2023 World Conference on Communication & Computing (WCONF)
Raipur, India. July 14-16, 2023

Personalized Medicine Recommendation and Disease Flagging Model Based on User's Previous Orders

Pritam Bhuiya
Computer Science and Engineering
Narula Institute of Technology
Agarpara, Kolkata
pritambhuiyanit25@gmail.com

Rohit Malik
Computer Science and Engineering
Narula Institute of Technology
Agarpara, Kolkata
rohitmaliknit25@gmail.com

Aritra Rana
Computer Science and Engineering
Narula Institute of Technology
Agarpara, Kolkata
aritrarana2@gmail.com

Md. Ashifuddin Mondal
Computer Science and Engineering
Narula Institute of Technology
Agarpara, Kolkata
ashifuddin.mondal@nit.ac.in

2023 World Conference on Communication & Computing (WCONF) | 979-8-3503-1120-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/WCONF58270.2023.10235023

***Abstract*—The Recommendation systems gaining importance every passing day and are frequently used in e-commerce websites, content streaming platforms, social media platforms, and other applications to analyze large amounts of data and recommend products or services to users based on their past actions, preferences, and other pertinent factors. This kind of system can effectively be used in healthcare to provide personalized recommendations to patients. In this paper, a recommendation model has been proposed that flags the users with the diseases that they may have based on analyzing their historical medicine order and recommends them with alternate related medicines and products. The proposed model is based on a machine learning technique that takes into consideration the user's medical history, disease diagnosis, and prescription orders and recommends other related medicine and products. The proposed model uses a combination of supervised and unsupervised learning algorithms. A sizable dataset of patient orders from a top pharmaceutical retailer served as the basis for training and validating the model. The simulation results show that the model can handle complex situations, such as multiple medications and medical conditions, and predict medication recommendations with high accuracy.**

***Keywords*—Recommendation system, Personalized medicine recommendation, Machine learning, eHealth**

I. INTRODUCTION

Digital healthcare [1], commonly referred to as eHealth [2,3], is the application of Information and Communication Technology (ICT) [4] in healthcare to enhance patient outcomes, boost productivity, and cut costs. Information and communication technologies can transform the way healthcare is provided and lead to better health outcomes for individuals all over the world. Through the use of medication recommendations, healthcare services can be enhanced by offering individualized and precise treatment plans, lowering medication errors, enhancing medication adherence, and lowering healthcare costs. Systems for recommending medications assess patient data using machine learning algorithms to offer personalized treatment alternatives, potentially increasing patient outcomes and lowering the possibility of problems or adverse drug responses. The current healthcare system has a number of issues, such as a lack of personalization, pharmaceutical errors, poor drug adherence, restricted access to healthcare, and rising healthcare expenses. Systems that recommend medications can provide individualized treatment regimens, lower medication errors, enhance drug adherence, broaden access to healthcare, and cut expenses associated with it.

Machine learning techniques have attracted a lot of attention in recent years due to their potential to improve patient outcomes and save healthcare expenditures. Medication recommendations based on a patient's medical history are one area where machine learning can be especially helpful. In this study, a machine-learning based strategy has been proposed for identifying users with diseases that they may have and alerting them to medication recommendations based on their previous orders. The model is intended to analyze user data and recommends other alternative medicine and products related to the user's disease by considering the user's medical history, disease diagnosis, and prescription orders. Here, user data refers to the past medicine orders by the user through e-commerce websites [5]. In order to analyze user data and make appropriate medication recommendations based on the user's medical history, the proposed model uses a combination of supervised and unsupervised learning algorithms [6]. A sizable dataset of patient orders from a top pharmaceutical retailer served as the basis for training and validating the model. Findings show that the model can handle complex situations, such as multiple medications and medical conditions, and predict medication recommendations with high accuracy. The suggested system has the potential to enhance patient outcomes, lessen medical errors, and simplify the process of providing healthcare. The ability of the proposed system to analyze a variety of medical data and offer individualized treatment recommendations based on the user's medical history are among its main benefits. By offering individualized treatment regimens based on previous drug orders, this model can benefit patients in the healthcare industry. With this tailored approach, medication adherence can be increased, medication errors can be decreased, and better health outcomes can result. This strategy can also save healthcare expenditures, increase general patient satisfaction, and enhance the effectiveness of healthcare services.

The rest of the paper is organized as follows. A review of related work is given in Section 2. A detailed explanation of the proposed methodology is described in Section 3. Section 4 demonstrates the simulation and experimental results of the suggested model. Finally, the paper is concluded and its future directions are discussed in Section 5. Overall, this study presents a novel strategy for disease flagging and personalized medicine recommendations that have the potential to enhance patient outcomes and greatly lower healthcare costs.

II. RELATED WORK

This section gives an overview of the related work about the recommendation model. Not many research papers have been carried out on the medicine recommendation model. In [7], such type of work has been done where author S. Garg introduced the Drug Recommendation System based on Sentiment Analysis of Drug Reviews using Machine Learning. But in other fields, some recommendation system research work has been found like movie recommendation system, book recommendation system, etc. Authors A. A. Joseph and A. M. Nair in their study [8] proposed Collaborative Movie Recommendation System. In [9], Book Recommendation Using Machine Learning Methods Based on Library Loan Records and Bibliographic Information is researched by authors K. Tsuji, F. Yoshikane, S. Sato, and H. Itsumura. In another study [10], authors propose an online pre-diagnosis doctor recommendation model by integrating ontology characteristics and disease text. The model takes into account the geographical location of patients and their information on symptoms, diagnosis, and geographical location. Results show that the model is of higher recommendation accuracy and feasibility in terms of department and doctor recommendation effectiveness. It also gives full consideration to patients' location factors, improving their online consultation experience and offline treatment convenience. In contrast to the above systems, in [11] the author proposes a recommender system for recommending relevant products in an e-shop using explanations. The proposed system consists of three recommender modules called VIEW, RATING, and PURCHASE. The recommender modules use a content-based filtering approach and a collaborative filtering approach where the proposed recommender system works with explanations that contain arguments about why the system recommended the specific product. Based on these explanations the user sees why specific products are recommended by the system. The proposed system was experimentally verified and the results of the experimental verification are discussed. Online shopping has become an important part of lifestyle, but users may find it difficult to narrow down the range of options. In another work [12], authors propose and implement an e-learning recommender system based on a logical approach and an ontology of material content based on different learning styles. This paper shows that there is a potential to implement a personalized recommender system in e-learning based on the student's learning style. The work presented in [13] discusses the design and development of mobile location-based recommender systems for social distancing in urban areas. The main goal is to enable better real-time recommendations for trip support in time-location amplitude with social distancing reality in urban areas. In another work [14], a recommender system is a filtering system that helps users find the optimal information from vast amounts of information. This paper introduces an architecture of a tourist support information system and proposes a system for gathering contents, repositories, and training data. Another similar work is carried out in [15] where the Application of a Collaborative Filtering Recommendation Algorithm is used in a pharmacy system. In the paper [16], the author proposes a keyword map-based learner profile for content-based recommender systems, which automatically creates a user profile based on visited learning materials and learning processes. It provides good accuracy while avoiding many of the problems of collaborative and keyword-based approaches.

As per the above discussion, it has been observed that many research works have been carried out related to recommendation systems in different application fields like in e-shopping portal, movie recommendation portal, book recommendation and e-learning system. But, not much work other than in [7] has been found in the field of medicine recommendation systems. In view of this, this paper proposed a machine learning-based medicine recommendation model which will not only suggest the related medicines to the users but also it will try to identify the kind of disease that may have. This research on personalized medicine recommendations and illness flagging shows how machine learning algorithms have the potential to offer precise and successful suggestions by utilizing electronic medical records and other healthcare information systems.

III. PROPOSED METHODOLOGY

This section describes a detailed methodology of the proposed model. The proposed model is based on machine learning techniques. This model recommends medicines to the users and flags users by disease based on the highest priority of previously ordered medicines. There are mainly three recommendation systems: Content-based recommendation systems, Collaborative filtering recommendation systems, and Hybrid recommendation systems. The proposed model is one kind of Content-based recommendation system.

A. Overview of the Model

In the current e-healthcare system, mainly in the e-commerce websites/apps where patients can order their medicines. There is a need to make these websites/apps to be more personalization in medicine recommendations and able to identify potential health risks. The proposed model is designed to improve the such website by providing personalized medicine recommendations and disease flagging based on the user's previous orders. Figure 1 shows the basic architecture diagram of a medicine order website/app where this proposed model is used as a content-based recommendation engine. In the medicine order website/app, firstly the user has to create an account if he/she is new otherwise the user logs into the website/app. And then the user goes to the purchase page to order medicines via the catalog page where the user can see different types of medicines for different diseases. From the purchase page, all the details are stored in the database like the user name, user Id, order date, and ordered medicines name. Then the content-based recommendation engine works to generate a recommended medicine list for the user and flag the user by disease based on previously ordered medicines whose data is stored in the database. The recommended medicine list is shown on the catalog page when the user visits this page next time to order some medicines. This total process is shown in Figure 1.

In the content-based recommendation engine, there are many machine-learning algorithms and mathematical concepts. Figure 2 describes the general working procedure of the proposed content-based recommendation engine. In Figure 2, Firstly a dataset of patient medicine orders is collected from a top pharmaceutical retailer to train the model. Another dataset that is required for this model is the disease-medicines list dataset. Then the datasets were read. In the pre-processing stage, datasets were cleaned, formatted, and scaled. Then two new data frames are created from the user order dataset and disease-medicine list dataset with some specific columns named Disease, Medicine Tag, and User name, User Id, Order Tag that are shown in Figure 2. After that, vectors of tags are created by the TF-IDF machine-learning algorithm. TfidfVectorizer is a class in the Python library scikit-learn used for text feature extraction. It is used to convert a collection of raw documents into a matrix of TF-IDF (term frequency-inverse document frequency) features. The term frequency-inverse document frequency (TF-IDF) method is a well-liked information retrieval methodology that assesses the significance of a word to a document within a collection of documents. Text data is frequently transformed into feature vectors for machine learning models in applications such as text mining, information retrieval, and machine learning.

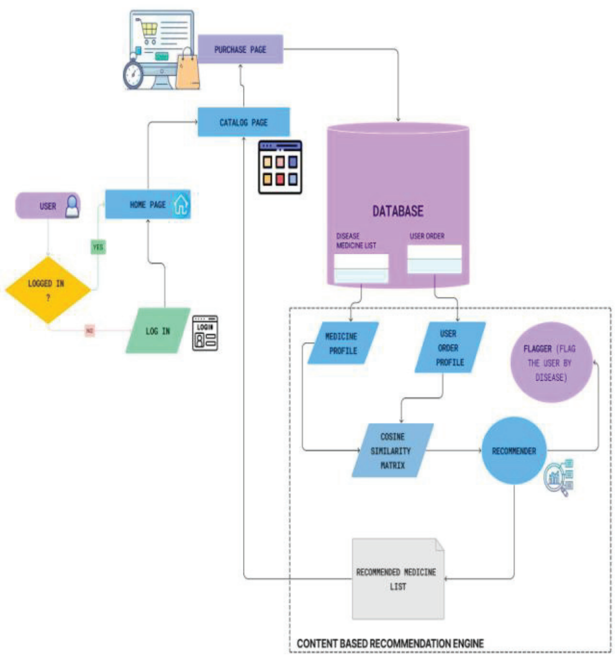

Fig. 1. Basic architecture diagram of the proposed medicine recommendation model

The fundamental tenet of TF-IDF is that words that are frequent in one text but uncommon in another are given greater weight in that document. The algorithm works in two main steps:

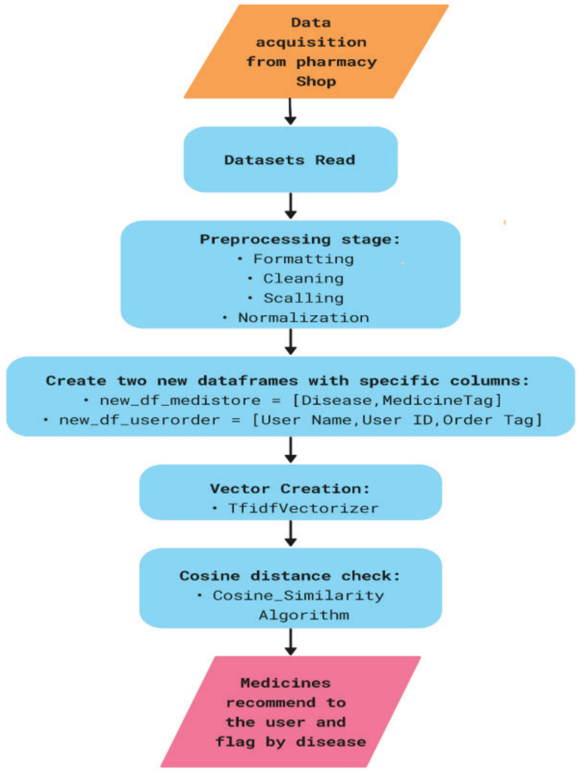

Fig. 2. General working procedure of the proposed recommendation engine

1. Term Frequency (TF): Determine how often each word or phrase appears in the text. The most typical way to determine a term's frequency is to count how many times it appears in the text and divide that total by the number of words.

2. Inverse Document Frequency (IDF): Determine each word's inverse document frequency, which gauges how uncommon a word is in all documents. The IDF is determined by calculating the logarithm of the quotient obtained by dividing the total number of documents by the number of documents containing the word.

$$w_{i,j} = tf_{i,j} \times idf_i \text{ .... } (1)$$

Where $w_{i,j}$ is the TF-IDF score for term i in document j, $tf_{i,j}$ is term frequency for term i in document j, and idf_i is the IDF score for term i.

After vector creation, two vectors are passed through a cosine-similarity algorithm to check the similarities (similarity distance) between two vectors which are shown in Figure 2.

Cosine similarity is a metric for comparing two non-zero vectors in an inner product space that calculates the cosine of the angle between them. To evaluate how similar two documents are, the cosine similarity technique is frequently used in text mining and information retrieval. The below formula is used to compute the cosine similarity:

$$cosine_{similarity} = S_c(A,B) := COS(\theta) = \frac{A.B}{|A||B|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \quad (2)$$

where A_i and B_i are the ith components of vectors A and B respectively.

The value of similarity ranges from -1 meaning exactly opposite, to 1 meaning the same, with 0 indicating orthogonality or decorrelation, while in-between values indicate intermediate similarity or dissimilarity. Based on the cosine similarity score a medicine list is recommended to the user and flagged the user by disease.

B. *Dataset Description*

Two datasets have been used in the proposed model. The first one is the user medicine order dataset which is collected from a pharmacy shop. Table 1 shows the part of the user medicine order dataset where user details are stored with their ordered medicines. In the user medicine order dataset shown in Table 1, the user name, user Id, purchase date, and ordered medicines are present. Another one is the disease medicine dataset, where medicine names are there for the proper treatment of different types of diseases. Table 2. shows the part of the disease medicine dataset that is created by doctors.

TABLE I. USER MEDICINE ORDER DATASET

User Name	User ID	Purchase Date	Ordered Medicines
Renee Glass	163740	28-Feb-12	Mirtazapine, Levofloxacin, Avelox, Escitalopram
Heather Lyons	206473	17-May-09	Monodox, Ziana, Isotretinoin, Macrobid
Amanda Vang	159672	29-Sep-17	Onabotulinumtoxi, Amitriptyline, Cymbalta, Bactrim
….	….	…..	……
Christopher Cross	215892	6-Jun-16	Lamotrigine, Etonogestrel, Dextromethorphan, Copper
Lori Cain	169852	21-Apr-09	Restoril, Bromfed DM, Mucinex DM, Amitriptyline
Michael Berry	23295	18-Oct-16	Biaxin XL, Geodon, Lamotrigine, Quetiapine
Jesus Mcpherson	71428	16-Apr-11	Nicoderm CQ, Levo-Dromoran, Bactrim, Levora

IV. SIMULATION RESULTS

This section presents the simulation results of the proposed model. The results show that the proposed model performed very well for medication recommendation and disease flagging. For the simulation, two data sets have been used as mentioned earlier and presented in Table 1 and Table 2. Figure 3 shows the similarity matrix between usernames and disease names based on previous medicine orders. Based on the highest similarity score, users are flagged by the disease shown in Figure 3.

A graph is shown in Figure 4 which describes the priority of diseases according to the user's medicine order. In the graph, the x-axis represents the user name and the y-axis represents the similarity score between disease and user-ordered medicines. In Figure 4 green circles in the green line represent top priority disease based on previous medicine orders. Based on the top priority disease, users are flagged by this disease and a recommended medicine list for that disease is created.

TABLE II. DISEASE MEDICINE DATASET

Disease Name	**Medicine Name**					
Depression	Escitalopram	Zoloft	Effexor XR	Venlafaxine	Bupropion	Desvenlafaxine
Asthma	Prednisone	AdvairHFA	Dulera	Montelukast	Singulair	Albuterol
Cough	Dextromethorphan	Etonogestrel	Mucinex DM	Bromfed DM	Codeine	Hycodan
….	…..	……	……	……	……	…….
Heart Attack	Metoprolol	Plavix	Aspirin	Metformin	Bayer Aspirin	Clopidogrel
Birth Control	Cyclafem 1 / 35	Copper	Levora	Blisovi Fe 1 / 20	NuvaRing	Ethinyl estradiol
High Blood Pressure	Ramipril	Atenolol	Hydrochlorothiazide	Benicar	Losartan	olmesartan

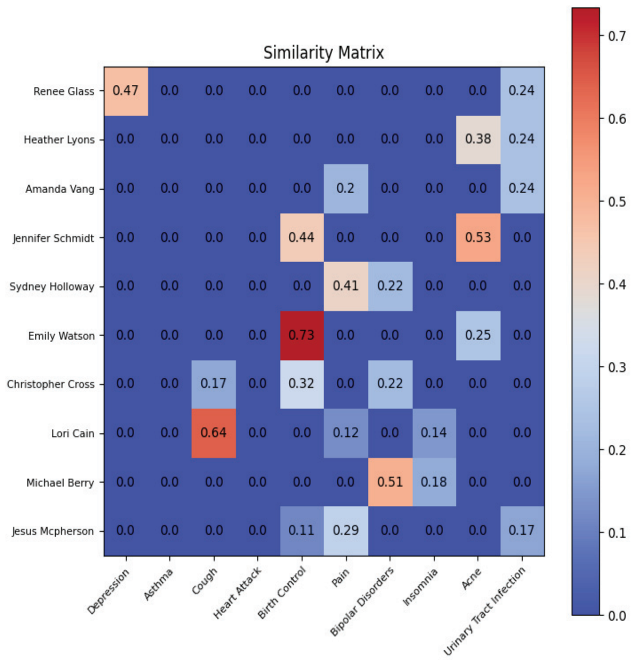

Fig. 3. Similarity Matrix of user disease

Table 3 shows the user details dataset after flagging the disease based on the previous order. Whereas Table 4 shows the recommended medicine dataset with the flagged disease of the user based on previously ordered medicines.

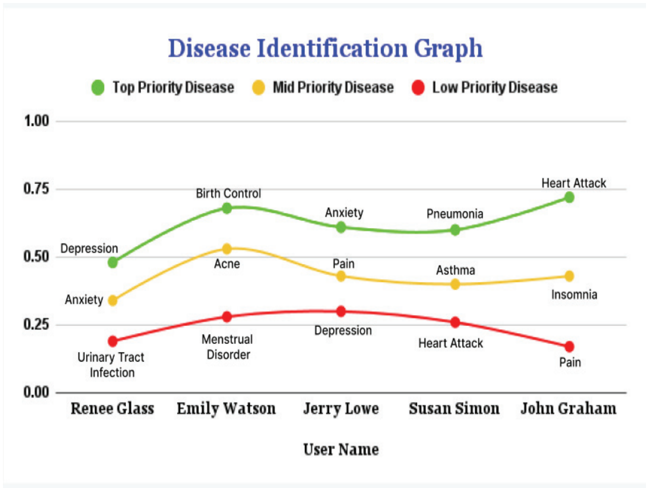

Fig. 4. Disease Identification Graph

Based on the experimental results, the accuracy of the proposed model is found to be above 90%, indicating a high level of accuracy in predicting the recommended medications and disease flags based on the user's previous orders. The model has the potential to enhance patient outcomes, lower prescription errors, and ultimately raise the standard of healthcare services due to its flexibility in adjusting to new patient data and offering personalized suggestions based on their medical history. To confirm the efficacy of the suggested approach and its possible effects on the healthcare sector, additional study and testing are required.

V. CONCLUSION

In conclusion, the proposed personalized medicine recommendations and disease flagging models based on users' previous orders are promising approaches for improving healthcare services. This model can offer precise and individualized medication recommendations as well as indicate prospective ailments based on a patient's prior orders by utilizing machine learning techniques and patient data. The suggested approach entails gathering patient data, pre-processing the data, creating a cosine similarity-based recommendation system, and putting a decision-tree-based disease flagging model into action. The findings of the experiment show that the model can effectively suggest medications to patients based on their prior purchases and can also precisely identify probable ailments. Additionally, this model can aid healthcare professionals in making better judgment calls and enhancing patient outcomes. However, there are still some issues that must be resolved, such as the need for more diverse and extensive users and their corresponding medicine order datasets. These issues can be addressed in future studies, and the model can be further improved to increase its efficacy and accuracy. Overall, there is great potential for the Personalized Medicine Recommendation and Disease Flagging Model Based on User's Previous Orders to revolutionize the healthcare sector and enhance patient care.

TABLE III. DISEASE FLAGGING

User Name	User ID	Purchase Date	Ordered Medicines	Flagged Disease
Renee Glass	163740	28-Feb-12	Mirtazapine, Levofloxacin, Avelox, Escitalopram	Depression
Heather Lyons	206473	17-May-09	Monodox, Ziana, Isotretinoin, Macrobid	Acne
Amanda Vang	159672	29-Sep-17	OnabotulinumtoxinA, Amitriptyline, Cymbalta, Bactrim	Migraine
…..	…..	…..	…..	…..
Jesus Mcpherson	71428	16-Apr-11	Nicoderm CQ, Levo-Dromoran, Bactrim, Levora	Pain

TABLE IV. RECOMMENDED MEDICINES

User Name	Previous Ordered Medicines	Flagged Disease	Recommendation Medicines
Renee Glass	Mirtazapine, Levofloxacin, Avelox, Escitalopram	Depression	'Escitalopram', 'Zoloft', 'Effexor XR', 'Venlafaxine', 'Bupropion', 'Desvenlafaxine', 'Bactrim', 'Uti-roz', 'Methenamine', 'Levaquin'
Heather Lyons	Monodox, Ziana, Isotretinoin, Macrobid	Acne	'Ziana', 'Isotretinoin', 'Ethinyl Estradiol', 'Benzoyl Peroxide', 'Monodox', 'Spironolactone', 'Bactrim', 'Uti-roz', 'Methenamine', 'Levaquin'
…..	…..	…..	…..
Sydney Holloway	Cyclafem 1 / 35, Ultram, Aripiprazole, Hydrocodone	Pain	'Ultram', 'Hydrocodone', 'Cymbalta', 'Atorvastatin', 'Oxycodone', 'Voltaren Gel', 'Lamotrigine', 'Geodon', 'Lithium', 'Asenapine'

REFERENCES

[1] J. Han and J. Lee, "Digital Healthcare Industry and Technology Trends," in 2021 IEEE International Conference on Big Data and Smart Computing (BigComp), Jeju Island, Korea (South), 2021 pp. 375-377. doi: 10.1109/BigComp51126.2021.00083

[2] Y. Wang and H. Chang, "Assessing the Performance of e-Health Service," in Service Sciences, International Joint Conference on, Taipei, Taiwan, 2011 pp. 232-236. doi: 10.1109/IJCSS.2011.53

[3] P. Stelmach, L. Falas, G. Kasiukiewicz, P. Kwaénicka, and P. Światek, "IoT modeling and runtime suite for e-Health," 2016 IEEE 18th International Conference on e-Health Networking, Applications, and Services (Healthcom), Munich, Germany, 2016, pp. 1-5, doi: 10.1109/HealthCom.2016.7749433.

[4] C. Renato and N. Maria, "Technologies' Application, Rules, and Challenges of Information Security on Information and Communication Technologies," in 2015 Asia-Pacific Conference on Computer Aided System Engineering (APCASE), Quito, Ecuador, 2015 pp. 380-386. doi: 10.1109/APCASE.2015.74

[5] X. Hao, L. Duo-lin and L. Zhi-jie, "The Research on E-commerce Website Success Mode," in Wearable Computing Systems, Asia-Pacific Conference on, Shenzhen, China, 2010 pp.299-302. doi: 10.1109/APWCS.2010.82

[6] S. Sapkal, P. Revankar, and S. Kakarwal, "Analysis of Classification by Supervised and Unsupervised Learning," in Computational Intelligence and Multimedia Applications, International Conference on, Sivakasi, Tamil Nadu, India, 2007 pp. 280-284.doi: 10.1109/ICCIMA.2007.237

[7] S. Garg, "Drug Recommendation System based on Sentiment Analysis of Drug Reviews using Machine Learning," 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2021, pp. 175-181, doi: 10.1109/Confluence51648.2021.9377188.

[8] A. A. Joseph and A. M. Nair, "A Comparative Study of Collaborative Movie Recommendation System," 2022 International Conference on Electronics and Renewable Systems (SEARS), Tuticorin, India, 2022, pp. 1579-1583, doi: 10.1109/ICEARS53579.2022.9752015.

[9] K. Tsuji, F. Yoshikane, S. Sato and H. Itsumura, "Book Recommendation Using Machine Learning Methods Based on Library Loan Records and Bibliographic Information," 2014 IIAI 3rd International Conference on Advanced Applied Informatics, Kokura, Japan, 2014, pp. 76-79, doi: 10.1109/IIAI-AAI.2014.26.

[10] C. Ju and S. Zhang, "Doctor Recommendation Model for Pre-Diagnosis Online in China: Integrating Ontology Characteristics and Disease Text Mining," 2021 IEEE 6th International Conference on Big Data Analytics (ICBDA), Xiamen, China, 2021, pp. 38-43, doi: 10.1109/ICBDA51983.2021.9402991.

[11] B. Walek and P. Fajmon, "A Recommender System for Recommending Suitable Products in Eshop Using Explanations," 2022 3rd International Conference on Artificial Intelligence, Robotics, and Control (AIRC), Cairo, Egypt, 2022, pp. 16-20, doi: 10.1109/AIRC56195.2022.9836983.

[12] N. N. Qomariyah and A. N. Fajar, "Recommender System for e-Learning based on Personal Learning Style," 2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Yogyakarta, Indonesia, 2019, pp. 563-567, doi: 10.1109/ISRITI48646.2019.9034568.

[13] O. Artemenko, V. Pasichnyk, and N. Kunanets, "Using Mobile Location-based Recommender Systems for Providing Real-Time Recommendations for Social Distancing Urban Route Planning," 2020 IEEE 15th International Conference on Computer Sciences and Information Technologies (CSIT), Zbarazh, Ukraine, 2020, pp. 305-308, doi: 10.1109/CSIT49958.2020.9321969.

[14] G. Hirakawa, G. Satoh, K. Hisazumi and Y. Shibata, "Data Gathering System for Recommender System in Tourism," 2015 18th International Conference on Network-Based Information Systems, Taipei, Taiwan, 2015, pp. 521-525, doi: 10.1109/NBiS.2015.78.

[15] Yihang Lv and Jingjing Kong 2021 *J. Phys.: Conf. Ser.* 1865 042113, doi: 10.1088/1742-6596/1865/4/042113.

[16] X. Wan, Q. Jamaliding, F. Anma and T. Okamoto, "Applying Keyword Map Based Learner Profile to a Recommender System for Group Learning Support," 2010 Second International Workshop on Education Technology and Computer Science, Wuhan, China, 2010, pp. 3-6, doi: 10.1109/ETCS.2010.439.