Journal of Industrial Information Technology and Application *ISSN(Online): 2586-0852*
Vol.7No.3

Online Disease Identification and Diagnosis and Treatment Based on Machine Learning Technology

Dr. P. Sumalatha[1)] and Dr. V.J. Chakravarthy[2,*)]
1)Dept. Of CS & IT, Central University of Andhra Pradesh, Anantapur , Andhra Pradesh, India
2)Arulmigu Kapaleeswarar Arts and Science College, Kolathur, Chennai, India

Abstract: The article uses machine learning algorithms to extract disease symptom keyword vectors. At the same time, we used deep learning technology to design a disease symptom classification model. We apply this model to an online disease consultation recommendation system. The system integrates machine learning algorithms and knowledge graph technology to help patients conduct online consultations. The system analyses the misclassification data of different departments through high- frequency word analysis. The study found that the accuracy rate of our machine learning algorithm model to identify entities in electronic medical records reached 96.29%. This type of model can effectively screen out the most important pathogenic features.

Keywords : Disease identification, ML disease treatment, Disease classification, Online consultation

1. Introduction

Background information for this study revealed that although the Internet medical service contains a substantial amount of structured information, such as diseases and hospitals, the service offered to users is relatively constrained. The majority of service websites solely provide options like doctor consultations and internet searches. Some websites provide straightforward self-diagnosis services. However, the majority restrict consumers to entering symptom terms. The system provides a lengthy list of ailments. So, using a related machine learning approach, this research suggests a technique for

* Corresponding author: chakkucksm1808@gmail.com
Received: May 10, 2023; Accepted: Jun 17, 2023; Published: Sep 30, 2023

extracting disease symptom keywords. By using this technology, information can be efficiently retrieved to help doctors and hospitals make informed recommendations, which will significantly increase the effectiveness of patient medical care. Although it is extremely important and has the potential to significantly improve the hospital's medical environment, one drawback is that not all patients are able to utilize the system properly, and learning how to use it efficiently requires extensive operation training. The approach used in this paper is that to assist patients with conducting online consultations, the system integrates knowledge graph and machine learning techniques. This strategy's strength lies in its ability to significantly raise the caliber of urban healthcare and services. High-frequency word analysis is used by the system to examine the misclassified data from several departments. The final stage of this method's development increases its overall accuracy to more than 90%, which has a very substantial impact [1]. This paper offers a design and implementation approach for developing an inquiry recommendation system based on knowledge graphs, deep learning, and social media in order to achieve this.

We create a "disease symptom" knowledge map by structuring the disease information on the website for those seeking medical guidance. We offer users services for disease self-diagnosis using a variety of informational sources. In order to improve the system's recommendation options, the system simultaneously mines the knowledge graph's structured data for probable diseases that the user may contract. We took a sample from the review information on the good doctor website. To analyses the service quality of doctors at various hospitals in Beijing, we merged it with the currently used indicators for evaluating service quality. We offer consumers recommendations for doctors and hospitals based on a variety of factors.

2. Related Work

A. Disease Diagnosis Algorithm

Two different sorts of methods based on algebraic operations and based on recommendation systems are used in traditional disease diagnosis. It is widely acknowledged that there are a lot of contradictory, unclear, and incomplete data in medical diagnosis. As a result, a lot of academics employ the FS (fuzzy set) and NS (neutron sophist) models to simulate the relationship between symptoms and illnesses. In order to diagnose the user's ailment, they compared the FS or NS of the patient's symptoms to those of various diseases. As machine learning becomes more and more popular, more researchers are attempting to use machine learning technology to diagnose diseases. A single-layer neural network has been trained using the back-propagation technique by some researchers to perform the task of illness diagnosis based on patient symptom keywords and laboratory data. In addition, some researchers employ patient-specific medical imagery to diagnose potential illnesses. To determine when a myocardial infarction occurs, some researchers employ a convolutional neural network (CNN).

For the classification of skin cancer, several academics also used CNN. It is necessary to model the conditions carefully in order to conduct a thorough examination of FS-based and NS-based disease diagnosis algorithms. The user finds it difficult to use this since it differs significantly from the natural language content they entered. Many users' prior diagnoses and treatments are needed for the recommender system method. Privacy of the user is involved. The two methods mentioned above also do not take into account how factors like age, gender, and others may affect how an illness is diagnosed.

B. Medical Knowledge Graph

A structured semantic information base is known as a knowledge graph. This approach uses symbols to represent ideas and how they relate to one another in the real world. The "entity- relation-entity" triplet serves as its fundamental unit. A networked knowledge structure can also consist of an entity and any associated attribute-value pairs that connect the entities through relationships. A portion of the knowledge graph is the medical knowledge graph. On its construction techniques and application scenarios, academia has conducted a significant amount of research. The creation of a specific medical knowledge graph is typically necessary for specific medical service demands. The technology for creating knowledge graphs from EHR data is currently extremely advanced. The extraction of medical concepts from the semi structured text is typically required by these techniques. Natural language processing tools like machine learning are used in the process. This approach is more difficult. We just need to create a "disease-symptom" knowledge map, which is in line with this paper's functional requirements for illness diagnosis. By utilizing a variety of current and trustworthy online resources for structured disease data, we can significantly reduce the complexity of creating a knowledge graph.

C. Analysis Technology Related to Medical Reviews

Sentiment analysis is the process of gathering and analyzing personal data, such as people's opinions, feelings, and emotions towards various products, services, entities, individuals, occasions, and traits. The two main types of sentiment analysis for medical reviews are sentiment polarity analysis and sentiment attribute extraction. Some academics base their learning models on the CSN's (cancer survivors' network) feedback. They employ binary classification to predict changes in the user's emotional state and help the community deliver better services by modelling the polarity of the sentiment in text. Some researchers coupled their understanding of the diabetes medical model with N-grams to categorise the sentiment polarity of tweets about diabetes. A latent Dirichlet allocation (LDA) model was created by some researchers using review data from the good doctor medical website between 2006 and 2014. Simple sentiment analysis was done after extracting topic models like "curative effect" and "seeking medical treatment." There are no sophisticated sentiment analysis methods for thorough study of medical reviews. None of them make use of the widely used deep learning model right now. The analysis of emotions themselves is where a lot of the work ends. They did not extract more information from the provided data based on the emotional data [2]. This paper will further mine emotional data together with markers for evaluating medical services. We create an evaluation model to measure the effectiveness of the doctor-hospital relationship and then use it to provide recommendation services.

D. Design and Implementation of Consultation Recommendation System

Figure 1 depicts the architecture design for the system for consultation and suggestion developed in this paper. Symptom data entered by the user is accepted by the recommender system. The content consists of the symptoms' descriptions, age, gender, and keywords (which are necessary). Three diseases and the associated recommended hospitals and physicians make up the final recommended outcomes. The two main services offered by the recommendation system are those that diagnose diseases and recommend doctors. Four steps overall make up the suggested course of action. They are creating and growing disease candidates, sorting and screening them, and making doctor-hospital recommendations. It will be discussed step-by-step how the consultation recommendation method was put into practise.

E. Building a Knowledge Graph

The information from the websites used to request medical care and order medications is used in this paper to build a knowledge map. A platform for online medical and health services, searching and seeking medicine is a website. This page makes use of details released on obtaining medical advice about 8802 disorders. This study's content mostly consists of information on the population most at risk for developing an illness, the incidence of that disease, the department that deals with that disease, its complications, and other data. A straightforward reprocessing is used in this work because the majority of the material on websites for seeking medical care and purchasing medications is structured. We primarily further retrieved data from the susceptible population, such as the gender and age of the susceptible individuals [3]. The age is derived from and somewhat modified from the new age segment data of 360 Encyclopaedia. Fetuses, newborns, children, teenagers, young adults, middle-aged people, and the elderly make up the seven age categories that make up the ageing stage. Tables 1 and 2 show the knowledge graph structure that was ultimately created in this study after taking into account the functional requirements of the work. Knowledge graphs are kept in this study's Neo4 graph database. 15,418 entities and 85,303 relationships will be present in the knowledge graph once construction is finished.

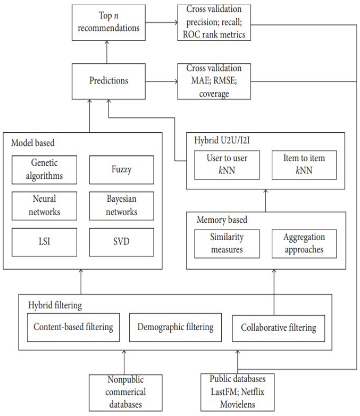

Figure 1. Framework diagram of consultation recommendation system.

F. Analysis of Medical Reviews

Reviews of good doctor websites were chosen as the data source for this study. In 2018 and 2019, this article used open data from the Good Doctor website. 133,667 online doctor reviews from renowned hospitals are included in the content. This study reprocesses the Good Doctor website's comment data to exclude invalid Chinese and English characters. The Python open- source third-party package zhconv is used to translate the comment text between traditional and simplified characters. In addition, we corrected a few frequent errors in the informal language of comment content. The linguistic mistake in the comment text in this article has been fixed using the Python third-party open-source module by corrector. This study begins by labelling the review data using the SERVQUAL model. Additionally, this work provides the fundamentals of comment annotation and refines the SERVQUAL model. The Tangibles evaluation dimension is disregarded because comments on the evaluation of hospital hardware facilities and the attire of medical staff are infrequently included in the data from the Good Doctor website. Additionally, this article merges responsiveness and empathy evaluation elements into one evaluation dimension because they are too comparable in the medical review data. R&E [4] is how it is noted. The review's description of efficacy matches reliability. The phrases "the condition has improved," "the condition has gotten worse," and similar ones are used frequently. The depiction of the doctor's medical attitude in the review is consistent with the R&E. The phrases "the doctor is very patient and kind," "the doctor is very impatient," etc. are common. Assurance reflects how the patient felt about the doctor's level of care overall during the evaluation. "Superior medical talents" and "average level" are frequent phrases. Each dimension is marked with three sentiment polarities: positive, neutral, and negative. There were 6019 marked comments in total.

Table 1. KNOWLEDGE GRAPH ENTITIES.

Entity	Attributes
Disease	Disease description, the incidence
Symptom keyword	None
Department	None
Age	None
Gender	None

Table 2. KNOWLEDGE GRAPH RELATIONSHIPS.

Relation	Head entity	Tail entity
Have symptoms	Disease	Symptom's keyword
Have complications	Disease	Disease
Department	Disease	Department
Susceptible age	Disease	Age
Predisposing gender	Disease	Gender

The BERT model is used in this study's patient review analysis. The annotation data set is split into training sets and validation sets in a 9:1 ratio. Tables 3-5 [5] show the final performance on the validation set. The BERT model has accuracy rates of 78.4%, 88.1%, and 93.4% in each of the three dimensions. The recall percentages are, respectively, 86.7%, 87.5%, and 97.2%. Since there are no ground-truth negative review data for the third dimension, the calculations for absolute precision and recall leave this out and are effective.

3. Disease Diagnosis Services

A. Construction of Disease Candidates

This section mostly compares the diseases in the knowledge graph using the user's keyword information. Based on disease incidence and keyword match, we rank potential diseases. In sorting, the degree of keyword matching is given higher priority. We evaluated the top 15 illnesses as potential candidates [6].

B. Expanding Disease Candidates

In order to increase the precision of disease diagnosis, this article opts to train the knowledge graph embedding with the Trans D model. Trans E, Trans H, and Trans R serve as the foundation for this concept. According to the Trans E model, h + r t represents the translation of the head entity vector h to the tail entity vector t. If the tail entity vector is represented by t, the relation vector is represented by r, and the head entity vector is represented by h. The Trans E model works well when dealing with one-to-one relationships since it presupposes that the same entity is represented under any relationship. For complex interactions like those in which there are many to many and one to many, this model is a little unreliable. The Trans H model contends that in order to properly represent the relation vector, the head entity vector and the tail entity vector must be projected on the relation vector's hyperplane, specifically:

$$h_{\perp} = h - \omega_r^T h \omega_r ; t_{\perp} = t - \omega_r^T t \omega_r . \tag{1}$$

Then, set up:

$$h_{\perp} + r \approx t_{\perp}, \tag{2}$$

Where ωr represents the normal vector of the hyperplane of relation vector r. The drawback of the Trans H model is that it still assumes that entities and relations are in the same semantic space. This limits its expressiveness. On this basis, the Trans R model further uses the projection matrix (M_r) to complete the projection operation.

$$h_{\perp} = h - M_r h; t_{\perp} = -M_r t. \tag{3}$$

For the same relation, the projection matrices provided for the Trans R model are identical. The distinctions between head and tail entities are not considered. The difficulty of the training is significantly increased by its projection operation, which requires matrix calculations. On the other hand, the Trans D model sees projections as the result of interactions between entities and relationships. Additionally, Trans D has a lower computational complexity than Trans R. It gives the projection matrix and h⊥ and t⊥ are calculated as follows:

$$\begin{gathered} M_r^1 = \omega_r \omega_h^T + I; M_r^2 = \omega_r \omega_t^T + I, \\ h_{\perp} = h - M_r^1 h; t_{\perp} = t - M_r^2 h. \end{gathered} \tag{4}$$

In order to provide the vector representation of relations and entities from the global graph structure information, this work implements the Trans D model using Open KE. According to how close the Euclidean distance is, we provide the disease that each disease in the candidate set has in common. It is added to the list of candidates [7]. The goal is to improve recommendation alternatives by mining probable ailments that users may have. The accuracy of disease diagnosis is increased as a result.

C. Filter Sort

According to the 4 dimensions given by user I this part rates and ranks the diseases in the candidate set. We pick the top three illnesses to suggest to users. This is how an illness is scored:

$$S_j = S_{age} + S_{sex} + S_{key} + S_{des}. \quad (5)$$

Sage, Ssex, Skey, and Sdes, among others, describe, in terms of the susceptible population's age, gender, symptom keywords, and illness description, respectively, the similarity between disease j and user I input. The comparisons of age and gender are made using straightforward string matching. The following is the formula for calculating Skey:

$$S_{key} = \frac{|\mathrm{Key}_i \cap \mathrm{Key}_j|}{|\mathrm{Key}_i|}. \quad (6)$$

Key_i, Key_j represent the symptom keyword set input by user i and the symptom keyword set owned by disease i, respectively [8]. The illness candidate set calculates Sdes. Text collection D is how it is categorised. We tokenize it and get rid of the stop words in Chinese. Assuming that the final vocabulary set is Formula, we then compute the TF-IDF values for each word in each text in the text set using the TF-IDF technique to produce the TF-IDF matrix: tfi dfij represents the TF-IDF value of word ti in a text Dj. Finally, the TF-IDF vector of the disease symptom description input by the user is also calculated by a similar method. Then, the row vector in the matrix M obtains the Sdes of each disease in the candidate set according to the cosine similarity calculation method.

$$M = \begin{bmatrix} tfi\,df_{11} & tfi\,df_{21} & \cdots & tfi\,df_{n1} \\ tfi\,df_{12} & tfi\,df_{22} & \cdots & tfi\,df_{n2} \\ \vdots & \vdots & \vdots & \\ tfi\,df_{1|D|} & tfi\,df_{2|D|} & \cdots & tfi\,df_{n|D|} \end{bmatrix}. \quad (7)$$

Table 3. RELIABILITY DIMENSION VERIFICATION RESULTS.

Emotional polarity	Positive (prediction)	Neutral (predicted)	Negative (prediction)
Positive (real)	76	12	3
Neutral (true)	15	344	59
Negative (true)	1	6	84

D. Doctor Referral Service

The BERT model was used in this study's sentiment polarity analysis of review data. In all doctor reviews, we counted the number of favourable, unfavourable, and neutral reviews for three dimensions. Based on the Wilson interval method, this study provides the ratings of physicians and even the associated departments. Here is Wilson's formula:

$$\text{Score} = \frac{p + \left(z_a^2/2n\right) - \left(z_a/2n\right)\sqrt{4n(1-p)p + z_a^2}}{1 + \left(z_a^2/2n\right)}. \tag{8}$$

n is the overall number of reviews, p is the positive rating, and za is the quantile. It is employed to convey the level of confidence in this score. Wilson's interval approach includes the following characteristics in addition to good discrimination:

(1) Score normalization.

(2) The slower the denominator's rate of decline is when p is constant, the slower the numerator's rate of decline is. The score now approaches p as n gets closer to infinity, and it gets higher as n gets bigger. In other words, when there are a lot of reviews, the rating system will consider the favourable rating to be trustworthy. The positive rating is also regarded as being unreliable when the overall number of reviews is minimal. With a modest modification to the Wilson interval approach, this paper chooses za to be 2 (i.e., the confidence level is around 95%). The Wilson interval method's underlying assumption is that there are only positive and negative comments, but in the real application scenario of this article, there are neutral remarks as well. In order to calculate the favourable rate, this paper treats half of the neutral ratings as positive. The remaining half are categorised as unfavourable reviews. We choose the appropriate department in accordance with the disease information after getting the doctor and hospital department scores. The system now suggests the top 4 hospitals. The top four doctors from each hospital are recommended.

4. Verification

A. Validation of the Disease Diagnosis Algorithm

From the four perspectives of age, gender, keywords, and symptom description, this article thoroughly describes the potential disorders. In addition, this paper uses the knowledge graph's structural data to provide probable diseases that the user could contract. This broadens the range of recommendations. Similar work is not currently being done in academia or industry. The method of creating a test set is therefore adopted in this paper in order to confirm the precision of the disease diagnosis algorithm [9]. This article chooses 50 prevalent illnesses from a variety of sections on the website for asking for and seeking medical advice, including otitis media, rhinitis, and colds. It was then reconstructed using natural language in combination with the data on sickness symptoms on the Baidu Baike website. It is entered as an illustrative user symptom description. We create user test scenarios using the keyword data from the website for seeking medical advice. This is test set T1 in question. The test set T2 is created by adding a specific amount of confusion-causing complication keyword data for each disease in the test set

T1. Information on complications can be found on the Seek for Medicine website. Table 6 displays the outcomes of applying the algorithm to the test set. The accuracy of the algorithm has drastically decreased after include the complexity information. On the one hand, it's due to the interference of extra information. However, due to their symptoms, some diseases do not have numerous keywords. The confounding effect is amplified by the inclusion of the complication term. Additionally, the descriptions of the same disease are not the same due to the varied sources of disease information on Baidu Baike and the website for seeking medical care and medication. For instance, the disease cheilitis is categorised in a number of ways by Baidu Encyclopaedia, including granulomatous cheilitis and actinic cheilitis. Only broad descriptions of cheilitis symptoms can be found on the website for obtaining medical help. This is another factor contributing to the accuracy rate's less-than-ideal performance. Although the algorithm sometimes fails to correctly identify the patient's illness, the suggested results often contain diseases that are comparable to the patient's illness or illnesses that belong to the same department [10].

B. Validation of Doctor Recommendation Algorithms

This section's verification is done by a questionnaire. The majority of the respondents are students at Peking University [11]. The poll chose three conditions that are generally common among students, including rhinitis, urticaria, and respiratory infections. The questions are the same for every illness. If the person has experienced the illness, inquire as to whether they received treatment in the Beijing region. Students are familiar with the doctors and hospitals that are suggested by the article's recommendation system. 57 questionnaires were ultimately found. Table 7 displays the results of the verification.

Table 4. R&E DIMENSION VALIDATION RESULTS.

Emotional polarity	Positive (prediction)	Neutral (predicted)	Negative (prediction)
Positive (real)	233	17	3
Neutral (true)	25	199	18
Negative (true)	3	3	179

Table 5. ASSURANCE DIMENSION VALIDATION RESULTS.

Emotional polarity	Positive (prediction)	Neutral (predicted)	Negative (prediction)
Positive (real)	148	3	0
Neutral (true)	16	421	0
Negative (true)	0	12	0

Table 6. VALIDATION ACCURACY % OF DISEASE DIAGNOSIS ALGORITHM.

Validation metrics	T1 test set	T2 test set	Average accuracy
Hit@3	80	68	74

Hit@5	88	74	81
Hit@10	94	80	87

Table 7. QUESTIONNAIRE SURVEY RESULTS.

Validation metrics	Respiratory infection	Hives	Rhinitis
Hospital recommendation error	7	1	3
Doctor recommends correct	54	61	65
Select the total number of doctors	61	68	69
Number of people who have had the disease	18	20	28
Number of people who have been treated for the disease in Beijing	11	15	18

There was only one instance of rhinitis among all the survey participants who had the condition and had been to the Beijing region. In the urticaria section, there were only 2 instances of hospital recommendation errors and 3 instances of doctor recommendation errors [12]. In the respiratory tract infection section, there was just one instance of a hospital recommendation error and one instance of a physician recommendation error. According to thorough questionnaire data, the consultation recommendation system for hospital referrals that is presented in this work is 93.57% accurate [13]. The accuracy of the system's recommendations for doctors is 90.91%. The system's application runs smoothly.

5. Conclusion

In this work, a recommendation system for medical consultations is designed and put into action. This study develops a "disease-symptom" knowledge map to offer users services for disease self- diagnosis. Additionally, in order to offer consumers better suggestion services, we apply the deep learning model to deliver the hospital doctor's service quality evaluation model. The method developed in this research enables users to enter different data, including gender. A realistic disease recommendation can be stated in full this way. This study makes use of knowledge graphs' structured data. In order to improve recommendation alternatives, the model can mine probable diseases that users may have. Additionally, this essay creatively connects patient feedback and the calibre of medical services. Users benefit from a more honest and reasonable recommendation service as a result.

References

[1] Alimadadi, Ahmad, et al. "Artificial intelligence and machine learning to fight COVID-19." Physiological genomics 52.4 (2020): 200-202.

[2] Goecks, Jeremy, et al. "How machine learning will transform biomedicine." Cell 181.1 (2020): 92-101.

[3] Tsikala Vafea, Maria, et al. "Emerging technologies for use in the study, diagnosis, and treatment of patients with COVID-19." Cellular and molecular bioengineering 13.4 (2020): 249-257.

[4] Ahn, Joseph C., et al. "Application of artificial intelligence for the diagnosis and treatment of liver diseases." Hepatology 73.6 (2021): 2546-2563.

[5] Yin, Zhijun, Lina M. Sulieman, and Bradley A. Malin. "A systematic literature review of machine learning in online personal health data." Journal of the American Medical Informatics Association 26.6 (2019): 561-576.

[6] Muhammad, L. J., et al. "Supervised machine learning models for prediction of COVID-19 infection using epidemiology dataset." SN computer science 2.1 (2021): 1-13.

[7] Shatte, Adrian BR, Delyse M. Hutchinson, and Samantha J. Teague. "Machine learning in mental health: a scoping review of methods and applications." Psychological medicine 49.9 (2019): 1426-1448.

[8] Bansal, Agam, et al. "Utility of artificial intelligence amidst the COVID 19 pandemic: a review." Journal of Medical Systems 44.9 (2020): 1-6.

[9] Khourdifi, Youness, Mohamed Bahaj, and M. Bahaj. "Heart disease prediction and classification using machine learning algorithms optimized by particle swarm optimization and ant colony optimization." International Journal of Intelligent Engineering and Systems 12.1 (2019): 242-252.

[10] Jamshidi, Afshin, Jean-Pierre Pelletier, and Johanne Martel-Pelletier. "Machine-learning-based patient-specific prediction models for knee osteoarthritis." Nature Reviews Rheumatology 15.1 (2019): 49-60.

[11] Balyen, Lokman, and Tunde Peto. "Promising artificial intelligence-machine learning-deep learning algorithms in ophthalmology." The Asia-Pacific Journal of Ophthalmology 8.3 (2019): 264-272.

[12] Samie, Farzad, Lars Bauer, and Jörg Henkel. "From cloud down to things: An overview of machine learning in internet of things." IEEE Internet of Things Journal 6.3 (2019): 4921-4934.

[13] Mincholé, Ana, and Blanca Rodriguez. "Artificial intelligence for the electrocardiogram." Nature medicine 25.1 (2019): 22-23.