Proceedings of the 4th International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/45/20241024

Personalized medical recommendation system supported by medical data

Winston Wang

Phillips Exeter Academy, 20 Main Street, Exeter, NH 03833, USA

wwang7@exeter.edu

Abstract. A personalized medical recommendation system is an intelligent system that utilizes medical data to provide targeted medical advice and services to individuals. With the lack of accumulation and development of medical data, personalized medical recommendation systems have great potential in improving medical effectiveness and saving medical resources. This article aims to explore the principles, methods, and applications of personalized medical recommendation systems based on medical data.

Keywords: Personalized Medical Recommendation System, Medical Data, Algorithms, Feature Extraction, Evaluation and Optimization, Data Security, Privacy Protection.

1. Introduction

1.1. Research Background

With the rapid development of medical technology and medical data, personalized medical recommendation systems have gradually become a hot research direction in the medical field. Traditional medical services usually provide medical advice and diagnosis based on the average level of the public, and cannot fully consider individual differences and special needs.

However, each individual's physical condition, genetic background, lifestyle habits, and other factors vary, and for the same condition or health issue, different individuals may require different treatment plans and medical services. Therefore, the emergence of personalized medical recommendation systems has made up for the shortcomings of traditional medical services, and can provide personalized medical advice and services based on individual characteristics and needs [1].

Medical data, as the foundation of personalized medical recommendation systems, covers various medical information, including disease diagnosis, treatment plans, drug use, patient feedback, etc. Through in-depth analysis and mining of medical data, potential laws and knowledge can be revealed, providing support for personalized medical recommendations.

Therefore, studying personalized medical recommendation systems based on medical data has important theoretical and practical significance. It can improve medical effectiveness, reduce medical burden, optimize resource allocation, and provide patients with more precise and personalized medical advice and services. At the same time, it also faces challenges in data privacy and security protection, and needs to seek reasonable solutions.

DOI: 10.54254/2755-2721/45/20241024

1.2. The significance and challenges of personalized medical recommendation systems

The emergence and development of personalized medical recommendation systems are of great significance. Firstly, a personalized medical recommendation system can provide more accurate and effective medical advice and treatment plans based on the individual differences and special needs of patients. This helps to improve treatment effectiveness and reduce misdiagnosis, missed diagnosis, and unnecessary treatment. Secondly, personalized medical recommendation systems can allocate and schedule resources reasonably based on patients' needs and resource availability. By optimizing resource utilization efficiency, the coverage and quality of medical services can be improved, and the problem of scarce medical resources can be alleviated. Finally, personalized medical recommendation systems can promote patients' active participation in medical decision-making, enhance their understanding and trust in the medical process and treatment plans. This helps to enhance patients' treatment compliance and satisfaction, and improve the quality of doctor-patient interaction [2].

At the same time, personalized medical recommendation systems also face some challenges in practical applications: the quality and credibility of medical data are crucial to the performance of personalized medical recommendation systems. However, medical data are often missing, wrong and biased, and Data cleansing and correction are needed. Medical data contains a large amount of sensitive information, such as patient personal identity, disease diagnosis, etc. Personalized medical recommendation systems require effective privacy protection measures to ensure the security and privacy of patient data. In addition, personalized medical recommendation systems use various complex algorithms and models to generate recommendation results, but these models are usually black boxes and lack interpretability and interpretability. Therefore, how to interpret and interpret recommendation results has become a challenge. Moreover, each individual's physiological characteristics, disease status, environmental background, etc. have differences, and personalized medical recommendation systems need to be able to fully consider these individual differences and complexities, providing personalized medical advice and services. The personalized medical recommendation system involves the recommendation of medical decisions and treatment plans, and needs to comply with relevant laws and ethical norms [3]. Therefore, how to consider these issues in system design is also a challenge. In response to these challenges, future research needs to combine professional knowledge in fields such as medical knowledge, data science, and privacy protection, continuously explore and innovate to improve the reliability, security, and interpretability of personalized medical recommendation systems.

2. Overview of Personalized Medical Recommendation System

2.1. Definition and Function

A personalized medical recommendation system is an intelligent system that provides personalized medical advice and services to patients based on medical data and individual characteristics, utilizing technologies such as machine learning and data mining.

Specifically, the functions of a personalized medical recommendation system include:

1. Disease diagnosis and prediction: By analyzing the symptoms, signs, and medical data of patients, disease diagnosis and prediction are carried out. The system can generate personalized diagnostic results and prognosis evaluations based on different medical characteristics and individual differences.

2. Recommendation of treatment plan: Based on the patient's disease situation and characteristics, recommend the most suitable treatment plan. The system can comprehensively consider factors such as treatment effectiveness, side effects, and individual differences among patients, and provide personalized treatment suggestions to patients.

3. Drug selection and dosage adjustment: Based on the patient's disease status, physiological characteristics, and drug sensitivity, recommend the most suitable drug for the patient and provide personalized dosage adjustment suggestions.

DOI: 10.54254/2755-2721/45/20241024

2.2. Related technologies and algorithms

Personalized medical recommendation systems involve various technologies and algorithms, and the following are some commonly used technologies and algorithms:

1. Collaborative filtering algorithm: Collaborative filtering algorithm makes recommendations based on the similarity between users or items. By analyzing the similarities between patients and other patients or medical resources, recommend medical choices or resource usage for similar patients.

2. Content based recommendation algorithm: Content based recommendation algorithm is to recommend medical services and treatment plans related to patients' individual characteristics by analyzing the features and attributes in medical data. For example, based on the patient's medical history, laboratory tests, and other information, recommend suitable treatment plans.

3. Hybrid recommendation algorithm: The hybrid recommendation algorithm integrates multiple recommendation algorithms to comprehensively consider the advantages of Collaborative filtering and content-based recommendation. By combining the results of different algorithms with weights or collaborating on training, more accurate and personalized recommendation results can be obtained.

3. Personalized Medical Recommendation System Data Collection and Processing

3.1. Sources and characteristics of medical data

The collection and processing of medical data is one of the important steps in personalized medical recommendation systems. The following are the sources and characteristics of medical data:

Source:

1. Electronic medical records: Electronic medical records that record patient medical history, diagnosis, treatment plans, and other information in hospitals or clinics are one of the important sources of medical data.

2. Medical imaging data: Medical imaging data includes X-ray, CT scanning, MRI and other image data, which can provide detailed disease information and disease characteristics.

3. Laboratory examination data: Laboratory examination data includes blood tests, urine tests, biochemical indicators, and other data, providing patients' physiological status and disease-related indicators.

4. Physiological monitoring data: Physiological monitoring data can provide real-time and continuous physiological information by monitoring physiological parameters of human body through medical equipment such as ECG, Electroencephalography, etc.

5. Health questionnaires and survey data: Health questionnaires and survey data are filled out by patients or participants to collect information on individual health behaviors, lifestyle habits, family history, etc.

Features:

1. Diversity: Medical data comes from different sources, including text, images, numerical values, and other types of data.

2. Scale: The scale of medical data is usually large, covering a large amount of patient information and medical records.

3. Multimodality: Medical data is usually multimodal, which combines comprehensive information from multiple data sources, such as diagnostic results combined with images and medical records.

3.2. Data Collection Methods and Privacy Protection

1. Active data collection

Actively obtain medical data through electronic medical record systems, laboratory databases, etc. in hospitals or clinics. This method can ensure the integrity and accuracy of the data, but it requires the consent and compliance of the patient.

2. Health monitoring equipment

DOI: 10.54254/2755-2721/45/20241024

Use health monitoring equipment such as Heart rate monitor and Sphygmomanometer to collect physiological parameters of patients. These devices can continuously monitor the health status of patients and transmit real-time data to personalized medical recommendation systems.

Utilize technologies such as mobile applications and sensors to collect patient behavior, activity, and environmental information. For example, recording a patient's steps, sleep status, etc. through sensors on a smartwatch or smartphone.

3.3. Data preprocessing and cleaning

Data preprocessing and cleaning are important steps in personalized medical recommendation systems, which can improve data quality and accuracy. The following are common data preprocessing and cleaning methods:

1. Missing value processing: Missing values often exist in medical data, and appropriate methods need to be taken for processing. Missing values can be processed through interpolation methods such as mean and median filling, as well as deleting samples or features with a large number of missing values.
2. Data smoothing and noise filtering: medical data may contain Outlier and noise, which will affect the accuracy of model training and results. You can use smoothing algorithms (such as moving average) or noise filtering algorithms (such as median filtering) to smooth data or remove Outlier and noise.
3. Data standardization and normalization: The scale and range of data for different features may vary, and standardization or normalization processing is needed to convert them into a unified scale. Common methods include Z-score normalization and maximum minimum normalization.
4. Feature selection and dimensionality reduction: medical data usually have high dimensional fea.

4. Personalized Medical Recommendation System Feature Extraction and Selection

4.1. Classification and extraction methods of medical features

The medical feature classification and extraction method of personalized medical recommendation system is to extract meaningful features from medical data for training and prediction of recommendation models. The following are some common classification and extraction methods for medical features:

1. Clinical features:

-Medical history information: including the patient's past medical history, allergy history, surgical history, etc.

-Physical indicators: physiological parameters such as body temperature, blood pressure, heart rate, etc.

-Laboratory test results: indicator values such as blood tests, urine tests, etc.

-Imaging features: features extracted from Medical imaging data, such as tumor size, density, etc.

2. Habits and behavioral characteristics:

-Dietary habits: including intake of nutrients, dietary types, dietary preferences, etc.

-Exercise habits: such as exercise intensity, frequency, type of exercise, etc.

-Smoking and alcohol consumption: including the frequency and quantity of smoking and alcohol consumption.

3. Genetic and gene expression characteristics:

-Genomic information: The genomic information of patients obtained through gene sequencing technology.

-Gene expression data: the level of gene expression in cells or tissues obtained by Transcriptome methods.

4.2. Feature selection and dimension reduction algorithm

In the personalized medical recommendation system, the medical Feature selection and dimension reduction algorithm is to extract the key features from the original medical data and reduce the feature dimensions for better recommendation and analysis.

DOI: 10.54254/2755-2721/45/20241024

A commonly used Feature selection algorithm is the correlation coefficient method, which can measure the correlation between features and target variables. By calculating the correlation coefficient between features and target variables, features that are highly correlated with the target variable can be selected as inputs.

Another commonly used Feature selection algorithm is variance threshold method, which can remove features whose variance is lower than a certain threshold. Features with low variance usually contain less information and have limited impact on the performance of recommendation systems.

In addition, the Feature selection algorithm based on machine learning is also widely used in personalized medical recommendation systems. For example, the decision tree algorithm can select important features by calculating the information gain or Gini coefficient of features. Algorithms such as Logistic regression and support vector machine can also be used for Feature selection [4].

After Feature selection, dimension reduction algorithm can further reduce the dimension of features. A commonly used dimensionality reduction algorithm is Principal Component Analysis (PCA), which maps original features to low dimensional spaces through linear transformations. PCA improves the efficiency of recommendation systems by retaining the majority of data information while reducing feature dimensions.

5. Personalized medical recommendation algorithms

5.1. Medical recommendation algorithm based on Collaborative filtering

Medical recommendation algorithm based on Collaborative filtering is one of the commonly used methods in personalized medical recommendation system. This algorithm makes recommendations based on the similarity between users or cases.

Collaborative filtering algorithms are mainly divided into two types: Collaborative filtering based on users and Collaborative filtering based on items (cases).

The user based Collaborative filtering algorithm is to find other users who are most similar to the current user by analyzing the behavior and preferences of users, and recommend the items they like to the current user. For example, in a medical recommendation system, users' medical record data, diagnostic results, and treatment plans can be analyzed to identify other users with the most similar conditions to the current user, and their treatment plans can be recommended to the current user.

The item based Collaborative filtering algorithm is to find other items (cases) most relevant to the current item (case) by analyzing the correlation between different items (cases), and recommend these related items (cases) to the current user. For example, in a medical recommendation system, case data, patient feedback, and treatment effectiveness can be analyzed to identify other cases that are most similar to the current case, and the treatment plans for these related cases can be recommended to the current user [5].

In practical applications, you can choose to use the user based Collaborative filtering algorithm, the project based Collaborative filtering algorithm, or a combination of both, according to the specific situation [6]. At the same time, other technologies, such as content filtering and Feature engineering, can also be combined to improve the recommendation effect.

It should be noted that when using the Collaborative filtering algorithm, sufficient user behavior data or case data are required to calculate the similarity. In addition, attention should also be paid to data sparsity and cold start issues, as well as considerations for protecting user privacy and data security [7].

5.2. Content based medical recommendation algorithm

Content based medical recommendation algorithms are another commonly used method in personalized medical recommendation systems. This algorithm performs recommendations by analyzing and utilizing content features in medical data.

Content based medical recommendation algorithms first require feature extraction and representation of medical data. These features can include diagnostic results, medical record descriptions, treatment plans, drug information, etc. Then, by calculating the similarity between different projects (cases),

DOI: 10.54254/2755-2721/45/20241024

identify other projects (cases) that are most relevant to the current project (case), and recommend these related projects (cases) to the current user [8].

When calculating similarity, Text mining and Natural language processing can be used to process the text features of medical data. For example, we can use bag-of-words model model, TF-IDF algorithm, word embedding technology (such as Word2Vec, BERT) and other methods to express and calculate the similarity of medical data.

In addition to text features, other types of content features can also be considered, such as image features and physiological signal features. By analyzing images or physiological signal data, key features can be extracted, similarity between different items (cases) can be calculated, and recommendations can be made.

It should be noted that when using content-based medical recommendation algorithms, it is necessary to fully utilize the rich information and features of medical data. At the same time, it is also necessary to consider the quality and accuracy of data, as well as the protection of user privacy and data security.

In addition, content-based medical recommendation algorithms can be combined with other algorithms, such as Collaborative filtering, deep learning and other methods to obtain more accurate and effective personalized recommendation results.

6. Personalized Medical Recommendation Systems

6.1. Recommended System Evaluation Indicators

When evaluating personalized medical recommendation systems, multiple indicators can be used to measure their performance and effectiveness. Here are some commonly used evaluation indicators:

1. Precision and Recall: Accuracy refers to the proportion of truly relevant items in the recommended project, while recall refers to the proportion of correctly recommended items among all truly relevant items. Accuracy and recall can measure the accuracy and comprehensiveness of the system for recommended items.

2. F1 score: F1 score is the harmonic average of accuracy and recall, which can comprehensively consider the performance of accuracy and recall [9].

3. Root-mean-square deviation (RMSE): For continuous numerical recommendation problems, RMSE can be used to measure the prediction error of recommendation results. The smaller the RMSE, the smaller the difference between the predicted results and the actual results.

4. Recommendation Coverage: Recommendation coverage reflects the system's ability to cover use

6.2. Optimization algorithms and methods

The optimization algorithms and methods of personalized medical recommendation systems can be considered from different perspectives. Here are some commonly used optimization algorithms and methods:

1. Improvement of Collaborative filtering algorithm: Collaborative filtering is one of the commonly used algorithms in personalized recommendation systems. You can improve the recommendation effect by improving its similarity calculation, neighbor selection, recommendation result weighting, etc. For example, model-based Collaborative filtering algorithms (such as Matrix decomposition) can be used to better capture the potential relationship between users and projects [10].

2. Solution to cold start problem: Cold start problem refers to the difficulty in recommending new users or projects. To address this issue, content-based recommendation methods or hybrid recommendation methods can be used to utilize the content characteristics of the project for recommendation in the absence of user behavior data.

3. Optimization of Sorting algorithm: The recommendation system usually needs to sort the recommendation results in order to present the most relevant and valuable items to users. The Sorting algorithm in machine learning (such as LambdaMART, RankNet) can be used to optimize the sorting of recommendation results.

DOI: 10.54254/2755-2721/45/20241024

7. Challenges and Future Development

With the accumulation of medical data and the continuous progress of technology, personalized medical recommendation systems will be able to provide personalized recommendations to each user in a more refined manner. The personalized medical recommendation system can better meet the needs of users and provide more personal medical advice and treatment programs by deeply mining the health status, genetic information, lifestyle and other personal characteristics of users, combined with Big data and artificial intelligence technology. Medical data includes not only textual data, but also various forms of data such as images and sounds. The future development direction is to effectively integrate multimodal data, comprehensively utilize different types of data to better understand users' conditions and needs, and provide more comprehensive personalized recommendations. The personalized medical recommendation system will be combined with sensor technology and mobile devices to obtain real-time user body data, environmental data, and other information. Through real-time data analysis and AI technology, users' health status can be monitored in a timely manner, providing real-time personalized medical recommendations and warnings, effectively promoting health management and disease prevention. Future personalized medical recommendation systems may draw inspiration from different recommendation algorithms and implement hybrid applications. For example, multiple algorithms such as Collaborative filtering, content-based recommendation methods, deep learning and Reinforcement learning are combined to obtain more accurate and comprehensive personalized recommendation results.

References

[1] Research on personalized search engine recommendation algorithms Chen Hua; Li Renfa; Liu Yufeng; Lianqi. Computer Application Research, 2010

[2] Overview of evaluation methods for personalized recommendation systems Liu Jianguo; Zhou Tao; Guo Qiang; Wang Binghong. Complex Systems and Complexity Science, 2009

[3] Research on personalized customization of Portal Liu Ying; Cai Wanjing. Computer Knowledge and Technology, 2009

[4] Research progress in personalized recommendation systems Liu Jianguo; Zhou Tao; Wang Binghong. Progress in Natural Science, 2009

[5] A Review of Chinese Word Segmentation Algorithms Zhang Qiyu; Zhu Ling; Zhang Yaping. Intelligence Exploration, 2008

[6] Research on personalized services based on neural network models Xiong Ying. Modern Computer (Professional Edition), 2008

[7] The application of Collaborative filtering technology in personalized recommendation Song Zhenzhen; Wang Hao; Yang Jing. Journal of Hefei University of Technology (Natural Science Edition), 2008

[8] Research on personalized service user model Chen Yuan; Gou Guanglei, Computer engineering and Design, 2008

[9] Design and implementation of personalized service recommendation system for medical information resources Zhang Wei; Mo Meiqi; Xia Zhiping; Xu Yixin. Library Magazine, 2006

[10] Analysis and Research on User Interest Models in Intelligent Search Engines Jiang Ping, Cui Zhiming. Microelectronics and Computers, 2004