Advances in Computer, Signals and Systems (2024)
Clausius Scientific Press, Canada
DOI: 10.23977/acss.2024.080515
ISSN 2371-8838 Vol. 8 Num. 5

An Expert System for Health Diagnosis Based on Natural Language Processing and Reasoning Engine

Yawen Deng[1,*]

[1]Aberdeen Institute of Data Science and Artificial Intelligence, South China Normal University, Nanhai, Foshan, China

[*]Corresponding author

Keywords: Expert System, Natural Language Processing, Inference Engine

Abstract: The paper discusses expert systems in the field of digital health, which provide services for health diagnosis. An expert system is an artificial intelligence technology that simulates a doctor's diagnostic process and provides diagnostic and treatment recommendations. The paper describes the techniques and methods related to expert systems, mainly including natural language processing and inference engines. Natural language processing is used to process and understand the natural language input provided by the patient, extract key information, and convert it into a machine-understandable form. The inference engine uses the medical knowledge base and rules to perform logical reasoning and inference to generate diagnostic results and treatment recommendations. The paper also describes related experiments and evaluations, as well as ethical issues and challenges. The aim is to explore the application and development of AI in expert systems for health diagnosis to inform and inspire the field of digital health.

1. Introduction

With the increase of various diseases and the aggravation of population aging, medical diagnosis is facing increasingly prominent challenges, such as uneven distribution of medical resources, overburden of hospital work, misdiagnosis and missed diagnosis, which seriously affect the quality and efficiency of medical services.

Expert systems provide new solutions to overcome these challenges. In the field of artificial intelligence, expert system is one of the most widely developed and fruitful branches, which is a procedural system that applies artificial intelligence technology, integrates a large amount of professional knowledge and experience, and simulates the decision-making process of human experts through reasoning and judgement, in order to solve those complex problems that require the participation of experts [1]. In the medical field, expert systems are mainly applied to assisted diagnosis and personalised treatment. Classified according to the different rules of reasoning, they are mainly divided into rule-based, case-based and artificial neural network-based expert systems. Classified according to functions, expert systems can be mainly classified into interpretation, prediction, diagnosis, design, planning, monitoring, control and debugging expert systems. Usually the expert system model consists of six modules: human-computer interface, knowledge acquisition procedure, knowledge base, interpreter, reasoning machine and comprehensive database [2], and the relationship of each module is shown in Figure 1.

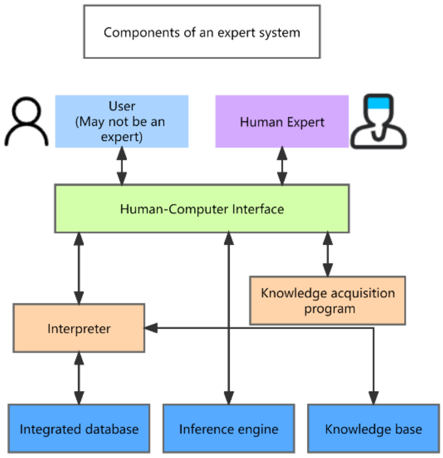

Figure 1: Components of an Expert System.

Natural language processing focuses on the interaction between computers and human language. Its goal is to enable computers to understand and generate natural language, from which key information can be extracted, entities can be identified, and sentence structure can be analysed [3]. In health diagnosis, NLP can be applied to information extraction from medical documents, automatic classification of patients' medical records and automatic generation of diagnostic reports, etc. This technology holds great potential for providing personalized diagnostic and treatment recommendations and enhancing patient interactions with the healthcare system. However, challenges in semantic understanding and data quality must be addressed to improve the accuracy and reliability of NLP in health diagnosis.

An inference engine is a logical reasoning system that makes logical inferences and deductions based on known rules and facts. In an expert system for health diagnosis, an inference engine can use existing medical knowledge, case banks, symptom data, etc., combined with the patient's symptom information, to derive the patient's diagnosis through inference and deduction, and help the system generate accurate diagnostic results and treatment recommendations.

2. Related work

Since the 1970s, Artificial Intelligence methods have been widely used in the medical field with the aim of improving the efficiency of disease diagnosis and treatment, thus driving the development of Artificial Intelligence in Medicine (AIM) [4]. With the proposal of various machine learning algorithms such as decision trees, random forests, and support vector machines after the 1980s, AIM has gradually become more mature. Figure 2 illustrates the common techniques of AIM.

Machine Learning (ML) is an algorithmic model whose concept was introduced by SAMUEL in 1959, implying that computers can learn from data without explicit programming [5]. Decision Tree (DT) algorithm was proposed by QUINLAN in 1986, which can classify data based on rules [6]. Support Vector Machine (SVM) is a common supervised ML algorithm proposed by VLADIMIR in 1995 and is widely used in classification and regression problems [7]. Random Forest (RF) is an algorithm proposed by HO in 1998 for effective feature extraction [8]. In the field of health diagnosis, machine learning can be used to learn and make predictions from patients' medical records, symptoms and medical data by training models. However, the "black box" problem of machine learning techniques makes it difficult for clinicians to understand the underlying mechanisms and provide recommendations based on specific situations. Therefore, advancing interpretable AI may be the key to solving this problem [9].

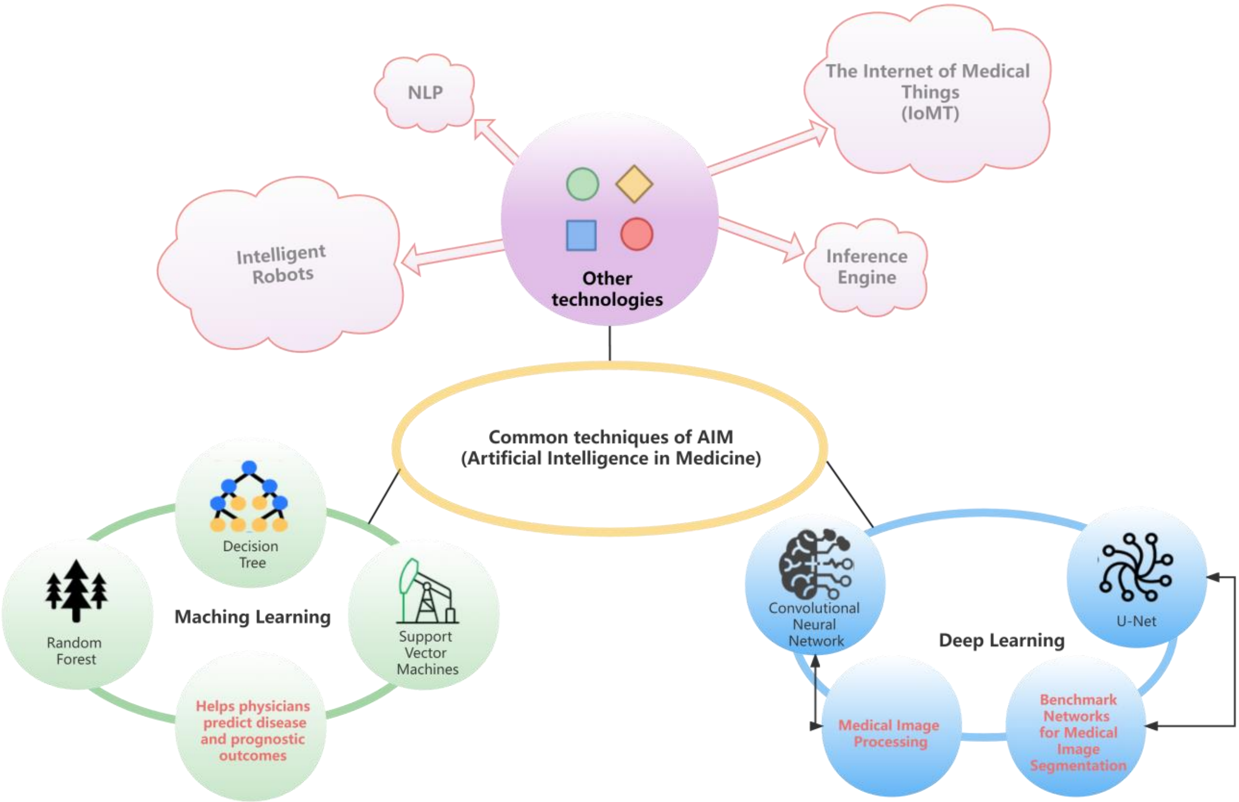

Figure 2: Common techniques of AIM.

Recent advancements in NLP have significantly enhanced its application in clinical healthcare. Large language models (LLMs) such as GPT-4 have shown promising capabilities in improving patient care, diagnostic assistance, and medical record analysis through advanced language understanding and generation [10]. Additionally, systematic reviews of NLP applications highlight the successful extraction of clinical information from electronic health records (EHRs) and the challenges associated with adverse event detection and patient-trial matching [11]. Furthermore, established NLP tools like MedLEE and cTAKES have demonstrated their utility in converting unstructured clinical narratives into structured data, thereby facilitating clinical research and improving the overall efficacy of healthcare practices [12].

Recent advancements in the field of inference engines for expert systems have focused on enhancing diagnostic reasoning capabilities. Reggia et al. (2023) introduced a novel frame-based inference model that addresses the complexity of managing multiple simultaneous disorders in medical diagnostics. This model improves on traditional set covering techniques by incorporating partial matching, making it more adaptable to real-world scenarios [13]. Benmimoune et al. (2015) presents a hybrid reasoning-based healthcare platform that utilizes clinical rules and experiences to assist clinicians in their decision-making process by combining rule-based reasoning and case-based reasoning [14].

3. Proposed Solution

Currently, the research on clinical diagnosis of diseases using AI technology mainly focuses on two aspects: one is to analyse and process massive medical data, and to quickly extract key information from a large amount of data by reasoning, analysing, comparing, summarising, and arguing to draw cognitive conclusions about the patient's physical condition and diseases [15][16]; the other is to analyse and understand the diagnostic data in the form of multimedia, such as texts, audios, images, videos, and so on, and to excavate and differentiate the characteristics of diseases to make diagnosis and evaluations [17].

Therefore, based on the review of the methodology in the previous section, this section proposes a solution for an expert system for health diagnosis by combining two contrasting techniques, namely natural language processing and inference engines, with the main flow shown in Figure 3.

As can be seen from Figure 3, NLP techniques are used to process and understand the natural language input provided by the patient, extract the key information and transform it into a machine-understandable form. The transformed information is then used as input to the inference engine. The inference engine uses the medical knowledge base and rules to perform logical reasoning and inference to generate diagnosis and treatment recommendations. The combination of NLP and inference engine enables the system to better understand patient inputs and provide accurate diagnosis and personalized recommendations.

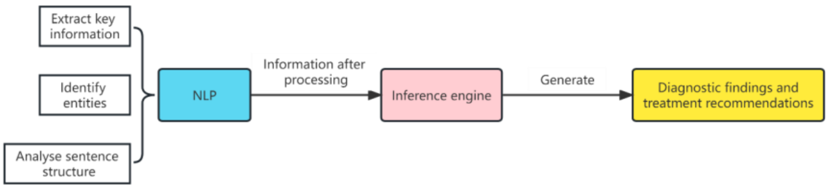

Figure 3: Flow chart of the proposed solution.

3.1. Natural Language Processing: Information Extraction and Understanding

In a natural language processing-based health diagnosis expert system, the information extraction process includes the following steps: First, the input text undergoes preprocessing (such as tokenization and part-of-speech tagging). Then, named entity recognition (NER) is used to extract key entities like diseases and symptoms. The system further extracts information related to symptoms and medical history, identifying relevant temporal details. Sentiment analysis is employed to assess the patient's emotional state, and relation extraction identifies relationships between entities. Finally, the extracted information is standardized into a medical knowledge base format, and contextual analysis is performed to enhance accuracy. The sample pseudo-code for medical document information extraction is as follows:

ALGORITHM 1: Example of medical document information extraction

```
from nltk import ne_chunk, pos_tag, word_tokenize
from nltk.tree import Tree
def extract_diagnosis(text):
    # Lexical annotation using natural language processing tools
    tagged_words = pos_tag(word_tokenize(text))
    # Disease extraction using named entity recognition
    named_entities = ne_chunk(tagged_words)
    diagnoses = []
    for entity in named_entities:
```

```
        if isinstance(entity, Tree) and entity.label() == 'DISEASE':
            diagnosis = ' '.join([child[0] for child in entity])
            diagnoses.append(diagnosis)
    return diagnoses
text = "The patient presented with fever and cough, which are consistent with a diagnosis of pneumonia."
diagnoses = extract_diagnosis(text)
print(diagnoses)  # ['pneumonia']
```

3.2. Inference Engine: Diagnosis and Treatment Recommendation

An inference engine is a logical reasoning system that uses known rules and facts to make inferences and deductions to produce new conclusions. Rules are generally represented by logical entailments between conditions and outcomes, while facts are representations of observed data or states. The inference engine is able to perform forward or backward reasoning depending on how well the rules and facts match. Forward reasoning starts from the facts and looks for applicable rules that lead to new conclusions. Reverse reasoning starts from the goal and looks for rules and facts that prove the goal in order to obtain evidence that supports the goal. The advantage of this method of reasoning is that it is highly interpretable and can simulate human thought processes. However, its disadvantage is that it requires manual definition or learning of rules, which may be incomplete or inconsistent.

The inference engine expresses this as a set of rules or logical expressions by utilising the knowledge and experience of domain experts. These rules describe the relationships and conditions between diseases, symptoms, medications, etc. When inputting the patient's symptom information after the previous step of natural language processing, the inference engine will reason based on these rules, applying the rules step-by-step, filtering and matching the symptoms, and ultimately arriving at a possible diagnosis.

4. Methodology

In order to evaluate the performance of the technique, this section uses case studies and python libraries as exploratory evaluation methods for related studies, respectively.

4.1. Natural Language Processing: CBLUE Research Summary

In April 2021, the Chinese Medical Information Processing Challenge List CBLUE [18], initiated by the Healthcare and Bioinformatics Processing Professional Committee of the Chinese Language Information Society of China, was launched on line, which contains eight classical medical natural language processing tasks, such as medical entity recognition, medical relationship extraction, medical terminology standardisation, medical text classification and medical Q&A, as shown in Table 1. These data are from real scenarios, such as medical textbooks, search engines and online Q&A, etc. The data are real and noisy, which puts higher requirements on the robustness of the model.

The study used pre-trained language models, a technique that learns generic language representations using large-scale unlabelled textual data to capture syntax, semantics and knowledge and provide effective features for downstream tasks. The study evaluated 11 Chinese pre-trained language models, analysed their results by testing them on CBLUE, and found that on some of the more challenging tasks, these models performed well below human levels. For the convenience of users, the study provides an online platform where users can submit their model predictions and have them evaluated, compared and analysed [18]. However, the method requires significant computational resources and time for pre-training and fine-tuning, which may have cost and efficiency issues. In addition, the pre-training data may be biased or noisy, thus affecting the quality and reliability of the mode.

Table 1: Subtasks of CBLUE.

Dataset	Task	Train	Dev	Test	Evaluation Metrics
CMeEE	NER	15,000	5,000	3,000	Micro F1
CMeIE	Relation Extraction	14,339	3,585	4,482	Micro F1
CHIP-CDN	Diagnosis Normalization	6,000	2,000	10,192	Micro F1
CHIP-STS	Sentence Similarity	16,000	4,000	10,000	Macro F1
CHIP-CTC	Sentence Classification	22,962	7,682	10,000	Macro F1
KUAKE-QIC	Sentence Classification	6,931	1,955	1,944	Accuracy
KUAKE-QTR	NLI	24,174	2,913	5,465	Accuracy
KUAKE-QQR	NLI	15,000	1,600	1,596	Accuracy
Short Title	Short title of article	History	Dates of article	/	/

4.2. Inference engine: python library experiments

In order to verify the feasibility of this method to diagnose health, this study used the experta library for python to build a rule-based inference engine that can infer possible diseases based on input symptoms and predefined rules. Its inference method is forward reasoning, i.e., it starts from known facts and looks for rules that can be applied to draw new conclusions. For example, if symptoms such as fever, cough, headache, etc. are entered, it will determine that it may have a cold or flu based on the rules and give advice accordingly. This code is an expert system shell which can modify or extend the rules according to different domains and knowledge bases to achieve different functions. Based on the operational results, it is clear that the inference engine has some advantages in the healthcare domain, but it also needs to address the challenges of knowledge updating, subjective judgement and data quality.

For example, a patient inputs: “I've been experiencing chest pain and shortness of breath for the past two days. I'm 50 years old and a long-time smoker.” The expert system first preprocesses this text, tokenizing it and tagging parts of speech. The system then uses named entity recognition to identify “chest pain” and “shortness of breath” as symptoms and “past two days” as the duration. Next, it extracts “50 years old” and “long-time smoker” as basic patient information and medical history. Sentiment analysis detects that the patient expresses concern about their health. Finally, the system standardizes this information and uses a reasoning engine to reference a medical knowledge base, considering the patient’s age and smoking history. It generates a potential diagnosis, such as "acute coronary syndrome" and recommends immediate medical attention.

5. Ethical Issues

The application of artificial intelligence in health diagnostic expert systems may involve ethical issues such as algorithmic unfairness, privacy and security, and data ethics. The World Health Organisation published a global report on AI in health with six guiding principles. Therefore, responsible AI needs to be considered when developing such applications, including patient privacy protection and dealing with uncertainty, erroneous results to ensure that it is beneficial to the patient.

6. Critical Review

The subject of this paper is the application of expert systems for health diagnosis in digital health. The paper describes relevant techniques and methods for expert systems, mainly including natural language processing and inference engines, and presents relevant experiments and evaluations, as well as ethical issues and challenges. The strengths of this paper are that it selects a research topic with practical significance and cutting-edge, introduces and analyses the related technologies and methods in detail, proposes an innovative and feasible solution, and conducts related experiments and evaluations to verify its validity and reliability. The drawback of this paper is that it does not fully consider the aspects of user needs and experience, data quality and security, model evaluation and optimisation.

7. Conclusion and Future Works

This paper explores the application and development of Artificial Intelligence in expert systems for health diagnosis to provide references and insights in the field of digital health, and proposes a solution that combines natural language processing and an inference engine. Future work can be carried out in the following aspects: first, designing a friendly, easy-to-use, and trustworthy human-computer interaction interface; second, ensuring the authenticity, integrity, and consistency of the data sources and preventing data leakage, tampering, and misuse; and third, selecting appropriate assessment indicators, methods, and standards, and adjusting, improving, and updating the model based on the assessment results.

References

[1] Jabbar H K, Khan R Z. Survey on development of expert system in the areas of Medical, Education, Automobile and Agriculture[C]//2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom). IEEE, 2015: 776-780.

[2] Ismaila L E, Ismail I M, Agwu N N. Relevance of User Data Completeness in Resolving Differential Diagnosis in Medical Expert System Optimization[C]//2018 14th International Conference on Electronics Computer and Computation (ICECCO). IEEE, 2018: 1-6.

[3] Thakur G K, Thakur A, Khan N, et al. The Role of Natural Language Processing in Medical Data Analysis and Healthcare Automation[C]//2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS). IEEE, 2024, 1: 1-5.

[4] Stefanelli M. The socio-organizational age of artificial intelligence in medicine[J]. Artificial intelligence in medicine, 2001, 23(1): 25-47.

[5] Samuel A L. Some studies in machine learning using the game of checkers[J]. IBM Journal of research and development, 1959, 3(3): 210-229.

[6] Quinlan J R. An empirical comparison of genetic and decision-tree classifiers[M]//Machine Learning Proceedings 1988. Morgan Kaufmann, 1988: 135-141.

[7] Huang S, Cai N, Pacheco P P, et al. Applications of support vector machine (SVM) learning in cancer genomics[J]. Cancer genomics & proteomics, 2018, 15(1): 41-51.

[8] Ho T K. The random subspace method for constructing decision forests[J]. IEEE transactions on pattern analysis and machine intelligence, 1998, 20(8): 832-844.
[9] Wu Y Z, Chen X C, Yi D. Advances and perspective of artificial intelligence in clinical area[J]. J. Army Med. Univ, 2022, 44: 89-102.
[10] Nassiri K, Akhloufi M A. Recent Advances in Large Language Models for Healthcare[J]. BioMedInformatics, 2024, 4(2): 1097-1143.
[11] Garg R, Gupta A. A Systematic Review of NLP Applications in Clinical Healthcare: Advancement and Challenges[C]//International Conference on Advances in Data-driven Computing and Intelligent Systems. Singapore: Springer Nature Singapore, 2023: 31-44.
[12] Wang Y, Tafti A, Sohn S, et al. Applications of natural language processing in clinical research and practice[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Tutorials. 2019: 22-25.
[13] Reggia J A, Nau D S, Wang P Y. A new inference method for frame-based expert systems[C]//Proceedings of the Third AAAI Conference on Artificial Intelligence. 1983: 333-337.
[14] Benmimoune L, Hajjam A, Ghodous P, et al. Hybrid reasoning-based medical platform to assist clinicians in their clinical reasoning process[C]//2015 6th International Conference on Information, Intelligence, Systems and Applications (IISA). IEEE, 2015: 1-5.
[15] Chen Y, Argentinis J D E, Weber G. IBM Watson: how cognitive computing can be applied to big data challenges in life sciences research [J]. Clinical therapeutics, 2016, 38(4): 688-701.
[16] Dilsizian S E, Siegel E L. Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment[J]. Current cardiology reports, 2014, 16: 1-8.
[17] Stoitsis J, Valavanis I, Mougiakakou S G, et al. Computer aided diagnosis based on medical image processing and artificial intelligence methods[J]. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2006, 569(2): 591-595.
[18] Zhang N, Chen M, Bi Z, et al. Cblue: A chinese biomedical language understanding evaluation benchmark[J]. arXiv preprint arXiv:2106.08087, 2021.