IJWR

A Recommendation System in the Medical Industry using SW-DBSCAN Algorithm

Reza Molaee Fard[a], Payam Yarahmadi*[b]

[a] Department of Computer Engineering, Dezful branch, Islamic Azad University, Dezful, Iran; molaefard@gmail.com
[b] Department of Computer engineering Taras National University of Kyiv; Yarahmadi.pay@gmail.com

ABSTRACT

A recommendation system is a system that, based on a limited amount of information provided by users as well as the feedback given to goods, persons, and locations by other users, provides appropriate suggestions to the user. Today, with the large number of physicians and specialists, it seems necessary to have a system for identifying the right specialist and experienced physician for the patient. We present in this study a system for medical recommendations that analyzes physicians and specialists. It uses collaborative filtering and scores provided by other users to suggest physician recommendations according to the area of expertise of the physician. Research conducted and evaluation of results show that this system can successfully recommend a specialist doctor to the user in 90% of cases.

Keywords: ***Recommender System, Medical Industry, Web Mining, Collaborative Filtering, SW-DBSCAN Clustering Algorithm.***

1. Introduction

Medicine is a science that aims to maintain and promote health, cure diseases and rehabilitate the injured. These goals are achieved by recognizing diseases, diagnosing, treating and preventing their occurrence. Doctors encounter many people from different sexual, social, and age groups on a daily basis, who, despite their differences, all carry the common name of the patient. Patients know that they must trust their doctor for treatment for their health and well-being, and as a result of this trust, they leave the decision about their health to the doctor. It has been recorded almost throughout history and in every part of the world, being a doctor has had a special meaning and people turn to doctors for help with their most urgent needs, relief from pain and restoration of health. Medicine existed in ancient Egypt, medical information on papyrus dates back to 3000 BC. Many sources consider Hippocrates to be the father of medicine, he laid the foundations for a rational approach to medicine. The Hippocratic Oath is one of the oldest laws of medical ethics still used by physicians today. He was the first to classify diseases into acute and chronic, endemic and epidemic. The World Health Organization has introduced the criteria of a good doctor in 5 areas, which is called a 5-star doctor. A) Provider of health care: a person who thinks of measures to prevent the occurrence of diseases themselves and to prevent the complications of diseases so that people can enjoy appropriate health care. B) Decision maker: A good doctor decides what technology and treatment methods are suitable for him according to the specific conditions of each disease and taking into account human issues. These decisions are unique to the situation, and strong decision-making power is one of the most basic characteristics of a physician. C) Communicate: A doctor is an example of someone who can promote a healthy lifestyle in the community through effective communication between individuals and society, and encourage people to follow health tips and maintain the health of themselves and their community. Good public relations is part of the character of a successful doctor. D) Social Leader: A physician with the trust of the community, understands the health problems of individuals and the community and can create and lead the right movement in this area. The ability to manage the community and group therapy is one of the main capabilities of a physician. E) Group member: A person who enters the field of medicine must coordinate with individuals and members of the medical staff, groups and organizations and achieve the set goals in the form of a collective work. Tourism is also very important in Iran, but the existence of a system that can introduce a specialist and reputable physician to the patient from among the large number of physicians in the country seems necessary. The best system for this is the use of recommender systems, which today are associated with great progress and have gained the trust of users in all fields. Referrer systems are systems that, with a range of data mining and web mining methods and information such as what doctors have searched for in the past, as well as obtaining limited information from users, can provide appropriate suggestions to the user.

2. Related work

Jianjia et al presented a method to recommend a specialist doctor to users in the field of medical supply. These researchers believed that in the context of industrial convergence, the business needs of users also show a trend of diversification and personalization. The phenomenon of the intersection of multi-service resources and organizational information barriers in the field of traditional medical supply leads to the delay of the recommendation of medical resources, therefore, the recommendation of a specialist doctor on a platform can improve the recommendation in the field of medicine. In this research, they used a single value decomposition algorithm based on time context to solve the model to achieve rational medical recommendations. These researchers' research showed that dynamic multi-service resources in the medical supply chain. The results showed that

http://dx.doi.org/10.22133/ijwr.2023.357731.1131

Citation R. Molaee Fard and P.Yarahmadi, "Provide a Recommendation System in the Medical Industry using SW-DBSCAN Algorithm," *International Journal of Web Research*, vol.5,no.1, pp.82-87, 2022, doi: http://dx.doi.org/ 10.22133/ijwr.2023.357731.1131.

*Coressponding Author
Article History: Received: 24 July 2022; 10 September 2022; Accepted: 20 September 2022

the proposed cloud-based dynamic medical business resource joint proposal model can effectively achieve medical business recommendations and ideas to reduce operational costs of alliances of companies. Provide medicine under the conditions of industrial connection and improve the efficiency of industrial resources [1].

Ponselvakumar et al reviewed recommender systems in the medical field. These researchers believed that the field of precision medicine has made great progress with big data, deep learning. Personalized health information provides greater insight into patient care in all aspects, leading to better treatment. Many researchers and people agree that the quality of diagnosis and personalized medicine is more provides an overview of methods, algorithms, frameworks developed for personalized health care. This paper focuses on two main cases, one is a recommendation system to recommend the doctor to the user, which provides an insight mechanism and a more approach to patient diagnosis, the second is deep learning mechanisms in various fields of healthcare, bioinformatics and genomics to provide results. Accurately focuses on improvements. In the algorithm of this article, it shows a combination of recommendation system and deep learning quality of accurate healthcare to patients [2].

Zhang et al used physiological data and a medical database to recommend a doctor to a user on a cloud network to recommend more valuable and reliable services to the target user. That is, diagnostic decision making and treatment recommendations, which provide subtle and accurate guidance for medical cases. Neighborhood-based recommendation (NBR) provides an efficient and straightforward approach to find like-minded nodes to generate recommendations, automating what is commonly known as word-of-mouth. In this encryption system, it uses BGN as a building block to encrypt data, and then creatively uses graph theory to expand neighbors. It searches for the best matching features and uses BLS signatures to ensure authentication, while also implementing the negligent transfer protocol to maintain the confidentiality of the recommendation. Several experiments have been conducted to evaluate the performance in terms of computational costs and communication costs, and show that the scheme is efficient and practical for medical recommendations [3].

Zhou et al discussed a technology on a platform for online medical advice and consultation. The rapidly developing Health 2.0 technology has provided people with more opportunities for online medical consultation than ever before. Understanding the contexts in various online medical communication and activities becomes an important issue to facilitate the medical decision making process of patients. As a subset of machine learning, neural networks have attracted increasing attention in natural language processing applications. In this paper, these researchers focused on modeling and analyzing patient-physician-generated data based on an integrated CNN-RNN framework to deal with situations where online patient questions are usually not too long. A so-called DP-CRNN algorithm with a novel neural network structure is designed to extract and highlight a combination of semantic and ordinal features in terms of patient questions. An intelligent recommendation method is then proposed to provide automatic physician guidance and pre-diagnosis suggestions to patients, where the clustering mechanism is used to refine the learning process with more accurate diagnosis scope and more representative features. Experiments based on collected real-world data demonstrate the effectiveness of the proposed model and method for intelligent physician recommendation services in online medical environments [4].

Guzman et al in [5] present a recommender system for medical care as well as physician referral to patients. This standard proposes a common measure of physiological changes to determine possible high-risk disease conditions, as well as a weighted ranking of specialized medical centers. This set consists of two steps: in the first step, a ubiquitous heart rate monitor is applied using the ID3 decision to classify the sensed data to detect the presence of a high heart disease. The second step proposes a recommender system that selects and clusters a set of hospitals, where medical specialties are ontologically defined in a functional definition. The clustering process is factored into consideration to determine the number of potential physicians and medical expertise required. As a result, it applies different decision-making proposals such as ID3, J48, NBTree and BFTree in order to evaluate and compare the segmentation performances. The effectiveness of the ID3 decision tree was 85.71%.

Ochoa and colleagues in [6] presented a method for evaluating patients by doctors in the context of a recommender system. These researchers believed that doctors could reduce their workload by using recommender systems and still be in control of every decision. In this way, they also gain insight into how other doctors make decisions in any given situation. In this work, the development of a novel recommender system that uses predictive outcomes based on continuous-valued logic and multi-criteria decision operators was reported. The advantage of this method is that it is transparent, because the model results mimic logical decision-making processes based on a hierarchy of relevant physiological parameters, and second, it is more secure against adversarial attacks than conventional deep learning methods, because the number of training greatly reduces the teachable. Finally, this system can meet the needs of patients and doctors can have less workload.

Rostam et al In [7], presented a method to improve medical recommender systems. The proposed approach consists of two modules, module-1 aims to train machine learning models using a dataset of disease symptoms and associated symptoms and precautions. Preprocessing and feature extraction are performed as prerequisite steps. In Module 1, several algorithms are applied to the disease dataset, such as support vector machine, random forest, additive tree classifier, logistic regression, simple polynomial Bayes, and decision tree. Module-2 interacts with the user (patient) through which the patient can describe the symptoms of the disease using a microphone.

Erdaniz et al [8] presented a method called QS to suggest medical services for patients. In this method, three recommendation approaches are proposed for QS applications: virtual doctor recommendation, virtual nurse recommendation, and virtual sleep regulator that helps QS users to improve their health conditions. The virtual trainer works like a real fitness trainer to recommend personalized exercise programs, while the virtual nurse takes into account the user's medical history and health goals to recommend the appropriate physical activity program. The virtual sleep regulator is specially designed for insomnia patients to improve their sleep quality

with the help of physical activity and recommended sleep schedules. This technology applied recommendation based on collected QS data to generate quality recommendations for user needs. An example of the recommended results of a virtual sleep regulator based on a dataset from a real-world QS application was presented, indicating the improved results of this method over other methods.

Hung et al presented a dental recommender system in [9]. In this research, feature selection was done using LASSO in R to determine the best regression model. Prediction models were developed using several supervised machine learning algorithms, including logistic regression, support vector machine, random forest, and classification and regression trees. Feature selection by LASSO combined with the inclusion of additional clinically relevant variables identified the top 8 features associated with recommendations for dental care. The top 3 characteristics were gum health, number of prescription medications taken, and race. Gum health shows a significantly higher relative importance than other characteristics. Demographics, access to health care, and general health variables were identified as the top characteristics associated with receiving additional dental care, consistent with previous research. Practicing dentists and other healthcare professionals can follow this model to practice precision dentistry by integrating our algorithms into a computerized screening tool or decision tree diagram to achieve more efficient and personalized preventive strategies and treatment protocols. Activate dental care.

2.1 Recommender System

Recommender systems are a subset of an information filtering system that seeks to predict the score or priority that a user will give to an item (data, information, product, etc.). In recent years, recommendation systems have become very common and have been used in various fields. Some popular applications of these devices include: music, web pages, news, books and articles, tourism industry, hospitality industry, medical industry, search and social networks. In fact, recommender systems act like a filter, a filter that shows only what the user wants, which is called personalizing information. Overall, the recommender system is a personalization support system that measures information for the three determinants, customization, interest, and usefulness, for specific users by analyzing their tastes and the content of the items. Recommended system is one of the tools that can guide users in electronic environments to find the information, services and items they want [10, 11]. Recommending systems, capable of detecting users' interests and predicting their priority, refine items that are likely to be of interest to the user from the high volume of data or offer them time-saving. On the other hand, these systems also infer the ability to store and analyze the user's past behaviors, services and information that are not of interest to users and may be interested in them, and provide interesting results to users. In fact, recommender systems are one of the main tools to overcome the problem of information redundancy and with the power to analyze user behaviors, it is an intelligent complement to the concepts of information retrieval and refinement. Today, recommender systems are used in various fields from refining the information available on the web according to the user's wishes to e-commerce, offering movies, music, books, articles, etc. [12, 13].

2.2 Data Mining

The U.S. Centers for Disease Control and Prevention's Accountability Agency says data mining requires the use of advanced tools to analyze and discover relationships and extract valuable data patterns in order to gain meaningful new rules. Discovering knowledge within data in the information age is one of the most exciting and key concepts that is becoming more and more important. According to the definitions and interpretations proposed from different perspectives, two basic components can be identified in data mining, the first is the discovery of hidden patterns in the data and the second is the use of these patterns to predict future results. In recent years, with the rapid and widespread development of networks and database technology, they have also faced a huge amount of information. With the proliferation of digital libraries and the proliferation of heterogeneous information resources, presents the enigma of information richness and lack of knowledge. In it, data mining technology with two important principles of predicting results and knowledge discovery in the web of global web measurement, and more importantly in attracting specific users with diverse information sources has found great practical value [14, 15].

2.3 Web Mining

The term web mining is synonymous with one of the terms knowledge extraction, information retrieval, data verification, and even data dredging, which actually describes the discovery of knowledge in databases. Thus, the idea underlying data mining is an important process of recognizing potentially useful and understandable data patterns. Data mining, or in other words, the discovery of knowledge in databases, is the obvious extraction of potentially useful information from events that are previously unknown. This includes some technical methods such as clustering, summarizing data, learning classification rules, finding network connections, analyzing changes, and detecting irregularities. Web mining refers to all data mining activities and related techniques used to automatically discover and extract knowledge from documents and web services. There is a lot of information and heterogeneity in the web environment that makes it more difficult to acquire knowledge contained in the content of web pages, so in such an environment it is necessary to use data mining tools and techniques to discover relevant information and knowledge [16, 17].

2.4 Clustering

In recommender systems, clustering can be used to group similar data together so that suggestions can be searched and constructed only on clusters that are expected to be more interesting to the user. As a result, clustering techniques have been used to increase the scalability of algorithms behind recommender systems. This issue has shown positive effects on the performance of the algorithm, especially in situations where it is necessary to search the entire problem space to suggest an item [18, 19].

3. Proposed Methods

In the proposed method, a new method is presented in order to improve the medical advisor system. In this way, first a list of all doctors and medical specialists in the country is collected. Then we have to pre-process the data on the available information because the data cannot be injected raw

into the data mining algorithms. In the next step, we need to clear the data of additional and unused data. Then we have to normalize the data, we use the Min-Max Normalization method to normalize the data. The purpose of normalization is to eliminate redundancy and maintain dependencies between data. Then in the next step we have to cluster the data. For this purpose we use the SW-DBSCAN algorithm and we will use the Euclidean method to measure the similarity of the clustered data. Finally, using a referral system based on participatory refinement, a list of specialist physicians is recommended to the user who may be of interest to him. Figure 1 shows an overview of the proposed method.

3.1 Data

The dataset used is a list of about 2420 doctors specializing in different medical fields, 700 of which were randomly selected. These 700 options were related to doctors with different medical fields, which were analyzed using data mining tools. In this way, the doctors who got the most points from the patients were placed in the highest recommended level for the use of the recommender system. The list of doctors using the rating table of the collaborative refinement recommender system and the points that were given to these doctors, and obviously the higher score was placed in the 1st rank of the suggested list and the lowest score was placed at the end of the recommended list. To evaluate the proposed method, 17 patients were invited to test the proposed system, out of 17 patients, 12 patients were 90% satisfied with the proposed system, 4 patients were 74% satisfied, and 1 patient was 50% satisfied with the performance of the proposed system.

3.2 Preprocessing Data

In the first step, we need to pre-process the data because it is usually not possible to inject raw data into data mining and machine learning algorithms. Preprocessing plays an essential role in the data processing process and their results.

3.3 Data Normalization

Data normalization is one of the most important pre-processing steps in data mining. The purpose of normalization is to remove the redundancy of data and keep the dependence between related data. We use the Min-Max Normalization method to normalize the data. Equ (1):

$$z = \frac{X - MIN(X)}{MAX(X) - MIN} \quad (1)$$

3.4 Web Mining Application

After the web logs pre-processing step, the web mining application is applied to the user access sessions. As an important tool for web mining applications, clustering helps to group sessions into clusters based on their common characteristics [17].

3.5 Data clustering with SW-DBSCAN algorithm

Next we need to cluster our data to find similarities between the data. The clustering method used in the research is the use of an improved version of the DBSCAN clustering algorithm, SW-DBSCAN. SW-DBSCAN is a network-based clustering method to reduce the time complexity and increase the efficiency of DBSCAN to obtain more accurate clusters, which we cluster in this research through this algorithm. First, we divide the network into different cells and then run the

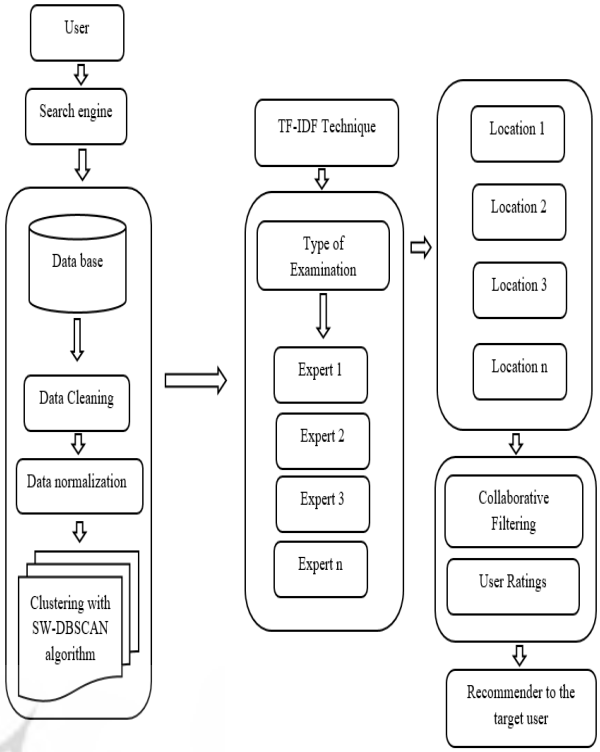

Figure. 1. View of the proposed method

DBSCAN algorithm on the other cells according to the number of points in each cell. Then we change the coordinates of the network. We now have a new network in which we will apply the same method as before. We now have two networks. The next step is to integrate the resulting clusters in the firstnetwork according to the second network. To do this, we create a matrix that helps us determine which categories should be merged. Finally, we integrate the first network clusters based on the generated matrix. As a result, it gives us clusters that are so accurate that this data can be used to identify the user's interest in a particular set of data [21].

3.6 Data Analysis

After clustering the data, we need to analyze the information based on the types of expertise of different physicians. There are many software's available for analyzing clustered information. In this study, we use R software to analyze clustered data and classify physicians into different categories.

3.7 Collaborative Filtering

Collaborative refinement is the process of refining or evaluating items using user feedback. Although only a little over a decade has passed since the formation of the term participatory refinement, the philosophy of this method is that it uses the opinions of others to make decisions, and has been used by humans for centuries. For example, if your friends describe a book, you will want to read it. Or conversely, if you describe a book badly, it is unlikely that you will buy that book. In addition, after a while you will find out which of your friends' opinions are closer to yours, and gradually you will only pay attention to the opinions of those friends who are similar to you. Computers and the Web have made it possible for us to go beyond personal communication and use the opinions of millions of users instead of making decisions based

on ten or more of our friends and relatives. Computer speed allows us to process these ideas in real time and to know what people like us think about a particular item of which we are unaware.

3.8 The Nearest Neighbor-Based Algorithm

These types of algorithms use the scores given by a similar user to predict a user's interest in a particular item. These similar users are called user neighbors. If *n* is similar to user *u*, *n* is said to be a neighbor of *u*. To predict user's interest in item , the average score given to by neighbor's *u* (including user *n*) should be equal to (1.2). In this equation, is the score that user *n* gave to item. We see how to calculate this algorithm in the following relation. Equ (2):

$$pred(\text{u. i}) = \frac{\sum \text{ncneighbors(u)r ni}}{\text{number of neighbors}} \qquad (2)$$

Predicting the score that user *u* will give to item *i* is obtained by calculating the weighted sum of user point's *u* on similar items *i*: Equ (3):

$$Pred(u,i) = \frac{\sum_{j \in rateditems(u)} itemsim(i,j).rui}{\sum_{j \in rateditems(u)} itemsim(i,j)} \qquad (3)$$

In the end, according to the performed operation, a list of specialized doctors is suggested to the user who may be of interest to her.

4. Evaluation of the Results of the Proposed Method

Accuracy is equal to the number of correct detections of the system over the number of retrieved sets. Equ (4):

$$Precision = \frac{relevant\ item \cap retervid\ item}{retrved\ item} \qquad (4)$$

The call is calculated using the following relationship:

Recall is equal to the ratio of the number of correct system diagnoses to the total number of benchmark system sets. Equ (5):

$$Recall = \frac{relevant\ item \cap retervid\ item}{relevant\ item} \qquad (5)$$

To evaluate the accuracy and recall rate of the proposed method with other existing methods, a comparison was made between the proposed method and the ID3 decision tree algorithms, Modul-1 method and QS method (references 5, 7, 8), and the evaluation results indicated that the proposed method is one of the methods The creature was able to obtain a higher level of accuracy and recall. In the comparison between the proposed method and other existing algorithms in the frequency section, the proposed method was able to obtain a recall rate of 95%, while the ID3 decision tree algorithms, Modul-1 method, and QS method were 79%, 92%, and 90% respectively were obtained. In terms of accuracy, the proposed method achieved a performance of 98% while ID3 decision tree algorithms, Modul-1 method and QS method achieved. (Figure 2, 3, 4).

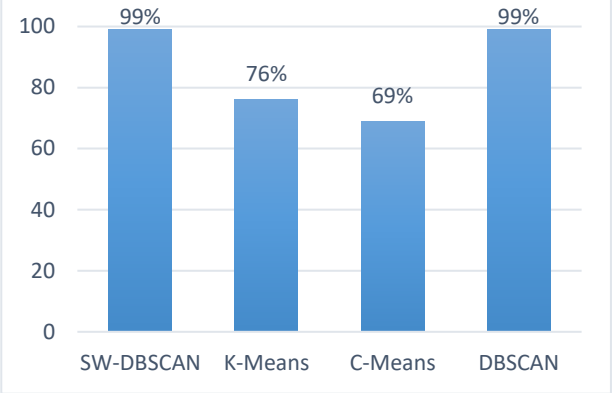

Figure. 2. Comparison of clustering algorithms used in the proposed method and

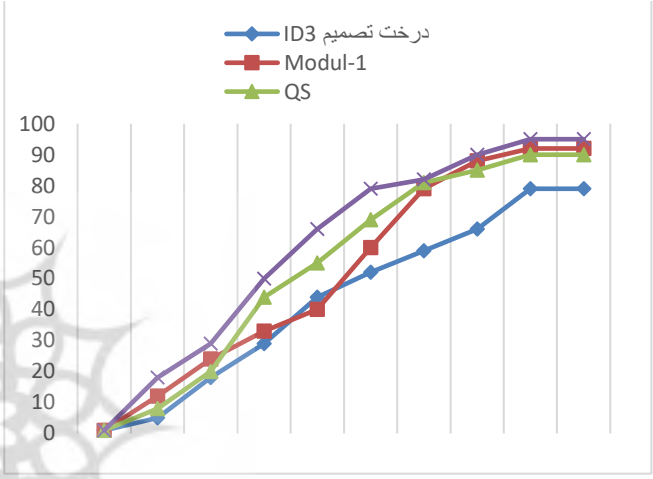

Figure. 3. Comparison of the call rate of the proposed method with other methods

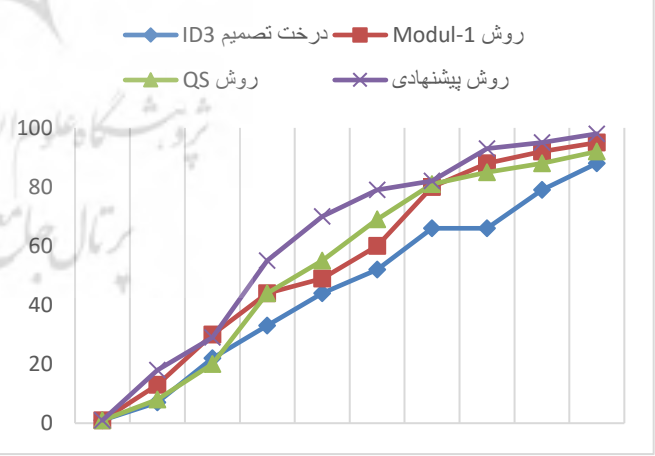

Figure. 4. Comparing the accuracy of the proposed method with other methods

88%, 95% and 92%, respectively, which according to the results it was found that the proposed method has achieved better performance than other similar algorithms. You can see the results of comparing the proposed method with other existing algorithms in the figure below.

5. Conclusion

In this study, a new method was proposed to improve medical recommendation systems. In this study, first a list of physicians specializing in the country was collected and after analyzing them in different medical categories and groups, then a list of physicians in all parts of the country was introduced to the patient using a collaborative filtering-based recommendation system. The fact that these doctors are rated by other users has a higher percentage of confidence and also has more experience and efficiency that the patient refers to the doctor with full confidence and puts himself under the supervision of a doctor with the relevant specialist. During the research, the SW-DBSCAN clustering used in the research was able to obtain a high efficiency score, and also this recommending system can identify up to 90% of the patient's needs and provide the doctor with the appropriate patient's disease.

For future work, other data mining algorithms such as neural networks and classification algorithms such as SVM can be used. In the clustering section, other clustering algorithms can be used for data clustering. Meta-heuristic algorithms can also be used to optimize data. For the recommendation part, other recommender system methods can be used, such as the content-based filter recommender system or the knowledge-based recommender system or the combined recommender systems.

Declarations

Funding

This research did not receive any grant from funding agencies in the public, commercial, or non-profit sectors.

Authors' contributions

PY: Study design, acquisition of data, interpretation of the results, statistical analysis, drafting the manuscript.
OF: Supervision, discussion of the result, revision of the manuscript.

Conflict of interest

The authors declare that there is no conflict of interest.

References

[1] Jianjia, H., Gang, L., Xiaojun, T., & Tingting, L. (2021). Research on collaborative recommendation of dynamic medical services based on cloud platforms in the industrial interconnection environment. Technological Forecasting and Social Change, 170, 120895.

[2] Ponselvakumar, A. P., Anandamurugan, S., Logeswaran, K., Nivashini, S., Showentharya, S. K., & Jayashree, S. S. (2021, February). Advancement in precision medicine and recommendation system for clinical trials using deep learning methods. In IOP conference series: materials science and engineering (Vol. 1055, No. 1, p. 012110). IOP Publishing.

[3] Zhang, M., Chen, Y., & Lin, J. (2021). A privacy-preserving optimization of neighborhood-based recommendation for medical-aided diagnosis and treatment. IEEE Internet of Things Journal, 8(13), 10830-10842.

[4] Zhou, X., Li, Y., & Liang, W. (2020). CNN-RNN based intelligent recommendation for online medical pre-diagnosis support. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 18(3), 912-921.

[5] G. Guzman, M. Torres-Ruiz, V. Tambonero, M. D. Lytras, B. López-Ramírez, R. Quintero, ... and W. Alhalabi, "A Collaborative Framework for Sensing Abnormal Heart Rate Based on a Recommender System: Semantic Recommender System for Healthcare". J. Med. Biol. Eng., Vol. 38, No. 6, pp. 1026–1045, 2018.

[6] J. G. D. Ochoa, O. Csiszár, and T. Schimper, "Medical recommender systems based on continuous-valued logic and multi-criteria decision operators, using interpretable neural networks". BMC medical informatics and decision making, Vol. 21, No. 1, pp. 1-15, 2021.

[7] F. Rustam, Z. Imtiaz, A. Mehmood, V. Rupapara, G. S. Choi, S. Din, and I. Ashraf, "Automated disease diagnosis and precaution recommender system using supervised machine learning". Multimed Tools Appl., Vol. 81, pp. 31929–31952, 2022.

[8] S. P. Erdeniz, A. Menychtas, I. Maglogiannis, A. Felfernig, and T. N. T. Tran, "Recommender systems for IoT enabled quantified-self applications". Evolving Systems, Vol. 11, pp.291–304, 2020.

[9] M. Hung, J. Xu, E. Lauren, M. W. Voss, M. N. Rosales, W. Su, ... and F. W. Licari, "Development of a recommender system for dental care using machine learning". SN Appl. Sci., Vol. 1, p. 785, 2019.

[10] S. Dhelim, N. Aung, M. A. Bouras, H. Ning, and E. Cambria, "A survey on personality-aware recommendation systems". Artificial Intelligence Review, Vol. 55, No. 3, pp. 2409-2454, 2022..

[11] Y. Ge, S. Liu, Z. Fu, J. Tan, Z. Li, S. Xu, ... and Y. Zhang, "A survey on trustworthy recommender systems". arXiv preprint arXiv:2207.12515., 2022.

[12] A. H. Khan, J. Siddqui, and S. S. Sohail, "A Survey of Recommender Systems Based on Semi-supervised Learning". In International Conference on Innovative Computing and Communications, Springer, Singapore, 2022.

[13] [13] B. Alhijawi, Y. Kilani, "The recommender system: a survey". Int. J. Adv. Intell. Paradigms, Vol. 15, No. 3, pp. 229-251, 2020..

[14] M. K. Gupta, and P. Chandra, "A comprehensive survey of data mining". International Journal of Information Technology, Vol. 12, No. 4, 1243-1257, 2020.

[15] Z. S. Ageed, S. R. Zeebaree, M. M. Sadeeq, S. F. Kak, H. S. Yahia, M. R. Mahmood, and I. M. Ibrahim, "Comprehensive survey of big data mining approaches in cloud systems". Qubahan Academic Journal, Vol. 1, No. 2, pp. 29-38, 2021.

[16] K. K. Ibrahim, and A. J. Obaid, "Web Mining Techniques and Technologies: A Landscape View". In Journal of Physics: Conference Series, IOP Publishing, May 2021, Vol. 1879, No. 3, p. 032125.

[17] S. S. Tandel, A. Jamadar, and S°Dudugu, "A survey on text mining techniques". In 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), IEEE, March 2019, pp. 1022-1026.

[18] P. Bhattacharjee, and P. Mitra, "A survey of density based clustering algorithms". Frontiers of Computer Science, Vol. 15, No. 1, pp. 1-27, 2021.

[19] M. Ren, J. Zhang, L. Khoukhi, H. Labiod, and V. Vèque, "A review of clustering algorithms in VANETs". Annals of Telecommunications, Vol. 76, No. 9, pp. 581-603, 2021.

[20] S. A. N. Alexandropoulos, S. B. Kotsiantis, and M. N. Vrahatis, "Data preprocessing in predictive data mining". The Knowledge Engineering Review, Vol. 34, 2019.

[21] N. Ohadi, A. Kamandi, M. Shabankhah, S. M. Fatemi, S. M. Hosseini, and A. Mahmoudi, "SW-DBSCAN: A grid-based DBSCAN algorithm for large datasets". In 2020 6th International Conference on Web Research (ICWR), IEEE, April 2020, pp. 139-145.

Reza Molaee Fard received his MSc degree in 2020 from Islamic Azad University, Dezful Branch. His research interests include data mining, Recommender systems, and wireless sensor networks, and he has conducted research in these areas.

Payam Yarahmadi received his M.SC in 2009 from Taras University of Ukraine and received his P.HD in 2015 from Taras University of Ukraine. His favorite research field is data mining, web mining and machine learning.