Radius: Journal of Science and Technology
Vol. 2, Issue 1, (2025) 251003 (8pp)

Debunking AI and Healthcare: The Connection

Krishna Vashistha[1] | Jayant Kataria[2] | Ekta Puri[1]

[1]St. Xavier College, Jaipur, Rajasthan (India)
[2]Indian Institute of Technology, Jodhpur, Rajasthan (India)

*Corresponding author: ekta@sxcjpr.edu.in

Abstract: Artificial intelligence has transformed the healthcare industry so much in terms of diagnosis, treatment, and patient care over the years. This paper outlines how AI was introduced into the healthcare industry from its humble beginnings with early tools that date back to the 20th century like decision support systems and various diagnostic programs. Recently, AI has affected the precision medicine area significantly with deep learning models, used for accurate analysis of medical images. AI also improved personalized treatments, particularly cancer therapies and drug development. Systems powered by AI also assist recovering patients. This paper explores how AI has evolved and its growing role in healthcare today.

Keywords: Artificial Intelligence; Healthcare; Precision medicine; Exoskeleton; Convolutional Neural Networks

Introduction

AI has been prospering for years leaving its mark on every field be it the automation of manufacturing, an improvement in operational efficiency in medicine, or even the management of records in more commercial fields. But what some people who use A.I., on a regular basis, do not realize is that A.I. which has seen quite an uphill trend in recent times has been developing and improving through many last decades. If you were told that A.I. was first coined in the 1960s it's possible that you might ponder how it was used back then.

One field that A.I. has advanced quite a bit is the medical field, be it breaking complex medicinal equations to proportioning the amount of medicine intake, A.I. does it all, A.I. Virtual chat-bots and assistants can answer medical questions and provide guidance on self-care. And so, the purpose of this paper is to debunk the origins of A.I. in the field of healthcare. “When” and “Why” was A.I. a breakthrough by medical standards? What were the factors which led to its boom in the medical industry? Through this research we hope to trace the path on which A.I. set its foot and how it replaced human efforts in the current world of healthcare. This would result in us looking into how A.I. has affected the precision medicine area and eventually highlighting the evolution of A.I. in the healthcare sector. And in the end, we would also be looking into the future possibilities of A.I. in the healthcare sector.

Some pre-written research can be used to give us a general idea of the current state of research on our topic, for example, “Artificial Intelligence and Healthcare: A Journey through History” gives us a general introduction to how A.I. has changed and evolved over the years in the given field and establishes how the telemedicine and personalization of medicine has improved as the years have gone by.

https://doi.org/10.5281/zenodo.14945533
Received: 22 January 2025 | **Revised:** 18 February 2025
Accepted: 20 February 2025 | **Published Online:** 28 February 2025

The issues "Artificial Intelligence in Medicine AIME 2009" and "Artificial Intelligence in Medicine AIME 2011" provide a general idea of the state of the integration of A.I. in medicine around 2010 which was basically the integration of A.I. in education through the use of various simulations to help students get a better graphical understanding of the biological bodies. From "Artificial Intelligence Applications in the Intensive Care Unit" it was established that A.I. was also being integrated into non-educational fields by using it for processes like data mining and image classification. "Artificial Intelligence in Medical Diagnosis" and "Artificial Intelligence in Radiology" give an insight into the very early use cases of A.I. through the use of decision support systems backed up by data-heavy databases to support the physicians of that time.

Research Methodology

The nature of the research conducted by us is majorly descriptive and aims at helping to get a better understanding of the roots of Artificial Intelligence in the healthcare field. We adopted a top-down design model of research for which we needed to analyze a lot of samples of different A.I. methodologies over the years and understand how the workings of said methods had changed. The data collection process is done using the PubMed database of the National Library of Medicine available under the National Health Institute (N.I.H.). The sampling of the data was done with respect to the date it was published to help us move through the years better. We categorized the data into different parts to make it easier to understand. Then it was further simplified into various plots and tabulations for the ease of presentation for the general public.

Results and Discussion

Fig. 1 represents a graph taking into consideration the amount of research and use cases of Artificial Intelligence on the y-axis and sorted by year on the x-axis. We can see that there is a significant rise in the slope at the 2000 and 2015 marks. Hence, it makes sense to review the works in three major periods, namely before the 2000s, 2000 to 2015, and 2015 to the present.

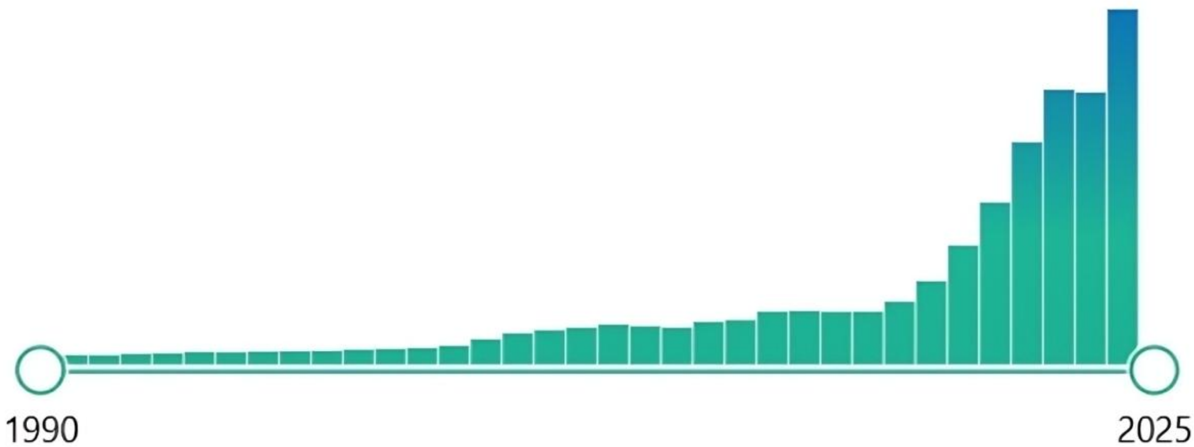

Fig. 1: Number of A.I. based research undertaken each year (**Source:** *pubmed.ncbi.nlm.nih.gov*)

The Current State of A.I.: In the last decade, there was a lot of shift towards individualization of medicine, meaning that the daily lifestyle, genetic makeup, living environment, and many more factors were being taken into consideration before the medicine was prescribed to a patient. And to no surprise, all of this data is managed by A.I. in real-time so that any change in the patient's condition can be catered to as soon as possible. Most of this is possible through the use of deep learning models like Convolutional Neural Networks (CNNs) which are also used to identify medical conditions from radiographic images, CT scans, MRIs, and ultrasounds with inhuman accuracy.

To get a better understanding of how CNNs achieve this, we can take a look at traditional neural networks like artificial neural networks (ANN). The difference between CNN and ANN is made clear if we take a look at the hidden layers of each. The hidden layer of a CNN model is composed further of three layers, namely the convolutional layer, sub-sampling layer (pooling) and fully-connected layer (Fig. 2).

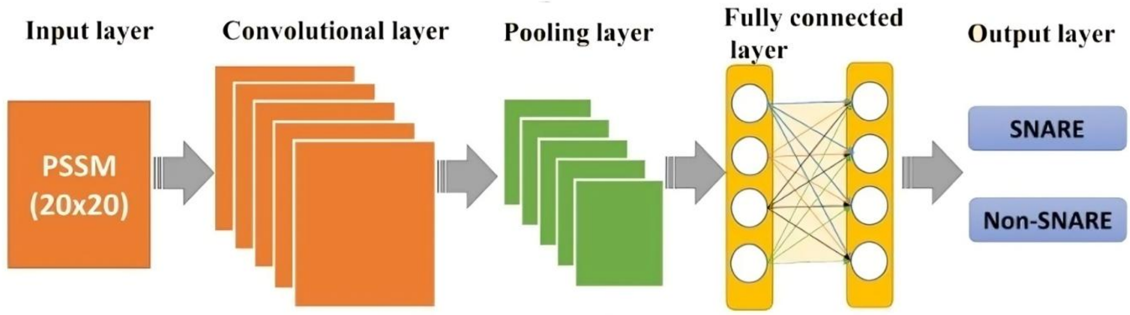

Fig. 2: A simple CNN flowchart (**Source:** *peerj.com*)

As we can see, each part of a hidden layer has a desiccated function that they work on exclusively to prove results. We can compare this to ANN in which there are many hidden layers that work together collectively to generate an output. This will be later explored in detail when we talk about ANNs to help give a better visualization of the improvement in the algorithm.

Table 1: Data representing the field of use for various A.I. algorithms

Study/Case	A.I. Technique used
Dermatologist-level classification of skin cancer	Convolutional Neural Networks (CNN)
Patient-centered pain care using A.I. and mobile health tools	Reinforcement Learning
Improved referral process for specialized medical procedures like epilepsy surgery	Natural Language Processing
Delineating ulcerative colitis from Crohn's disease	Guided Image Filtering (GIF)
Chemotherapy selection for gastric cancer	Random Forest Machine Learning Model

Table 1 **[1]** depicts the various use cases of A.I. algorithms in many recent medical fields from which we can conclude just how much A.I. has been integrated into the healthcare field. And there are even more use cases of A.I. that are currently being developed. Let's take Precision Medicine, where A.I. was being used to evaluate the genetic data of an individual to check for vulnerability and resistance to specific diseases so that that treatment can be more effective. One such example of precision medicine is the challenge regarding overcoming T-cell exhaustion **[2]** in infection and cancer in which A.I. is used to monitor the T-cells in a human body and determine the precise timing of using the inhibitors of Programmed Cell Death to keep the T-cells from dissolving. Such inhibiting drugs are also developed through the use of A.I. in which the slowest steps of a biological interaction are identified and stimulated as per the need to get the desirable quality of the drug and the speed of production. Another field that has come up in the past few years is the use of A.I.-controlled exoskeletons to help regain movement after a long period of limb inactivity that can be due to, but not limited to, stroke patients. In this, an exoskeleton is attached to the affected limb and controlled through another system to help support and recuperate the flow of commands into the limb without any interruption to the patient.

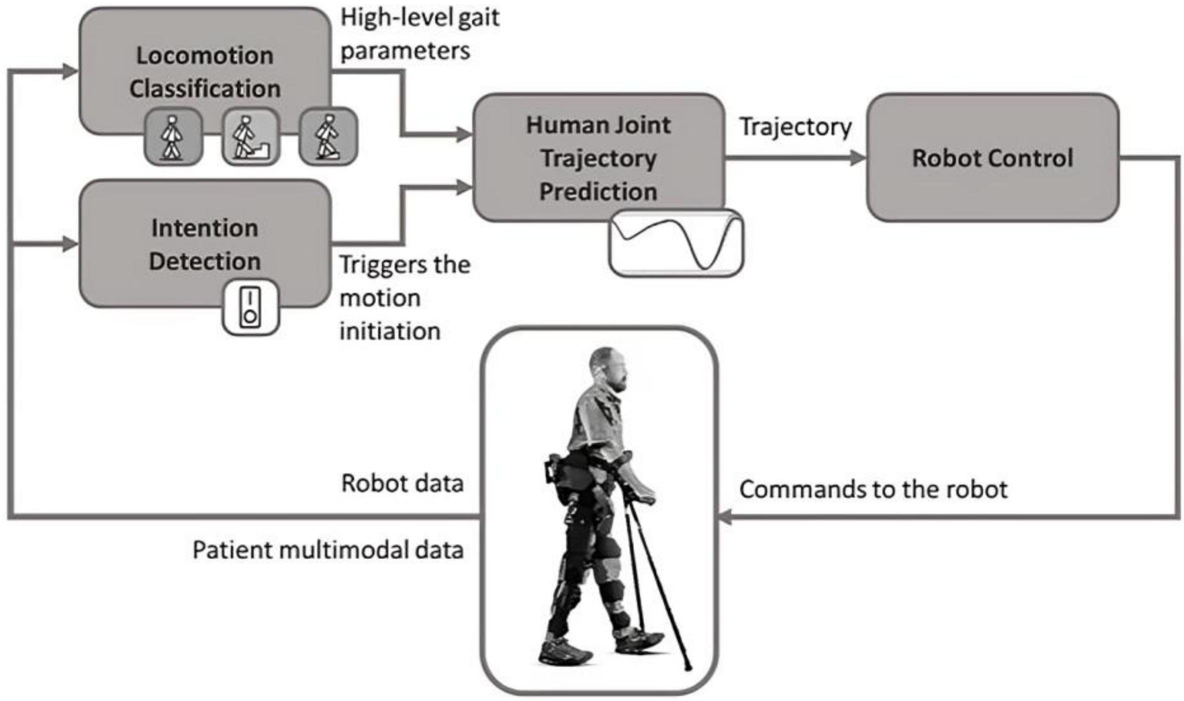

Fig. 3: Command flow of a prediction based limb exoskeleton (**Source:** *ncbi.nlm.nih.gov*)

Fig. 3 **[3]** represents how A.I. can be used in the working of the said exoskeleton by using predictive analysis to evaluate and support the motion of a patient in the required direction of limb motion.

The Early 2000s [4, 5]: The shift from the 1990s to the 2000s saw the rise of machine learning techniques, which marked a shift from professional expert systems to a more data-driven approach. It was at this time that Artificial Neural Networks (ANNs) were being developed and customized for medical image classification and identification (Fig. 4). Contrasting to the earlier CNN discussion, we can get the picture that the ANNs only work on a feedback loop and contrast it to the 3 sub-layers of a CNN (Fig. 2) that work on both sampling, elimination and feedback. If we consider the input here to be the symptoms shown by a patient, the output will be the diagnosis. While at one hand, such algorithms were starting to be used to develop even more complex uses that we see in the current day, there were already many other existing use cases of A.I. in those days, such as text-based mining of biomedical literature, virtual reality dental simulator with a haptic interface and even CBR with the patients.

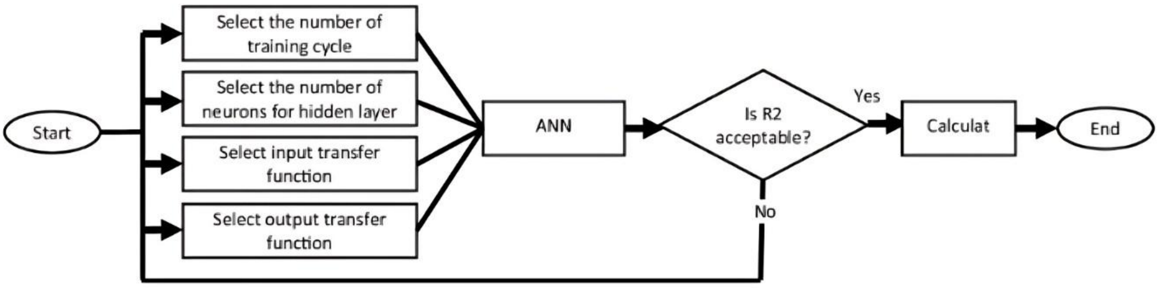

Fig. 4: A simple ANN flowchart (**Source:** *researchgate.net*)

One such simple yet tedious use case of A.I. in the 2000s was in an Intensive Care Unit (ICU) **[6]** be it intelligent assistants to clinicians or constant monitoring of electronic data streams for important trends. An ICU is a place with a lot of data to be managed, some specific sources of information include physiologic data, laboratory data, outcomes information and prediction systems, critical care research, and quality and cost information, while simultaneously catering to the patients so automating the data in an ICU will free up a lot of hands that can be used in other necessary spaces. And so, CBR-based assistants were introduced that simplified the work schedule of doctors by simplifying the stream of data that needed to be monitored throughout the day and put out alerts in case of emergencies. The bedside devices could now also be remotely monitored and controlled significantly reducing the response timing and disregarding the need to always be present on site.

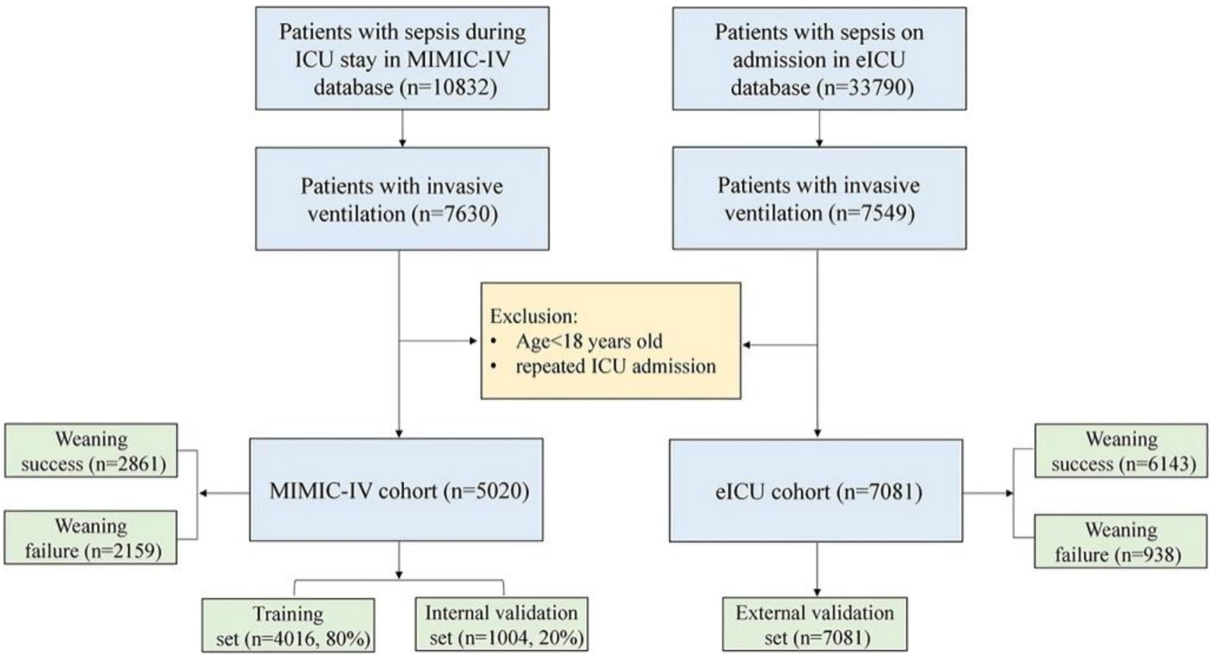

Fig. 5: Dataflow in an eICU (**Source:** *eicu-crd.mit.edu*)

Fig. 5 can be used to realize just how data-heavy an ICU can be and the scope of automation converting it to an eICU. It was due to such considerations that we saw the rise of eICUs in the current day and age. In order to contextualize biases, heterogeneity, and validation issues in medical AI research, benchmark datasets like MIMIC-III (for ICU data), ImageNet (for medical imaging), and PubMed (for medical NLP) are essential. A thorough ICU dataset called MIMIC-III helps create fairer healthcare models by highlighting clinical irregularities, demographic differences, and potential biases in patient outcomes. Although ImageNet, a popular tool for picture classification, offers a basis for medical imaging applications, it may induce biases because of variations in data representation and distribution among populations. NLP applications in healthcare are supported by PubMed, a substantial archive of biological literature; yet, there are issues with linguistic diversity, domain specialization, and possible disinformation. These benchmark datasets allow researchers to evaluate and reduce biases, and improve the generalizability of their models.

To the Origins: In the 20th century, A.I. was a term much scarce than we have seen in any of the two time periods we've talked about till now, yet there were still places where we could see the very early ideas of artificial intelligence come into play. These included decision support systems, medical informatics and diagnosis.

In the field of Radiology **[7]**, there were algorithms that helped physicians choose appropriate radiologic procedures and choose diagnoses accurately. These systems were based on rule-based techniques, reasoning, artificial neural networks, hypertext, Bayesian networks, and case-based reasoning, some of which were still in the early stages of development but still accurate enough to be helpful in the process. It was also during this period that diagnostic consultants like INTERNIST/CADUCEUS **[8]** were put to use.

Table 2: Data taken from the accuracy testing done on the INTERNIST-I

	Number of Instances		
Category	**INTERNIST-I**	**Clinicians**	**Discussants**
Definitive, correct	17	23	29
Tentative, correct	8	5	6
Failed to make diag.	18	15	8
Definitive, incorrect	5	8	11
Tentative, incorrect	6	5	2
Total incorrect	11	13	13
Total errors	29	28	21

INTERNIST-I was an experimental program for computer-assisted diagnosis in general internal medicine that used a heuristic computer program. Given a patient's initial history, results of a physical examination, or laboratory findings, INTERNIST-I was designed to make multiple complex diagnoses. INTERNIST-I was also the first model to be applied to a single broad medical term namely general internal medicine and was accompanied by a well-stocked knowledge base which it used to give out its diagnosis. However, since the classification procedures that were needed to work well with the database were still in development, INTERNIST-I was still quite inaccurate.

Table 2 explains the accuracy test of Internist-I A.I. program. The accuracy testing of INTERNIST-I compared its diagnostic performance with clinicians and discussants. INTERNIST-I made 17 definitive correct diagnoses, fewer than both clinicians (23) and discussants (29). It also failed to diagnose 18 cases, more than the others. While its total incorrect diagnoses (11) were slightly lower than clinicians (13) and discussants (13), its total errors (29) were the highest, indicating room for improvement in diagnostic precision. We can observe that while INTERNIST-I still made the most errors out of the three ways that are being used for diagnosis, it does not lag behind too much considering the limiting technology of the time. We can also see that the errors made by the algorithm mostly lie in the fact that it failed to reach a diagnosis in most cases, it was due to an indecision being prioritized instead of an incorrect diagnosis.

Another prominent A.I. model of that time was an ANN model for the diagnosis of lower back pain **[9]** which worked by examining the pain drawing of low back pain patients and classifying it into one of five clinically significant categories. The model was a really simple image classification model but was quite effective practically and simplified the treatment process quite a bit.

Machine Learning Models

It may be interpreted using SHAP (Shapley Additive Explanations) values, which give each feature a significance score that indicates how much it contributed to a certain prediction. SHAP evaluates various feature combinations based on game theory to ascertain how they affect the model's output. In the medical field, SHAP highlights important variables like age, test findings, or symptoms to assist explain why a model indicates a patient has a high risk of illness. It increases the transparency of AI-driven judgments by offering both local (individual prediction) and global (overall model behavior) interpretability.

Deep learning algorithms, especially convolutional neural networks (CNNs), frequently employ Grad-CAM (Gradient-weighted Class Activation Mapping) to show which aspects of an image affected the model's prediction. In order to contextualize biases, heterogeneity, and validation issues in medical AI research, benchmark datasets like MIMIC-III (for ICU data), ImageNet (for medical imaging), and PubMed (for medical NLP) are essential.

A thorough ICU dataset called MIMIC-III helps create more fair healthcare models by highlighting clinical irregularities, demographic differences, and potential biases in patient outcomes. Although ImageNet, a popular tool for picture classification, offers a basis for medical imaging applications, it may induce biases because of variations in data representation and distribution among populations. NLP applications in healthcare are supported by PubMed, a substantial archive of biological literature; yet, there are issues with linguistic diversity, domain specialization, and possible disinformation. These benchmark datasets allow researchers to evaluate and reduce biases, and improve the generalizability of their models.

Limitations:

a) **Data Limitations:** The data availability may affect generalizability.
b) **Methodological Constraints:** Research design, measurement tools, and biases can impact accuracy.
c) **Theoretical and Scope Restrictions:** Findings are limited to a specific context or framework.
d) **Time, Resource, and Ethical Constraints:** These constraints also affect the outcomes.

Future Possibilities: We've seen how Artificial intelligence at the moment is really accurate and ever-improving. However, a very major drawback to using A.I. systems in a medical environment instead of human interactions is the lack of human empathy in A.I. leading to a lot of ethical issues hindering the further integration of A.I. to make healthcare fully automated. So, a future possibility for the betterment of A.I. is improvement in algorithms so that A.I. can take empathy into consideration with its logical side to generate better results leading to an accurate and ethical environment in the healthcare department.

Conclusion

Studying through various readily available literature, we were able to date back the origin of AI in the medical field to the late 20th century with simple but effective diagnostic systems like INTERNIST and creating more streamlined data flow systems for ease of use. These very simple algorithms laid the framework for the much more complex algorithms that are used nowadays and continue to help develop the healthcare systems better and better each day be it the development of new inhibitor drugs for treatment or helping people recuperate.

References

1. Hirani, R., Noruzi, K., Khuram, H., Hussaini, A. S., Aifuwa, E. I., Ely, K. E., ... & Etienne, M. (2024). Artificial intelligence and healthcare: A journey through history, present innovations, and future possibilities. *Life, 14*, 557. https://doi.org/10.3390/life14050557
2. Pauken, K. E., & Wherry, E. J. (2015). Overcoming T cell exhaustion in infection and cancer. *Trends in Immunology, 36*, 265–276. https://doi.org/10.1016/j.it.2015.02.008
3. Shahar, Y., & Combi, C. (2011). Artificial Intelligence in Medicine AIME 2009. *Artificial Intelligence in Medicine*, *52*, 57-58. https://doi.org/10.1016/j.artmed.2011.04.006
4. Peleg, M., Lavrač, N., & Combi, C. (2011). *Artificial intelligence in medicine: 3th Conference on Artificial Intelligence in Medicine, AIME 2011, Bled, Slovenia, July 2-6, 2011, Proceedings*. Springer. https://doi.org/10.1007/978-3-642-22218-4
5. Banks, G. (1986). Artificial intelligence in medical diagnosis: The INTERNIST/CADUCEUS approach. *Critical Reviews in Medical Informatics, 1*, 23-54.
6. Hanson III, C. W., & Marshall, B. E. (2001). Artificial intelligence applications in the intensive care unit. *Critical Care Medicine, 29*, 427-435. https://doi.org/10.1097/00003246-200102000-00038
7. Kahn Jr, C. E. (1994). Artificial intelligence in radiology: Decision support systems. *Radiographics, 14*, 849-861. https://doi.org/10.1148/radiographics.14.4.7938772
8. Mann 3rd, N. H., & Brown, M. D. (1991). Artificial intelligence in the diagnosis of low back pain. *The Orthopedic Clinics of North America, 22*, 303-314. https://doi.org/10.1016/S0030-5898(20)31655-2
9. Nguyen, K. T., & Nguyen, H. D. (2022, July). Designing the robot arms for upper limb recovery movements of post-stroke patients. In *2022 6th International Conference on Green Technology and Sustainable Development (GTSD)*. IEEE, 607-611. https://doi.org/10.1109/GTSD54989.2022.9989183