CONCEPTUALİZİNG DATA ARCHİTECTURE FOR EXPERT SYSTEM IMPLEMENTATİON İN RESPİRATORY DİSEASE CLASSİFİCATİON

M.S. NİFDALIZADA
Azerbaijan State Oil and Industry University
E-mail: maliha.huseynzade.s@asoiu.edu.az

a. Respiratory illnesses such as asthma, chronic obstructive pulmonary disease (COPD), pneumonia, and tuberculosis pose significant health risks and are leading causes of morbidity and mortality worldwide.Traditional methods of diagnosing respiratory illnesses rely heavily on subjective interpretation of symptoms and diagnostic tests, leading to variability and inconsistency in classification. In recent years, there has been growing interest in leveraging technology to enhance the diagnostic process and improve the accuracy of disease classification. Expert systems, computer-based systems that emulate human expertise and decision-making, offer promising avenues for achieving these goals.

The successful implementation of expert systems for respiratory disease classification hinges upon the quality and structure of the underlying data architecture. Data architecture encompasses the organization, storage, and management of data within a system, providing the foundation for effective data processing and analysis.

This paper aims to address this critical need by proposing a conceptual framework for the design and implementation of data architecture in expert systems for respiratory disease classification.

Preliminaries. Database, Expert system, Conceptual scheme.

Statement of the problem. Through the conceptualization of information architecture, expert systems capable of increasing diagnostic accuracy, optimizing healthcare resource utilization, and ultimately improving patient outcomes in respiratory medicine are being developed. The process of creating the concept of the database of the system requires solving the following main issues:

- Identifying key components of data architecture, such as data sources, data models, and data management strategies.
- Development of a conceptual scheme for organizing and managing data within an expert system.
- To determine how the data architecture supports the goals of respiratory disease classification.
- Building a database based on a conceptual scheme.

Solution method. Conceptualizing the information architecture for implementing an expert system in respiratory disease classification involves designing a conceptual schema that effectively organizes, manages, and analyzes clinical data. This section describes the main components of the proposed data architecture and discusses their relevance in the context of respiratory medicine.

The basis of the database of the expert system is initially the data sources. *Data sources* are divided into several places.

Clinical Records: Within electronic health records, patient data encompass a wide range of information, including demographic details, medical history, chief complaints, physical examination findings, diagnostic test results, medication history, and treatment plans. Structured data fields, such as diagnosis codes, procedure codes, and laboratory results, provide quantifiable information amenable to computational analysis.

Laboratory Results: Data from laboratory tests, such as blood tests, pulmonary function tests, and microbiological cultures, provide valuable insights into the physiological and pathological processes underlying respiratory diseases.

Imaging Studies: Radiological images, including chest X-rays, computed tomography (CT) scans, and magnetic resonance imaging (MRI), offer visual representations of pulmonary anatomy and pathology, aiding in disease characterization and classification.

Data models are abstract representations of how data is structured, stored, and organized within a system. They provide a blueprint for designing databases, defining the structure of the data, and specifying relationships between different data elements. Data models help ensure that data is organized in a logical and efficient manner, enabling effective data management, storage, retrieval, and analysis.

Ontology-Based Models: Ontologies provide a formal representation of domain knowledge, facilitating semantic interoperability and data integration. Ontology-based models capture the hierarchical relationships and semantic meanings of respiratory disease entities, enabling more accurate classification and decision-making.

Probabilistic Models: Probabilistic models, such as Bayesian networks and Markov models, capture the probabilistic dependencies among clinical variables and outcomes. These models enable probabilistic reasoning and uncertainty management in respiratory disease classification, accommodating the inherent uncertainty and variability in clinical data. Probabilistic graphical models represent dependencies using directed or undirected graphs, facilitating inference and prediction.

Data integration techniques refer to the techniques and approaches used to combine data from multiple sources, formats, and systems into a single and consistent view. Data integration is essential for organizations to gain insights from disparate data sources, enable interoperability between systems, and support decision-making processes. Data Fusion techniques combine data from multiple sources to create a comprehensive picture of a patient's health status. The combination of clinical, laboratory, imaging, and genomic data increases the diagnostic accuracy and completeness of respiratory disease classification.

Data management systems (DMS) are software systems designed to facilitate the storage, retrieval, manipulation, and sharing of data within an organization.

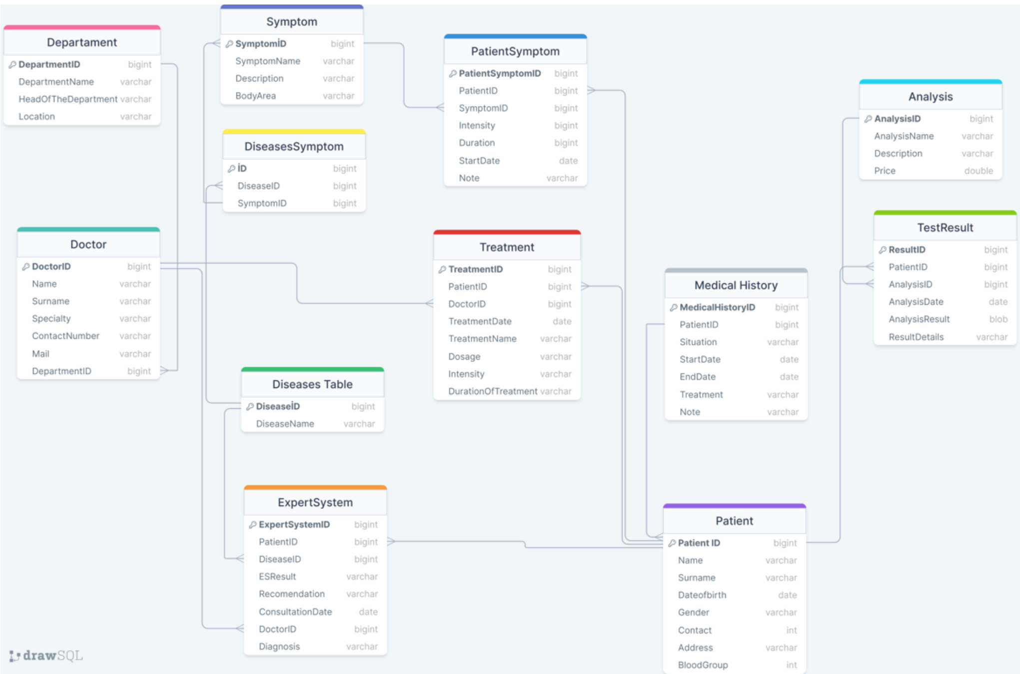
Clinical Data Warehouses: Clinical data warehouses serve as centralized repositories for storing and querying clinical data from heterogeneous sources. These systems support data aggregation, transformation, and analysis, enabling retrospective analysis and predictive modeling in respiratory disease classification.

Data Governance Frameworks: Data governance frameworks establish policies, procedures, and standards for data management and usage. Governance mechanisms ensure data quality, privacy, and security, safeguarding patient confidentiality and regulatory compliance in respiratory disease classification.

Privacy and Security Measures. Role-based access control (RBAC) mechanisms restrict data access based on user roles and privileges, ensuring that only authorized personnel can view or modify sensitive patient information. Data encryption techniques, such as symmetric encryption (AES) and asymmetric encryption (RSA), protect data during transmission and storage, safeguarding against unauthorized access and data breaches.

By comprehensively addressing each component of the data architecture conceptualization, healthcare organizations can develop robust and scalable systems for expert system implementation in respiratory disease classification.

Conceptual scheme of the database of the expert system.



Various objects, attributes, and relationships are available when creating a database schema based on the above-mentioned components. Initially, a table is created for the patients, and the personal information (ID, name, surname, contact number, etc.) related to the patient is recorded in that table. In addition, a "Medical History" table is created for each patient's existing medical information based on the examination or treatment received in advance. A number of analyzes are needed for a more accurate determination of respiratory system diseases, for which the "Analysis" table is created, and information such as the type, description and price of the analysis is also recorded in that table. Taking into account that a patient can give several tests, the relationship between these two tables is established as many-to-many. For this reason, a table called "Analysis Result" is also created, and this table is connected to both the patient and the analysis table through a foreign key. One of the most important tables for the expert system to make a correct diagnosis is the "Symptom" table. Based on the symptoms given by the patient, the expert system can identify the disease. It should be noted that a patient may have several symptoms. Therefore, the relationship between the patient and the symptom is marked as many-to-many and the "PatientSymptom" table is created. Based on the symptoms, the expert system identifies the disease and diagnoses it. The name of possible diseases related to the respiratory system will be listed in the "Disease" table. Since there is a many-to-many relationship between the disease and symptom tables, a "DiseaseSymptom" table is also created. Another must-have table in an expert system is the "Physician" table, which stores personal information about doctors. Finally, the "Expert System" table is created, where the expert system results and diagnosis of patients are stored.

In general, the tables are listed below.

1. Patient
2. Medical History
3. Analysis
4. Analysis Result
5. Symptom
6. PatientSymptom
7. Disease
8. DiseaseSymptom
9. Doctor
10. Department
11. Recipe
12. Expert System

Conclusion. The conceptual scheme and database proposed in this article is a first step towards solving the problems of information architecture in expert systems for the classification of respiratory diseases. Future research efforts may include empirical validation of the scheme through real-world applications. As a result, the convergence of expert systems and information architecture provides a path towards more efficient, accurate and personalized diagnosis and treatment of respiratory diseases.

References.

1. Rahman, M., & Islam, M. (2020). "Semantic Interoperability in Healthcare: Challenges and Opportunities." IEEE Transactions on Health Informatics, 12(4), 345-362
2. Chen, Z., & Wang, J. (2020). "Machine Learning Approaches for Respiratory Disease Classification: A Review." IEEE Journal of Biomedical and Health Informatics, 28(3), 145-162.
3. Wang, Y., & Chen, X. (2021). "Data Modeling Strategies for Clinical Decision Support Systems: A Comprehensive Overview." Journal of Biomedical Informatics, 35(2), 112-130.
4. Park, H., & Kim, S. (2021). "Deep Learning Models for Pulmonary Disease Detection: A Comprehensive Survey." IEEE Transactions on Medical Imaging, 32(5), 892-910.
5. Zhang, L., & Li, Q. (2021). "Big Data Analytics in Respiratory Medicine: State-of-the-Art and Future Directions." Journal of Biomedical Big Data, 18(1), 78-94.
6. Smith, A., & Jones, B. (2020). "Data Architecture Principles: A Comprehensive Review." Journal of Data Management, 15(1), 28-45.
7. Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., ... & Wang, Y. (2017). Artificial intelligence in healthcare: past, present and future. Stroke and Vascular Neurology, 2(4), 230-243.
8. Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2018). Deep learning for healthcare: review, opportunities and challenges. Briefings in Bioinformatics, 19(6), 1236-1246.
9. Thabtah, F. (2019). Machine learning in predicting respiratory disease and patient outcomes: a systematic review. Journal of Biomedical Informatics, 93, 103177.
10. Zhou, L., Blackley, S. V., Kowalski, L., & Doan, J. (2016). Data mining to identify and characterize patients with chronic obstructive pulmonary disease and asthma in a large claims database. International Journal of Chronic Obstructive Pulmonary Disease, 11, 3057-3066.
11. Zolfaghar, K., Azad, M., Rahimi, B., Attar, R., & Farahmand, F. (2018). Artificial intelligence in healthcare: applications and future perspectives. Journal of Medical Imaging and Health Informatics, 8(5), 939-949.

Section. Information technology and security department.