Integrative approach for modern health risk modeling and predicting in patients through artificial intelligence method

Anusha Ampavathi¹*, P. Tamiiselvi², Raghu Kumar Lingamallu³, Abhisek Sethy⁴, Sreenivasulu Gogula⁵, M.N. Sharath⁶, and K. Vijayalakshmi⁷

¹Department of Artificial Intelligence, Vidya Jyothi Institute of Technology, Hyderabad, Telangana, India.
²Department of Computer Science and Business Systems, Rajalakshmi Engineering College, Thandalam, Chennai, India.
³Department of CSE, KG Reddy College of Engineering and Technology, Chilkur Village, Hyderabad, Telangana, India.
⁴Department of Computer Science & Engineering, Silicon Institute of Technology, Bhubaneswar, Odisha, India.
⁵Department of Computer Science and Engineering (Data Science), Vardhaman College of Engineering, Shamshabad, Hyderabad, Telangana, India.
⁶Rajeev Institute of Technology, Hassan, Karnataka, India.
⁷SRM IST Department of EEE Ramapuram Campus, Chennai, India.

Abstract. This research suggests a modeling approach for health risk prediction that utilizes an ambient environment and Artificial Intelligence (AI). The proposed AI-based Health Risk Modeling and Predicting System (AI-HRMP) included gathering medical records from chronic illness patients, including Electronic Health Records (EHR), Personal Health Records (PHR), medical records, and environmental variables from a health portal. Diverse data is combined via choosing, cleaning, modeling, and assessing raw data, followed by data production. Sensor data is standardized by converting the time-domain details to frequency-domain details. The standardized input is processed using an AI to provide an ambient environment. A health risk prediction system has been proposed to analyze specific health issues about environmental factors. The program utilizes ambient context patterns identified via metadata and AI. The risk prediction framework is integrated into a person's risk alert/prevention mechanism. The system might substantially influence healthcare and AI studies, ultimately enhancing the future society's standard of life.

1 Health risk modeling overview

Ubiquitous healthcare enabled by ambient intelligence is crucial in a healthcare system. An intelligent medical platform has been designed using many modern Internet of Things (IoT) technologies [1]. A personalized context awareness technology benefits medical and

* Corresponding author: anuampavathi@gmail.com

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).
emergency response, serving as a fundamental paradigm for quickly caring for individuals with chronic diseases in a health system. Chronic conditions are illnesses that develop gradually and need extended treatment and recovery [2]. Medically, chronic conditions are diseases with signs that last for longer than six months to over one year—such as Chronic Respiratory Disease (CRD) resulting from industrialization as a contagious illness [3]. Due to the growing geriatric population, the amount of seniors with chronic diseases has risen. The medical expenses of older people are growing as a proportion of the overall medical expenditures. Cancer is a chronic illness associated with a high mortality rate. This chronic condition results in a prolonged sickness, making it difficult to diagnose accurately. High cholesterol levels and chronic disease have prolonged disease duration and lead to different consequences based on the level of treatment [4]. Thus, when a disease occurs, treatment should be administered in an initial medical facility with easy access, and ongoing care should be given to the patients. The self-care of a patient with a chronic condition significantly impacts the course of the disease. Artificial Intelligence (AI) is crucial for medicine to manage chronic diseases and reduce medical expenses for people with chronic conditions [5]. An enhanced context awareness system with excellent security and utility is vital due to the growing needs of chronic illness sufferers [6].

The sections are organized as follows: Section 2 covers the related research on Health Risk Modeling. Section 3 presents the design and discussion of the proposed AI-based Health Risk Modeling and Predicting System (AI-HRMP). Section 4 demonstrates the software analysis and outcomes of various algorithms. Section 5 summarizes the research and explores the future scope of the study.

2 Related works and findings

The healthcare service records consist of fundamental data, including medical big data from hospitals, government agencies, and healthcare groups, as well as bio big data and lifelog big data. Lifelog data involves saving personal life experiences by gathering sensor knowledge and storing what was learned. Data is collected, documented, and held daily without knowing the location or moment. The study suggested a strategy for extending life utilizing ambient intelligence inside a wireless sensor network [7]. Gyrard et al. created a proficient IoT recommendation engine and a graphics-based recommending method [8]. Various ambient sensing techniques are employed for collecting everyday living details such as location, movements, time, location, and biosignals. Users’ interests, preferences, lifestyles, and other relevant data are input into an inference machine and stored continuously. The integration of electronic systems in major medical facilities and the evolution of IoT have led to the collection of a significant volume of lifelog big data from connected devices [9].

Furthermore, the data provided by the Korean Meteorological Department and National Healthcare Services grows increasingly varied and detailed [10]. Integrating and analyzing diverse large datasets is crucial for developing healthcare services for chronic illness patients. This technology mainly focuses on combining Electronic Medical Records (EMR) [11] and Personalized Health Records (PHR) [12], offering services based on public health data. The diverse large datasets will be combined and analyzed using AI and data mining techniques to create the health database. Integrating organized and unorganized information is essential to derive a relevant knowledge foundation in diverse big data [13]. Utilizing a situation awareness system requires integrating individually collected IoT data, bio-logging big data, and other health records.

Additionally, doing smart algorithm-based research and providing tailored information is essential [14]. It is necessary to integrate ensemble filtering [15], deep learning [16], and other methods, using their respective strengths to achieve data integration. The integrated
hybrid method enhances the learning system, resulting in an optimal health risk evaluation algorithm with a higher predicting accuracy.

3 Proposed AI-based health risk modeling and predicting system

This study suggests an AI-HRMPs for predicting chronic disease. The proposed strategy reduces the need for extensive pre-processing by incorporating an epidemiological knowledge foundation, unlike the typical AI method that requires a lengthy training period and significant pre-processing. This involves using a validated training dataset endorsed by a medical professional and algorithms to extract health information from raw sensor information. The concept outlines processes for acquiring an evolving test database during the testing period, created by combining sensory and non-sensory information to match the training data architecture.

3.1 Training phase

3.1.1 Dataset generation

The critical step to creating a training database in AI-HRMPs is establishing an epidemiology library based on illness risk variables identified from patient information. Patient information is gathered using a direct pre-screening survey or other permitted methods, monitored by medical professionals who validate the data’s classification level. The processed information is kept in a knowledge library. Validated data from practitioners might enhance the credibility of risk forecasting for non-communicable diseases. Using one database and constructing an epidemiological library utilizing EHR is a promising option for predicting different chronic diseases. EHR have been used in several tele-healthcare services.

3.1.2 Knowledge base

A knowledge base comprises validated datasets, ontology, and procedures for data analyses. For instance, risk ontology, symptoms, illness ontology, healthcare rules for attribute value determination, and other reusable data for dataset query purposes. Utilizing an epidemiology knowledge base helps expedite the efficiency of the categorization method. The suggested method utilizes the knowledge base throughout the training and assessment phases. During the training stage, the validated database from the knowledge database was utilized to train the classification using several AI categorization techniques. The knowledge database established rules in the testing stage, allocated data labels, and retrieved characteristics for forecasting chronic diseases from sensor information.

3.1.3 Trained classifier creation

The validated training dataset is employed to train the classification. Several widely used AI methods are employed for chronic disease forecasting. The result of the training step is the trained classifications. These classifications are used to assess sensory input during the testing stage.
3.2 Testing phase

The suggested technique introduces new procedures for the AI-HRMPS testing stage. The objective is to create a dynamic testing database from the unprocessed sensor information and categorize it to forecast chronic diseases. The same procedure predicts other chronic diseases based on the system’s recorded data and classification algorithms. Figure 1 illustrates the methods in this stage, with the information flow shown by numbers to provide a closed-loop system.

![Fig. 1. Testing model of the proposed AI-HRMPS](image)

3.3 System architecture

EMR data usually consists of an individual health information managing component and a medical record management component. The PHR data utilizes ZigBee, using the IEEE 101073 standard for connectivity. Every standard encrypts information to enhance security and address the most recent privacy concerns. To securely gather and store data, a data collection technique is necessary. The collecting system includes an advanced big data gateway architecture that operates on standard data modeling and utilizes decentralized file handling. Structured and unstructured data are mined using a data repository. Using the established XML standard, a gateway device gathers and analyses the decentralized data repository. Figure 2 depicts the system design.

![Fig. 2. System design of the proposed AI-HRMPS](image)

An integrative health system is built to optimally use ambient context by including prepared individual physical characteristics and environmental context details from public health data. The initial preprocessed data comprises unstructured information. The existing unorganized information processing technique for decision-making is applied to the database.
3.4 Risk factor extraction

The PHR analyses the trends of bio-data generated by an ambient biosensor in an individual’s wearable gadget or smart home security technology. The diverse data is combined using a context-aware computing method. Ambient environment recognition is the system that identifies an individual’s environment and delivers personalized information services based on the user’s surroundings. A context-aware computer system does context categorization based on the circumstances. PHR data include a range of elements gathered as big data. The significance of each component varies based on the specific chronic condition. An approach that assesses the weight of each item and produces a decision is implemented appropriately. A deep neural network is an advanced technique that effectively handles large amounts of data. Therefore, developing a learning model that incorporates an interconnected file architecture and training information is essential. The learning system constructed modifies a weight during data collecting and undergoes iterative learning.

![Fig. 3. Architectural diagram of the AI-based pattern detection](image)

Figure 3 depicts the deep neural network model designed for ambient environment monitoring. A typical deep neural network model consists of input layers that receive information, concealed layers that adjust connection weights for training, and output layers that provide the outcome of the information. Learning is divided into controlled learning, such as categorization or analysis, and autonomous learning, such as grouping or organization, based on whether the outcome information is available. Reinforcement learning is employed for the iterative feedback assessment process. A deep neural network allows for the representation of complex nonlinear interactions. Eight nodes were created in this investigation based on the frequency inputs. A concealed layer consists of hidden nodes that have a similar structure.

4 Simulation analysis and outcomes

The experimental setting in this study is separated into two stages. During the first phase, the training process utilizes a validated chronic database. The dataset was obtained ethically and with expressed agreement from patients at an established disease treatment hospital. The data is gathered from the patient’s authorization when a medical officer recognizes a patient with a possible chronic disease. Patients suggested for clinical testing are classed as positive. This dataset is used to forecast the probability of chronic disease in its early stages based on typical symptoms and indicators to reduce the risk of premature mortality due to chronic disease. During the second stage, test information was generated from modeling and prototyping to assess the effectiveness of several categorization methods. A sensor network for wearable devices was created by upgrading instances of the Cooja simulator with the Contiki OS. The detectors have been altered using the Python coding platform. The findings have been contrasted with other studies on predicting chronic risk using machine learning techniques.
Fig. 4. Accuracy analysis of chronic disease prediction models

Figure 4 shows the results for the accuracy of the chronic disease prediction models. The dataset about chronic diseases is used to compute the accuracy of several machine learning models, including but not limited to Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), Linear Discriminant Analysis (LDA), Principal Component Analysis (PCA), Naïve Bayes (NB), and the suggested AI-HRMPS. The findings indicate that the proposed AI-HRMPS technique achieves more significant outcomes in terms of accuracy than the other methods. AI-HRMPS demonstrates an accuracy rate of 89.4%, with a standard deviation of 1.65% in the findings.

Fig. 5. RMSE analysis of chronic disease prediction models

Figure 5 presents the results of an investigation of chronic illness prediction using the Root Mean Squared Error (RMSE) method. The standard deviation for the RMSE is shown alongside the average results of the training and testing phases. With the assistance of an AI model, the suggested AI-HRMPS can decrease the total error, which in turn leads to an improvement in the accuracy of the forecast. There is a standard deviation of 3.4% and an RMSE of 86.3% for the AI-HRMPS.

Fig. 6. Computation time analysis of chronic disease prediction models
When calculating the amount of time required for the training and testing phase of the chronic illness prediction model, the entire amount of time is considered. Figure 6 displays the overall outcomes of training and testing for various models. With the assistance of an AI model, the suggested AI-HRMPS can decrease the total amount of time spent on computing while simultaneously improving the accuracy of predictions. The AI-HRMPS demonstrates an overall processing time of 45.3 milliseconds, with a standard variation of 2.65%.

5 Conclusion and findings

This research suggested a modeling approach for health risk assessment for chronic diseases that utilizes the ambient environment using a deep neural network. The AI system under consideration is a medical device designed to provide tailored assistance to those with chronic illnesses. AI for chronic disease patients can consolidate diverse big data sources, including medical data from hospitals and government agencies, individual medical histories, and lifelong. The dispersed file processing-based chronic disease prediction model incorporates personal context elements like nutrition, surroundings, and weather data. It was essential to integrate and analyze diverse unstructured and structured information from different sources to gather valuable new information and use it as the knowledge foundation for Ambient Intelligence. The system preprocessed the organized temporal domain data by converting it to the frequency domain information using an altered Fourier transform approach. AI technique was employed to learn contextual knowledge for context understanding. The efficacy of the health risk evaluation system was evaluated by analyzing how it varied with the development rate of the AI. Optimally applying the method to the combined data enhances the system’s efficiency. This research demonstrated the significant potential of the health risk warning service for people with chronic diseases.

While the data sets employed in this study need to be revised to offer trustworthy information, the prediction’s accuracy is expected to improve if the system established during this study is implemented in an area where the data will quickly increase. Future research must integrate information from several sources and create effective technologies to enhance present studies. Furthermore, the suggested AI-HRMPS model is a medical decision-support method for preventing chronic diseases and detecting early signs of illness in patients receiving surgical care.

References


