International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Healthcare Predictions Using Machine Learning and Artificial Intelligence Algorithms

Sarah j mohammed*[1], Shaymaa k mohsin[2], Maysoon A. Mohammed[3]

Submitted: 10/09/2022 **Accepted:** 20/12/2022

Abstract: Machine Learning (ML) combined with Artificial Intelligence (AI) are essential implements for enhancing medical care. Medical problems such as pulse rate, respiratory rate, oxygen level, blood pressure, falls, diabetes level, human body temperature and diagnosis blunders are prominent adverse occurrences in healthcare. This proposal aimed to employed AI's ability besides machine learning improve patient care in this eight high-risk areas to predict, avoid, or diagnose undesirable outcomes. Healthcare-associated Infections to determine if AI can improve safety, the literature was analyzed regarding incidence, cost, prevention, and treatment. The paper included 100 different samples, provided numerous cases of how intelligence was used in all eight damage categories. In several fields, AI and new data sources can help reduce damage rates. The proposed achieved 81% accuracy in diagnosing the tested cases. The treatment plan's reliability rate increased to 91% compared to the traditional treatment by the doctor to 58%. So, it includes adverse medication effects, hypertension, and diagnostics errors to mention several.

Keywords: *CNN, Sensitivity Analysis algorithm, Machine Learning, AI, Healthcare.*

1. Introduction

There are many main reasons for death and failure internationally, and a following to a good portion of them appear to be avoided. Investing in reducing harm saves money and improves patient outcomes. Despite patient-centered safety measures, including inpatient checklists, digital ordering, and bar-coding, care issues persist 25 years after the Institute of Health's "Error Is Individual" [1] research. Although care is quickly moving outside hospitals, outside hospital safety has gotten significantly less attention than within hospital safety [2]. These technologies can help forecast dangers, collect and analyze data (both old and new), and improve care both outside and inside the hospital .For example, AI can guide prevention strategies and early intervention activities by recognizing patients at patient role risk of injury in the clinic. Ambulatory, neighborhood, and home healthcare settings can use AI. These tools, when integrated with digital techniques, can improve patient-provider communication, and reduce preventable risks. Available data will be beneficial, but more data from sensors will be required to improve forecasts [3,4].

A number of scholars have made proposals for healthcare development, using computers and intelligent applications. Research from University of Tsukuba Hospital proposed use deep convolutional neural network (DCNN) for diagnosing skin tumors for 4867 images and these are from 2003 to 2016 [5]. Aditi B. and Priyanka G. presented a proposal to analyze healthcare data using Dynamic-Slot-Allocation in Hadoop [6]. Van-Dai Ta, Chuan M. L., Goodwill Wandile N. based on Real-Time-Analytics of Big Data Stream Computing [7]. Jagreet K., Kulwinder S. presented Health-Care Platform for Real-Time based on AI application and Reactive-Programming.ML can predict clinical risk and increase patient safety [8]. Other research testing for diabetic retinopathy using Retina image to make a decision at the illness stage by deep learing techinqes[9,10].The goal is to identify determinants and consequences. Data-driven ML algorithms beat rule-based approaches. Healthcare organizations are quickly adopting machine learning and other AI technologies to enhance patient outcomes. Adopting these solutions through the system, monitoring groups, and the market is required to achieve major gains in safety and lower costs. Infections, adverse pharmaceutical events [11], venous thrombolytics, medical disorders [12], pressure sores, falls, insufficient retrograde motion sensors, and diagnostic errors such as missing and delayed diagnosis cause most healthcare repercussions. At the same time, other factors play a part; harmful actions currently account for most of the clinic damage.

Most AI study in this field has been on exploiting sensor facts for rapid recognition; like a response, employing AI to guess coming risk stays a probable issue. An arbitrary model was created using Electronic Health Record (HER) records to establish important care patients with a high chance of getting pressure ulcers and can show in assessment (AUC = 0.80 versus 0.74 for the Neurological Assessment), according to a recent study. Researchers previously investigated the prospect of employing smart beds and wheelchairs cushions to identify pressure ulcers using artificial neural networks and machine learning algorithms. These algorithms analysed data from attached sensors that detect a lack of urgency and pinpoint specific patches of skin on the verge of ulceration. Despite the models' ability to detect up to 89.75% of instances under experimental conditions, their utility and effectiveness is alerting

[1] *Mechanical Engineering Department, University of Technology, Baghdad,, IRAQ.*
ORCID ID : 0000-0003-0458-7569
[2] *Mechanical Engineering Department, University of Technology, Baghdad, IRAQ..*
ORCID ID : 0000-0002-6748-2636
[3] *Mechanical Engineering Department, University of Technology, Baghdad,, IRAQ.*
ORCID ID : 0000-0002-4892-6772
** Corresponding Author Email: sarah.j.mohammed@uotecnology.edu.iq*

healthcare professionals and encouraging early diagnosis remain uncertain.

Many articles examined artificial intelligence's use to predict or detect healthcare-Associated Infections HAIs early in their development. In the past, machine learning with fuzzy logic [13, 14] (rational and logic models established on inadequate or unclear information) was used to aid in the early diagnosis of HAIs. Most algorithms were created using claims-based knowledge/analysis from health records, like laboratory data and diagnostic imaging [15,16]. Machine intelligence analytics could speed up detection and enhance diagnostic accuracy by merging new and complex data. AUC = 0.97), discriminate between five generic wound pathogens (accuracy = 78%), and identify Clostridium difficult straining (sensitivity > 79.5 percent; specificity > 70%) using data from e Noses. AI can assist in infection management by predicting HAI risk in real-time and guiding patient-specific therapy before the occurrence of an infection [17,18]. Another research based on machine learning models in a study concerned with environmental pollution and its impact on children's perceptions, the problem was the number of huge data and the way it was collected from the community [19,20]. This study aimed to determine if AI could enhance healthcare safety by lowering adverse occurrences in these eight core risk domains, by building a model consisting of artificial intelligence and machine learning algorithms to reach a meaningful decision in diagnosing the patient's health.

2. Methods

The aim of proposal is to predict the diagnosis of the diseased condition of the individual and to make a decision in his health future. The proposal uses the CNN and the SAA respectively as show in figure (1). The data entry into the CNN, and the output is a report that includes a detail of the patient's condition, but the large number of information and need to extract the strongest and most related characteristics reasons we will use the SAA, through training repetition gets rid of the case Input uncertainty. When using the results of the model outputs, the sensitivity measurement process reaches the benefit, thus making the decision to diagnose the patient's condition more accurately.

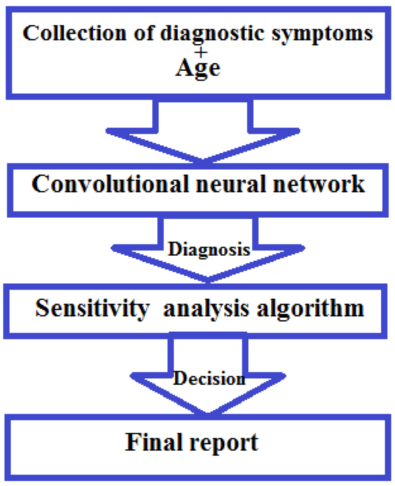

Figure 1. General structure of the proposed system.

To achieve this goal, the patient's data (according to the doctor and patient) is determined and arranged numerically in n*n matrix, which is the symptoms and medical history. Some inputs include medical images that will be entered as digital numbers (after image processing performing). All relationships are simplified with a single matrix for the individual. This matrix enters the convolutional neural network model. The proposal includes building a filter kernel that includes the eight approved diseases (pulse rate, respiratory rate, oxygen level, blood pressure, fall, diabetes level and human body temperature) in addition to the age, which was classified into the well-known age groups (children, young adults, middle-aged adults and the last old-age adults) as shown in Figure 2.

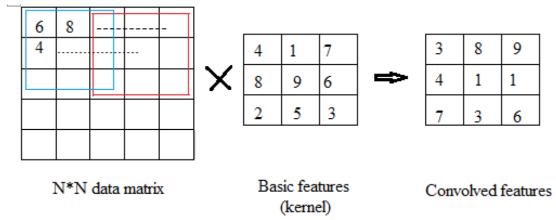

Figure 2. Convolved features production.

The function of artificial neuron that has weights W with bias B and activation function K is define as below:

$$F(X_1, X_2, \dots, X_m) = K\big((X,W) - B\big) = K\left(\sum_{i=1}^{m} XiWi - B\right) \quad (1)$$

The activation function (basicly predict output probabilities) is define as:

$$K(x) = \frac{1}{1+e^{-x}} \quad (2)$$

Convolutional layers wrap the inputs to produce (attribute learning) and classify the data, then pass the results to the next layer where the data is quantified to extract the best features and nested in hyperlinks to produce an individual diagnosis report. The proposed network includes a level for side effects related to diagnostic errors and drug effects to go back to the beginning (feedback) as shown in Figure 3.

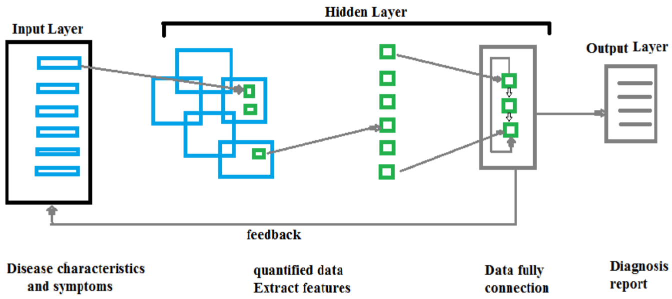

Figure 3. The CNN stages.

Here employed feature importance in conjunction with the branch classifier to reduce the number of features. Then we organized the features into categories. Each category is assigned a feature

<u>Algorithm :: Sensitivity Features Analysis</u>

Input:: n: number of Important features // features resulting from CNN
  C: convergence criteria
Output:: D: nominated feature
  S:dataset
  Repeat x
   k= n/i   // k represent one feature of sensitive analysis
    S= k ∪ [C_sensitive]
    If (S = Categories [Ci]) then // validation dataset
     D_sensitive = S   //amount of predicts
   Else
    Ci[] ← Shift (S)//shifting kernel in array for rearrangement
  Until x> n
End

sensitivity score. The proposed sensitivity algorithm's steps are depicted bellow.

In the context of data mining, feature selection is the process of identifying and selecting the most significant qualities from a dataset. Machine learning is a complex process in which it plays an important role. Irrelevant properties degrade the performance of a machine learning model by slowing training, making model interpretation more difficult, and, most critically, diminishing efficiency on the test set [21]. While this may appear to be a straightforward procedure, it is actually one of the more difficult aspects of constructing a new machine learning model to master. The initial step of the model should comprise the selection of features and the purification of data. The significance of characteristics is established through the use of feature selection procedures, which consider the variance of the features and their relationship to the target attribute. In which was before step of the data, significant features are identified, and the features that are picked are then used to train a model [22].

The age is the most important variable, accounting for roughly 14% of the total, and that the other features each account for about 10% of the total. The contribution of a feature in the model is measured by its significance. Adding more features to the model involves adding more data to the model. As a result, the learning rate is not influenced by the feature selection process at all. It is a presentation style that displays how the properties of a model influence the model's forecasting ability. There are a variety of ways that may be used to decipher these black-box models. Depending on its significance to the output variable, each data item is assigned a numerical value; the higher the aggregate score, the more critical or important the feature is of the result variable [23].

3. Results and Discussion

Sensitivity Feature analysis is a machine learning technique that consists of a set of independent assessment and is commonly used for categorization tasks [24]. Each tree is created using samples and features that are chosen at random [25,26]. To forecast the test results, we constructed a model. As a result, we fit the model with training data and test it with testing data. 75% of the data was utilized for training, while 25% was used for testing. For handling dataset imbalance, the oversampling technique have been used on the instruction set. We also tweaked the algorithm's hyperparameters, such as n estimators, max intensity, and arbitrary stage, which were set to 40, 16, and 18 respectively. On the testing data, the model had 81% an accuracy rate, table (1) shows a comparison with some works that serve healthcare by using CNN.

Table 1. Accuracy rate of the proposed works

Disease	*algorithm*	*Accuracy rate %*	*Reference*
tuberculosis	NN	78	[27]
skin cancer classification system	DCNN	76.5	[5]
skin diseases	CNN	87.1	[28]
(healthcare system)The proposed system	CNN+SAA	81	

Diagnostic error is the most challenging of the eight injury categories to identify, even with different types of data and artificial intelligence. Diagnosticians frequently make mistakes in pattern detection, biased minimization, and infinite capacity, among other things. The number of attempts in the training process is determined to find the minimum acceptance and error determination. Initial tests showed that the training process reaches a stage of stability after identifying all the variable data, and this in turn is reflected in the stability of the error rate as shown in the figure 4.

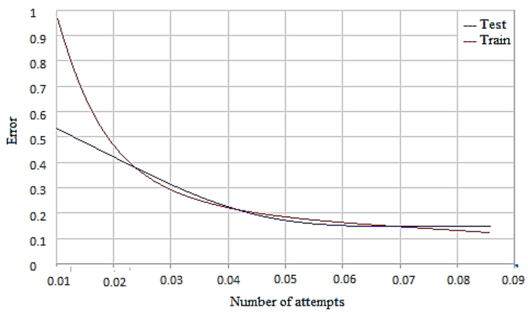

Figure 4. Error regression under training and testing.

The Large databases with reliable error reporting are required to develop practical machine learning approaches that aid in the reduction of diagnostic errors. Many adverse drug reactions can be avoided when persons in danger are appropriately recognized prior to medication delivery or prescription. AI has mostly assisted in VTE identification by analyzing diagnostics or radiologic reports. ML may blend traditional patient data (like medical history and lab outcomes) with unique data (like bioactivity of genomic polymorphisms) to generate customized ADE hazard maps and therapy recommendations to help with patient care decisions. Patients between the ages of 0 and 10 are clearly the most likely to have machine learning, with an 80% chance, while those between the ages of 11 and 30 are in the vital zone. Because we work in the healthcare industry, the FS is approximately 50%, hence the absence of a Age feature values is critical for knowledge. Patients over 50, on either hand, have a lower chance. As a result, they may not be in the important decision-making area. Table (1) shows the success rate of the proposal relied on the age.

Table 2. Success rate of the proposal according to age groups.

The proposed algorithm	*age groups*	*success rate*
CNN+ SAA	Children (0-14)	71%
	young adults (15-30)	87%
	middle-aged adults (31- 45)	83%
	old-age adults above 46	58%

4. Conclusion

Intelligence can considerably improve patient safety. Transparent population-based datasets will be required to create robust and egalitarian models, as well. Patients' safety is the primary goal of AI, and businesses must set up, support, and iterate clinical, group, and system procedures for AI to be effective. The research found that the proposed sensitivity feature analysis algorithm can evaluate feature sensitivity within features. When a feature value is absent, the feature sensitivity score helps clinicians make accurate decisions and direct them to real and dangerous alerts. In medicine, we need to be precise. Clinicians' decisions are a major factor in deciding the cost and quality of health services. Decisions help determine prevention strategies, assessment, testing, and treatment options. The convolutional neural network is the best structure for training the given big data. According to the test, we notes the following:-

- Avoid wasting the use of medical materials and equipment.
- Manual analysis to discover hidden relationships between the given data decreased to 70%.
- Reducing waiting times and delaying the patient's turn to suit his situation.
-After testing and trying the proposal, it achieved 81% accuracy in the diagnosis, and it is worth noting that the accuracy rate of the proposal is closely related to the accuracy of the given data.
- Determining the risks of medicines by 68% for the tested samples.
-It is preferable to use interactive machine learning with data mining to reach the best results.
- In health care cases, there is ambiguity in information compared to other cases of deep learning such as natural language processing.

Acknowledgements

We would like to express our gratitude to the management and people at the hospital for giving us with the dataset, as well as the medical staff and the statistics department. The Ministry of Health (MOH) was also contacted for ethical consent to utilize the dataset in this study.

Author contributions

Sarah built the idea of the research and laid down the basic lines of the proposal and wrote part of the research and worked on programming the proposal in the C-Sharp language and extracting the results of training and testing the proposal and analyze it. **Shaymaa** adopted the process of collecting samples, reviewing the health departments, writing down the opinions of the medical staff and patients who participated in the test, in addition to writing part of the research. **Maysoon** did an extensive electronic search on similar research, test the proposal, make arithmetic modifications, in addition to writing part of the research; all authors had approved the final version.

Conflicts of interest

The authors declare no conflicts of interest.

References

D. . Biau, "In Brief: Standard Deviation and Standard Error," *Clinical Orthopaedics and Related Research,* vol. 469, no. 9, pp. 2661-2664, 2011.

D. F. Sittig and H. . Singh, "Electronic Health Records and National Patient-Safety Goals," The New England Journal of Medicine, vol. 367, no. 19, pp. 1854-1860, 2012.

"The State Hospital," , . [Online]. Available: http://www.tsh.scot.nhs.uk. [Accessed 16 9 2021].

Ahmad, P., Qamar, S., & Rizvi, S. Q. A. ,"Techniques of data mining in healthcare: a review. nternational Journal of Computer Applications"(2015).

Fujisawa Y, Otomo Y, Ogata Y, Nakamura Y, Fujita R, Ishitsuka Y, Watanabe R, Okiyama N, Ohara K, Fujimoto M. Deep-learning-based, computer-aided classifier developed with a small dataset of clinical images surpasses board-certified dermatologists in skin tumour diagnosis. British Journal of Dermatology. 2019; 180(2): 373-381.

Aditi Bansal and Priyanka Ghare, "Healthcare Data Analysis using Dynamic Slot Allocation in Hadoop" in Proc. Conf. International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-3 Issue-5, November 2014.

Aditi Bansal and Priyanka Ghare, "Healthcare Data Analysis using Dynamic Slot Allocation in Hadoop" in Proc. Conf. International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-3 Issue-5, November 2014.

D. Wendell, The Ethics of Clinical Research, ed., vol. , E. N. Zalta, Ed., , : Metaphysics Research Lab, Stanford University, , p.

Wang, Fei, Anita Preininger. AI in health: State of the art, challenges, and future directions. Yearbook of medical informatics. 2019;28(1):16-26.DOI:10.1055/s-0039-1677908

Available: https://www.healthcareitnews.com/ai-powered-healthcare/ai-tappedalleviate-postpartumdepressiondeveloping-regions , May 2, 2019.

"Healthcare-Associated Infections (HAI) data," , . [Online]. Available: https://www.medicare.gov/hospitalcompare/Data/Healthcare-Associated-Infections.html. [Accessed 16 9 2021].

J. M. Currier, J. M. Currier, J. M. Holland, K. D. Drescher and D. W. Foy, "Initial psychometric evaluation of the Moral Injury Questionnaire--Military version.," Clinical Psychology & Psychotherapy, vol. 22, no. 1, pp. 54-63, 2015.

A. . Cavallo, "Scraped Data and Sticky Prices," The Review of Economics and Statistics, vol. 100, no. 1, pp. 105-119, 2016.

H. . Schwellnus and H. . Carnahan, "Peer-coaching with health care professionals: What is the current status of the literature and what are the key components necessary in peer-coaching? A scoping review," Medical Teacher, vol. 36, no. 1, pp. 38-46, 2014.

F. . Niu, C. . Zhang, C. . Ré and J. W. Shavlik, "Elementary: Large-Scale Knowledge-Base Construction via Machine Learning and Statistical Inference," International Journal on Semantic Web and Information Systems, vol. 8, no. 3, pp. 42-73, 2012.

McGhin, T., Choo, K. K. R., Liu, C. Z., & He, D., "Blockchain in healthcare applications: Research challenges and opportunities", Journal of Network and Computer Applications, (2019),135, 62-75.

A. T. Ali and K. A. Adam, "A Multi-Machine Power System Stabilizer Using Fuzzy Logic Controller," , 2012. [Online]: Available: http://ijceronline.com/papers/vol2_issue6/f026028032.pdf. [Accessed 17 9 2021].

M.L. Kothari, Ravi Segal, Bhushan K. Ghodki. "Adaptive conventional power system stabilizer based on artificial neural network" Power Electronics, Derives and Energy Systems, vol.2, Jan 1996, pp 1072–1077

Stingone JA, Pandey OP, Claudio L, Pandey G Environ Pollut "Using machine learning to identify air pollution exposure profiles associated with early cognitive skills among U.S. children".. 2017 Nov; 230:730-740.

Murali N, Sivakumaran N.," Artificial Intelligence in Healthcare", A Review; 2018.

A. G. D. o. Health, "About the Diagnostic Imaging Accreditation Scheme (DIAS) Standards fact sheet," , . [Online]. Available: http://www.health.gov.au/internet/main/publishing.nsf/Content/diagnosticimaging-accred2. [Accessed 17 9 2021].

K. R. Foster, R. . Koprowski and J. D. Skufca, "Machine learning, medical diagnosis, and biomedical engineering research - commentary," Biomedical Engineering Online, vol. 13, no. 1, pp. 94-94, 2014.

J. R. Robbins and A. I. Bakardjiev, "Pathogens and the placental fortress," Current Opinion in Microbiology, vol. 15, no. 1, pp. 36-43, 2012.

K. A. Wetterstrand, "DNA Sequencing Costs: Data from the NHGRI Large-Scale Genome Sequencing Program," , . [Online]. Available: http://www.genome.gov/sequencingcosts. [Accessed 17 9 2021].

R. J. Sweeney, Y. . Zhang, Q. . An and P. H. Thakur, "METHODS AND APPARATUS FOR DETECTING HEART FAILURE," , 2014. [Online]. Available: https://patents.google.com/patent/ep2999395a1/en. [Accessed 17 9 2021].

"Hypoglycemia (Low Blood Glucose)," , . [Online]. Available: http://www.diabetes.org/living-with-diabetes/treatment-and-care/blood-glucose-control/hypoglycemia-low-blood.html. [Accessed 17 9 2021].

D Burnet,"Numbers of Approaches of Data Mining Are Being Used By Various Researchers in Healthcare Sector", Tennessee Research International of Social Sciences, Vol. 1, Issue 1, January-June 2019.

Kim MS., Han SS, Lim W., Park GH, Park I, Chang SE. Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm. J Invest Dermatol. 2018 Jul;138(7): 1529-1538.