/app13179574.
[33] R. Alabduljabbar, M. Alshareef, and N. Alshareef, “Time-Aware Recommender Systems: A Comprehensive Survey and Quantitative Assessment of Literature,” *IEEE Access*, vol. 11, no. May, pp. 45586–45604, 2023, doi: 10.1109/ACCESS.2023.3274117.
[34] N. Nikzad-Khasmakhi, M. A. Balafar, and M. Reza Feizi–Derakhshi, “The state-of-the-art in expert recommendation systems,” *Eng. Appl. Artif. Intell.*, vol. 82, no. June, pp. 126–147, 2019, doi: 10.1016/j.engappai.2019.03.020.
[35] T. Chen and R. C. Wong, “An Efficient and Effective Framework for Session-based Social Recommendation,” pp. 400–408, 2021.
[36] H. Wang, M. Zhao, W. Li, and Z. Wang, “Knowledge-aware Graph Neural Networks with Label Smoothness Regularization for Recommender Systems,” pp. 968–977, 2019.
[37] M. Naghiaei, H. A. Rahmani, and M. Dehghan, “The Unfairness of Popularity Bias in Book Recommendation,” *Commun. Comput. Inf. Sci.*, vol. 1610 CCIS, pp. 69–81, 2022, doi: 10.1007/978-3-031-09316-6_7.
[38] Z. Huang, C. Yu, J. Ni, H. Liu, C. Zeng, and Y. Tang, “An efficient hybrid recommendation model with deep neural networks,” *IEEE Access*, vol. 7, pp. 137900–137912, 2019, doi: 10.1109/ACCESS.2019.2929789.
[39] S. Bhaskaran and B. Santhi, “An efficient personalized trust based hybrid recommendation (TBHR) strategy for e-learning system in cloud computing,” *Cluster Comput.*, vol. 22, pp. 1137–1149, 2019, doi: 10.1007/s10586-017-1160-5.
[40] K. Kondo, “Recommendation system for CINEMA,” no. May, 2008, doi: 10.1007/978-981-13-9714-1.
[41] S. Alzu’bi, A. Zraiqat, and S. Hendawi, “Sustainable Development: A Semantics-aware Trends for Movies Recommendation System using Modern NLP,” *Int. J. Adv. Soft Comput. its Appl.*, vol. 14, no. 3, pp. 153–173, 2022, doi: 10.15849/IJASCA.221128.11.
[42] X. Zhu, J. Fu, and C. Chen, “Matrix Completion of Adaptive Jumping Graph Neural Networks for Recommendation Systems,” no. August, pp. 88433–88450, 2023.
[43] X. Chen, “Serendipitous Page Recommendation on Web IndeX System with Potential Preferences,” 2020.
[44] S. Kang, C. Jeong, K. Chung, and S. Member, “Tree-Based Real-Time Advertisement Recommendation System in Online Broadcasting,” vol. 8, pp. 192693–192702, 2020.
[45] Y. Pan, F. He, and H. Yu, “Learning social representations with deep autoencoder for recommender

system," *World Wide Web*, vol. 23, no. 4, pp. 2259–2279, 2020, doi: 10.1007/s11280-020-00793-z.

[46] W. Li and S. Wang, "STAN : Stage-Adaptive Network for Multi-Task Recommendation by Learning User Lifecycle-Based Representation," pp. 602–612, doi: 10.1145/3604915.3608796.

[47] M. Nilashi, O. Ibrahim, and K. Bagherifard, "A recommender system based on collaborative filtering using ontology and dimensionality reduction techniques," *Expert Syst. Appl.*, vol. 92, pp. 507–520, 2018, doi: 10.1016/j.eswa.2017.09.058.

[48] W. E. I. Wang *et al.*, "Attributed Collaboration Network Embedding for Academic," vol. 15, no. 1, 2020.

[49] C. Iwendi, S. Khan, J. H. Anajemba, A. K. Bashir, and F. Noor, "Realizing an Efficient IoMT-Assisted Patient Diet Recommendation System Through Machine Learning Model," *IEEE Access*, vol. 8, pp. 28462–28474, 2020, doi: 10.1109/ACCESS.2020.2968537.

[50] A. Fields, "A Survey of Recommendation Systems : Recommendation," 2022.

[51] D. B. Guruge, R. Kadel, and S. J. Halder, "The state of the art in methodologies of course recommender systems—a review of recent research," *Data*, vol. 6, no. 2, pp. 1–30, 2021, doi: 10.3390/data6020018.

[52] C. L. Sanchez Bocanegra, J. L. Sevillano Ramos, C. Rizo, A. Civit, and L. Fernandez-Luque, "HealthRecSys: A semantic content-based recommender system to complement health videos," *BMC Med. Inform. Decis. Mak.*, vol. 17, no. 1, pp. 1–10, 2017, doi: 10.1186/s12911-017-0431-7.

[53] Y. Sun and Y. Zhang, "Conversational recommender system," *41st Int. ACM SIGIR Conf. Res. Dev. Inf. Retrieval, SIGIR 2018*, pp. 235–244, 2018, doi: 10.1145/3209978.3210002.

[54] Y. Li, Y. Ge, and Y. Zhang, "CIKM 2021 Tutorial on Fairness of Machine Learning in Recommender Systems," *Int. Conf. Inf. Knowl. Manag. Proc.*, pp. 4857–4860, 2021, doi: 10.1145/3459637.3483280.

[55] I. H. Sarker and A. S. M. Kayes, "ABC-RuleMiner: User behavioral rule-based machine learning method for context-aware intelligent services," *J. Netw. Compfile///C/Users/ayat_PC/Desktop/ml recommrndaton.pdfuter Appl.*, vol. 168, no. July 2020, 2020, doi: 10.1016/j.jnca.2020.102762.

[56] M. Attaran and P. Deb, "Machine Learning: The New 'Big Thing' for Competitive Advantage," *Int. J. Knowl. Eng. Data Min.*, vol. 5, no. 1, p. 1, 2018, doi: 10.1504/ijkedm.2018.10015621.

[57] V. Kuleto *et al.*, "Exploring opportunities and challenges of artificial intelligence and machine learning in higher education institutions," *Sustain.*, vol. 13, no. 18, pp. 1–16, 2021, doi: 10.3390/su131810424.

[58] N. W. P. Y. Praditya, "Literature Review Recommendation System Using Hybrid Method (Collaborative Filtering & Content-Based Filtering) by Utilizing Social Media as Marketing," *Comput. Eng. Appl. J.*, vol. 10, no. 2, pp. 105–113, 2021, doi: 10.18495/comengapp.v10i2.368.

[59] M. Lotfian, J. Ingensand, and M. A. Brovelli, "The partnership of citizen science and machine learning: Benefits, risks and future challenges for engagement, data collection and data quality," *Sustain.*, vol. 13, no. 14, 2021, doi: 10.3390/su13148087.

[60] A. Dey, R. Rajkumar, and J. Masih, "An analytical study and visualisation of human activity and content-based recommendation system by applying ml automation," *Int. J. Mech. Prod. Eng. Res. Dev.*, vol. 9, no. 3, pp. 75–88, 2019, doi: 10.24247/ijmperdjun20198.

[61] P. K. Roy, S. S. Chowdhary, and R. Bhatia, "A Machine Learning approach for automation of Resume Recommendation system," *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 2318–2327, 2020, doi: 10.1016/j.procs.2020.03.284.

[62] Z. Hussain, B. Mago, A. Khadim, and K. Amjad, "An Intelligent Data Analysis for Recommendation Systems Using Machine Learning," *2nd Int. Conf. Bus. Anal. Technol. Secur. ICBATS 2023*, vol. 2019, 2023, doi: 10.1109/ICBATS57792.2023.10111411.

[63] C. Iwendi, E. Ibeke, H. Eggoni, S. Velagala, and G. Srivastava, "Pointer-Based Item-to-Item Collaborative Filtering Recommendation System Using a Machine Learning Model," *Int. J. Inf.*

Technol. Decis. Mak., vol. 21, no. 1, pp. 463–484, 2022, doi: 10.1142/S0219622021500619.
[64] L. Duan, W. Wang, and B. Han, “A hybrid recommendation system based on fuzzy c-means clustering and supervised learning,” *KSII Trans. Internet Inf. Syst.*, vol. 15, no. 7, pp. 2399–2413, 2021, doi: 10.3837/tiis.2021.07.006.
[65] V. B. P. Tolety and E. V. Prasad, “Hybrid content and collaborative filtering based recommendation system for e-learning platforms,” *Bull. Electr. Eng. Informatics*, vol. 11, no. 3, pp. 1543–1549, 2022, doi: 10.11591/eei.v11i3.3861.
[66] S. Zhang, L. Yao, A. Sun, and Y. Tay, “Deep learning based recommender system: A survey and new perspectives,” *ACM Comput. Surv.*, vol. 52, no. 1, pp. 1–35, 2019, doi: 10.1145/3285029.
[67] R. Dash, C. Rebman, and U. K. Kar, “Application of Artificial Intelligence in Automation of Supply Chain Management,” *J. Strateg. Innov. Sustain.*, vol. 14, no. 3, pp. 43–53, 2019, doi: 10.33423/jsis.v14i3.2105.
[68] N. Patel and S. Trivedi, “Leveraging Predictive Modeling, Machine Learning Personalization, NLP Customer Support, and AI Chatbots to Increase Customer Loyalty,” *Empir. Quests Manag. Essences*, vol. 3, no. 3, pp. 1–24, 2020, [Online]. Available: https://researchberg.com/index.php/eqme/article/view/46
[69] M. F. Dacrema, P. Cremonesi, and D. Jannach, “Are we really making much progress? A worrying analysis of recent neural recommendation approaches,” *RecSys 2019 - 13th ACM Conf. Recomm. Syst.*, pp. 101–109, 2019, doi: 10.1145/3298689.3347058.
[70] Z. Batmaz, A. Yurekli, A. Bilge, and C. Kaleli, “A review on deep learning for recommender systems: challenges and remedies,” *Artif. Intell. Rev.*, vol. 52, no. 1, pp. 1–37, 2019, doi: 10.1007/s10462-018-9654-y.
[71] S. Sahu, R. Kumar, M. S. Pathan, J. Shafi, Y. Kumar, and M. F. Ijaz, “Movie Popularity and Target Audience Prediction Using the Content-Based Recommender System,” *IEEE Access*, vol. 10, pp. 42030–42046, 2022, doi: 10.1109/ACCESS.2022.3168161.
[72] M. Fu, H. Qu, Z. Yi, L. Lu, and Y. Liu, “A Novel Deep Learning-Based Collaborative Filtering Model for Recommendation System,” *IEEE Trans. Cybern.*, vol. 49, no. 3, pp. 1084–1096, 2019, doi: 10.1109/TCYB.2018.2795041.
[73] X. Wang, J. Liu, and J. Duan, “Improved conversational recommender system based on dialog context,” *Nat. Lang. Eng.*, pp. 1–19, 2023, doi: 10.1017/S1351324923000451.
[74] P. Chithra and M. Henila, “International Journal of Computer Sciences and Engineering Open Access,” *Int. J. Comput. Sci. Eng.*, vol. 6, no. 10, pp. 628–632, 2019, doi: 10.26438/ijcse/v6i1.161167.
[75] M. B. Otto, “Advances,” *Kerntechnik*, vol. 67, no. 5–6, p. 280, 2002, doi: 10.1515/kern-2002-675-614.
[76] G. Ritchie, “Survey of natural language processing,” *ACM SIGART Bull.*, vol. 5, no. 80, p. 61, 1982, doi: 10.1145/1056176.1056184.
[77] A. Ly, B. Uthayasooriyar, and T. Wang, “A survey on natural language processing (nlp) and applications in insurance,” pp. 1–34, 2020, [Online]. Available: http://arxiv.org/abs/2010.00462
[78] Z. Kastrati, F. Dalipi, A. S. Imran, K. P. Nuci, and M. A. Wani, “Sentiment analysis of students’ feedback with nlp and deep learning: A systematic mapping study,” *Appl. Sci.*, vol. 11, no. 9, 2021, doi: 10.3390/app11093986.
[79] Jon Ezeiza Alvarez and H. Bast, “A review of word embedding and document similarity algorithms applied to academic text,” *Bachelor’s Thesis*, 2017, [Online]. Available: http://ad-publications.informatik.uni-freiburg.de/theses/Bachelor_Jon_Ezeiza_2017.pdf
[80] I. Harrando, R. Troncy, I. Media, C. Recommendation, I. Harrando, and R. Troncy, “Improving Media Content Recommendation with Automatic Annotations Ismail Harrando , Raphaël Troncy To cite this version : HAL Id : hal-03555169 Improving Media Content Recommendation with

Automatic Annotations," 2022.

[81] A. Fosset *et al.*, "Docent: A content-based recommendation system to discover contemporary art," pp. 1–11, 2022, [Online]. Available: http://arxiv.org/abs/2207.05648

[82] O. Yadav, K. Mishra, D. Patil, E. Braganza, and C. Finny, "Design and Implementation of Movie Recommendation System based on NLP And Content-based Filtering algorithm," no. June, pp. 2323–2327, 2020.

[83] K. Appadoo, M. B. Soonnoo, and Z. Mungloo-Dilmohamud, "JobFit: Job Recommendation using Machine Learning and Recommendation Engine," *2020 IEEE Asia-Pacific Conf. Comput. Sci. Data Eng.*, pp. 1–6, 2020, [Online]. Available: https://ieeexplore.ieee.org/document/9411584/

[84] R. Alabduljabbar, H. Almazrou, and A. Aldawod, "Context-Aware News Recommendation System: Incorporating Contextual Information and Collaborative Filtering Techniques," *Int. J. Comput. Intell. Syst.*, vol. 16, no. 1, 2023, doi: 10.1007/s44196-023-00315-5.

[85] J. C. Sanguino, O. Mariño, N. Cardozo, R. Manrique, and M. Linares-Vásquez, "A course hybrid recommender system for limited user information scenarios," *J. Educ. Data Min.*, vol. 14, no. 3, pp. 162–188, 2022, doi: 10.5281/zenodo.7304829.

[86] M. He, B. Wang, and X. Du, "HI2Rec: Exploring Knowledge in Heterogeneous Information for Movie Recommendation," *IEEE Access*, vol. 7, pp. 30276–30284, 2019, doi: 10.1109/ACCESS.2019.2902398.