

literature Review prediction of chronic diseases1

by Mohammed Anbar

Submission date: 21-Nov-2024 09:50AM (UTC+0800)

Submission ID: 2376672743

File name: literature_Review_prediction_of_chronic_diseases.docx (30.86K)

Word count: 3045

Character count: 17964

literature Review prediction of chronic diseases

Table of contents:

- Summary research paper's
- Critical Analysis
- Connections
- Conflicts between the findings
- Synthesis Matrix for Literature Review
- References

Summary research paper's:

Paper1: Identification and Prediction of Chronic Diseases Using Machine Learning Approach.

In light of environmental challenges and contemporary ways of living, humans face many chronic diseases, and early detection and prediction of these diseases is crucial to reduce their spread. Doctors have difficulties in accurately identifying diseases, which reinforces the need to use sophisticated technologies such as machine learning. This paper aims to develop predictive models for identifying patients with common chronic diseases. Techniques such as convolutional neural networks (CNN) are used to extract features and predict diseases, as well as the "nearest neighbour" algorithm (KNN) to calculate the distance and find matches in the dataset. A data set has been prepared containing a set of symptoms associated with the disease, as well as living habits and details about doctor consultations. A comparative study was conducted between the proposed system and a number of other algorithms, such as Naïve Bayes, decision trees, and logistic regression, to prove the effectiveness of the model. (Alanazi, Rayan, 2022)

Paper2: Feature selection and classification systems for chronic disease prediction: A review.

This paper discusses the importance of predicting chronic diseases in healthcare informatics, emphasising the need for early diagnosis. The study provides a survey of the use of feature selection and classification techniques for the diagnosis of diseases, where proper feature selection contributes to improving the accuracy of systems. Classification algorithms have been applied to chronic disease datasets, yielding promising results in the development of intelligent and adaptable diagnostic systems. It also reviews the impact of parallel classification systems in speeding up processes and enhancing efficiency. The paper presents a comprehensive overview of feature selection methods, explaining the pros and cons of each of them, and analyses adaptive taxonomic systems in the context of disease forecasting. The results highlight the importance of integrating machine learning technologies to improve healthcare outcomes. (Divya Jain, Vijendra Singh, 2018)

Paper3: A method for the early prediction of chronic diseases based on short sequential medical data.

Noncommunicable diseases are the leading cause of death globally, and controlling these diseases through increased patient awareness is crucial but challenging. With the increase in Electronic Health Record data, it has become possible to use machine learning techniques to predict these diseases early. However, these efforts face difficulties due to the irregularity of the samples, the shortness of the data sequences, and their imbalance, which hinders the effective use of data. This paper outlines the characteristics of short serial data necessary for early prediction of non-communicable diseases and highlights the importance of their use in machine learning technologies. The study proposes a new method known as the "short sequence early

prediction method based on medical data (SEB)¹", which includes two stacked subnets to improve multi-marker prediction. The network uses long-term short-term memory (LSTM) and attention layers to extract temporal and non-temporal features. In addition, the process of increasing data through previous clinical knowledge is applied to bypass data limitations. Experiments conducted using real medical data show that the system outperforms the latest algorithms in predicting Noncommunicable Diseases, shows effectiveness in processing short serial data. The results suggest that using short sequential data directly may be more efficient than formatting traditional datasets with time exclusion constraints. (Chengkai Wu, Tianshu Zhou, Yu Tian, Junya Wu, Jingsong Li, Zhong Liu, 2022)

Paper4: Chronic Disease Prediction Using Deep Learning.

Data are rapidly increasing in the fields of biosciences and healthcare, providing opportunities for improving the early identification of infections and social insurance of patients. Forecasting is an important aspect of health care. This paper aims to develop deep learning models and algorithms for predicting chronic diseases in patients. The study uses a standard dataset to experiment with a modified prediction model. The method is based on machine learning techniques, such as the "nearest neighbour" algorithm, a decision tree, as well as deep learning using various activation functions (such as ReLU, linear activation function, X-function) and the "Adam" optimizer. The results showed that the proposed prediction system enhances the accuracy of the prediction, with deep learning algorithms achieving an accuracy of about 98.3%

when compared with other algorithms. These techniques have been particularly ⁷ applied to predict heart disease, breast cancer and chronic diabetes. (Mishra, J., Tarar, S.2020)

Paper5: Prediction of Chronic Disease by ¹⁴ Machine Learning.

Accurate ¹⁸ prediction of disease risk in patients is vital in the biomedical and healthcare communities, as it contributes to the early detection of diseases. However, the quality of medical data significantly affects the accuracy of the forecast, as incomplete or inaccurate data leads to less effective results.

This study uses a specific machine learning algorithm to determine disease rates by processing a dataset drawn from specific hospitals. The process involves pre-processing the data and reconstructing the missing values to ensure the accuracy of the analysis. After comparing several algorithms, the RNN (iterative neural network) algorithm proved to be the most accurate, achieving up to 97.6% with a noticeable convergence speed. (P. Anandajayam, S. Aravindkumar, 2019)

Paper6: Disease ³ Prediction using Machine Learning Algorithms

The development of data mining technologies has led to their increasing use in multiple fields such as industry, healthcare, and biosciences, contributing to the improvement of medical information analysis. This research focuses on the use of machine learning technologies in the early prediction of diseases and the enhancement of patient care.

The development of a classifier system using machine learning algorithms is aimed at helping doctors diagnose diseases in their early stages. Data from the registries of 4,920 patients diagnosed with 41 diseases were used, with 95 out of 132 independent variables (symptoms) selected as relevant.

The paper reviews the disease prediction system developed using machine learning algorithms such as decision tree classifiers, random forest classifiers, and naive Bayes classifiers. The study also presents a comparison of the performance results of these algorithms, highlighting the effectiveness of each of them in improving the accuracy of disease prediction. (S. Grampurohit and C. Sagarnal, 2020)

Critical Analysis:

A critical analysis of the presented research papers demonstrates the importance of using machine learning technologies in enhancing the ability of healthcare systems to predict chronic diseases, a topic of great importance in the fields of health informatics and medicine. Research is characterized by the use of diverse algorithms such as convolutional neural networks (CNN), deep learning algorithms, and techniques such as KNN, reflecting the diversity of methods used to achieve prediction goals. Despite the strengths, such as practical applications and promising results, there are many weaknesses that should be focused on and addressed.

Despite the strengths, such as practical applications and promising results, there are several weaknesses that need to be addressed. Firstly, some papers lack sufficient details about the quality of the data used and how to process the missing values, which negatively affects the reliability of the final results. In addition, the lack of sufficient clarification of how the models are applied in clinical settings limits the ability of researchers and practitioners to benefit from them in practice.

Secondly, some studies suffer from the limited scope of comparison of the algorithms used, since it is preferable to expand this scope to include newer and more sophisticated algorithms than those that have been employed. Such an expansion would provide a deeper and more comprehensive insight into the performance of algorithms and contribute to improving the accuracy of the proposed models. In addition, enhanced transparency and clarity in the analysis methodology and method of application can lead to improved reliability of the results. This transparency facilitates the process of verifying the results and repeating the experiments, which makes the proposed models more applicable and helps in achieving broader and more effective health benefits.

Finally, while these papers provide valuable insights, focusing on applying the findings in practical clinical contexts may enhance the usefulness of research and help improve healthcare outcomes. In general, the research reflects a deep understanding of the current challenges in the field of chronic disease forecasting, which suggests the need to develop better methodologies to enhance the effectiveness of these models and bring tangible benefits to patients.

Connections

1. its shows the importance of using artificial intelligence to support the diagnosis of chronic diseases at an early stage, which is associated with global efforts aimed at improving the accuracy of clinical diagnosis using machine learning.
2. highlights the importance of dimensionality reduction in improving the performance of models, and is related to other studies concerned with the selection of optimal features to enhance prediction accuracy and avoid noise in medical data.
3. This research is related to the development of models capable of handling incomplete or short-sequence data, reflecting a growing interest in machine learning applications in dealing with unstructured data in the medical field.
4. corresponds to numerous attempts to integrate deep learning into the prediction of diseases such as heart and cancer, as it seeks to develop accurate models based on modern algorithms.
- 5 . This study is in line with other research aimed at improving the accuracy of disease prediction through the use of generalization algorithms that deal with missing data

Conflicts between the findings:

1. Research is unanimous in its opinion that the integration of CNN and CNN algorithms contributes to enhancing the accuracy of forecasting chronic diseases, but there are studies indicating that some alternative algorithms, such as repetitive neural networks, may be more accurate and effective in analyzing temporal data of patient records
2. While this study showed that dimensionality reduction enhances the efficiency of models, other studies have suggested that dimensionality reduction may lead to the loss of some vital information, which may affect the accuracy of the prediction.
3. The study shows that the use of short sequence data leads to high accuracy in forecasting, although some other studies believe that such data may be insufficient, and models need longer time data to ensure forecast accuracy.
4. ²⁰ algorithms such as decision tree ²² and random forest provide high accuracy in predicting chronic diseases, while other studies show that deep learning algorithms such as CNN excel in improving accuracy, showing a clear divergence in algorithm preference.
5. Some researchers focus on filling in missing data as a way to improve accuracy, but other studies warn that these missing data may lead to biases that affect the results of models, highlighting inconsistencies in the effectiveness of dealing with missing data.

Synthesis Matrix for Literature Review:

Study	Authors	Year	Research Design	Key Findings/Results	Methodology	Limitations
5 Identification and Prediction of Chronic Diseases Using Machine Learning Approach	Rayan Alanazi	2022	Experimental	3 The study proved the effectiveness of the use of convolutional neural networks (CNN) and the nearest neighbour algorithm (KNN) in the early detection of chronic diseases and achieved high accuracy in forecasting based on the symptoms and health habits of patients.	The study was based on the collection of comprehensive data covering symptoms and health habits, where the model was trained using convolutional neural networks (CNN) and nearest neighbour (KNN) to extract and classify features. A large dataset comprising information from patients, including symptoms and health behaviours, was used to develop the model and improve its	Specifically, the study focuses on specific aspects of the patient's data, namely only symptoms and health habits, which may be an oversimplification of the patient's condition. It can be better if the data includes other factors, such as family history and environmental factors, to get a more accurate overall picture of the patient's condition. In addition, the performance of some algorithms, such as CNN and KNN, depends on

					ability to predict chronic diseases.	the presence of large and balanced data to achieve high accuracy, and such data may not always be available in some medical environments. Therefore, it is important to take these factors into account when designing studies and applying models to predict diseases and provide a medical diagnosis.
² Feature selection and classification systems for chronic disease prediction: A review.	Divya Jain , Vijendra Singh	2018	Experimental	The study explained that the careful selection of features enhances the accuracy of classification, and the reduction of dimensions improves overall performance, which contributes to the	We have used feature selection techniques to reduce dimensionality and applied multiple classification algorithms to a dataset that includes many variables related to chronic	Focusing on improving classification accuracy by reducing dimensions can sometimes lead to ignoring some less common but important symptoms. This can affect the ability of the

				construction of adaptive and rapid diagnostic systems for chronic diseases.	diseases. A large sample size of multidimensional data from a large number of patients was included with the aim of improving the accuracy of disease classification and enhancing the ability of the model to predict pathological conditions with higher accuracy.	model to recognise rare or complex cases. In addition, the variety in the classification algorithms used may limit the ability of the study to achieve results generalisable to different data sets. So, these points should be taken into account when designing studies and choosing appropriate methods for analysing medical data and predicting diseases.
¹ A method for the early prediction of chronic diseases based on	¹ Chengkai Wu, Tianshu Zhou, Yu Tian, Junya Wu, Jingsong Li, Zhong Liu	2022	Experimental	The study provided a model for early prediction of diseases using short serial data from medical records, achieving high accuracy	The study was based on short-sequence electronic medical data, where a network (SIB) was developed using LSTM layers and	The model is based on short-sequence data, which may lead to challenges in dealing with diseases with variable symptoms over

short sequential medical data.				results through the use of short memory layers (LSTM) and attention mechanisms.	attention mechanisms. Data were collected from patients with short-sequence symptoms of a variety of chronic diseases, which helped to develop an accurate model for predicting those diseases.	time. Also, the SIB model requires structured sequential data, which is a hindrance when using incomplete or unbalanced data from electronic health records.
7 Chronic Disease Prediction Using Deep Learning.	Mishra, J., Tarar, S.	2020	Experimental	A comparison of algorithms such as KNN, decision tree, and deep learning showed that the latter outperformed with 98.3% accuracy in predicting diseases such as heart, breast cancer, and diabetes, demonstrating the effectiveness of deep learning with standard medical datasets.	Available benchmarking data involving heart disease, breast cancer, and diabetes were used, comparing KNN and decision tree algorithms with deep learning models such as ReLU and Sigmoid. A standard data sample size for multiple diseases has been included with the aim of achieving high	The study relies heavily on previously available data, which may lack updating and may be biased towards certain samples. The use of a certain number of deep learning algorithms may also lead to limitations in the results, as other models may be more compatible with certain data.

					accuracy in predicting those diseases and improving the performance of the models used.	
5 Prediction of Chronic Disease by Machine Learning.	P. Anandajayam, S. Aravindkumar, P. Arun and A. Ajith	2019	Experimental	The recent study showed promising results as the random forest algorithm achieved an amazing accuracy of up to 97.6% in predicting chronic diseases. And not only that, but also highlights the crucial importance of addressing the missing data in the process of analysis and forecasting of these diseases, as this is considered 16 decisive factor in improving the accuracy and speed of the scientific model.	The study relied on medical data from specific hospitals, and the data set was pre-processed to deal with the missing values. The random forest algorithm was used to achieve high accuracy, and the results were based on real patient data that includes a large number of dimensions and health features. This approach is considered important in improving medical diagnostics and disease forecasts	The study requires intensive processing of missing data to ensure the accuracy of the forecast, and this may cause inaccuracies in some cases if the data is not adequately processed. Relying only on random forests may also be insufficient in some cases where higher complexity or deeper models are needed to handle the data.

					with greater accuracy and effectiveness.	
4 Disease Prediction using Machine Learning Algorithms	S. Grampurohit and C. Sagarnal,	2020	Experimental	Comparing decision tree classifiers, random forests, and knife bays, the study showed that random forests outperformed the other with strong and reliable performance in predicting multiple diseases. This outstanding performance enhances the reliability of the use of random forests in health systems and makes it an ideal choice for improving the predictability of chronic diseases and improving the quality of health care.	Decision tree, random forests, and knife Bayes algorithms were used to develop a classified system based on health record data of 4,920 patients with 41 different diseases. 95 relevant variables were selected from 132 variables to analyse the results and improve the accuracy of disease prediction. This approach contributes to the improvement of medical diagnosis and prediction of diseases with higher accuracy and improved effectiveness.	The study lacks the use of more sophisticated algorithms that may be more effective in processing high-dimensional data. Also, it has not been clarified how to choose independent variables from 132, which may lead to a bias in the prediction results.

References:

Alanazi, Rayan, Identification and Prediction of Chronic Diseases Using Machine Learning Approach, *Journal of Healthcare Engineering*, 2022, 2826127, 9 pages, 2022. <https://doi.org/10.1155/2022/2826127>

Divya Jain, Vijendra Singh, Feature selection and classification systems for chronic disease prediction: A review, *Egyptian Informatics Journal*, Volume 19, Issue 3, 2018, Pages 179-189, ISSN 1110-8665, <https://doi.org/10.1016/j.eij.2018.03.002>

Chengkai Wu, Tianshu Zhou, Yu Tian, Junya Wu, Jingsong Li, Zhong Liu, A method for the early prediction of chronic diseases based on short sequential medical data, *Artificial Intelligence in Medicine*, Volume 127, 2022, 102262, ISSN 0933-3657, <https://doi.org/10.1016/j.artmed.2022.102262>

Mishra, J., Tarar, S. (2020). Chronic Disease Prediction Using Deep Learning. In: Singh, M., Gupta, P., Tyagi, V., Flusser, J., Ören, T., Valentino, G. (eds) *Advances in Computing and Data Sciences*. ICACDS 2020. Communications in Computer and Information Science, vol 1244. Springer, Singapore. https://doi.org/10.1007/978-981-15-6634-9_19

P. Anandajayam, S. Aravindkumar, P. Arun and A. Ajith, "Prediction of Chronic Disease by Machine Learning," *2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN)*, Pondicherry, India, 2019, pp. 1-6, <https://doi.org/10.1109/ICSCAN.2019.8878724>

S. Grampurohit and C. Sagarnal, "Disease Prediction using Machine Learning Algorithms," *2020 International Conference for Emerging Technology (INCET)*, Belgaum, India, 2020, pp. 1-7, <https://doi.org/10.1109/INCET49848.2020.9154130>

literature Review prediction of chronic diseases1

ORIGINALITY REPORT

13%

SIMILARITY INDEX

11%

INTERNET SOURCES

11%

PUBLICATIONS

2%

STUDENT PAPERS

PRIMARY SOURCES

1	pubmed.ncbi.nlm.nih.gov Internet Source	3%
2	doaj.org Internet Source	2%
3	Arvind Dagur, Karan Singh, Pawan Singh Mehra, Dharendra Kumar Shukla. "Artificial Intelligence, Blockchain, Computing and Security", CRC Press, 2023 Publication	1%
4	www.coursehero.com Internet Source	1%
5	www.semanticscholar.org Internet Source	1%
6	nitter.grimneko.de	

Internet Source

1 %

7

link.springer.com

Internet Source

1 %

8

www.hindawi.com

Internet Source

<1 %

9

www.sciencepubco.com

Internet Source

<1 %

10

cajotas.centralasianstudies.org

Internet Source

<1 %

11

ebin.pub

Internet Source

<1 %

12

Duo Xu, Zeshui Xu. "Machine learning applications in preventive healthcare: A systematic literature review on predictive analytics of disease comorbidity from multiple perspectives", Artificial Intelligence in Medicine, 2024

Publication

<1 %

13

centaur.reading.ac.uk

Internet Source

<1 %

14	iris.unica.it Internet Source	<1 %
15	publish.mersin.edu.tr Internet Source	<1 %
16	storage.googleapis.com Internet Source	<1 %
17	wjbphs.com Internet Source	<1 %
18	www.ijisae.org Internet Source	<1 %
19	Divya Jain, Vijendra Singh. "Feature selection and classification systems for chronic disease prediction: A review", Egyptian Informatics Journal, 2018 Publication	<1 %
20	Pawan Singh Mehra, Dharendra Kumar Shukla. "Artificial Intelligence, Blockchain, Computing and Security - Volume 2", CRC Press, 2023 Publication	<1 %

21	"Mobile Radio Communications and 5G Networks", Springer Science and Business Media LLC, 2024 Publication	<1 %
----	---	------

22	S.J. Xavier Savarimuthu, Sivakannan Subramani, Alex Noel Joseph Raj. "Artificial Intelligence for Multimedia Information Processing - Tools and Applications", CRC Press, 2024 Publication	<1 %
----	---	------

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		